Disentangling the Evolution of a Developmental Switch Network that Regulates Phenotypic Plasticity in a Morphological Novelty: a Nematode Evo-Devo Perspective

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von
Tobias Theska
aus Grevesmühlen

Tübingen
2023

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	21.12.2023
Dekan:	Prof. Dr. Thilo Stehle
1. Berichterstatter/-in:	Prof. Dr. Ralf J. Sommer
2. Berichterstatter/-in:	Prof. Dr. Alfred Nordheim

Acknowledgements

First, I would like to thank my supervisor, Ralf J. Sommer, for not only providing me with the opportunity, but also fully endorsing me, to explore my passions from morphology to developmental genetics in such a stunning group of organisms. I deeply enjoyed my endeavors to explore the molecular processes that make these "endless forms most beautiful" in our favorite worms.

Next, a huge thank you goes to the many people in the lab that helped me to develop into the scientist that I wanted to be, especially Ziduan Han and Bogdan Sieriebriennikov. I deeply appreciate the many hours you guys invested to get me up to speed and ready to go explore worm genetics. Similar things can be said about Mike Werner, Jim Lightfoot, Christian Rödelsperger, and Adrian Streit who always provided open ears and helping hands.

I'm also greatly indebted to our wonderful technicians Metta Riebesell, Hanh Witte, Walli Röseler, Tobias Loschko, Chris Weiler, and Heike Haussmann for the many tips and tricks they taught me over the years, ever so patiently.

Thanks to Matthias Herrmann and Natsumi-Sensei for showing me that nematodes are more than just tiny tubes filled with copious amounts of genes, but also fantastic little beasts of great diversity. I really enjoyed our expeditions into the Schönbuch and our collaborations.

Much of my work would not have been possible without my lovely friends, who always kept me sane and jolly throughout my time here in Tübingen: Devansh R. Sharma, Catrin Bitter, and Martin Karn, you guys are the best!

I also want to thank my family, especially my parents, for their constant care and endorsement. You made my journey to become a biologist possible in the first place.

Lastly, I wholeheartedly want to thank you, Tess, for your unwavering love and support over all these years. I am in awe with your positivity, your passion, and your perspicacity, and I could not be happier that we experienced this journey together. Just like my beloved nuclear receptors can only do what they are supposed to do once they find and bind their ligand, I was only able to get this far and do all this work because of you. You are my ligand.

Table of contents

I. Abstract

Phenotypic plasticity describes the ability of a single genotype to produce different phenotypes in response to distinct environmental conditions. It has been proposed to be a diving force of morphological evolution and a trailblazer for the emergence of individualized novelties - new traits which progressively evolved from one or more preexisting ones and acquired a new qualitative dimension. Yet, in order to reconstruct such character transformations and the facilitator role phenotypic plasticity may have played, we need model systems that allow comparative morphological and genetic investigations in organisms that display plastic novelties, as well as in their out-group relatives that do not exhibit these novel traits. Nematodes belonging to the Diplogastridae, such as *Pristionchus pacificus*, and their paraphyletic out-group "Rhabditidae", such as *Caenorhabditis elegans*, fulfill these requirements. "Rhabditids" have tube-shaped mouths with immovable cuticular feeding structures, called flaps, that act as valves preventing the regurgitation of food. In contrast, Diplogastridae, which emerged from "rhabditids", evolved moveable cuticular teeth from these flaps. Intriguingly, the flap-to-tooth transformation that accompanied the "rhabditid"-to-diplogastrid transition was tightly linked to phenotypic plasticity, and most Diplogastridae display stable polyphenisms in their tooth-like feeding structures. Previous research identified that a modular switch network underlies the development of these plastic feeding structures in *P. pacificus.* In my thesis work, I aimed to expand the knowledge of the switch network by identifying molecular players involved in the production of the cuticular teeth, and to reconstruct its evolutionary history by studying the functions of its conserved constituents in the "rhabditid" *C. elegans*, which has non-plastic flaps. First, to facilitate quantitative morphological studies in nematodes, I established a computational pipeline that combines landmark-based geometric morphometrics with unsupervised clustering methods. Using this pipeline, I demonstrated that DPY-6 is required for proper feeding-structure development in "rhabditids" and diplogastrids, confirming this mucin-type protein as the first known structural component of the nematode mouth. Furthermore, I discovered that feeding-structure development, which is known to be controlled by the two transcription factors NHR-1 and NHR-40 in *P. pacificus*, is regulated by a different genetic module in *C. elegans.* I also revealed that NHR-10, unlike its sister paralogs NHR-1 and NHR-40, is not involved in mouth morphogenesis, but required for starvation resistance in *P. pacificus.* Together, this research suggests that NHR-1 and NHR-40 were co-opted for mouth morphogenesis during the evolution of novel teeth in Diplogastridae. Altogether, my work highlights the complexities of the developmental genetic processes which accompany the evolution of new traits.

II. Zusammenfassung

Phänotypische Plastizität beschreibt die Fähigkeit eines Genotyps, in Reaktion auf unterschiedliche Umwelteinflüsse, verschiedene Phänotypen zu generieren. Dieser Prozess gilt als eine der treibenden Kräfte in der morphologischen Evolution und als Wegbereiter für die Entstehung von individualisierten Neuheiten. Letztere sind evolutionär neue anatomische Strukturen, die aus bereits bestehenden Strukturen hervorgingen und eine neue qualitative Dimension erlangten. Um derartige Merkmalstransformationen und die Rolle, die phänotypische Plastizität in diesen Prozessen spielt, studieren zu können, brauchen wir jedoch geeignete Modellsysteme. Idealerweise handelt es sich dabei sowohl um eine Gruppe von Organismen deren Mitglieder eine phänotypisch plastische Neuheit aufweisen, als auch um ihre Außengruppen-Verwandten ohne diese Neuheit. Des Weiteren, müssen quantitativ-morphologische und funktionell-genetische Untersuchungen in diesen Organismen möglich sein, um die molekularen Grundlagen dieser Entwicklungsprozesse aufschlüsseln zu können. Fadenwürmer, die zu den "Rhabditiden" und den Diplogastridae gehören, erfüllen all diese Kriterien. "Rhabditiden" wie *Caenorhabditis elegans* besitzen röhrenförmige Münder, an deren Basis einfache kutikuläre Klappen zu finden sind, welche die Nahrung daran hindern aus dem Rachen zurück in die Mundhöhle zu gelangen. Diplogastridae wie *Pristionchus pacificus* hingegen, welche aus den paraphyletischen "Rhabditiden" hervorgingen, haben aus den "Rhabditiden"-artigen Klappen bewegliche Zähnen evolviert, welche sie zum Fangen von Beutetieren nutzen können. Interessanterweise war diese Merkmalstransformation eng mit der Entstehung von phänotypischer Plastizität in den Mundstrukturen dieser Tiere verbunden. Frühere Studien zeigten, dass der Entwicklung der plastischen Mundstrukturen von *P. pacificus* ein modulares genetisches Netzwerk zugrunde liegt. In meiner Dissertation zielte ich darauf ab, das Grundwissen über dieses Netzwerk zu erweitern und dessen evolutionäre Geschichte zu rekonstruieren, indem ich die Funktionen konservierter Gene in *C. elegans* untersuchte. Um quantitativ-morphologischen Studien zu ermöglichen, habe ich zuerst eine Pipeline entwickelt, die Methoden der geometrischen Morphometrie mit verschiedenen Clusteringprozessen verbindet. Mithilfe dieser Pipeline konnte ich zeigen, dass DPY-6 für die Entwicklung der Mundstrukturen erforderlich ist, was dieses Protein als erste bekannte Strukturkomponente des Mundes bestätigte. Darüber hinaus entdeckte ich, dass die Mund-Morphogenese, die in *P. pacificus* von den Transkriptionsfaktoren NHR-1 und NHR-40 kontrolliert wird, in *C. elegans* durch ein anderes genetisches Modul reguliert werden muss. Dazukommend konnte ich demonstrieren, dass NHR-10 (im Gegensatz zu seinen Schwesterparalogen NHR-1 und NHR-40) nicht an der Kontrolle der Mund-Morphogenese beteiligt ist, dafür aber für das Überleben von Hungerstress in *P. pacificus* benötigt wird. Meine Ergebnisse deuten darauf hin, dass NHR-1 und NHR-40 ihre Funktionen in der Mundentwicklung während der Evolution der Zähne erwarben. Insgesamt unterstreicht meine Dissertation die Komplexität der entwicklungs-genetischen Prozesse, welcher der Evolution neuer anatomischer Merkmale zugrunde liegen.

III. List of publications

Rödelsperger C, Athanasouli M, Lenuzzi M, **Theska T,** Sun S, Dardiry M, Wighard S, Hu W, Sharma DR, Han Z. Crowdsourcing and the feasibility of manual gene annotation: A pilot study in the nematode *Pristionchus pacificus.* **Scientific Reports. 2019**; 9:1-9.

Theska T, Sieriebriennikov B, Wighard SS, Werner MS, Sommer RJ. Geometric morphometrics of microscopic animals as exemplified by model nematodes. **Nature Protocols. 2020**; 15:2611–2644.

Kanzaki N, Herrmann M, Weiler C, Röseler W, **Theska T**, Berger J, Rödelsperger C, Sommer RJ. Nine new *Pristionchus* (Nematoda: Diplogastridae) species from China. **Zootaxa. 2021**; 4943:001-066.

Sun S*, **Theska T***, Witte H, Ragsdale EJ, Sommer RJ. The oscillating Mucin-type protein DPY-6 has a conserved role in nematode mouth and cuticle formation. **Genetics. 2022**; 220:iyab233. †

* **co-first authors (equal contribution)**

† highlighted by the journal editors

Werner MS, Loschko T, King T, Reich S, **Theska T**, Franz-Wachtel M, Macek B, Sommer RJ. Histone 4 lysine 5/12 acetylation enables developmental plasticity of *Pristionchus* mouth form. **Nature Communications. 2023**;14:2095.

Theska T, Renahan T, Sommer RJ. Starvation resistance in the nematode *Pristionchus pacificus* requires a conserved supplementary nuclear receptor.
bioRxiv: https://doi.org/10.1101/2023.08.21.554071
Submitted to **Zoological Letters (2023).**

Theska T, Sommer RJ. Feeding-structure morphogenesis in "rhabditid" and diplogastrid nematodes is not controlled by a conserved genetic module.

bioRxiv: https://doi.org/10.1101/2023.11.07.565949

Submitted to **Evolution & Development (2023).**

IV. Introduction

1. The challenge of morphological evolution

"NOTHING IN BIOLOGY MAKES SENSE EXCEPT IN THE LIGHT OF EVOLUTION."
- THEODOSIUS DOBZHANSKY (1973)

For the majority of the last century practitioners of organismal biology were deeply divided: there were those who almost exclusively focused on how populations of organisms change during evolution and those who mainly concentrated on how individual organisms change during development [1]. Only by the late 1970s, following great methodological progress in molecular genetics and substantial conceptual advancements in evolutionary thinking [2-4], did these two research programs join hands to give rise to what is today known as "evolutionary developmental biology", or evo-devo [5]. This discipline investigates both the evolution of organismal development and the ways by which developmental processes guide organismal evolution, ultimately aiming to create a mechanistic understanding of phenotypic change and morphological evolution [6,7]. Thus, evo-devo shifted the target of evolutionary inquiry away from the almost monolithic and abstract mathematical equations capturing associations of population-level natural selection and phenotypic transformations that characterized most of 20th century biology [6]. Instead, it re-introduced the organism to the center of attention by experimentally disentangling the developmental processes which actually generate phenotypic diversity and novelty in the first place [5-7].

While evo-devo as we know it today is a young discipline, its intellectual roots run deep [5]. In his magnum opus *Ontogeny and Phylogeny* [3], Stephen Jay Gould reminded us that the thought behind the fundamental tenant of evo-devo - that phenotypic change and organismal development are intricately and inseparably linked on a causal level - can be traced back for millennia [5]! This duality is also reflected in evo-devo's terminology, which constitutes a dense thicket of old concepts intertwined with re-interpreted patterns, re-labeled processes, and newly-discovered mechanisms, which, by now, fills entire compendia [8]. Thus, it seems worthwhile to take a step back and (re-)introduce some evo-devo's fundamental terms.

1.1 The form-function dualism in evolutionary biology

Just as most other scientific disciplines, evolutionary biology has historically seen a deep philosophical schism. At the heart of its conceptual divide sits an ontological distinction between organismal form and function [1,9], and the importance individual researchers assign to either of the two in their pursuit to understand phenotypic transformations. In simple terms, form describes what a trait *is* (i.e. its structure and type), and function describes what a trait *does* (i.e. its use).

One of the historically most relevant debates concerning the primary role of form or function in biological inquiry took place in 1830, when the two famous French naturalists George Cuvier and Étienne Geoffroy Saint-Hilaire fenced an intellectual duel following many years of mutual correspondence. Cuvier was a convinced "functionalist" who argued that the vast diversity of phenotypes among animals is a reflection of the many ways in which they adapted to the environments they live in [9]. Similarities in phenotypes, he suggested, were due to similarities in how animals conformed to comparable environments. Geoffroy was an equally-convinced "structuralist" who argued that the phenotypes of animals are independent from their environmental uses, and that phenotypic similarities would be explained by their "unity of type" [9,10]. He reasoned that animal phenotypes come as modifications from an underlying common "type" and that an unknown natural law defines said structural "type" [9]. Cuvier emerged as the debate winner and thus his arguments for functionalism led the way of biological thinking for the years to come. Yet, amongst the morphologists of that time, structuralism found increasing support [9], especially after Karl Ernst von Baer formulated his empirical rules of animal embryogenesis, which later became known as his "Laws of Development" [9-12]. In comparing representative species of highly-divergent animal taxa, von Baer showed that, starting from a point of high phenotypic similarity, animal embryos continuously diverge phenotypically during their ontogeny. He proposed that general phenotypes (those shared by many animal taxa) develop first, and that more specific phenotypes (those shared by fewer taxa) develop later [12]. These patterns fit nicely with the "types" described by Geoffroy and thus structuralism gained momentum in morphology [9].

From this, we can see that the form-function dualism has affected the way phenotypic traits and their correspondence among organisms, their "sameness" across species, have been understood in evolutionary biology. The functionalist claims that, for a phenotypic trait to be the same in any two species, it must serve the same function.

For the structuralist, the correspondence of a phenotype across two or more species is completely independent of its functional uses. Instead, it is understood as "the same trait" based on its type alone. In 1843, the morphologist Richard Owen formalized this structuralist idea of types underlying the same traits by coining the term "homology". Homologs, he said, are "*the same organ in different organisms under every variety of form and function*" [13]. This conceptualization of sameness clearly emphasizes that phenotypic traits can, and in fact do, remain the same, even if their form or their function diversified across species. Charles Darwin heavily referred to this concept of homology in his *Origin of Species* [14] and put it into an evolutionary context. His fundamental insight - that species evolve by decent with modification via natural selection - directly relates to phenotypes and homology: the sameness of traits in different species, irrespective of any structural or functional modifications, is due to them originating from the same trait in a common ancestor.

Almost paradoxically, even though Owen's structuralist conceptualization of homology and Darwin's theory of evolution complement each other, the form-function dualism was not resolved [9]. This is reinforced by the fact that Darwin, despite his insights into genealogical descent and his understanding of homology and phenotypic transformations, was an enthusiastic functionalist [9]. His most significant contribution lies in the identification of natural selection as the driving force of evolutionary change; more specifically as the process that produces adaptation - the functional fit between the organisms' phenotypes and environments [9,10,14]. Interestingly, while this aspect of his work was celebrated by functionalists, most of the early Darwinians were practicing morphologists (and thus structuralists) who had little interest in studying adaptation [9]. Therefore, the form-function dualism continued to persist in evolutionary biology, essentially unaltered, for the better part of a century. By the 1930s, population genetics was incorporated into standard evolutionary theory, which had acquired a firm mathematical foundation. This integration, the "Modern Synthesis" of evolutionary theory [15], heavily shifted the emphasis towards functionalist interpretations of phenotypic change. Most of the architects of the Modern Synthesis, especially Ernst Mayr, strongly criticized any form of "typological thinking", such as 19th century structuralism, and instead argued in favor of "population thinking", i.e. the notion that the only way to properly study phenotypic transformations is through the lens of population genetics, natural selection, and adaptation [16] - that is through function. Additionally, embryology and developmental biology, disciplines highly

structuralist in nature, were largely excluded from the Modern Synthesis [1,3,5,17], further undermining the credibility of non-functionalist insights. Resulting was a lineage of evolutionary thinking that had become known as "adaptationism". Essentially a rebranding of 19th century functionalism, it describes the view that phenotypic change over geological timespans is best understood by focusing on the functions of phenotypes and their adaptive values [10]. In the late 1970s, Stephen Jay Gould and Richard Lewontin famously pointed out that adaptationism had become a mind-cuffing doctrine [4,18]: both defining individual organismal traits based on the specific functions they serve and explaining them as a product of natural selection optimizing their functional fit to the environment had become blind habits of the evolutionists of that time, leaving little to no space for non-adaptive forces in nature or explanations in evolutionary inquiry [4]. Their critique of adaptationism triggered a radical shift in thinking: structuralist reasoning started to be appreciated and welcomed again, non-adaptive processes were broadly incorporated as important explanations of phenotypic transitions, and, together with a general movement that stressed the importance of developmental processes for evolutionary research, lead to the formation of evo-devo as the discipline that seeks to holistically explain phenotypic change [5,6].

The contemporary evo-devo researcher proposes that phenotypic transformations of organisms are best understood by focusing on changes to the types underlying these phenotypes and the developmental underpinnings of the transforming processes [7,9,10,17]. For them, the population genetics and natural selection of the Modern Synthesis yield crucially important, but certainly not all-powerful, explanations of how phenotypic traits change over time on the level of populations [9,10,17]. Yet, in their pursuit of understanding phenotypic innovation mechanistically, they focus on bodies, on individual organisms, more than they do on populations [10]. They concentrate on developmental processes which generate phenotypes, on how developmental systems evolve, and on the means by which development lays out the path of evolutionary trajectories, more than they do on natural selection maintaining or diversifying existing phenotypes [7-9,17]. But in order to understand how they study phenotypic change, one first has to ask: what exactly do they mean by "novel" phenotypes in the first place?

1.2 Homology and novelty - sameness and newness in evo-devo

While the Modern Synthesis made great progress in explaining how organismal traits diversify by the means of natural selection [15,17], it offered few insights into how novel organismal traits are generated in the first place [17]. Merely the question of what constitutes "innovation" in evolution, or what precisely qualifies a trait as "novel" still ignites much debate amongst evolutionists to this day. Arguably, this debate, too, is rooted in the aforementioned form-function dualism, since defining what a novelty is requires us to specify two things: what constitutes a trait, and what delineates an "old" (or rather "the same" trait) from a "new" trait.

We've touched on what constitutes a phenotypic trait, and saw that functionalists and adaptationists define traits based on their functions and their adaptive fit to the environment, while structuralists define them by their underlying type. For the latter, sameness - that is homology - is conveyed not by similarity in function or form, but by descent with modification [9,10]. This concept of sameness through homology has essentially remained unaltered since its inception, and it now constitutes a central column of evo-devo thought. However, it has been expanded to include phenotypes other than morphological traits, like genes, proteins, and even developmental processes [19-22]. It should be noted, however, that by defining all these phenotypes based on homology, one needs to restrict homology to a specific level of biological organization, as homology of structure does not necessarily mean homology of the underlying genes and vice versa - an idea that was first proposed by Sir Gavin de Beer in the 1930s and more explicitly formulated in the 1970s [19,20]. Today, it is well established that no two organismal traits are likely to be *entirely* homologous [20-23], since genes and developmental processes can become dissociated from structural characters and since phenomena like developmental systems drift are ubiquitous in nature [22].

Now, let's turn towards evolutionary innovations, the different ways in which they can be defined, and, ultimately, what they mean to the evo-devo practitioner. Just like sameness, "newness" has been defined in various ways, and, again, usually based on the individual researcher's emphasis on either form or function.

Ernst Mayr provided one of the most frequently cited definitions of evolutionary novelty when he coined it "*…any newly acquired structure or property that permits the performance of a new function, which, in turn, will open a new adaptive zone.*" [24]. Here, it is immediately clear that Mayr puts emphasis on the functions of phenotypes, in order to

separate "old" from "new" traits. In fact, he proposed that there are only two different ways by which evolutionary novelties arise: "*…by an intensification of function or by the adoption of an entirely new function…*" [16 (p.225)]. Additionally, in the classical adaptationist manner of his time, he linked the emergence of new traits to their adaptive value by demanding that they are connected to adaptive radiation [16 (p.229)], and he offers only one plausible causal mechanism for the evolution of novel traits: natural selection [16 (pp.224-229)]. This definition is appealing to many biologists; it makes the assignment of novelty seemingly straightforward and intuitive [25]. A new trait is one that is present in one group of organisms but not in another, it fulfills a function for these organisms which is absent in the others, and it provides an adaptive advantage. Take the insect wing as an example: it exists in pterygotes but not in their closest relatives (zygentomes and entognathes), it functionally allows pterygotes to fly, and it is conceivable that it bestowed some adaptive advantage when it first evolved, as pterygotes are amongst the most successful and diverse animals on the planet. However, there are severe limitations to function-based definitions of novelty, such as Mayr's, especially when they are linked to adaptive radiation and natural selection. First, it is highly unlikely that every novelty can be linked directly to a specific event of adaptive radiation [26]. Second, this definition ignores the possibility that structural novelties may arise without any initial adaptive advantage. Third, it neglects the fact that new traits can evolve by combining other pre-existing ones [9,10,26-28]. Forth, linking novelties to adaptive radiation and defining them purely based on new functions implies that they evolve under selection for these new functions, rendering the whole argument perilously circular [25]. Moreover, in order for natural selection to play a role in the generation of novelty, there would have to be heritable variation in this novel trait or its new function in the first place [10,25]. Selection itself cannot act on non-existing traits and therefore it is difficult to see how it could generate novelties. Realizing this leaves the researcher with practically no framework to explore the mechanistic basis of the origin of new traits. Interestingly, this problem has already been noted by some of Darwin and Wallace's contemporaries, e.g. St. George J. Mivart, who wrote in 1871: "*'Natural selection,' simply and by itself, is potent to explain the maintenance or the further extension and development of favourable variations, […] But […] utterly fails to account for the conservation and development of the minute and rudimentary beginnings […] of structures, however useful those structures may afterwards become.*" [29]. This criticism persisted, as even the Modern Synthesis could not fix the problem. In

1977, the famous (or infamous depending on one's philosophical proclivities) Karl Popper noted that the Modern Synthesis "*…never gives us a full explanation of anything's coming into being in the course of evolution…*" and that "*…it is strictly a theory of genes, yet the phenomenon that has to be explained in evolution is that of the transmutation of form…*" [30].

So, how can this blatant limitation of the Modern Synthesis be circumvented? Structuralism! In evo-devo, novelties are neither defined by their functions alone, nor based on any adaptive notion. Instead, they are approached from the perspective of homology and defined relative to ancestral conditions [10,13,21,22,28]. Currently, two different categories of phenotypic novelties are broadly recognized under many different names and, for simplicity, I choose to refer to them as discrete novelties (similar to "type I novelty" in [10]; identical to "type 2 novelty" in [27] and "discretizing novelty" in [28]) and individualized novelties (roughly similar to "type II novelty" in [10] and "type 3 novelty" in [27]; identical to "individualizing novelty" in [28]). **Discrete novelties** are defined as "*structural elements newly inserted into an existing body plan, with no homologous counterpart in the ancestral species*"; **individualized novelties** are regarded to be "*...major variations of an existing body plan element through progressive individualization, with a new quality or functional capacity*" [28]. As we can see, a discrete novelty is very intuitively identified, in that it is a new structure present in one organism but not in its closest relatives, nor in the common ancestor it shares with them. An example of a discrete novelty is the aforementioned insect wing: while all pterygotes share this trait (with the exception of a handful of species that secondarily lost them), it is just as non-existent in modern-day zygentomes and entognathes, as it was in the last common ancestor that pterygotes share with them. In contrast, individualized novelties capture a different type of newness, in that they allow a new structure to evolve from one or more pre-existing ones [25-28]. These uniquely specialized traits emerge as their precursor trait(s) are progressively refined and transformed; as they develop into an extreme form of variation from their original state, they assume a new qualitative dimension or function [28]. A beautiful example of an individualized novelty is the narwhale's tusk. Technically, this tusk is "just a modified tooth" [25]. Since teeth, homologous to the tusks of narwhales, exist in other whales and many other vertebrates, including all shared common ancestors, narwhale tusks do not constitute new structures in the sense of discrete novelties. Yet, the narwhale's tusk is quite unlike any other structure in terms of its impressive length, its iconic shape, and its idiosyncratic sensory capabilities [25].

Individualized novelties conceptually capture exactly that: the narwhale's tusk evolved from a pre-existing trait (a tooth) which reached an extreme point of variation from its original state. It developed into a uniquely specialized structure that gained the quality of a sensory organ and now serves functions which its precursor did not [25].

Conceptualizing novelty the evo-devo way thus helps avoid the aforementioned pitfalls that adaptation-based definitions of novelty are plagued by. First, this homology-based approach to novelty explicitly emphasizes that new traits can arise *de novo* or continuously by modification of pre-existing elements [26,28]. Second, it does not force any long-term evolutionary effects (e.g. adaptive values) onto newly emerging traits, which circumvents the circular reasoning that arises from invoking natural selection as the causal agent generating novelty [25,26]. Third, it still captures the fact that phenotypic novelties often acquire new functions without making this a necessity for their emergence. Forth, and most importantly, it does not invoke any specific underlying mechanisms that generate these novelties, leaving the question of causality open for empirical investigation [10,23,28,31]. By defining novelty specifically on the level of phenotypic homology, evo-devo facilitates the appreciation of changes in the underlying gene regulatory networks (which bestow a unique developmental and evolutionary identity on these traits) as causal mechanisms [10]. This makes it possible to integrate knowledge of how new genes evolve (e.g. *de novo*, duplication, horizontal transfer) and how new regulatory networks emerge (e.g. *de novo,* co-option of pre-existing genes into new regulatory networks, diversification of transcription factor function) [23] into the empirical quest to reveal the processes generating new phenotypes. And that is precisely what novelty means to the evo-devo researcher. It is conceptualizing and contextualizing phenotypic change in a way that allows scientists to investigate the exact molecular and developmental mechanisms which bring it about.

1.3 Phenotypic plasticity - a facilitator of novelty

While a host of developmental processes has been proposed to influence or explain phenotypic change, one has received unparalleled interest by evo-devo researchers: phenotypic plasticity[1]. Generally understood to be "*…the ability of a genotype to produce different phenotypes in response to distinct environmental conditions…*" [32], phenotypic

[1] some researchers prefer the term "developmental plasticity"; here, these terms are considered synonymous

plasticity has been cherished as an essential trailblazer in the evolution of new phenotypes [10,26,33-35], especially of individualized novelties [28].

Just like the form-function and sameness-newness dualisms, the concept of phenotypic plasticity has a long history in evolutionary biology. The first experimental investigations of phenotypic plasticity date back to 1909, when Richard Woltereck studied how varying levels of nutrients in the environments of small crustaceans called *Daphnia* impacted the development of their head anatomy [32,36]. In plotting these nutrient levels against *Daphnia's* head dimensions, he obtained correlational trend lines which he found to differ across independent strains of these organisms. In calling these associations "reaction norms" he created one of the first conceptualizations of phenotypic plasticity [32]. Later, in 1965, the botanist Anthony Bradshaw expanded on this notion after finding that extreme environmental conditions (such as strong variations in light intensity or temperature) can cause plants to develop alternative phenotypes [32,37]. Just like Woltereck did for reaction norms, Bradshaw argued that phenotypic plasticity is a trait in itself, one which varies across closely-related organisms [32,37]. He deduced that the *variation in the plasticity* of a trait is independent of the *variation in the trait* that exhibits the plasticity, and poured the foundation of our modern understanding of phenotypic plasticity when he concluded that this observation is "*…difficult to explain unless [the plasticity of a character] is under its own specific genetic control…*" [37]. While these contributions demonstrated that phenotypic plasticity can be an important mediator of phenotypic differences in response to environmental factors, its general importance for organismal evolution remained questioned (or ignored) for decades, largely due to a perceived lack of empirical cases demonstrating the pervasiveness of plasticity in nature [32-34].

This lasted until 2003, when Mary Jane West-Eberhard published her magnum opus on *Developmental Plasticity and Evolution* [34] and finally ridded evolutionary biology of its phantom pain. In her book, she presents an immensely comprehensive collection of polyphenisms - alternative phenotypes which are brought about by environmental, rather than genetic, variation - demonstrating their ubiquity across all domains of life [34], and providing important insights into the universal responsiveness of organisms to the world surrounding them [32]. The omnipresence of polyphenisms, she argues, proves the importance of both the environment and organismal development for the generation of phenotypic novelty and, thus, evolutionary theory in an evo-devo era [33,34]. She argues that these alternative phenotypes constitute

developmentally and functionally independent targets of natural selection, which allow "*…the elaboration of a new trait without elimination of an established one, thereby facilitating the evolution of new adaptive specializations…*" and she reasons that they "*…can represent a crucial phase in the evolution of […] novelties, allowing macroevolutionary changes to occur via microevolutionary processes within species, and without disruption of preexisting adaptations…*" [34 p.377]. Considering the points discussed before, this idea fits neatly into evo-devo's research program and theoretical framework: it postulates that the evolution of novel phenotypes sits at the very heart of evolutionary inquiry [33,34], it defines novelty in a structuralist manner relative to established traits (i.e. homologs) [34], it acknowledges that novelties commonly arise by modifications or combinations of preexisting traits (i.e. individualized novelties) [34 pp.145,197-198], it accepts that the natural selection and adaptation of the Modern Synthesis are essential forces in evolution [34 pp. 140-142], and it demands the incorporation of the environment and development as causal factors into evolutionary theory [32,34].

But how does phenotypic plasticity facilitate the evolution of novelties exactly? West-Eberhard proposes a step-wise sequential model [34, pp.140-141]: First, a novel input, which can be genetic or environmental, occurs and affects a preexisting responsive phenotype in one or more organisms, causing phenotypic change. Second, the affected organisms respond by mutual adjustment of their constituting parts (i.e. traits) and, thus, express a novel phenotype [38]. Third, the initiating novel input (genetic or environmental) reoccurs and leads to a spread of the novel phenotype in a subpopulation of the affected organisms. Forth, natural selection acts on heritable variation in the regulation and/or the form of the novel phenotype in the subpopulation of organisms expressing it [34,38]. While these four steps cover the fundamental sequence of adaptive evolution, West-Eberhard also acknowledges that, in a further step, plastic phenotypes usually enter a phase of evolutionary stability in the form of switch-regulated polyphenisms. This may happen even when an alternative phenotype is expressed only infrequently, as it may still be under positive selection and thus considered an adaptation rather than an anomaly [34, p.141]. Furthermore, West-Eberhard reminds us that evolutionary modifications will affect the form of the plastic trait or its switch-regulatory basis, allowing alternative phenotypes to change morphologically, and to eventually become fixed (if one alternative becomes the only one expressed) or lost entirely (if the inducing conditions change completely). She also predicted that the secondary fixation of alternative phenotypes would be associated

with a burst of morphological diversification [32,34]. She reasoned that, during phases of stable polyphenisms, selection would focus on maintaining their underlying switches and thus necessarily be more relaxed on the genes underlying the formation of either alternative phenotype (i.e. the genes downstream of the switches). This, in turn, would allow genetic variation to accumulate in those downstream components of the switch network [34]. Ultimately, once plasticity is lost, these accumulated changes will become exposed to selection and act as rich source of variation, facilitating rapid evolution of form [34]. This idea has since become known as the "facilitator hypothesis" [32]. Thus, in West-Eberhard's model of novelty evolution, organismal development marks the beginning of the causal chain [34,38]: environmental or genetic factors lead development to produce phenotypic variation within subpopulations of organisms. This phenotypic variation leads to differences in reproductive success and, thus, natural selection. If (or better: once) phenotypic variation has a genetic foundation, evolution - the intergenerational shift of phenotypes and gene frequencies - will follow [34 pp.140-142].

Phenotypic plasticity hence demonstrates that development is the first-order cause of phenotypic change, and that natural selection is only a second-order cause which shapes and maintains the distribution of new phenotypes on the population level depending on their influence on reproductive success [34,38]. This also shows that, while mutations are the undisputed source of all *genetic* novelty, they are not necessarily the source of all *phenotypic* novelty - a crucial distinction [34 p.144]. In fact, evo-devo researchers now work based on the premise that the most common source of phenotypic innovation is environmental induction of developmental variation (rather than mutations) [10,17,34,38]. The environment affects all individuals in a local subpopulation and thus environmentally-induced novelties are likely to appear in many individuals at once. Hence, the chances for an environmentally-induced novelty to become established in the population and exposed to natural selection are much higher than for novelties induced by mutations [10]. Additionally, this premise also circumvents the frequently uttered concern that individuals with novel phenotypes may not find a suitable partner for reproduction, as the novelty might render its carrier sexually incompatible with those individuals that lack it [10,34]. If novelties do indeed commonly originate from environmental induction affecting many organisms at once, then carriers of novelties will have decent chances finding compatible mates, even if they are sexually incompatible with conspecifics that lack the new trait [10].

Acknowledging the important roles development and phenotypic plasticity play in evolution resulted in three major predictions that can be empirically studied by evo-devo researchers [32]: First, and most important, the evolution of novelties is tightly linked to phenotypic plasticity and it begins with environmentally-induced phenotypic variation caused by development [10,17,32,34]. This link is especially relevant for those interested in the origin of individualized novelties [28], as both the evolution of individualized novelties and the facilitator hypothesis explicitly start with preexisting traits. Second, phenotypic plasticity (especially a stable polyphenism) requires the presence of a genetically-encoded developmental switch that senses environmental information and controls the generation of alternative phenotypes [32,34]. Third, phenotypic change is associated with pulses of phenotypic plasticity over evolutionary time, as an initially environmentally-induced alternative phenotype will ultimately come under genetic control, leading to its fixation [32,34].

Surprisingly, despite the wealth of examples for polyphenisms, our current understanding of the molecular basis for phenotypic plasticity is still very limited [32]. This arguably constitutes one of the biggest challenges in evo-devo research today, as a full appreciation of phenotypic plasticity's role in the evolution of novelties hinges on a comprehensive understanding of its genetic underpinnings. How are the predicted developmental-switch networks that govern plasticity structured? How did they evolve in the first place, and what kind of genes are involved? Are they also involved in the morphogenesis of novelty? Which molecular mechanisms ensure the stabile execution of the plastic response to the environment? And which processes lead to canalization?

To address all of these questions in due detail, evo-devo researchers must focus on a model organism [7], ideally one in which comparative functional genetic interventions and ecological manipulations of a plastic trait are possible. Additionally, to study the role of plasticity in the evolution of novelties, this model organism should not only display any polyphenism, but ideally one associated with a novel trait. And, to make it even more complicated, in order to fully address the evolution of novelties in general, this model organism also needs to allow the researcher to comprehensively study the evolutionary history of its novelty (i.e. phylogenetic reconstructions of character transitions), the reasons why the novelty is maintained in its populations (i.e. functional ecological studies), and the means by which the novelty is generated during its development (i.e. functional genetic investigations) [26]. Thus, one might wonder: is there an organism that brings all of these perks to the table (or rather the lab bench)?

2. Opening a flask of worms

"EVOLUTION MAKES MORE SENSE IN THE LIGHT OF DEVELOPMENT."
- KOSTAS KAMPOURAKIS & ALESSANDRO MINELLI (2014)

Nematodes, colloquially known as "roundworms", are an incredibly abundant and amazingly diverse group of invertebrates which occupy nearly all ecosystems across the globe. The most recent estimates predict that they likely originated some 500 million years ago (~100 myr before the first nematode appeared in the fossil record) [39], that they make up approximately 80% of all terrestrial animals alive today [40], and that the number of yet-to-be-identified species ranks in the millions [41,42]. Nematodes also constitute one of the most economically and medically relevant groups of organisms, considering that plant-parasitic species tend to cause devastating damages to global crop yields amounting to billions of US dollars every year, and given that certain animal-parasitic species are responsible for some of the most detrimental tropical diseases afflicting millions of humans [43,44]. In terms of size, these worms range from microscopic free-living sea creatures spanning less than a tenth of a millimeter to eight-meter long beasts parasitizing the placentas of sperm whales [45].

In laboratories around the world, nematodes - especially the microscopic soil-inhabitants - have been (and continue to be) used to study all aspects of biology [46]. This is because they come with some invaluable features: they are generally small and easily maintainable on agarose plates inoculated with bacteria as a food source (allowing handy storage of large populations), they mostly produce large broods and often have short generation times (allowing scalable population genetic studies), and they mostly produce isogenic populations via self-fertilization (facilitating detailed functional genetic studies). Additionally, they can be cryopreserved [46] for extraordinarily long periods of time, as demonstrated by a recently discovered species that could be revived effortlessly from a 46,000-year long slumber in Siberian permafrost [47]. Furthermore, most of these worms have a highly simplified mode of development, in which a small number of cells generated from an invariable and predictable lineage make up the animals' bodies (a phenomenon called eutely) [46]. This allows scientists to homologize the worms' traits down to the level of individual cells [48], rendering them prime subjects for mechanistic studies of phenotypic change.

In considering these perks, eutelic nematodes, such as the "rhabditid" *Caenorhabditis elegans* and the diplogastrid *Pristionchus pacificus*, have been used in

comparative evo-devo studies [48-50] that aimed to dissect some of the fundamental processes of animal development and evolution, including heterochrony [51], developmental systems drift [52], and - excitingly - phenotypic plasticity [32,48,53-55].

2.1 Phenotypic plasticity in nematodes - the dauer larva

Given their renown as animals with invariable and predictable somatic development, it seems contradictory that nematodes currently offer one of the most well-described examples of phenotypic plasticity known to evo-devo [48]: the dauer larva.

Many nematodes can opt for one of two alternative developmental trajectories: the "direct" or the "indirect" life cycle [56]. Under normal (optimal) conditions, e.g., where bacterial food is abundantly available and population densities are low [57], nematodes like *C. elegans* and *P. pacificus* develop into reproducing adult hermaphrodites via a direct life cycle with four larval stages (L1-L4; or, if "juvenile" is the preferred term, J1-J4). The completion of the direct life cycle takes three to four days [46,49] and transitions between the life history stages, including the transition from L4/J4 larvae into adult worms, are marked by distinct molting events in which the worms shed their old cuticles to replace them with new ones [46]. However, under stressful (i.e. suboptimal) conditions, when food sources are depleted or populations reach extreme densities [57], young larvae may enter the indirect life cycle and develop into dauers - an alternative to the L3/J3 larvae of the direct life cycle [56,57].

Dauers are developmentally arrested and non-feeding larvae adapted for dispersal and stress survival [57]. Some of their stress-protective anatomical features include a thickened body wall cuticle, a buccal plug sealing the opening of their alimentary channel, and a constricted pharynx that is not actively used for pumping [46,56,57]. These features aid dauers during their dispersal through varying environments, when they actively search for new habitats with conditions suited for the establishment of new populations. The physiology of dauers is highly specialized for survival too, as they (unlike feeding larval stages) enter a hypometabolic state in which they generate energy by breaking down stored neutral lipids into acetyl-CoA (via lipolysis and β-oxidation), followed by a CO_2-efficient conversion of these molecules into glucose and trehalose (via glyoxylate shunting and gluconeogenesis) [57-60]. While glucose is subsequently metabolized to produce energy, trehalose is transported into somatic cells, where it protects the worms' cell membranes from desiccation-related

damages [60,61]. The starvation-resistance conveyed by this hypometabolic state, in combination with the activation of biological pathways that protect against molecular damages and slow down aging [58], allows dauers to survive for many months [62,63], rendering them incredibly long-lived compared to adult worms which tend to live for only up to two weeks [46,49].

The genetic switch regulating dauer plasticity in *C. elegans* (especially dauer entry and dauer exit in response to their inducing environmental conditions) is incredibly complex and involves multiple deeply-conserved signaling pathways [57]. Simply, dauer plasticity integrates the TGF-β-like and insulin-like signaling pathways, which, under favorable environmental conditions, are up-regulated and together promote the synthesis of dafachronic acid from cholesterol [64]. Dafachronic acid is a steroid hormone and the activating ligand of the nuclear receptor DAF-12 [65]. Upon ligation, DAF-12 assembles a putative co-activator complex that by-passes dauer entry and promotes reproductive development via the direct life cycle. Conversely, in suboptimal environments, TGF-β-like and insulin-like signaling pathways are down-regulated and dafachronic acid biosynthesis is inhibited. In the absence of this ligand, DAF-12 interacts with another protein (DIN-1) to form a putative co-repressor complex that blocks reproductive development and instead promotes dauer entry [64].

Although this fountain of knowledge makes dauer a beautifully illustrated case of phenotypic plasticity, it is actually not the best example to study the evolution of phenotypic novelty and the facilitator hypothesis. This is mainly due to two obstacles: first, and most importantly, the case of dauer plasticity does not seem to meet all of the aforementioned criteria that define a good model system to study morphological novelty. Indeed, as discussed above, utilizing nematodes like *C. elegans* did allow evo-devo researchers to shine light on the genetic and developmental bases of dauer formation (i.e. functional genetics studies) [57,64,65]. It also allows them to ask why this novel phenotype was and/or is maintained on a population level (i.e. functional ecological studies) [63,66,67]. Yet, the evolutionary history of this novel larval stage is not easily reconstructed. While dauers exist in a broad spectrum of nematodes, covering large aspects of their phylogeny, it is currently still unclear where (or how often) dauer larvae evolved [68,69]. Additionally, the predictions offered by the facilitator hypothesis go beyond the establishment of evolutionarily stable polyphenisms. It postulates that, eventually, alternative phenotypes become fixed and modified [34]. Yet, to my knowledge, no known nematode has a "fixed" indirect life

cycle that would be characterized by the loss of the L3 larva and a canalization of the ontogenetic trajectory via dauer[2]. Second, these two alternative phenotypes effectively constitute two near-completely distinct (larval) animals, which combine plenty of alternative morphological and physiological traits. Thus, as a whole, dauer larvae may simply be too complex of a model to study the plasticity, or the plasticity-led evolution, of morphological novelty. Hence, while dauer formation demonstrates that nematodes are promising model organisms to study the genetic and developmental bases of phenotypic plasticity, a different polyphenism is needed to link phenotypic plasticity to the evolution of a specific novel trait. Fortunately, such a case exists.

2.2 Linking plasticity and novelty - a feeding polyphenism

Despite some frequently mentioned superficial similarities, nematodes are far from morphologically uniform. A closer look immediately reveals a stunning diversity in their mouth structures and associated feeding strategies [70,71], which are arguably the reason why nematodes were able to conquer almost every ecosystem.

The mouth of roundworms (also known as their "stoma") is the major opening in their body wall and, when talking about the stoma, nematologists classically differentiate the "buccal cavity" - the lumen of the mouth - from the "buccal capsule" - the inner cuticular lining of the mouth (Fig. 1a,b,d) [72]. Although both the outer lining of the body wall and the buccal capsule are made of cuticle, they have historically been treated as separate structures [70]. Yet, ultrastructural analyses utilizing transmission electron microscopy (TEM) revealed that, based on the underlying cells which secrete it, the buccal capsule can be divided into six cuticular feeding structures (from anterior to posterior): cheilostom, gymnostom, prostegostom, mesostegostom, metastegostom, and telostegostom [73]. Interestingly, the cheilostom was found to be secreted by skin cells (hypodermis), just like the body-wall cuticle. This lead nematologists to hypothesize that these two structures are in fact continuous, and that the cheilostom is a feeding structure that is qualitatively different from the remainder of the buccal capsule [73] - an idea that has so far not been corroborated by molecular data. Next, the gymnostom,

[2] Some parasitic nematodes might represent an exemption from this, as their infective juvenile stage (IJ) (which is homologous to the dauer stage of free-living species [69]) is a fixed aspect of their life histories. However, these nematodes also have a host switch and a generation switch as integral parts of their life cycles, and IJs are expressed only in some of the generations but not all. Thus, IJ plasticity is arguably a fixed aspect of the parasite's life history, but it is not a fixed phenotype that constitutively expressed in every of the parasite's generations, which is exactly what I mean when I refer to a "canalization of the ontogenetic trajectory via dauer".

which forms the middle part of the buccal capsule, was found to be associated with the two arcade syncytia [73,74] - groups of cells which are unique to nematodes and whose functions are enigmatic. The four posterior-most feeding structures are collectively referred to as the stegostom - the part of the buccal capsule that is made by pharyngeal cells. Its first two elements, the prostegostom and mesostegostom, are each produced by a set of "e" cells (pharyngeal epithelia); its last two elements, the metastegostom and telostegostom, are each secreted by a set of "pm" cells (pharyngeal muscles) [73-75]. Notably, in each of the investigated species thus far, all six feeding structures have very different electron-densities in TEM sections, as well as varying degrees of rigidity in live worms [74,75]. This suggests that, although each of these structures is cuticular in nature, their exact biochemical compositions can vary dramatically. Given this, one can easily imagine the many potential functional uses these feeding structures (or combinations thereof) may offer, and what a rich substrate they provided for the evolution of the amazingly diverse feeding morphologies and strategies in nematodes. Interestingly, such feeding-related diversity can sometimes be found even *within* individual species.

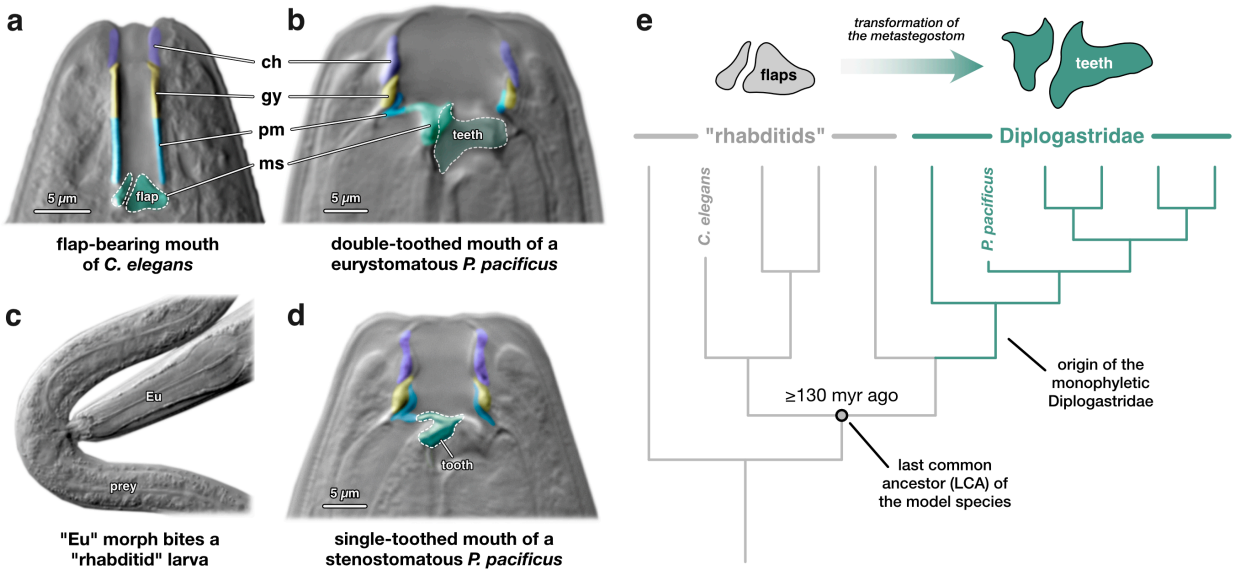


Figure 1. Evolution of the cuticular feeding structures in "rhabditid" and diplogastrid nematodes. **a.** Buccal capsule of the "rhabditid" model nematode *Caenorhabditis elegans*. **b,d.** Alternative buccal capsules of diplogastrid model nematode *Pristionchus pacificus* (ch = cheilostom, gy = gymnostom, pm = promesostegostom, ms = metastegostom). **c.** Eurystomatous ("Eu") morph of *P. pacificus* preying on a "rhabditid" larva. **e.** Schematic phylogeny depicting the flap-to-tooth transformation of the metastegostom during the "rhabditid"-to-diplogastrid transition (myr = million years). Figure was modified from Theska *et al.* (2023) and Theska & Sommer (2023), see Results (pp. 30,31).

Pristionchus pacificus and other members of the Diplogastridae exemplify this beautifully, as these soil-dwelling and often arthropod-associated worms display phenotypic plasticity in their feeding structures and strategies [50,54]. In response to environmental inputs such as pheromone signals [76], temperature changes [77], and/or compositional differences in the media surrounding them [78], these animals can develop into either narrow-mouthed bacterivores ("stenostomatous" morph; St) or wide-mouthed predators ("eurystomatous" morph; Eu) [50,72]. The buccal capsules of both morphs are characterized by the presence of a moveable cuticular tooth (which is flint-shaped in St and hook-shaped in Eu) at the anterior tip of the dorsal pharynx (Fig. 1b,d). The predatory feeding mode of Eu animals is facilitated by the presence of a second hook-shaped tooth that sits at the anterior tip of the right subventral pharynx (Fig. 1b), where St animals only have a flat cuticle ridge [50,72]. By moving their two cuticular teeth against each other, in a motion that is reminiscent of inverse scissors, Eu animals can actively rip through the body walls of other nematodes (sometimes referred to as "biting") and subsequently ingest their prey's internal organs (Fig. 1c) [50,79]. Thus, the feeding plasticity of *P. pacificus* and other diplogastrids constitutes another bimodal polyphenism, like dauer plasticity. Excitingly, the very structures that display this plasticity - the moveable cuticular teeth - are a morphological trait that cannot be found in any of the diplogastrids' closest relatives (i.e. the paraphyletic grade "Rhabditidae" from which the clade Diplogastridae emerged), which have immovable cuticular "flaps" in their mouths instead (Fig. 1a,e) [50,72,80].

Do these teeth constitute a novelty according to evo-devo's definitions? As mentioned before, the fact that these worms are eutelic allows us to homologize their anatomical traits down to the level of individual cells, including those which secrete the buccal capsule. This also means that, as long as homology of the underlying mouth cells can be ascertained across species, we can also directly homologize the individual feeding structures produced by them. Fortunately, detailed TEM-based ultrastructural analyses and three-dimensional reconstructions of the mouth cells and feeding structures exist for *C. elegans* and *P. pacificus* as representatives of the "Rhabditidae" and Diplogastridae, respectively. These studies revealed that the flaps of *C. elegans* and the teeth of *P. pacificus* correspond to the metastegostom[3] of other nematodes (including the last common ancestors they share with them), and are thus homologous (Fig. 1a,b) [72-75]. Yet, it seems obvious that the movable teeth of diplogastrids are unlike any

[3] in Eu morphs of *P. pacificus* the telostegostom may contribute a tiny amount of cuticle to the dorsal tooth, but this is still contested

other feeding structure in nematodes. They evolved from a pre-existing structure (the metastegostom) and reached an extreme point of phenotypic deviation from its ancestral state (flaps); they developed into uniquely specialized traits that gained the quality of raptorial structures (hook-shaped teeth) which now serve a function their precursors did not (predation). Hence, the diplogastrid's teeth are essentially "the narwhale's tusk in a nematode's mouth" - a textbook example of individualized novelties (Fig. 1a,b,e).

Taking everything together, the diplogastrid's feeding polyphenism offers a perfect model case to link phenotypic plasticity to the origin of novelty. It comes with all the perks that were previously discussed for dauer plasticity in nematodes, but with none of its associated shortcomings. Where phenotypic differences between dauers and L3/J3 larvae are insurmountable, the differences of St and Eu morphs are neatly restricted to feeding structures. Where the novelty of dauer larvae is difficult to define due to its uncertain point(s) of origin, the plastic teeth can clearly be defined as individualized novelties that are linked to the origin of diplogastrids [50,80]. Lastly, where the facilitator hypothesis could not be tested for dauer plasticity (due to a lack of fixed/canalized alternative phenotypes), it has already found its first empirical support in the feeding plasticity of diplogastrids. A macroevolutionary study reconstructed the evolution of the feeding structures and their phenotypic plasticity across more than 90 nematode species, including *C. elegans* and many diplogastrids, and found that both the morphological novelty and the plasticity co-originated in the last common ancestor of diplogastrids [80]. This event accelerated the rate of morphological diversification in this lineage and resulted in a boost of structural disparity, just as predicted by West-Eberhard. Additionally, while most investigated diplogastrids evolved stable polyphenisms, multiple independent lineages secondarily lost feeding plasticity and genetically encoded only one of the two morphs (Eu or St). Excitingly, these linages showed another burst of morphological diversification and structural innovation within their canalized morphs, which was even stronger than the initial one that characterized the origin of the plastic novelty [80]. Again, this observation is exactly in line with the predictions of the facilitator hypothesis, leaving evo-devo researchers with a wonderful starting point in their pursuit to study the molecular mechanisms which bring these phenotypic transitions about. In fact, diplogastrids can be used to coherently study all major aspects of plasticity-led evolution, by elucidating: (I) the molecular architecture of the developmental switch network which controls the plasticity and morphogenesis

of their feeding structures, (II) the origin of the developmental switch network, and (III) the molecular processes which are involved in the canalization and subsequent diversification of a single morph [32].

The first of these points was the major focus of extensive molecular studies over the last decade, which have resulted in a detailed draft of a modular developmental switch network underlying *P. pacificus's* feeding plasticity (Fig. 2a).

The upstream-most components of the developmental switch network, such as ion channels and intraflagellar transport proteins, act in the worm's sensory neurons [77,81,82]. These proteins are involved in the reception of environmental inputs and their subsequent transduction to the remainder of the network [32]. Therefore this upstream aspect of the developmental switch has been referred to as the "environmental-perception module" (Fig. 2a). Interestingly, although epigenetic mechanisms are frequently hypothesized to be involved in this kind of environmental perception [17,34], since they could encode received information into molecular signals, their role in developmental switch networks remains essentially unexplored [32].

Downstream of the environmental-perception module sits the "plasticity-switch module" which controls the feeding plasticity, as it integrates the received environmental inputs and initiates the development of a corresponding feeding morphology [32]. It is mostly comprised of an X-chromosomal multi-gene locus that is expressed in the central nervous system [83-86]. The major switch gene within this locus, *eud-1*, is a sulfatase that presumably governs the chemical modification of ligands for downstream transcription factors of the developmental switch network [87]. While mutation-based inactivation of *eud-1* causes all worms to develop as St morphs, its overexpression leads to all-Eu populations [83]. The other two genes of the locus, *nag-1* and *nag*-2, encode N-acetyl-glucosaminidases either of which, when mutated, cause phenotypic effects exactly opposite of *eud-1* mutations [84]. The same holds true for the sulfotransferase SULT-1/SEUD-1 (which is expressed in the head hypodermis and the pharynx), implying that it too is a part of the plasticity-switch module (Fig. 2a) [85,86].

Completing the developmental switch network is a downstream "phenotype-execution module", which controls the morphogenesis of the alternative mouths, including the evolutionarily novel teeth [32]. At the heart of this module sit two transcription factors (TFs), the nuclear receptors *Ppa*-NHR-1 and *Ppa*-NHR-40, which control the expression of multiple enzyme classes involved in the modification and

degradation of extracellular matrix and cuticle (Fig. 2a) - the very materials that make up the feeding structures [32,88,89].

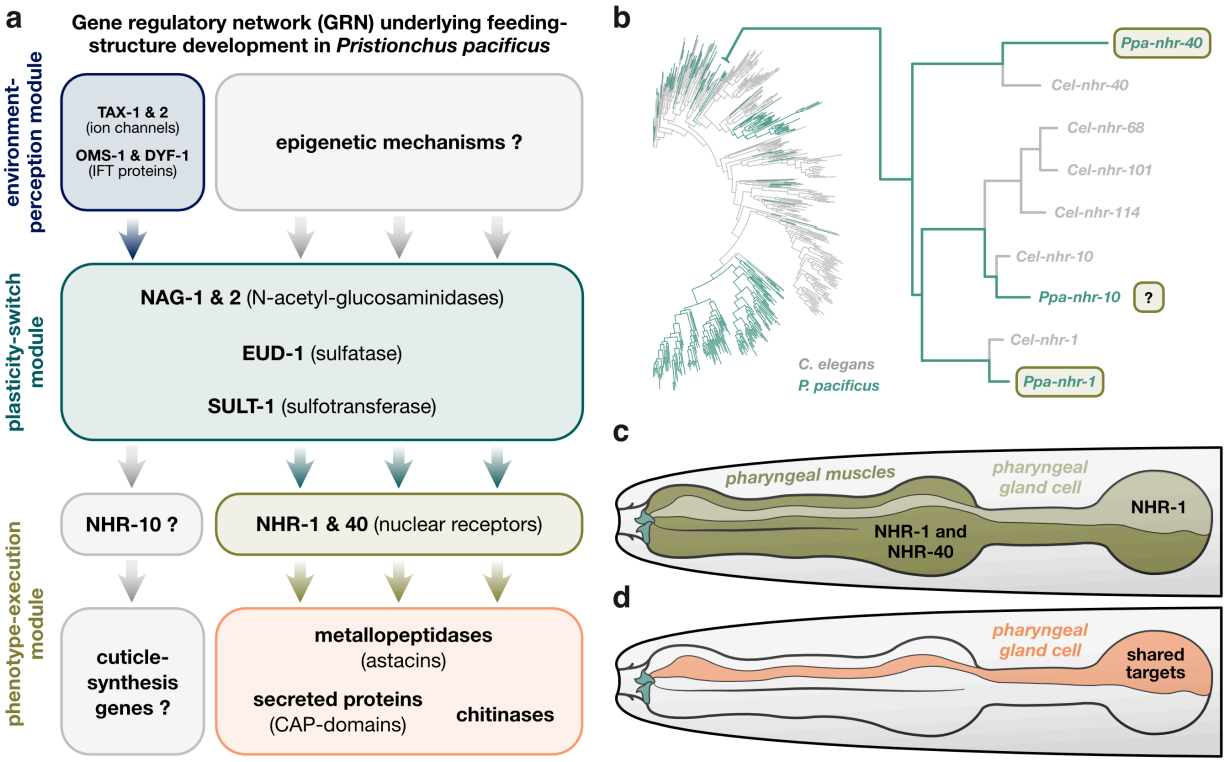


Figure 2. A modular developmental switch network underlies the plasticity and morphogenesis of feeding structures in *Pristionchus pacificus*. a. The currently identified modules and molecular constituents of the gene regulatory network (GRN). Note that the epigenetic mechanisms involved in environmental perception are currently unknown. The same applies to the aspect of the phenotype-execution module which governs the synthesis of the cuticular feeding structures, including structural constituents and their transcription factors. **b.** Phylogeny of all *nhr* genes encoded in the genomes of *C. elegans* and *P. pacificus* (re-drawn form [89]). Note that most *nhr* genes are species-specific and that only a few (mostly basal) branches contain *nhrs* which are conserved across species. Right side of the panel shows a zoom-in on the branch that harbors the conserved NHR-1/-10/-40 cluster. Note that both *Ppa-nhr-1* and *Ppa-nhr-40* are regulators of phenotype-execution and that the function of their sister paralog *Ppa-nhr-10* is currently unknown. **c,d.** Schematic representation of the expression patterns of NHR-1 and NHR-40 and their shared regulatory target genes in *P. pacificus*. Depicted is the head region of the worm in lateral position (buccal cavity is on the left; teeth in green). Figure was modified from Theska *et al.* (2023) and Theska & Sommer (2023), see Results (pp. 30,31).

Nuclear hormone receptors (NHRs) are a peculiar group of TFs which integrate cellular metabolism with physiological function in animals [90]. They are restricted to metazoans [91] and generally share an archetypical protein-domain organization. Their two most important domains are the DNA-binding domain (DBD), which recognizes specific DNA motifs in the *cis*-regulatory regions of their target genes, and the ligand-binding domain (LBD), which regulates the activity state of the NHR depending on the presence or absence of a specific ligand [90,91]. Recent genomic studies revealed that

nematodes, in contrast to other animals, often possess hundreds of *nhr* genes [89,92]. The overwhelming majority of these genes were found to have originated via a nematode-specific explosive burst of duplications of the ancient receptor HNF4 [93,94]. Thus, only a mere handful of the nuclear receptors found in these worms are deeply-conserved across all animals. The remainder (and vast majority) of *nhr* genes is unique to nematodes, functionally uncharacterized, and referred to as "supplementary nuclear receptors" (*supnrs*) [94]. Another striking observation was that, when comparing distantly-related species with similarly high numbers of NHRs (e.g. *C. elegans* [n=266] and *P. pacificus* [n=254]), even the vast majority of the nematode-specific *supnrs* is completely lineage-specific with no identifiable one-to-one orthologs between the two species (Fig. 2b) [89]. Intriguingly, *Ppa-nhr-1* and *Ppa-nhr-40*, the two aforementioned phenotype-execution regulators in *P. pacificus's* feeding plasticity, are representatives of these nematode-specific *supnrs* [89] and belong to the very few which are conserved across distantly-related species, such as *P. pacificus* and *C. elegans* (Fig. 2b).

Expectedly, these two TFs are expressed in the target tissues (the pm cells) which secrete the teeth in *P. pacificus* (Fig 2c) [88,89]. Additionally, their shared regulatory targets are expressed in a single cell, the dorsal pharyngeal gland g1D (Fig. 2d), which runs through the pharyngeal muscles and opens into the buccal cavity via a channel in the dorsal tooth (Fig. 2c,d). Worms which lack a functional copy of *Ppa-nhr-1* constitutively develop intermediate mouth morphologies combining features of St and Eu morphs, suggesting a disruption of proper phenotype execution [32,75,89]. Interestingly, worms with *gain-of-function* mutations in *Ppa-nhr-40* become all-Eu, while *loss-of-function* mutations lead to all-St phenotypes [88,89]. Therefore, *Ppa*-NHR-40 shows switch-gene characteristics in addition to its role in the regulation of stoma morphogenesis [32], which render it the putative integrator of the plasticity-switch and phenotype-execution modules. Interestingly, all components of the plasticity-switch module are evolutionarily young and derived from diplogastrid-specific duplication events [32,84]. In stark contrast to these genes, the two regulators of the phenotype-execution module are ancient. Both *Ppa-nhr-1* and *Ppa-nhr-40* have deeply-conserved one-to-one orthologs in "Rhabditidae", such as *C. elegans* [89]. Excitingly, although their specific target genes are fast-evolving [32,89], the enzyme classes to which these targets belong (astacins, chitinases, and cysteine-rich secreted proteins) are ancient too, and found across all ecdysozoans.

2.3 Something old? Something new? Something borrowed? Something to pursue!

Bringing all of the aforementioned points together, the establishment of the detailed, albeit still incomplete, draft of a developmental switch network now opens up the unique opportunity to study how it was assembled during the "rhabditid"-diplogastrid transition. We have already seen that the plasticity-switch module is made of new genes that emerged alongside the teeth and immediately gained their new plasticity-associated functions in diplogastrids [32,83,84]. However, we do not yet understand how the phenotype-execution module was assembled, and we still need to identify the specific structural proteins of the stoma that would form the downstream-most molecules of this module (Fig. 2a) [32]. Knowing that both NHR-1 and NHR-40, as well as their regulatory target classes (cuticle-modifying enzymes), already existed *before* the new feeding structures evolved in diplogastrids, together with the fact that clear homologs of these molecules exist in the "rhabditid" *C. elegans*, renders them exciting targets for comparative evo-devo studies. Especially in light of the fact that the teeth of diplogastrids are individualized novelties that evolved from "rhabditid" flaps (Fig 1e), one might now be eager to ask: is the phenotype-execution module an ancient genetic cassette that is also *functionally* conserved as a mouth morphogenesis module in "rhabditids", such as *C. elegans*? This would imply that the developmental switch network of diplogastrids evolved by building a new plasticity-switch module (made of evolutionarily young genes) upon an old mouth-morphogenesis program. Alternatively, the two old NHRs may have controlled a different biological function in "rhabditids". In this case, these ancient TFs would have been co-opted into a new regulatory context and borrowed for stoma morphogenesis in diplogastrids.

By identifying the exact structural molecules that constitute the downstream-most "base" of the phenotype-execution module, and by investigating the principal ways in which the phenotype-execution module may have been assembled during the "rhabditid"-diplogastrid transition, would yield important insight into the evolutionary history of this developmental switch network. So would the identification of any epigenetic mechanisms involved in environmental perception upstream of the plasticity-switch module, and the reconstruction of their evolutionary history in context of the feeding plasticity in *P. pacificus*. Is this network a complex mix of something old, something new, and something borrowed, in terms of its molecular makeup? I think this possibility is an interesting one to pursue.

3. Thesis aims

NOTHING IN EVOLUTIONARY MORPHOLOGY MAKES SENSE EXCEPT IN THE LIGHT OF DEVELOPMENTAL PLASTICITY.

In pursuit to contribute to our understanding of phenotypic plasticity and the ways in which it relates to the evolution and development of novel structures, I set out to study how a modular developmental switch network controlling the plasticity of novel feeding structures in *Pristionchus* nematodes evolved.

I sought to expand the existing knowledge of the switch network and reconstruct its evolutionary history by (I) identifying and functionally characterizing molecular players which are likely, yet unvalidated, members of the switch network, and (II) studying the functions of the deeply-conserved molecular constituents of the switch network in a "rhabditid" species, *Caenorhabditis elegans*, which lacks the morphological novelty and the associated feeding plasticity that characterize *Pristionchus pacificus*. To do so, I:

1) established a standardized, pliable, and repeatable approach to quantify anatomical differences within and across species, facilitating well-rounded comparative evo-devo studies focusing on the development and evolution of morphological novelties in microscopic animals such as nematodes.

2) demonstrated that experimental inactivation of the gene *dpy-6* leads to consistent malformations of the cheilostom (the anterior-most feeding structure of nematodes) in both *P. pacificus* and *C. elegans*, and thus helped to identify the first known deeply-conserved protein component of the nematode mouth.

3) participated in a study which revealed that differences between the alternative feeding structures in *P. pacificus* emerge in late larval stages, and that a specific epigenetic mechanism (H4K5/12ac), which governs developmental speed in nematodes and other ecdysozoans, evolved to regulate feeding plasticity in *P. pacificus*.

4) characterized the function of *nhr-10*, a deeply-conserved *supnr* and the sister-paralog of the two downstream plasticity regulators in *P. pacificus* (*nhr-1* and *nhr-40*), and demonstrated that, unlike its paralogs, it is not involved in the regulation of feeding plasticity but broadly required for starvation resistance in adult worms and dauer larvae.

5) examined the functions of *nhr-1* and *nhr-40* in *C. elegans* and found that they are involved in the regulation of metabolic pathways and stress responses, suggesting that these receptors acquired their plasticity-related roles in *P. pacificus* and other diplogastrids during the evolution of their novel feeding structures.

V. Results

1. Community-based re-annotation of a reference genome

Rödelsperger C, Athanasouli M, Lenuzzi M, **Theska T**, Sun S, Dardiry M, Wighard S, Hu W, Sharma DR, Han Z. *Crowdsourcing and the feasibility of manual gene annotation: A pilot study in the nematode Pristionchus pacificus.*
Scientific Reports (2019), 9:18789.

1.1 Synopsis

Nematodes such as *Caenorhabditis elegans* and *Pristionchus pacificus* can be used as powerful model systems to study the basic aspects of biology in a comparative manner. This includes the genetic basis of phenotypic differences that evo-devo researchers are primarily interested in. The tremendous knowledge of nematode genetics derived from decades of *C. elegans* research facilitates functional genetic studies using reverse genetics. However, the translatability of such approaches to less well-established model nematodes, like *P. pacificus*, heavily depends on having genome annotations of approximately similar quality to *C. elegans*, especially since erroneous annotations of orthologs hamper comparative functional genetic assessments across species.

To improve the current gene annotations for *P. pacificus* to a qualitative level comparable to those of *C. elegans*, we performed a genome-wide screen for *C. elegans* genes with potentially incorrectly annotated *P. pacificus* orthologs. We took a community-based approach to manually inspect more than 2000 candidate loci and to propose corrected gene models based on recently generated Iso-seq and RNA-seq datasets. We found that most erroneous annotations were due to artificial fusion of gene predictions or absence of gene models. Our joint re-annotation effort raised the total gene count for *P. pacificus* from 25,517 to 28,036 and increased the single copy ortholog completeness level of its genome from 86% to 97%. This pilot project drastically improved the overall quality of gene annotations for *P. pacificus*, demonstrating the value of even small-scale crowdsourcing projects in genomics, and facilitating more precise comparative evo-devo studies involving functional genetic manipulations of orthologous genes across nematode species.

1.2 Own contribution

I manually re-annotated roughly one fifth of the candidate loci in this study. My contribution is 10%.

2. Standardizing geometric morphometrics for nematode evo-devo

Theska T, Sieriebriennikov B, Wighard SS, Werner MS, Sommer RJ. *Geometric morphometrics of microscopic animals as exemplified by model nematodes.*
Nature Protocols (2020), 15:2611-2644.

2.1 Synopsis

The main objective of evolutionary developmental (evo-devo) biologists is to study phenotypic change mechanistically. To do this, they started to combine comparative developmental approaches that identify candidate genes underlying morphological change with functional genetic experiments that aim to demonstrate causative relationships between the observed morphological change and the identified candidate genes. While a wide variety of molecular techniques such as transcriptomics and CRISPR-based mutagenesis are now routinely used in such studies, sophisticated tools which allow the quantification of morphological differences among organisms are still largely under-appreciated. This is especially true for evo-devo studies focusing on small animals such as nematodes.

To foster the integration of quantitative morphological approaches into evo-devo research, we provide a highly pliable, easy-to-use, and robust protocol for the analysis of morphological differences in microscopic animals. Our protocol combines modern geometric morphometrics with unsupervised clustering algorithms to allow users to quantify highly multivariate phenotypes (including shape, size, and allometry) and to identify hidden group structures in their morphological data. We demonstrate the use and broad applicability of our protocol by quantifying intra- and interspecific differences in mouth morphology of the model nematodes *Caenorhabditis* and *Pristionchus*. Additionally, to foster repeatability, we include a pipeline to systematically assess measurement errors in the obtained data. Thus, for the first time, our protocol standardizes these kinds of analysis for microscopic model organisms and adds an important tool to the toolkit of evo-devo practitioners.

2.2 Own contribution

I was involved in conceptualizing the study. I designed, programmed, and tested major aspects of the computational pipeline, collected the raw data for the study, conducted all exemplary analyses, interpreted all results, and wrote the majority of the manuscript. My contribution is 70%.

3. Description of nine new species from China

Kanzaki N, Herrmann M, Weiler C, Röseler W, **Theska T**, Berger J, Rödelsperger C, Sommer RJ. *Nine new Pristionchus (Nematoda: Diplogastridae) species from China.* **Zootaxa (2021),** 4943:001-066.

3.1 Synopsis

Pristionchus nematodes have been intensively studied in the last two decades with contemporary work focusing on their development, evolution, ecology, behavior, and genomics. Studies which focused on the developmental plasticity and the facultative predatory life style that characterize these worms provided exciting insights into the mechanistic associations between plasticity and the origin of novel traits. These investigations followed a comparative evo-devo research program that leverages a total of 39 available species of *Pristionchus*, all of which can be studied in laboratory settings. While previous sampling efforts revealed that Asia is a hotspot of *Pristionchus* biodiversity, initial collections were biased towards Northern and island areas, largely for logistical reasons.

Here, we report on two extensive sampling trips to the Yunnan and Shaanxi provinces in mainland China. We were able to isolate and subsequently describe nine new *Pristionchus* species from these areas based on morphological features, mating experiments, and genome-wide sequence analyses. Excitingly, we found that some of the newly identified species display morphological characteristics that were unknown from any of the previously described species, indicating that the phenotypic diversity in these nematodes is yet to be fully explored. Therefore, this study drastically increased the total number of known *Pristionchus* species that can be used in future comparative studies focusing on the developmental and evolutionary peculiarities characterizing these organisms.

3.2 Own contribution

I obtained all morphometric measurements for the species descriptions. My contribution is 10%.

4. Identification of the first structural protein in the nematode mouth

Sun S*, **Theska T***, Witte H, Ragsdale EJ, Sommer RJ. *The oscillating Mucin-type protein DPY-6 has a conserved role in nematode mouth and cuticle formation.*
Genetics (2022), 220:iyab233.
*** co-first authors (equal contribution)**

4.1 Synopsis

Nematodes are well-known for their stunning diversity in feeding structures and strategies, which enable them to invade most ecosystems. Yet, little is known about the structural composition of the nematode mouth.

In this study, we used an evo-devo approach, combining a large-scale genetic screen with quantitative morphological analyses, to identify candidate constituents of *P. pacificus's* feeding structures. We found that *Ppa-dpy-6,* a gene predicted to encode a mucin-type protein, is involved in the formation of the cheilostom - the anterior-most cuticle element of the nematode mouth. Using our recently established protocol for geometric morphometrics, we demonstrated that *Ppa-dpy-6* mutants show general defects in cheilostom shape and size, resulting in overall malformations of the mouth. Furthermore, these mutants display the classical "Dumpy" phenotype which results in an extremely shortened body, indicating an additional role of *Ppa-dpy-6* in the formation of the body-wall cuticle. Together, these concomitant phenotypes provide the first molecular support for two old morphological hypotheses: that the cheilostom and the body-wall cuticle are continuous structures, and that the cheilostom is a mouth element that is qualitatively different from the rest of the feeding structures. We also demonstrated that mutations in the one-to-one ortholog of *dpy-6* in *Caenorhabditis elegans*, a distantly-related "rhabditid" nematode, cause the same malformations in mouth and body-wall structures we initially identified in *P. pacificus*. Thus, we were able to describe the first protein constituent of the nematode mouth, one which has a structural function that has been conserved through at least 100 million years of evolution. Additionally, these findings extend our knowledge of the switch network that controls the development of the plastic feeding structures in *P. pacificus.*

4.2 Own contribution

I designed and conducted all morphometric analyses of this study, interpreted and visualized the phenotypic data, and wrote the corresponding parts of the manuscript. My contribution is 40%.

5. Histone acetylation was co-opted to regulate feeding plasticity

Werner MS, Loschko L, King T, Reich S, **Theska T**, Franz-Wachtel M, Macek B, Sommer RJ. *Histone 4 lysine 5/12 acetylation enables developmental plasticity of Pristionchus mouth form.*
Nature Communications (2023), 14: 2095.

5.1 Synopsis

One of the most pressing challenges when studying phenotypic plasticity concerns the identification of the exact molecular mechanisms that facilitate the sensing of the environmental inputs which are subsequently relayed into divergent developmental trajectories. While epigenetic modifications are often invoked to explain this process, it remains unclear to what extent such modifications may be involved and to what degree they may encode memories of environmental exposures on the molecular level.

In this study, we demonstrate that the feeding plasticity of *Pristionchus pacificus* nematodes is determined by histone 4 lysine 5 and 12 acetylation (H4K5/12ac). In early larval stages the acetylation of H4K5/12 provides a permissive chromatin state that is susceptible to induction during a critical window of sensitivity to environmental stimuli. Later in development, the deacetylation of H4K5/12 represses the expression of the genetic switch which underlies feeding plasticity, terminating the window of environmental sensitivity. Interestingly, by using geometric morphometrics to track the development of feeding structures throughout the worm's life, we were able to show that the alternative mouth anatomies start to emerge exactly when this window of environmental sensitivity is terminated in late larval stages. Additionally, we found that the inhibition of deacetylases causes prior developmental trajectories to become fixed, demonstrating that such epigenetic modifications provide a form of environmental memory, since inputs that were perceived and encoded by young juveniles are carried on into their adulthood. Furthermore, we provide evidence that these regulatory modifications derived from an ancient mechanism that controls developmental speed. Altogether, our results show that H4K5/12ac enables epigenetic regulation of feeding plasticity, greatly extending our knowledge of *Pristionchus's* developmental switch network and yielding important insights into its evolutionary history.

5.2 Own contribution

I was involved in designing and executing the morphometric analyses in this paper and wrote a small corresponding aspect of the manuscript. My contribution is 5%.

6. *nhr-10* does not control mouth development in *P. pacificus*

Theska T, Renahan T, Sommer RJ. *Starvation resistance in the nematode Pristionchus pacificus requires a conserved supplementary nuclear receptor.*
Submitted to *Zoological Letters* (2023).

6.1 Synopsis

Nuclear hormone receptors (NHRs) are a deeply-conserved superfamily of metazoan transcription factors which underwent a massive expansion in nematodes. Many of these worms possess hundreds of *nhr* genes for the majority of which the biological functions remain completely unexplored. However, recent studies showed that two sister receptors, *Ppa*-NHR-1 and *Ppa*-NHR-40, are crucial down-stream regulators of feeding plasticity in the diplogastrid model organism *Pristionchus pacificus*.

In this study, we functionally characterized *Ppa*-NHR-10, the closest paralog to *Ppa*-NHR-1 and *Ppa*-NHR-40, aiming to demonstrate that it too regulates aspects of feeding plasticity. We used CRISPR/CAS9 to create knock-outs of this receptor and applied geometric morphometrics and unsupervised clustering to characterize mutant phenotypes. Surprisingly, we found that *Ppa*-NHR-10 does not regulate phenotypic plasticity in feeding structures. Instead, comparative transcriptomics revealed that its regulatory target genes are involved in lipid catabolism. We hypothesized that their mis-regulation could affect the survival of mutant worms during starvation, where lipid catabolism is often essential. Using survival assays, we found that mutants indeed show decreased starvation resistance, both as adults and as dauer larvae. We also, for the first time, characterized genome-wide changes to the transcriptional landscape in *P. pacificus* when exposed to 24hrs of acute starvation, and found that *Ppa*-NHR-10 contributes to their regulation. Taken together, we demonstrate that *Ppa*-NHR-10 is broadly required for starvation resistance throughout the *P. pacificus* life cycle and that it regulates different biological processes than its closest paralogs *Ppa*-NHR-1 and *Ppa*-NHR-40. This also corroborated the hypothesis that its sister receptors were specifically co-opted into the feeding-structure development GRN during diplogastrid evolution.

6.2 Own contribution

I conceptualized the study, designed and conducted most of the experiments, analyzed and visualized all data, and wrote the manuscript. My contribution is 90%.

7. *nhr-1* and *nhr-40* do not control mouth development in *C. elegans*

Theska T, Sommer RJ (2023). *Feeding-structure morphogenesis in "rhabditid" and diplogastrid nematodes is not controlled by a conserved genetic module.*
Submitted to *Evolution & Development* (2023).

7.1 Synopsis

Nematodes are characterized by extraordinarily high numbers of nuclear hormone receptors (NHRs) in their genomes. Most of these NHRs are entirely nematode-specific (called *supnrs*) and the vast majority of them remains uncharacterized in terms of the biological functions they control. Additionally, only a handful of these *supnrs* is deeply conserved across distantly related nematode species, such as *C. elegans* and *P. pacificus*. Excitingly, previous studies revealed that two of the very few conserved *supnrs* (NHR-1 and NHR-40) regulate the morphogenesis of novel and phenotypically plastic feeding structures in *P. pacificus* and other diplogastrids. It is therefore possible that these receptors already controlled feeding-structure morphogenesis, before moveable cuticular teeth evolved from immovable flaps in the "rhabditid"-diplogastrid transition.

To investigate this possibility, we created knockout mutants for both of these receptors in the "rhabditid" *C. elegans*, which has immovable flaps in its mouth. Geometric morphometric analyses revealed that the mouth morphology of these mutants was not affected by the loss of either receptor, indicating that NHR-1 and NHR-40 are not involved in "rhabditid" stoma morphogenesis. We confirmed that this observation is not due to species-specific expression patterns. In fact, in both *C. elegans* and *P. pacificus*, each of these receptors is expressed in the pharyngeal muscles cells which secrete the flaps and the teeth. Comparative transcriptomics revealed that, in *C. elegans*, NHR-1 and NHR-40 do not regulate any of the cuticle-modifying enzymes which are known to be their targets in *P. pacificus*. Instead, they broadly regulate metabolic pathways and stress responses in "rhabditids". These results confirm previous observations that conserved *supnrs* can functionally diversify across distantly related species, and they suggest that NHR-1 and NHR-40 acquired their mouth-morphogenesis related functions only during the evolution of teeth in diplogastrids.

7.2 Own contribution

I was involved in conceptualizing the study and I solely conducted all experiments and analyses. I wrote the manuscript and created all visualizations. My contribution is 95%.

VI. Discussion

The molecular mechanisms that control phenotypic plasticity in the novel feeding structures of *P. pacificus* have been the focus of extensive genetic research over the last decade [32,54]. As a result, we are equipped with a detailed draft of the modular developmental switch network which underlies the polyphenism (Fig. 2a). While this is a stunning achievement for evo-devo, the draft requires extension on two fronts. The exact structural components making up the nematodes' feeding structures (i.e. the most downstream components of the phenotype-execution module) were unknown, just like the epigenetic mechanisms involved in the integration of environmental information upstream of the plasticity-switch module (Fig. 2a). Yet, addressing both of these issues is of great importance if we strive to comprehensively understand the switch network from its beginning (perception and molecular encoding of environmental information) to its end (morphogenesis of the feeding structures). Further, the evolutionary history of this switch network has only begun to be explored: while members of the plasticity-switch module were found to be young and diplogastrid-specific genes [32] that evolved their plasticity-related functions in association with the origin of the novel feeding structures, it remained unclear whether the ancient members of the phenotype-execution module acquired their morphogenesis-related functions in a similar manner, or whether they already performed these functions *before* the novel feeding structures emerged. Additionally, since neither the putatively involved epigenetic mechanisms nor the precise structural molecules of the buccal capsule were known, nothing could be deduced about their evolutionary incorporation into the developmental switch network either.

The work I present in this thesis (which was often collaborative) addresses these two aspects of the feeding plasticity in *P. pacificus*: it fills some of the gaps in the architectural draft of the developmental switch network and it explores how this network came to be (Fig. 3).

First, I established a standardized and repeatable approach to quantify anatomical differences (within and across nematode species) by combining geometric morphometrics with unsupervised clustering. The integration of this approach into the methodological toolkit of evo-devo facilitated well-rounded comparative studies of phenotypic changes in nematode development and evolution. In the first study of this kind, we used high-resolution developmental transcriptomics to identify *dpy-6* - a gene

with oscillating expression patterns that coincide with the molting cycle (and thus cuticle generation) of the worm - as a likely constituent of the feeding structures in *P. pacificus*. We combined CRISPR-based gene knockouts with geometric morphometric quantification of mutant phenotypes, and were able to reveal that the loss of *dpy-6* leads to an altered body-wall cuticle morphology and, more importantly, consistent malformations of the cheilostom (Fig. 1a,b). Hence, we described the first protein constituent of nematode feeding structures and thus identified the first structural target of the phenotype-execution module - the downstream-most part of the developmental switch network controlling feeding plasticity in *P. pacificus* (Fig 3a). Excitingly, the observed phenotypic effects of *dpy-6*-inactivation were consistent in both *P. pacificus* and *C. elegans*, indicating that DPY-6's function as a structural constituent of the buccal capsule predates the "rhabditid"-diplogastrid transition and the evolution of the novel feeding morphologies that characterize the latter (Fig. 3b). Additionally, *dpy-6* is the first member of the developmental switch network in *P. pacificus* that is both phylogenetically old and functionally conserved. As it constitutes the most downstream element of the network known so far, it appears that the remainder of the phenotype-execution module (to which it belongs) and the upstream modules were built "on top" of *dpy-6* and its pre-existing function in mouth formation (Fig. 3a,b).

Next, we wondered whether (and which) epigenetic mechanisms may be involved in the integration of environmental information upstream of the plasticity-switch module. Via various developmental-resolved biochemical assays we were able to identify histone 4 lysine 5/12 (H4K5/12ac) acetylation as the crucial epigenetic regulator of switch gene expression (Fig. 3a). In combining the results with geometric morphometric analyses of the phenotypic transitions in mouth morphology throughout *P. pacificus's* post-embryonic ontogeny, we were able to demonstrate that the alternative mouth morphologies start to emerge in late larval stages, exactly around the time when deacetylation of H4K5/12 starts to repress the activity of the plasticity-switch module, thereby terminating the phase of sensitivity to environmental information and initiating the phase of phenotype-execution. Additionally, we demonstrate that H4K5/12ac is an epigenetic mechanisms that is also present in *C. elegans* as well as other ecdysozoans, such as *Drosophila melanogaster*, and found it to be an ancient regulatory mechanism that governs developmental speed. Thus, similar to *dpy-6*, H4K5/12ac predates the "rhabditid"-diplogastrid transition and has a general function that is deeply-conserved across hundreds of million years of evolution (Fig. 3b). Yet, our findings also imply that

H4K5/12ac gained an additional function specifically during diplogastrid evolution: the developmental regulation of environmental sensitivity upstream of the plasticity-switch module. Thus, in the context of the developmental switch network, H4K5/12ac is an old mechanism that was co-opted into a novel regulatory context during the evolution of novel feeding structures in diplogastrids (Fig. 3a,b).

Lastly, I focused on the two transcription factors which sit at the center of the phenotype-execution module in *P. pacificus* (NHR-1 and NHR-40) [32,89] and form a deeply-conserved cluster of *supnrs* with NHR-10 (Fig. 2a,b). My interest in these receptors was two-fold: first, I wondered whether NHR-10, and therefore an entire conserved *supnr* cluster, is involved in the phenotype-execution in *P. pacificus's* feeding plasticity (Fig. 2a,b). I found that genetic inactivation of NHR-10 does not lead to any quantitative or qualitative differences in overall mouth morphology, and only minimally - and probably indirectly - impacts the ratio of predatory animals formed in the population. Using exploratory transcriptomics, I found that its regulatory target genes are involved in lipid catabolism and, using condition-specific survival assays, I was able to demonstrate that young adult *nhr-10* mutants show drastically reduced survival rates specifically under starvation - a condition where lipid catabolism is crucial. I further confirmed that not only adults, but also dauer larvae were negatively affected by the loss of NHR-10 in terms of longterm starvation survival (as well as dauer entry rates and dauer exit speed). Thus, I demonstrated that NHR-10 is broadly required for starvation resistance and involved in dauer physiology in *P. pacificus* (Fig. 3c). Interestingly, neither of these functions is known to be controlled by the one-to-one ortholog of NHR-10 in *C. elegans*, which regulates a metabolic shunt pathway instead [95]. Thus, neither its sister paralogs in *P. pacificus*, nor its ortholog in *C. elegans,* were adequate predictors of NHR-10's biological roles - a finding that was unexpected, since these receptors are some of the very few *supnrs* which are deeply-conserved across "rhabditids" and diplogastrids. Intrigued by this observation, I then wondered whether the lack of functional conservation across species may also hold true for NHR-1 and NHR-40. Quantitative analyses of the mouth morphology of *C. elegans* which lacked functional copies of NHR-1 or NHR-40 revealed no discernible mutant phenotypes - the opposite to the stark effects such mutations have in *P. pacificus*. I was able to show that this was not due to interspecific differences in expression patterns, as both receptors are broadly expressed in the pharyngeal target tissues (amongst others) in both species. Comparative transcriptome analyses showed that both of the NHRs co-regulate the

expression of targets genes which are highly species-specific, and I found that they regulate genes involved in metabolic and immune responses in the "rhabditid" in *C. elegans* (Fig. 3c). Thus, this investigation confirmed my original observation of interspecies differences in the biological function of *supnrs*. Again, neither the closest paralogs of these receptors in the same species, nor the one-to-one orthologs in another, were suitable predictors of their biological functions. More importantly, these two studies together confirmed that it was not the entire NHR-1/-10/-40 cluster of *supnrs* that acquired plasticity-related functions in *P. pacificus*, but only NHR-1 and NHR-40 (Fig. 3a,c). The absence of any mouth-morphogenesis related functions of these NHRs in *C. elegans* indicates that they did not regulate a phenotype-execution module in the last common ancestor of "rhabditids" and diplogastrids, meaning that this module likely evolved during the evolution of the phenotypically plastic teeth in diplogastrids (Fig. 3b).

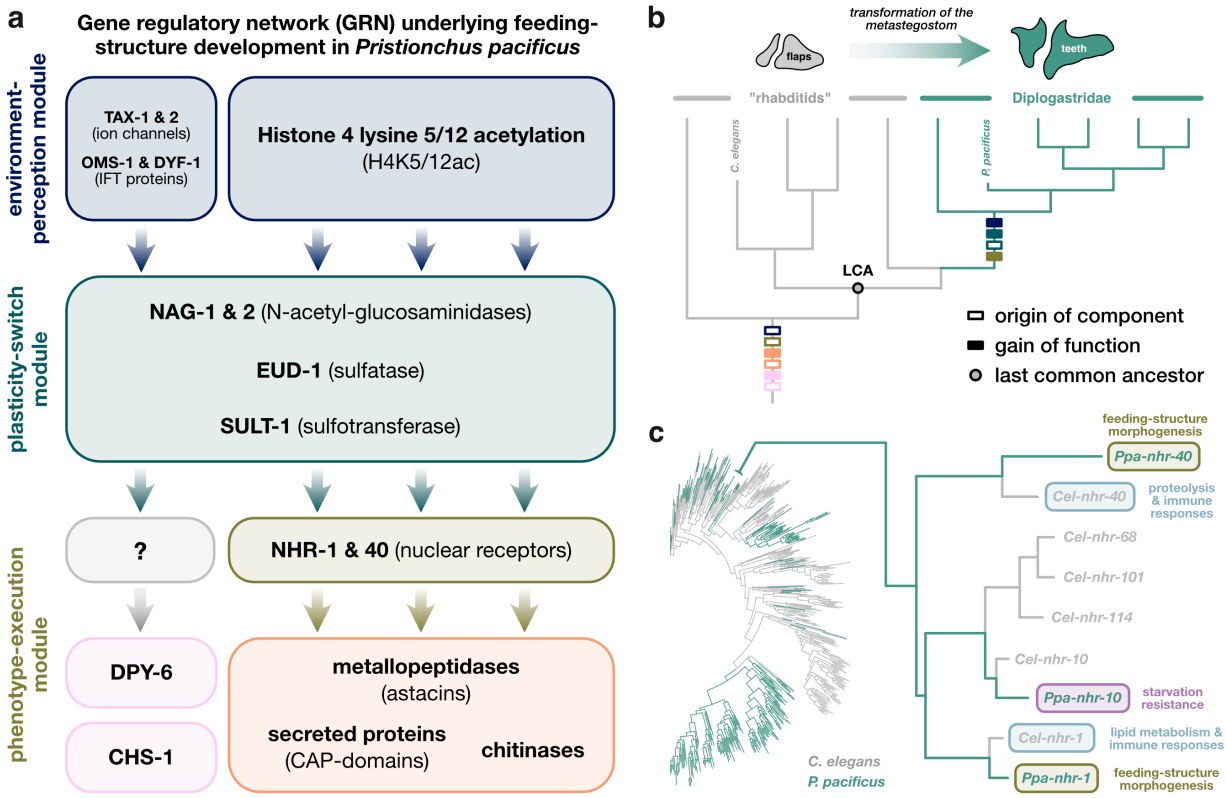


Figure 3. Expansions of the GRN that control feeding-structure development in *P. pacificus* and insights into their evolutionary history. a. Updated version of the GRN (compare to Figure 2) based on the research presented in this thesis. Note that CHS-1 (chitin-synthase 1) was identified as a member of the phenotype-execution module in work that is unrelated to this thesis. **b.** Schematic representation of the evolutionary history of the feeding-structure development GRN mapped onto the phylogeny of "rhabditid" and diplogastrid nematodes. Empty boxes on the branches of the phylogeny indicate the origin of the molecular players; filled boxes indicate the gain of their functions related to feeding-structure development in Diplogastridae. Color coding of the boxes corresponds to modules depicted in panel a. **c.** Unlike its sister paralogs, *Ppa-nhr-10* does not regulate phenotype execution. Figure was modified from Theska *et al.* (2023) and Theska & Sommer (2023), see Results (pp. 30,31).

Taken together, the developmental switch network controlling feeding plasticity of *P. pacificus* assembled via a range of principles. The environmental-perception module (which integrates environmental information and relays it towards downstream aspects of the switch network) works largely through H4K5/12 acetylation, an ancient epigenetic mechanism which adopted a novel regulatory function during the evolution of teeth in the lineage leading to today's diplogastrids. The plasticity-switch module (which interprets the received environmental inputs and initiates the development of a corresponding feeding morphology) is made of young (i.e. novel) genes that appeared alongside the morphological novelty and immediately gained their plasticity-related functions [32,84]. Lastly, the phenotype-execution module (which governs mouth formation) includes a mix of old proteins that acquired new biological functions (NHR-1 and NHR-40), and old proteins that maintained old functions (DPY-6). This also demonstrates that none of these three modules - not even the most downstream one - were already pre-made for diplogastrids to "just built upon" when their developmental switch network was evolutionarily assembled.

This is indeed surprising as the classical literature in evo-devo proposes that the downstream aspects of gene regulatory networks (GRNs) - usually comprising the structural components of the phenotype and their immediate transcriptional regulators - are more likely to be preserved (both structurally *and* functionally) in evolution than the more upstream aspects of GRNs [96,97]. The latter are thought to undergo frequent compositional or functional changes due to phenomena like developmental systems drift [48,52,96]. Thus, it seems that some of our current presuppositions of how phenotypes and their associated GRNs evolve turned out to be - at least in the case of diplogastrid feeding structures - oversimplified premises rather than general guiding principles underlying phenotypic transformations.

This leads me to address the elephant (or major worm) in the room: we currently know neither which of the classified GRN-assembly principles (*de novo* evolution of molecules with new functions, co-option of preexisting molecules into novel functional contexts, integration of preexisting molecules with conserved functions), nor which of the identified molecular players that make up the switch network in diplogastrids (genes, gene classes, epigenetic mechanisms, etc.), are *generally* involved in the evolution of phenotypic plasticity and morphological novelty. We just identified them; how could we? But this could be considered a grand challenge ahead of nematode evo-devo. If we truly want to identify general principles in the plasticity-led evolution of

novelties, such as the molecular processes which facilitated the flap-to-tooth transformation during the "rhabditid"-diplogastrid transition, we need another example to compare our new results to. Only this way will we be able to distinguish the common from the unique. This however necessitates that the candidate satellite system allows for similarly-detailed investigations of these phenomena. Fortunately, we don't need to search far-away branches of the animal phylogeny to find a suitable contender.

During a recent field trip, nematologists found a new worm species on logs of dead Japanese red pine. *Bursaphelenchus sinensis* displays an exciting behavior [98]: using their stylets - syringe-like cuticular feeding structures that characterize these tylenchid nematodes - individuals can suck out fluids from fungal hyphae running through plant matter. However, some animals in the population - notably ones with more massive stylets - prefer a different food source: nematodes! In further lab experiments the researchers were able to confirm that *B. sinensis* is indeed another nematode with a bimodal feeding polyphenism (mycophagous *vs.* predatory) [98]. Both morphs have a movable stylet as the feeding structure used to penetrate foods. Yet the predatory animals have enlarge stylets with engorged stylet "knobs" that act as attachment sites for massive pharyngeal muscles. Additionally, the stylet itself is more massive and has a bigger lumen in the predatory morphs, allowing them to penetrate more sizable and robust prey (such as nematodes). Excitingly, tylenchid stylets have already been studied in great detail and it is known that they evolved from the same six feeding structures that also make up the buccal capsules of "rhabditids" and diplogastrids [99]. Based on this homology we can deduce that the stylets are another, but very different, case of individualized novelties. The posterior base of the stylet - the "knob" - is derived from the stegostom [99], and is thus largely homologous to the movable teeth of diplogastrids. The shaft and the cone of the stylets - the anterior part which penetrates the food - consists of a thickened and elongated gymnostom [99]. The anterior-most feeding structure of tylenchids is a cuticular channel that directs the stinging motion of the stylet toward the prey. This structure - the "guiding apparatus" - is homologous to the cheilostom [99]. Additionally, while many tylenchids have monomorphic (that is non-plastic) mouths, polyphenic species similar to *B. sinensis* have been reported before, yet have not been carefully studied [98]. And even better, multiple species are known to be monomorphic - possible secondarily.

Thus, the tylenchid stylet may represent another example of plasticity-led evolution of an individualized novelty [80]. Potentially, a phenotypically plastic stylet

evolved at the base of this clade, leading to alternative feeding phenotypes. Subsequently, some species may have canalized one morph, leading to plant-parasitic and mycophagous species on one hand, and predatory species on another (similar to the phenotypically plastic *vs.* the secondarily monomorphic bacterivores or predators in diplogastrids) [98]. Future studies of the feeding structures and strategies in tylenchids will show whether they too support the facilitator hypothesis, and, eventually, molecular genetic studies will reveal the underlying developmental switch network.

Here, I propose that by elucidating the architecture and evolutionary history of the switch network in *B. sinensis,* and comparing these findings to the case of the flap-tooth transformation of diplogastrids, we will learn more about the general processes that guide the evolution of plastic novelties. Does *B. sinensis's* switch network also reflect a colorful modular mix of ancient processes, novel genes, and freshly-acquired regulatory capabilities? Do these worms encode environmental information in the language of H4K5/12ac? Did *B. sinensis's* switch genes also derive from recent duplications or are they ancient for a change? Are NHRs - possibly even non-conserved ones - involved in regulating phenotype execution? And would *dpy-6* mutants of *B. sinensis* starve to death due to grossly malformed guiding apparatuses that render their stylets useless? All of these insights will inform our mechanistic understanding of phenotypic change in these animals, and truly support evo-devo's stance that nothing in evolutionary morphology makes sense except in the light of developmental plasticity.

VII. References

1. Amundson R. The changing role of the embryo in evolutionary thought: roots of evo-devo. Cambridge University Press; 2005.
2. Jacob F. Evolution and tinkering. Science. 1977;196:1161–1166.
3. Gould SJ. Ontogeny and Phylogeny. Harvard University Press; 1977.
4. Gould SJ, Lewontin RC. The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. Proceedings of the Royal Society B: Biological Sciences. 1979;205:581-598.
5. Hall BK. Evolutionary developmental biology (Evo-Devo): Past, present, and future. Evolution: Education and Outreach. 2012;5:184–193.
6. Nuño de la Rosa L, Müller GB. A Reference Guide to Evo-Devo. In: Nuño de la Rosa L, Müller GB, editors. Evolutionary Developmental Biology: A Reference Guide. Springer, Cham.; 2021, pp. 3–12.
7. Sommer RJ. The future of evo–devo: model systems and evolutionary theory. Nature Reviews Genetics. 2009;10:416–422.
8. Nuño de la Rosa L, Müller GB (editors). Evolutionary Developmental Biology: A Reference Guide. Springer; 2021.
9. Amundson R. Form and Function in Evo-Devo. In: Nuño de la Rosa L, Müller GB, editors. Evolutionary Developmental Biology: A Reference Guide. Springer, Cham.; 2021, pp. 457–467.
10. Wagner GP. Homology, Genes, and Evolutionary Innovation. Princeton University Press; 2014.
11. Abzhanov A. von Baer's law for the ages: lost and found principles of developmental evolution. Trends in Genetics. 2013;29:712–722.
12. Richardson MK. Theories, laws, and models in evo-devo. Journal of Experimental Zoology Part B: Molecular and Developmental Evolution. 2022;338:36–61.
13. Owen R. Lectures on comparative anatomy and physiology of the invertebrate animals. Delivered at the Royal College of Surgeons in 1843. Longman, Brown, Green and Longman; 1843.
14. Darwin CR. On the Origin of Species. John Murray; 1859.
15. Huxley J. Evolution: The Modern Synthesis. Allen and Unwin; 1942.
16. Mayr E. What Evolution is. BasicBooks; 2001.
17. Laland KN, Uller T, Feldman MW, Sterelny K, Müller GB, Moczek A, et al. The extended evolutionary synthesis: its structure, assumptions and predictions. Proceedings of the Royal Society B: Biological Sciences. 2015;282:20151019.
18. Lewontin RC. Adaptation. Scientific American. 1978;3:156–169.
19. de Beer GR. Embryos and Ancestors. Oxford University Press; 1940.
20. de Beer GR. Homology, an unsolved problem. Oxford University Press; 1971.
21. DiFrisco J. Developmental Homology. In: Nuño de la Rosa L, Müller GB, editors. Evolutionary Developmental Biology: A Reference Guide. Springer, Cham.; 2021, pp. 85–97.
22. Sommer RJ. Homology and the hierarchy of biological systems. BioEssays. 2008;30:653–658.
23. Wagner GP, Lynch VJ. Evolutionary novelties. Current Biology. 2010;20:R48–52.
24. Mayr E. Animal species and Evolution. Havard University Press; 1963.
25. Moczek AP. On the origins of novelty in development and evolution. BioEssays. 2008;30:432–447.

26. Pigliucci M. What, if anything, is an evolutionary novelty? Philosophy of Science. 2008;75:887–898.

27. Peterson T, Müller GB. What is evolutionary novelty? Process versus character based definitions. Journal of Experimental Zoology Part B: Molecular and Developmental Evolution. 2013;320:345–350.

28. Müller GB. Developmental Innovation and Phenotype Novelty. In: Nuño de la Rosa L, Müller GB, editors. Evolutionary Developmental Biology: A Reference Guide. Springer, Cham.; 2021, pp. 69–84.

29. Mivart GJ. On the Genesis of Species. Cambridge University Press; Reissue Edition 2010.

30. Platnick NI, Rosen DE. Popper and evolutionary novelties. History and Philosophy of the Life Sciences. 1987;5–16.

31. Wagner GP. Evolutionary innovations and novelties: Let us get down to business! Zoologischer Anzeiger - A Journal of Comparative Zoology. 2015;256:75–81.

32. Sommer RJ. Phenotypic plasticity: from theory and genetics to current and future challenges. Genetics. 2020;215:1–13.

33. West-Eberhard MJ. Phenotypic plasticity and the origins of diversity. Annual Review of Ecology and Systematics. 1989;20:249–278.

34. West-Eberhard MJ. Developmental plasticity and evolution. Oxford University Press; 2003.

35. Moczek AP, Sultan S, Foster S, Ledón-Rettig C, Dworkin I, Nijhout HF, et al. The role of developmental plasticity in evolutionary innovation. Proceedings of the Royal Society B: Biological Sciences. 2011;278:2705–2713.

36. Paaby AB, Testa ND. Developmental Plasticity and Evolution. In: Nuño de la Rosa L, Müller GB, editors. Evolutionary Developmental Biology: A Reference Guide. Springer, Cham.; 2021, pp. 1073–1086.

37. Bradshaw AD. Evolutionary significance of phenotypic plasticity in plants. Advances in Genetics. 1965;13:115–155.

38. Uller T, Feiner N, Radersma R, Jackson IS, Rago A. Developmental plasticity and evolutionary explanations. Evolution & Development. 2020;22:47–55.

39. Howard RJ, Giacomelli M, Lozano-Fernandez J, Edgecombe GD, Fleming JF, Kristensen RM, et al. The Ediacaran origin of Ecdysozoa: integrating fossil and phylogenomic data. Journal of the Geological Society. 2022;179:jgs2021-107.

40. Van Den Hoogen J, Geisen S, Routh D, Ferris H, Traunspurger W, Wardle DA, et al. Soil nematode abundance and functional group composition at a global scale. Nature. 2019;572:194–198.

41. Lambshead PJD, Boucher G. Marine nematode deep-sea biodiversity–hyperdiverse or hype? Journal of Biogeography. 2003;30:475–485.

42. Blaxter M. Nematodes: the worm and its relatives. PLOS Biology. 2011;9:e1001050.

43. Kandoth PK, Mitchum MG. War of the worms: how plants fight underground attacks. Current Opinion in Plant Biology. 2013;16:457–463.

44. Puthiyakunnon S, Boddu S, Li Y, Zhou X, Wang C, Li J, et al. Strongyloidiasis—an insight into its global prevalence and management. PLOS neglected tropical diseases. 2014;8:e3018.

45. Giribet G, Edgecombe GD. The Invertebrate Tree of Life. Princeton University Press; 2020.

46. Corsi AK, Wightman B, Chalfie M. A transparent window into biology: a primer on *Caenorhabditis elegans*. Genetics. 2015;200:387–407.

47. Shatilovich A, Gade VR, Pippel M, Hoffmeyer TT, Tchesunov AV, Stevens L, et al. A novel nematode species from the Siberian permafrost shares adaptive mechanisms for cryptobiotic survival with *C. elegans* dauer larva. PLOS Genetics. 2023;19:e1010798.

48. Haag ES, Fitch DH, Delattre M. From "the worm" to "the worms" and back again: the evolutionary developmental biology of nematodes. Genetics. 2018;210:397–433.

49. Hong RL, Sommer RJ. *Pristionchus pacificus*: a well-rounded nematode. BioEssays. 2006;28:651–659.

50. Sommer RJ (editors). *Pristionchus pacificus*: a nematode model for comparative and evolutionary biology. Brill; 2015.

51. Moss EG. Heterochronic genes and the nature of developmental time. Current Biology. 2007;17:R425–434.

52. Sommer RJ. Evolution of regulatory networks: nematode vulva induction as an example of developmental systems drift. In: Soyer OS, editor. Evolutionary Systems Biology. Springer; 2012, p. 79–91.

53. Viney M, Diaz A. Phenotypic plasticity in nematodes: evolutionary and ecological significance. Worm. 2012; 1:98–106.

54. Sommer RJ, Dardiry M, Lenuzzi M, Namdeo S, Renahan T, Sieriebriennikov B, et al. The genetics of phenotypic plasticity in nematode feeding structures. Open Biology. 2017;7:160332.

55. Sommer RJ, Ogawa A. Hormone signaling and phenotypic plasticity in nematode development and evolution. Current Biology. 2011;21:R758–766.

56. Cassada RC, Russell RL. The dauerlarva, a post-embryonic developmental variant of the nematode *Caenorhabditis elegans*. Developmental Biology. 1975;46:326–342.

57. Hu PJ. Dauer. In: The *C. elegans* research community, editor. WormBook: The Online Review of *C. elegans* Biology. 2007.

58. Braeckman B, Houthoofd K, Vanfleteren J. Intermediary metabolism. In: The *C. elegans* research community, editor. WormBook: The Online Review of *C. elegans* Biology. 2009.

59. Erkut C, Gade VR, Laxman S, Kurzchalia TV. The glyoxylate shunt is essential for desiccation tolerance in *C. elegans* and budding yeast. Elife. 2016;5:e13614.

60. Hibshman JD, Doan AE, Moore BT, Kaplan RE, Hung A, Webster AK, et al. daf-16/FoxO promotes gluconeogenesis and trehalose synthesis during starvation to support survival. Elife. 2017;6:e30057.

61. Hibshman JD, Clegg JS, Goldstein B. Mechanisms of desiccation tolerance: themes and variations in brine shrimp, roundworms, and tardigrades. Frontiers in Physiology. 2020;11:592016.

62. Klass M, Hirsh D. Non-ageing developmental variant of *Caenorhabditis elegans*. Nature. 1976;260:523–525.

63. Renahan T, Lo W, Werner MS, Rochat J, Herrmann M, Sommer RJ. Nematode biphasic 'boom and bust' dynamics are dependent on host bacterial load while linking dauer and mouth-form polyphenisms. Environmental Microbiology. 2021;23:5102–5113.

64. Antebi A. Nuclear receptor signal transduction in *C. elegans*. In: The *C. elegans* research community, editor. WormBook: The Online Review of *C. elegans* Biology. 2015.

65. Motola DL, Cummins CL, Rottiers V, Sharma KK, Li T, Li Y, et al. Identification of ligands for DAF-12 that govern dauer formation and reproduction in *C. elegans*. Cell. 2006;124:1209–1223.

66. Renahan T, Sommer RJ. Nematode interactions on beetle hosts indicate a role of mouth-form plasticity in resource competition. Frontiers in Ecology and Evolution. 2021;9:752695.

67. Renahan T, Sommer RJ. Multidimensional competition of nematodes affects plastic traits in a beetle ecosystem. Frontiers in Cell and Developmental Biology. 2022;10:985831.

68. Schmidt-Rhaesa A (editor). Handbook of Zoology Gastrotricha, Cycloneuralia and Gnathifera. Volume 2: Nematoda. De Gruyter; 2014.

69. Sudhaus W. Preadaptive plateau in Rhabditida (Nematoda) allowed the repeated evolution of zooparasites, with an outlook on evolution of life cycles within Spiroascarida. Palaeodiversity. 2010;3(Suppl):117–130.

70. Wright KA. Functional organization of the nematode's head. In: Croll NA, editor. The Organization of Nematodes. London Academic Press; 1976, p. 71–106.

71. Malakhov VV. Nematodes: Structure, Development, Classification and Phylogeny. Smithsonian Books; 1994.

72. Von Lieven AF, Sudhaus W. Comparative and functional morphology of the buccal cavity of Diplogastrina (Nematoda) and a first outline of the phylogeny of this taxon. Journal of Zoological Systematics and Evolutionary Research. 2000;38:37–63.

73. De Ley P, Van de Velde MC, Mounport D, Baujard P, Coomans A. Ultrastructure of the stoma in Cephalobidae, Panagrolaimidae and Rhabditidae, with a proposal for a revised stoma terminology in Rhabditida (Nematoda). Nematologica 1995;41:153–182.

74. Burr A, Baldwin JG. The nematode stoma: homology of cell architecture with improved understanding by confocal microscopy of labeled cell boundaries. Journal of Morphology. 2016;277:1168–1186.

75. Harry CJ, Messar SM, Ragsdale EJ. Comparative reconstruction of the predatory feeding structures of the polyphenic nematode *Pristionchus pacificus*. Evolution and Development, 2022;24:16-36.

76. Werner MS, Claaßen MH, Renahan T, Dardiry M, Sommer RJ. Adult influence on juvenile phenotypes by stage-specific pheromone production. iScience. 2018;10:123–134.

77. Lenuzzi M, Witte H, Riebesell M, Rödelsperger C, Hong RL, Sommer RJ. Influence of environmental temperature on mouth-form plasticity in *Pristionchus pacificus* acts through daf-11-dependent cGMP signaling. Journal of Experimental Zoology Part B: Molecular and Developmental Evolution. 2023;340:214–224.

78. Werner MS, Sieriebriennikov B, Loschko T, Namdeo S, Lenuzzi M, Dardiry M, et al. Environmental influence on *Pristionchus pacificus* mouth form through different culture methods. Scientific Reports. 2017;7:7207.

79. Lightfoot JW, Wilecki M, Rödelsperger C, Moreno E, Susoy V, Witte H, et al. Small peptide–mediated self-recognition prevents cannibalism in predatory nematodes. Science. 2019;364:86–89.

80. Susoy V, Ragsdale EJ, Kanzaki N, Sommer RJ. Rapid diversification associated with a macroevolutionary pulse of developmental plasticity. Elife. 2015;4:e05463.

81. Moreno E, Lenuzzi M, Rödelsperger C, Prabh N, Witte H, Roeseler W, et al. DAF-19/RFX controls ciliogenesis and influences oxygen-induced social behaviors in *Pristionchus pacificus*. Evolution & Development. 2018;20:233–243.

82. Serobyan V, Xiao H, Namdeo S, Rödelsperger C, Sieriebriennikov B, Witte H, et al. Chromatin remodelling and antisense-mediated up-regulation of the developmental switch gene eud-1 control predatory feeding plasticity. Nature Communications. 2016;7:12337.

83. Ragsdale EJ, Müller MR, Rödelsperger C, Sommer RJ. A developmental switch coupled to the evolution of plasticity acts through a sulfatase. Cell. 2013;155:922–933.

84. Sieriebriennikov B, Prabh N, Dardiry M, Witte H, Röseler W, Kieninger MR, et al. A developmental switch generating phenotypic plasticity is part of a conserved multi-gene locus. Cell Reports. 2018;23:2835–2843.

85. Namdeo S, Moreno E, Rödelsperger C, Baskaran P, Witte H, Sommer RJ. Two independent sulfation processes regulate mouth-form plasticity in the nematode *Pristionchus pacificus*. Development. 2018;145:dev166272.

86. Bui LT, Ivers NA, Ragsdale EJ. A sulfotransferase dosage-dependently regulates mouthpart polyphenism in the nematode *Pristionchus pacificus*. Nature Communications. 2018;9:4119.

87. Igreja C, Sommer RJ. The role of sulfation in nematode development and phenotypic plasticity. Frontiers in Molecular Biosciences. 2022;9:838148.

88. Kieninger MR, Ivers NA, Rödelsperger C, Markov GV, Sommer RJ, Ragsdale EJ. The nuclear hormone receptor NHR-40 acts downstream of the sulfatase EUD-1 as part of a developmental plasticity switch in *Pristionchus*. Current Biology. 2016;26:2174–2179.

89. Sieriebriennikov B, Sun S, Lightfoot JW, Witte H, Moreno E, Rödelsperger C, et al. Conserved nuclear hormone receptors controlling a novel plastic trait target fast-evolving genes expressed in a single cell. PLOS Genetics. 2020;16:e1008687.

90. Tao LJ, Seo DE, Jackson B, Ivanova NB, Santori FR. Nuclear hormone receptors and their ligands: metabolites in control of transcription. Cells. 2020;9:2606.

91. Miglioli A, Canesi L, Gomes ID, Schubert M, Dumollard R. Nuclear receptors and development of marine invertebrates. Genes. 2021;12:83.

92. Sural S, Hobert O. Nematode nuclear receptors as integrators of sensory information. Current Biology. 2021;31:4361–4366.

93. Sluder AE, Mathews SW, Hough D, Yin VP, Maina CV. The nuclear receptor superfamily has undergone extensive proliferation and diversification in nematodes. Genome Research. 1999;9:103–120.

94. Robinson-Rechavi M, Maina CV, Gissendanner CR, Laudet V, Sluder A. Explosive lineage-specific expansion of the orphan nuclear receptor HNF4 in nematodes. Journal of Molecular Evolution. 2005;60:577–586.

95. Bulcha JT, Giese GE, Ali MZ, Lee YU, Walker MD, Holdorf AD, et al. A persistence detector for metabolic network rewiring in an animal. Cell Reports. 2019;26:460–468.

96. Mallarino R, Abzhanov A. Paths less traveled: evo-devo approaches to investigating animal morphological evolution. Annual Review of Cell and Developmental Biology. 2012;28:743–763.

97. Carroll SB. Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. Cell, 2008;134:25-36.

98. Kanzaki N, Ekino T, Giblin-Davis RM. Feeding dimorphism in a mycophagous nematode, *Bursaphelenchus sinensis*. Scientific Reports. 2019;9:13956.

99. Baldwin J, Ragsdale E, Bumbarger D. Revised hypotheses for phylogenetic homology of the stomatostylet in tylenchid nematodes. Nematology. 2004;6:623–632.

VIII. Appendix

SCIENTIFIC REPORTS
nature research

OPEN

Crowdsourcing and the feasibility of manual gene annotation: A pilot study in the nematode *Pristionchus pacificus*

Christian Rödelsperger*, Marina Athanasouli, Maša Lenuzzi, Tobias Theska, Shuai Sun, Mohannad Dardiry, Sara Wighard, Wen Hu, Devansh Raj Sharma & Ziduan Han

Nematodes such as *Caenorhabditis elegans* are powerful systems to study basically all aspects of biology. Their species richness together with tremendous genetic knowledge from *C. elegans* facilitate the evolutionary study of biological functions using reverse genetics. However, the ability to identify orthologs of candidate genes in other species can be hampered by erroneous gene annotations. To improve gene annotation in the nematode model organism *Pristionchus pacificus*, we performed a genome-wide screen for *C. elegans* genes with potentially incorrectly annotated *P. pacificus* orthologs. We initiated a community-based project to manually inspect more than two thousand candidate loci and to propose new gene models based on recently generated Iso-seq and RNA-seq data. In most cases, misannotation of *C. elegans* orthologs was due to artificially fused gene predictions and completely missing gene models. The community-based curation raised the gene count from 25,517 to 28,036 and increased the single copy ortholog completeness level from 86% to 97%. This pilot study demonstrates how even small-scale crowdsourcing can drastically improve gene annotations. In future, similar approaches can be used for other species, gene sets, and even larger communities thus making manual annotation of large parts of the genome feasible.

How well can biological knowledge be transferred across species? Are biological functions carried out by the same genes in different organisms? How fast do regulatory networks diverge? In order to address these fundamental questions, more than 20 years ago, the nematode *Pristionchus pacificus* has been introduced as a so-called "satellite" model organism to one of the most successful animal model systems, *Caenorhabditis elegans*[1,2]. Since then, several comparative studies in developmental and ecological contexts have highlighted the importance of developmental system drift as a concept in evolution[3] and have demonstrated that the divergence between *Pristionchus* and *Caenorhabditis* was accompanied by extensive chemical[4–6], genic[7–9], and morphological[10–12] innovations. The establishment of multiple genetic[13,14] and genomic tools and resources[15,16] by Sommer and colleagues motivated an increasing number of independent groups to adapt *P. pacificus* as a model system for comparative studies at a mechanistic level[17–21]. However, reverse genetic approaches based on candidate genes with known functions in *C. elegans*[22,23] have been hampered not only by the huge amount of lineage-specific duplications[23–26], but also by missing and incorrect gene annotations. Traditionally, protein-coding genes are annotated by gene prediction algorithms that model general sequence features of transcription and translation start and end sites, as well as splicing signals[27–29]. This can be complemented with evidence based approaches using transcriptomic and protein homology data[30,31]. While automated annotation pipelines perform reasonably well to be useful for genetic screens[32–34] and evolutionary genomic analyses[35–37], their outcomes by far do not meet the standards of the gene annotations from classical model organisms such as *C. elegans*, *Drosophila melanogaster*, and *Mus musculus* that have been curated over decades by a large research community[38]. In order to make the *P. pacificus* system more tractable for researchers without extensive genomic and phylogenetic expertise, we need to minimize the discrepancy in gene annotation quality between *C. elegans* and *P. pacificus*. To this end, we employed an integrative approach using comparative genomic and transcriptomic data combined with crowdsourcing to improve the *P. pacificus* annotations of *C. elegans* homologs and orthologs. First, we carry out a comparative assessment of 22

Max Planck Institute for Developmental Biology, Department for Integrative Evolutionary Biology, Max-Planck-Ring 9, 72076, Tübingen, Germany. *email: christian.roedelsperger@tuebingen.mpg.de

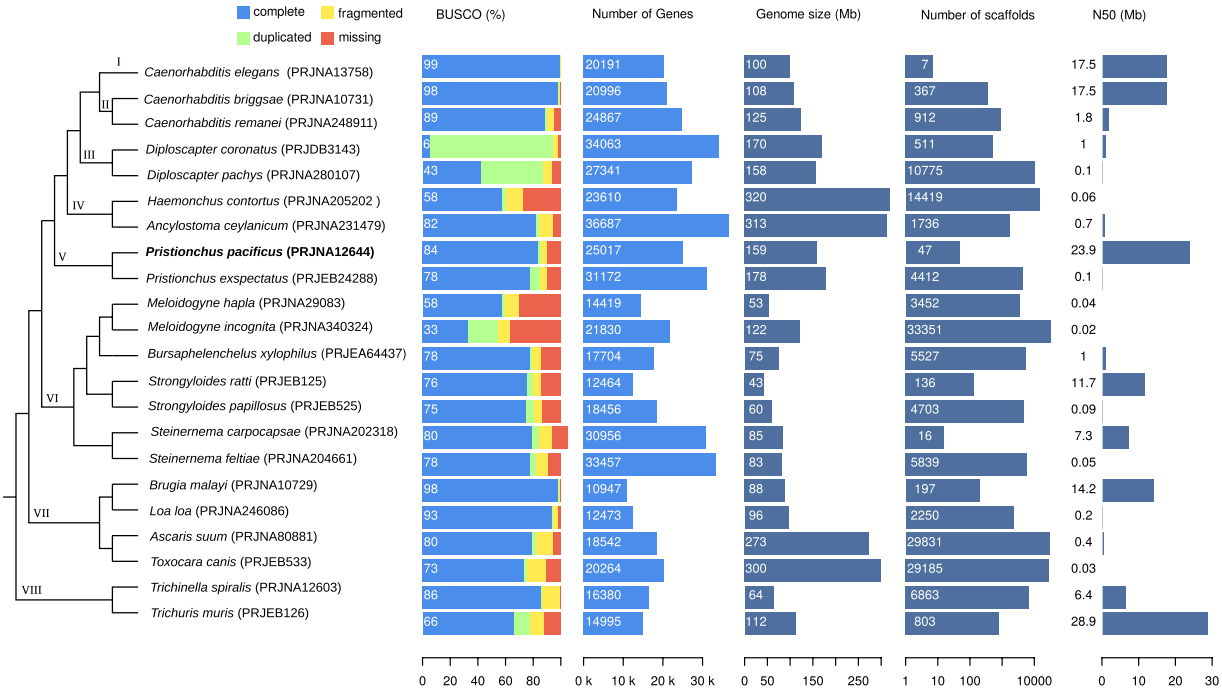


Figure 1. Comparative assessment of nematode genome quality. Genomic data for 22 nematode species was obtained from WormBase ParaSite (release WBPS13) and evaluated based on completeness level of gene annotations and genome assembly contiguity. The barplots show the results of a benchmarking of single copy orthologs (BUSCO[40]) analysis, the number of genes, genome sizes, number of scaffolds, and the N50 measure of assembly contiguity. The genome and annotations of *P. pacificus* exhibit an overall comparatively high quality. The schematic phylogeny is based on phylogenomic analysis of 108 nematodes[39], Roman numerals indicate phylostrata that are used for further analysis.

nematode genomes and demonstrate that *P. pacificus* has one of the best available nematode genomes. Second, we perform a genomewide screen for *C. elegans* genes where homologs and orthologs are not or incorrectly annotated in the *P. pacificus* genome. Third, a community-based manual curation of suspicious gene models reveals thousands of hidden orthologs and missing homologs. This pilot study can be extended to even larger gene sets and communities possibly employing citizen scientists, which would raise the quality of gene annotations to the next level[38].

Results

The quality of nematode draft genomes is highly heterogeneous. To obtain a general overview of the current status of nematode genome quality, we analyzed assemblies and gene annotations of 22 species (Fig. 1). The species were arbitrarily selected to span the diversity of the nematode phylum[39]. We will further use this taxon sampling to perform an analysis of gene age, i.e. phylostratigraphic analysis where each phylostratum is defined by at least two outgroup species to minimize the effect of species-specific gene loss. Nematode genomes range in size between 43 and 320 Mb and contain between 11 and 37 thousand annotated protein-coding genes (Fig. 1). Analyses of assembly features and gene annotations indicate a wide range of qualitative variability. Some genomes are assembled and scaffolded to the level of chromosomes with high degrees of contiguity (the N50 value which is a measure of genome assembly contiguity is up to 29 Mb) whereas others are largely fragmented into up to 33 thousand scaffolds with N50 values below 0.1 Mb (Fig. 1). Similarly, analyses of completeness levels based on benchmarking univeral single copy orthologs (BUSCO[40]) reveal substantial amount of either missing or duplicated genes and it is not totally clear to what extent these differences are of biological or technical nature[41]. In the case of *Diploscapter coronatus*, the apparent high fraction of duplicated genes could either be explained by hybridization of two divergent lineages or a whole genome duplication[42]. The genome of *P. pacificus*, which was generated by assembly from single-molecule, long-read sequencing data and scaffolding with the help of a genetic linkage map[15], shows one of the highest levels of contiguity (47 scaffolds, N50 = 24 Mb). Gene annotations were generated by the MAKER2 pipeline[30,31] which combined gene prediction algorithms, transcriptome data, and protein homology data from other *Pristionchus* species[11,15,43]. The completeness level of gene annotations (BUSCO completeness: 84%) is in the upper range when compared to most other nematode genomes (median 78%, interquartile range (IQR): 68–85%, Fig. 1). This demonstrates the relatively high quality of the current *P. pacificus* assembly and gene annotations.

Complementary genome and transcriptomes reveal potentially missing gene models. The completeness analysis as implemented in the software BUSCO[40] can also be applied to the raw genome assembly of *P. pacificus*. This yielded a combined completeness value of 93% (complete single copy and duplicates) as

Data set	BUSCO (%)				Ref
	Complete Single Copy (+Duplicates)	Duplicate	Fragmented	Missing	
Genome assembly (El Paco assembly)	91.6 (92.9)	1.3	3.1	4.0	[15]
El Paco annotation v1/WS268	84.0 (85.8)	1.8	4.3	9.9	[15]
de novo transcriptome assembly	59.1 (97.1)	38.0	2.6	0.3	[16]
Iso-Seq assembly	48.0 (73.3)	25.3	10.9	15.8	[44]
El Paco annotation v2	95.4 (97.1)	1.7	2.0	0.9	this study

Table 1. Completeness analysis of different *P. pacificus* data set. The high level of duplicates in the two transcriptomic data sets is due to the presence of isoforms.

compared to 86% for the *P. pacificus* gene annotations and indicates towards the presence of incorrectly annotated or missing *C. elegans* orthologs in the genome of *P. pacificus*. Moreover, the fact that a recent *de novo* transcriptome assembly that was based on a strand-specific RNA-seq data set exhibited an even higher combined completeness level of 97% (Table 1) demonstrates even further room for improvement[16]. Finally, single-molecule, long-read transcriptome sequencing data were recently generated for *P. pacificus* which allows a much more accurate definition of gene structures from reference alignments of single reads[44]. However, neither transcriptomic data set was available when the existing gene annotations (version: El Paco annotation v1/WormBase release: WS268) were generated and they could still be used for further improvement.

To systematically identify potentially missing genes in the *P. pacificus* genome, we searched for *C. elegans* genes lacking homologs in the current *P. pacificus* gene annotations (BLASTP e-value $< 10^{-5}$) but having a matching open reading frame in the *de novo* transcriptome assembly (Fig. 2a). While 12,504 (62%) *C. elegans* genes had BLASTP hits in both data sets, 634 (3%) *C. elegans* genes showed only BLASTP hits against the current gene annotations suggesting that these genes are properly annotated but are expressed so weakly that they were not captured in the transcriptome assembly of mixed-stage cultures[45]. Similarly, we identified 526 (3%) *C. elegans* genes that were only found in the transcriptome assembly and therefore represent candidates for missing gene annotations.

Community-based curation identifies missing genes in the *P. pacificus* genome. In order to improve the existing gene annotations, we chose to manually inspect and classify all 526 missing gene candidates in the *P. pacificus* genome browser (http://www.pristionchus.org). Thereby, we recruited and trained colleagues as community annotators, who would be capable to classify a genomic locus and to propose a correction to the existing gene models (see *Methods*). Lists of missing gene candidates were shared in online spreadsheets and documents, which allowed multiple annotators to inspect and correct candidate loci in parallel. 119 (25%) of the 486 non-redundant *P. pacificus* loci were classified as missing genes in predicted UTRs of annotated genes (Fig. 2b). We would speculate that this is caused by the fact that nematode genomes are compact and UTR regions can frequently overlap[45]. This can cause artificial fusion of transcripts during the assembly of RNA-seq data. Consequently, only the largest ORF of such a gene is annotated as protein-coding and the rest is classified as 3′ and 5′ UTR. Alternatively, this problem could arise when a fused gene prediction from the sister species is used as homology information but MAKER2 fails to generate a complete gene model out of it. The *C. elegans* gene C29H12.2 is one example of a missing gene model residing in the UTR of a *P. pacificus rars-2* homolog (Fig. 2c). The corresponding *P. pacificus* locus is spanned by two assembled transcripts that are homologous two C29H12.2 and *rars-2*, respectively. Both transcripts are also well supported by Iso-seq data and exhibit different expression levels[44,46]. In such a case, we would propose a replacement of the old *P. pacificus* gene model by the two distinct transcripts.

After manual inspection of all 526 missing gene candidates, 201 (41%) of the 486 non-redundant *P. pacificus* loci were classified as missing genes (Fig. 2b). Presumably this kind of error could arise when the gene annotation pipeline is mostly dependent on gene prediction algorithms which fail to predict all genes in gene dense regions (e.g. operon structures) as the intergenic distances might span only a few hundred nucleotides, which could be too small for triggering the initiation of a new gene model. The *C. elegans* gene *apn-1* is one example of a missed gene model in a gene dense region (Fig. 2d). Given that the *P. pacificus* homolog of *apn-1* has good transcriptomic support, the correction in this case would simply add the transcript to the existing gene models. Other instances of missing homologs are due to borderline cases in the BLASTP searches where one search resulted in an e-value slightly below the e-value threshold (10^{-5}) and the result of the other BLASTP search was slightly above the threshold. In total, we encountered 87 of such cases which we termed 'weak similarity' (Fig. 2b). For such cases no correction was proposed. In summary, we compiled corrections for 280 *P. pacificus* genes which were replaced by 714 new gene models. All these changes were submitted to WormBase and were incorporated in the release WS272.

Artificial gene fusions mask thousands of hidden orthologs. A small number of *C. elegans* genes with missing homologs in the current gene annotations (version: El Paco v1/WS268) of *P. pacificus* were classified as located in fused gene models (Fig. 2b). One potential explanation could be that an artificially fused gene prediction from the sister species is taken as homology data to annotate the orthologous locus in *P. pacificus*, but small errors cause parts of the gene model to be either incompletely or incorrectly annotated in *P. pacificus* resulting in a loss of detectable homology (Fig. 2c). Even if the homolog of a *C. elegans* gene is incorporated in the correct ORF within an artificially fused gene model, this could still cause a loss of one-to-one orthology as the

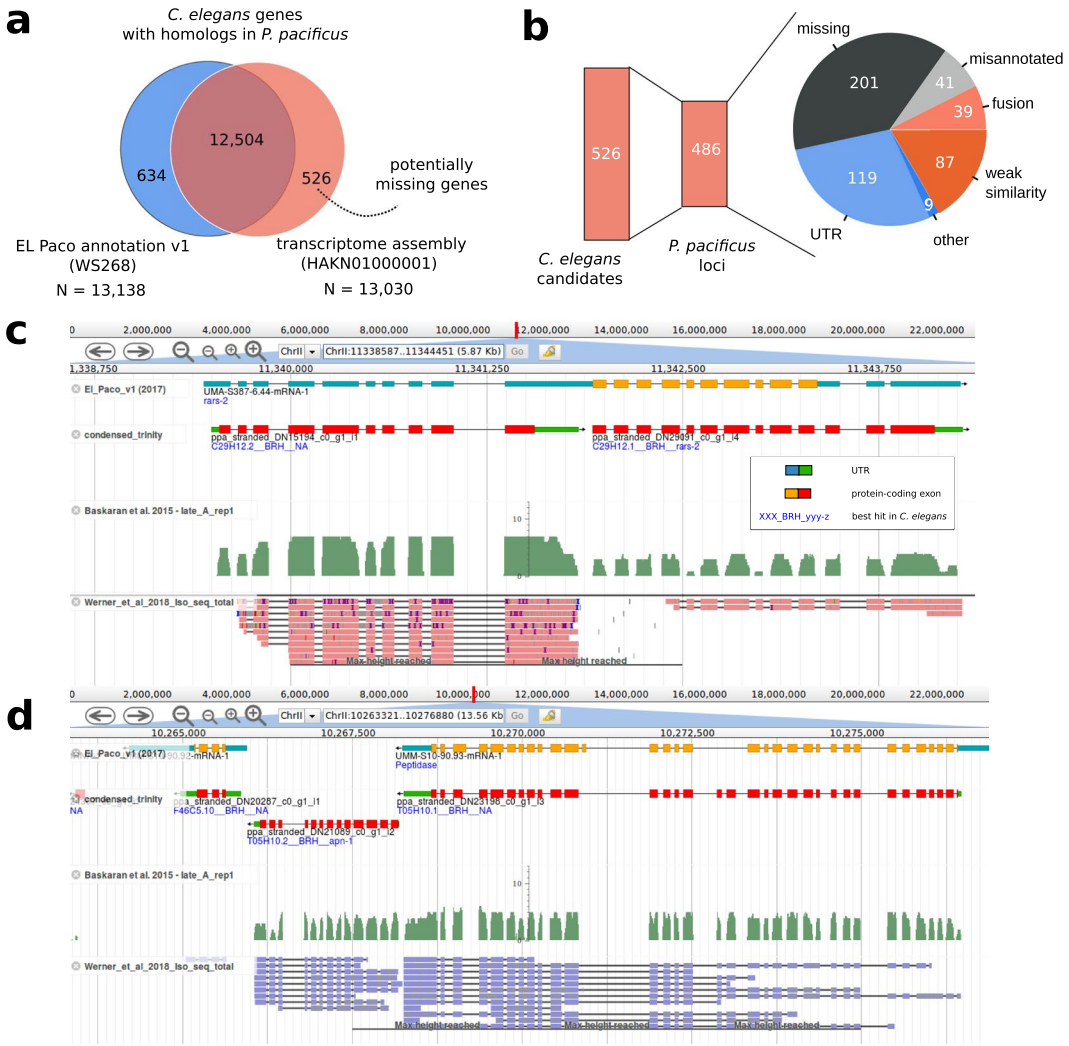


Figure 2. Identification of missing genes. (**a**) 526 potentially missing genes were identified based on *C. elegans* genes with homologs in the transcriptome assembly but not in current gene annotations. (**b**) The 526 missing gene candidates were located in 486 *P. pacificus* loci that were classified based on community annotators. (**c**) The genome browser screenshot shows a homolog of *C. elegans* C29H12.2 which is located in the annotated 5′UTR of a *P. pacificus* gene. This locus harbors two *P. pacificus* transcripts with different expression levels and well supported as non-overlapping transcripts based on RNA-seq and Iso-seq data. (**d**) A homolog of apn-1 is completely missing from current gene annotations.

corresponding *P. pacificus* gene can only be identified as one-to-one ortholog of a single *C. elegans* gene. Thus, we performed a second screen for *C. elegans* genes that had a predicted one-to-one ortholog (best-reciprocal hit) in the transcriptome assembly but not in current gene annotations (Fig. 3a). In total, 6075 (93%) of *C. elegans* genes with a predicted one-to-one ortholog (based on best-reciprocal hits) in current gene annotations, also had a predicted one-to-one ortholog against the *de novo* transcriptome assembly (Fig. 3a). Nevertheless, we found 2075 *C. elegans* genes that only had predicted one-to-one orthologs in the *de novo* transcriptome assembly. Excluding *C. elegans* genes that were identified already in the previous screen for missing homologs, this resulted in 1692 *C. elegans* genes with predicted one-to-one orthologs in the *de novo* transcriptome assembly but not in the current set of gene annotations (version: El Paco v1/WS268). Community-based classification and curation of the 1281 corresponding *P. pacificus* loci classified 912 (71%) cases as artificial gene fusions (Fig. 3b). One such an example is the *C. elegans* gene D1053.3. Its putative ortholog is fused with the *P. pacificus mvb-12* ortholog (Fig. 3c). Apart from being orthologous to two different *C. elegans* genes, both *P. pacificus* genes are supported as non-overlapping transcripts by RNA-seq and Iso-seq, and are expressed at different levels. This confirmed the interpretation of an artificially fused annotation. The proposed correction in this case would be a replacement of the old gene model by the two non-overlapping transcripts. In total, we updated 1241 *P. pacificus* gene models and replaced them with 3305 new models. These updates were submitted to WormBase and will be released following curation. The new *P. pacifcus* gene annotation (version: El Paco v2) with 28,036 gene models is also available on http://www.pristionchus.org/download. The results of the BUSCO analysis (Complete and Single Copy: 95.4%, Duplicated:

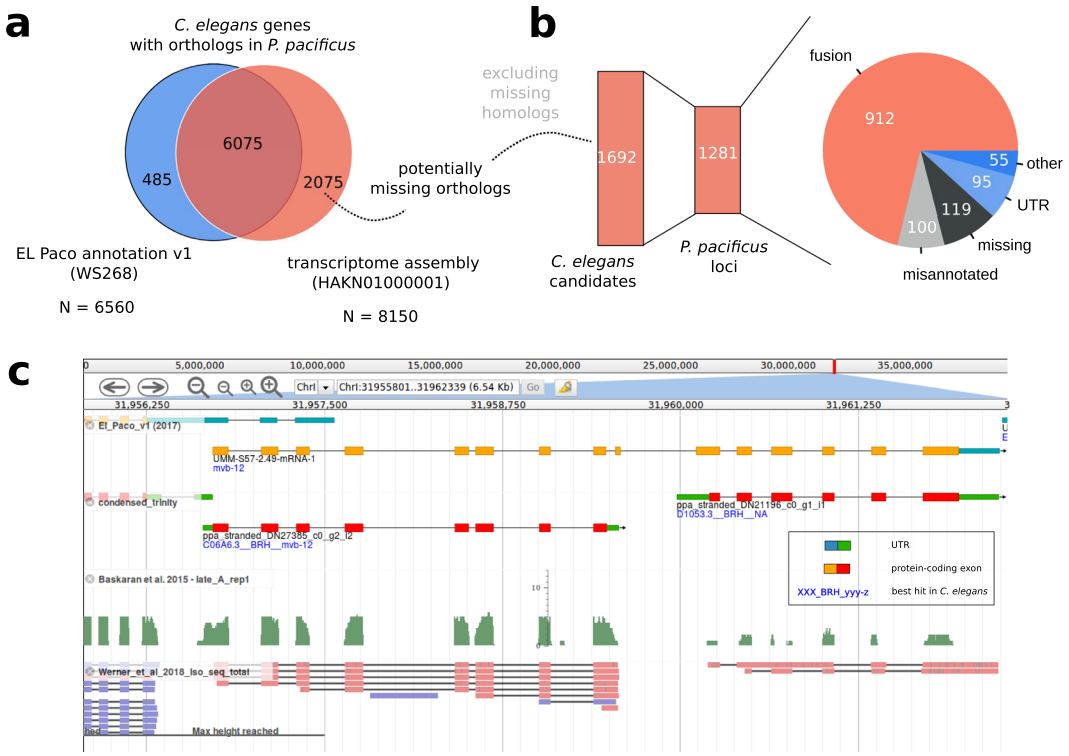


Figure 3. Community-based curation of hidden orthologs. (**a**) We identified 2075 putative *C. elegans* one-to-one orthologs that were specific to the *P. pacificus* transcriptome assembly. (**b**) Community-based curation classified most of the corresponding gene loci as artificial gene fusions. (**c**) Non-overlapping transcripts corresponding to *P. pacificus* orthologs of *mvb-12* and D1053.3 are artificially fused in a current gene model. This prohibits the detection of a one-to-one ortholog of D1053.3 based on a genome-wide approach such as best reciprocal hits.

1.7%, Fragmented: 2.0%, Missing: 0.9%) indicate that the new annotation represents a substantial improvement over the previous annotations[15] (Table 1).

Improved gene annotations facilitate the establishment of a catalog of *C. elegans* homologs and orthologs in the *P. pacificus* genome. Since our primary focus was to improve the annotation of *C. elegans* orthologs in the *P. pacificus* genome, we wanted to use the updated gene annotation to generate a catalog of predicted orthologs between *C. elegans* and *P. pacificus*. As the identification of orthologs typically requires sufficient genomic and phylogenetic knowledge to retrieve relevant protein data sets and to perform reconstruction of gene trees[24,45,46], a genome-wide catalog of orthologs would be highly useful as a starting point for researchers without sufficient expertise. Previous comparisons between *C. elegans* and *P. pacificus* identified putative one-to-one orthologs for roughly 6000–8000 genes[44,46]. To further characterize *C. elegans* genes without orthologs in *P. pacificus*, we additionally carried out a phylostratigraphic analysis[47] to estimate their relative age. Basically, phylostratigraphy uses absence-presence patterns of a gene to map its origin to an internal branch in a species tree[47]. Our analysis revealed that 5258 (26%) of *C. elegans* genes do not have BLAST hits in *Pristionchus* or more distantly related species (Phylostrata I–IV, Supplemental Table 1). This strongly suggests that they are younger than the common ancestor between *C. elegans* and *P. pacificus* and consequently have no orthologs. Next, we applied two different approaches to predict orthologs between *C. elegans* and *P. pacificus*: best reciprocal hits and Markov clustering as implemented in the software orthAgogue[48]. Computation of best reciprocal hits is a standard approach for predicting one-to-one orthologs across species[49,50]. In order to capture more complex orthology relationships (e.g. many-to-many), more general approaches such as Markov clustering have been widely applied[48,51]. Based on best reciprocal hits, we identified 8348 predicted one-to-one orthologs between both species (Supplemental Table 1) whereas the orthAgogue pipeline identified 7643 orthologous clusters, of which only 3345 corresponded to one-to-one orthologs. The large majority (98%) of these predicted one-to-one orthologs was also identified as best reciprocal hits and in 3260 (99%) cases, the same *P. pacificus* gene was predicted as one-to-one ortholog. The large discrepancy between the total number of best reciprocal hits and one-to-one orthologs defined by orthAgogue could be explained by the fact that best reciprocal hits do not take inparalogs into account[49]. However, only 1049 (21%) of *C. elegans* genes that were not identified as one-to-one orthologs by orthAgogue could be explained by the presence of lineage-specific inparalogs, suggesting that orthAgogue with default settings might be too conservative for this analysis. This is further supported by the reanalysis of 57 one-to-one orthologous pairs that were previously confirmed by phylogenetic analysis[46]. While 53 of the previously confirmed one-to-one orthologs were captured as best reciprocal hits, only 33 were also identified

by orthAgogue. Taken together, the improved gene annotation facilitated the prediction of substantially more one-to-one orthologs (Fig. 3a, Supplemental Table 1). This resource can be taken as a starting point to identify candidate genes in *P. pacificus*.

Discussion

With *C. elegans*, *C. briggsae*, and *P. pacificus*, three genetically tractable and free living nematode model organisms have been well established and can be used to study the evolution of gene function at various time-scales[2,3,52]. For example, recent reverse genetic approaches in *P. pacificus* have revealed functional divergence of genes with known roles in *C. elegans* dauer formation[22,23,53]. In addition, mutant screens in *P. pacificus* for social behaviours have uncovered multiple orthologous *C. elegans* genes for which a behavioral phenotype had been overlooked previously[33,54]. Together with complementary studies of the functional importance of novel genes[7,32,55], this makes nematodes an extremely powerful system to study genome evolution and gene function at a mechanistic level.

In order to facilitate fruitful functional studies across multiple model organisms, it is crucial to generate genomic resources (e.g. assemblies, annotations) and experimental genetic toolkits (e.g. forward and reverse genetics) of comparable quality. The chromosome-scale assembly of the *P. pacificus* genome[15] was a major step towards making this species more tractable for other groups. In our study, we aimed to minimize the discrepancy between automatically generated gene annotations for *P. pacificus* and heavily curated annotations for *C. elegans*. To this end, we incorporated recently generated Iso-seq and RNA-seq data into current gene annotations by manual curation of suspicious candidate loci that were identified by comparative genomic analysis. While application of alternative annotation pipelines can generate overall better gene annotations[29,41], they cannot guarantee that gene annotations will only improve. In certain cases, new annotation pipelines will also cause new errors. In contrast, during manual inspection, each community curator has the choice to not propose any change of gene models in case of uncertainty. Thus, manual inspection should only lead to removal of errors and thus improve annotation quality without introducing biases elsewhere. While manual annotation is an incredibly tedious task that is probably not scalable to complete genomes[38], we minimized the workload by focusing on a small gene set of *C. elegans* orthologs, recruiting colleagues as community curators, and restricting the task just to the selection of alternative gene models that were generated from transcriptomic data[16,44] or previous rounds of gene prediction[56,57]. In our opinion, the most crucial aspect of this community project is a good training of new annotators. We achieved this by personal training sessions between experienced and new annotators and the possibility to always discuss cases of uncertainty with other curators. For larger projects, initial training could be achieved by comprehensive online tutorials and communication via email, but this will likely be less efficient. In the case of the *P. pacificus* gene annotations, our study raised the gene count from 25,517 to 28,036 and increased the single copy ortholog completeness level from 86% to 97%. In the *P. pacificus* genome, the greatest source of error was the artificial fusion of neighboring genes. This type of error might be more prevalent in nematodes where genomes are compact[9] and genes frequently overlap[37,45]. Consequently, manual annotation of restricted gene sets has been proposed and applied previously to circumvent this problem[58]. Given that nematode genomes tend to be pretty compact (Fig. 1), we anticipate that misannotation due to overlapping gene models should be much less pronounced in large vertebrate or plant genomes. Nevertheless, it would be interesting to apply similar screens for gene annotation artifacts to other systems and eventually this could reveal some incorrect annotations in the genomes of classical model organisms.

While this study was restricted to *P. pacificus* genes with putative orthologs in *C. elegans*, we cannot reliably estimate the fraction of erroneous gene models across the whole genome. Our results would suggest that the fraction of missing genes is around one percent (Fig. 2a,b) and the amount of gene models affected by artificial fusions may be up to 15% (Fig. 3a). However, as the *P. pacificus* genome has a higher gene density and a higher concentration of old genes at the chromosome centers[8,15], we hypothesize that errors due to artificial gene fusions should be much less pronounced at chromosome arms. To test this, an unbiased quantification of error rates across genomic segments would be needed. In future, we also plan to focus on large gene families and lineage-specific orphan genes[55] that were not explicit subjects of this study. Artificial fusions in these classes of genes could be identified by screens for unexpectedly long gene models or unusual protein domain content. For orphan genes abundant RNA-seq studies of different developmental stages[22,46], tissues[10,46], environmental conditions[59], sexes[16], and genetic backgrounds[60,61] could be used to detect non-overlapping transcripts that exhibit anticorrelated expression within a single locus. Thus, while our study has demonstrated that community-based curation of gene annotations is feasible and can lead to substantial improvements, continued effort is needed to lift its quality to a level that would be similar to classical model organisms.

Methods

Comparative assessment of nematode genomes. We downloaded 22 nematode genomes and corresponding protein sequences from WormBase ParaSite (release WBPS13). For *Steinernema carpocapsae*, the latest version at WBPS14 was used. In case of multiple isoforms, we selected the longest isoform for further analysis. We ran BUSCO (version 3.0.1) in protein mode (option: -m prot) against the nematode_odb9 data set (N = 982 genes) to evaluate the completeness level of available protein sequences.

Genome browser integration of transcriptomic resources. To allow community annotators to propose alternative gene models, we integrated recent raw read alignments and reference guided transcript assemblies of Iso-seq data[44] and a *de novo* assembly of strand-specific RNA-seq data[16] into our genome browser (implemented in jbrowse[62]) on our webserver (http://www.pristionchus.org). Genomic coordinates for the *de novo* transcriptome assembly were generated by alignment to the *P. pacificus* reference genome (version El Paco) with the program exonerate[63] (version: 2.2.0, options: -m est2genome – dnawordlen 20 – maxintron 20000). To reduce the complexity of this data set, a condensed version of the *de novo* transcriptome assembly (selection of the isoform with the longest ORF as single representative isoform per gene, minimum peptide length of 60 amino acids, removal of single exon transcripts) with annotated best-reciprocal hits and best hits (BLASTP, e-value $< 10^{-5}$) in

C. elegans was also incorporated into our jbrowse instance. In addition, predicted protein sequences of previous versions of *P. pacificus* annotations (Hybrid1[56] and TAU2011[57]) were mapped against the *P. pacificus* assembly by exonerate (version: 2.2.0, options: -m protein2genome – dnawordlen 20 – maxintron 20000). All data sets are available under the gene annotation track of the El Paco reference assembly in our genome browser. To evaluate the quality of the two recent transcriptome assemblies, we ran BUSCO (version 3.0.1, options -m trans) against the nematode_odb9 data set (N = 982 genes) for completeness assessment (Table 1).

Identification of missing and fused gene models in current gene annotations. We ran bidirectional BLASTP (e-value $< 10^{-5}$) searches between *C. elegans* (version: WS260, longest isoform per gene) and two different *P. pacificus* data sets: the annotated proteins (version: El Paco v1, WS268) and the *de novo* transcriptome assembly[16]. For the *de novo* transcriptome, we reduced the redundancy resulting from different isoforms by selecting the longest ORFs per gene. Based on the different BLASTP searches, we first screened for *C. elegans* proteins with BLASTP hits against ORFs in the *de novo* transcriptome assembly but not against the currently annotated proteins. This yielded 526 candidate genes. In a second phase, we screened for *C. elegans* proteins with putative orthologs, defined by best-reciprocal BLASTP relationships, in the *de novo* transcriptome assembly but not in the annotated proteins, resulting in 2075 candidate genes.

Community-based manual curation of candidate loci. All *C. elegans* genes together with their candidate homologs and orthologs in the *P. pacificus de novo* transcriptome assembly were stored in a shared online spreadsheet. Community annotators were trained to find the corresponding locus in the genome browser by entering the transcript identifier and to manually inspect the surrounding regions that were defined by the encompassing *P. pacificus* gene model. The candidate locus was then classified as untranslated region (UTR) (the query transcript overlapped exons that were annotated as UTR), missing gene (the query transcript did not overlap any annotated exon), gene fusion (the query transcript did overlap protein-coding exons and homology was detected by BLASTP), misannotation (the query transcript did overlap protein-coding exons but no BLASTP hit was found due incorrect reading frame annotation or minimal overlap) or inconclusive. After classification, a correction was proposed that either added new genes (identifiers could be selected from the *de novo* assembled transcripts, Iso-seq assemblies, or previous versions of gene annotations) or replaced an existing gene model by one or more new genes. In such a case the objective was to lose as little annotated coding sequence as possible. Thus, new genes were selected from the above mentioned data set in order to cover as much coding sequence of the initial gene model as possible. If parts of the old gene model were not covered, BLAST searches against *C. elegans* and other *Pristionchus* species were used to split the old gene model into several parts with sequence matches to distinct *C. elegans* genes, or to extract partial protein sequences of the old gene model that were not covered. Such protein sequence stretches were given a pseudo identifier and were stored in a shared online document. All these sequences were later automatically reannotated by mapping them against the reference genome with the help of exonerate. In case that an existing gene model was replaced by multiple new gene models, we additionally selected one of the new gene models to inherit the WormBase identifier of the old gene model to allow WormBase to record the history of a given gene model. Usually, either the most conserved or the longest new gene model was chosen. Due to the fact that a single artificially fused gene could cause missing homologs and orthologs for multiple *C. elegans* genes, some loci were curated multiple times. We randomly picked some of these cases to compare the classifications and the corresponding corrections from multiple curators, which turned out to be largely consistent. In case of redundant curations, one out of many possible curations for a given locus was chosen based on the following criteria: preference towards higher number of new models, experience of the curator (number of curated loci), and transcriptional evidence over gene prediction.

Phylostratigraphy and orthology predictions. Outgroup data sets were defined by concatenating all protein sequence data from different species in the ladder-like phylogeny leading to *C. elegans* (Fig. 1). More precisely, we pooled all data from species in an induced subtree defined by branches with roman numbers in Fig. 1. We then ran a BLASTP search (e-value < 0.001) of *C. elegans* proteins (longest isoform per gene) against each of the outgroup data sets. Starting from the *C. elegans* genes with homologs in the most distant outgroup set (VIII), we iteratively defined phylostrata by comparison with the next, more closely related outgroup set. The results of this analysis are summarized in Supplemental Table 1. *C. elegans* specific genes are assigned to phylostratum I, whereas genes that are present in the most divergent outgroups are assigned to phylostratum VIII. Orthologs were defined after performing all pairwise BLASTP searches including self-searches (e-value $< 10^{-5}$) between *C. elegans* and *P. pacificus* and extracting best reciprocal hits from the BLAST output files. Simultaneously, the program orthAgogue was run with default setting on the same input files[48].

Data availability

The strand-specific *de novo* transcriptome was submitted to the European Nucleotide Archive under the accession HAKN01000001[16] and the Iso-seq data was submitted to the European Nucleotide Archive under the accessions ERX2315712 and ERX2315713[44]. All data sets are also available at http://www.pristionchus.org/download. The initial set of *P. pacificus* gene annotations corresponds to WormBase WS268. Corrections from this study were submitted to WormBase and will be released following curation.

Received: 1 October 2019; Accepted: 20 November 2019;
Published online: 11 December 2019

References

1. Sommer, R. J., Carta, L. K., Kim, S.-Y. & Sternberg, P. W. Morphological, genetic and molecular description of *Pristionchus pacificus* sp. n. (Nematoda: Neodiplogastridae). *Fundam. Appl. Nematol.* **19**, 511–521 (1996).
2. Sternberg, P. W. Why *Caenorhabditis elegans* is great and *Pristionchus pacificus* might be better. In *Pristionchus pacificus* (ed. Sommer, R. J.) **11**, 1–17 (BRILL).
3. Sommer, R. J. Evolution of regulatory networks: nematode vulva induction as an example of developmental systems drift. *Adv. Exp. Med. Biol.* **751**, 79–91 (2012).
4. Bose, N. *et al*. Complex small-molecule architectures regulate phenotypic plasticity in a nematode. *Angew. Chem. Int. Ed Engl.* **51**, 12438–12443 (2012).
5. Yim, J. J., Bose, N., Meyer, J. M., Sommer, R. J. & Schroeder, F. C. Nematode signaling molecules derived from multimodular assembly of primary metabolic building blocks. *Org. Lett.* **17**, 1648–1651 (2015).
6. Falcke, J. M. *et al*. Linking genomic and metabolomic natural variation uncovers nematode pheromone biosynthesis. *Cell Chem Biol* **25**, 787–796.e12 (2018).
7. Mayer, M. G., Rödelsperger, C., Witte, H., Riebesell, M. & Sommer, R. J. The orphan gene *dauerless* regulates dauer development and intraspecific competition in nematodes by copy number variation. *PLoS Genet.* **11**, e1005146 (2015).
8. Prabh, N. *et al*. Deep taxon sampling reveals the evolutionary dynamics of novel gene families in *Pristionchus* nematodes. *Genome Res.* **28**, 1664–1674 (2018).
9. Rödelsperger, C., Streit, A. & Sommer, R. J. Structure, function and evolution of the nematode genome. In: *eLS*. John Wiley & Sons, Ltd: Chichester, https://doi.org/10.1002/9780470015902.a0024603 (2013).
10. Lightfoot, J. W., Chauhan, V. M., Aylott, J. W. & Rödelsperger, C. Comparative transcriptomics of the nematode gut identifies global shifts in feeding mode and pathogen susceptibility. *BMC Res. Notes* **9**, 142 (2016).
11. Susoy, V. *et al*. Large-scale diversification without genetic isolation in nematode symbionts of figs. *Sci Adv* **2**, e1501031 (2016).
12. Bumbarger, D. J., Riebesell, M., Rödelsperger, C. & Sommer, R. J. System-wide rewiring underlies behavioral differences in predatory and bacterial-feeding nematodes. *Cell* **152**, 109–119 (2013).
13. Witte, H. *et al*. Gene inactivation using the CRISPR/Cas9 system in the nematode *Pristionchus pacificus*. *Dev. Genes Evol.* **225**, 55–62 (2015).
14. Srinivasan, J. *et al*. A bacterial artificial chromosome-based genetic linkage map of the nematode *Pristionchus pacificus*. *Genetics* **162**, 129–134 (2002).
15. Rödelsperger, C. *et al*. Single-molecule sequencing reveals the chromosome-scale genomic architecture of the nematode model organism *Pristionchus pacificus*. *Cell Rep.* **21**, 834–844 (2017).
16. Rödelsperger, C. *et al*. Phylotranscriptomics of *Pristionchus* nematodes reveals parallel gene loss in six hermaphroditic lineages. *Curr. Biol.* **28**, 3123–3127.e5 (2018).
17. Namai, S. & Sugimoto, A. Transgenesis by microparticle bombardment for live imaging of fluorescent proteins in *Pristionchus pacificus* germline and early embryos. *Dev. Genes Evol.* **228**, 75–82 (2018).
18. Lo, T.-W. *et al*. Precise and heritable genome editing in evolutionarily diverse nematodes using TALENs and CRISPR/Cas9 to engineer insertions and deletions. *Genetics* **195**, 331–348 (2013).
19. Bui, L. T. & Ragsdale, E. J. Multiple plasticity regulators reveal targets specifying an induced predatory form in nematodes. *Mol. Biol. Evol*, https://doi.org/10.1093/molbev/msz171 (2019).
20. Ishita, Y., Chihara, T. & Okumura, M. Serotonergic modulation of feeding behavior in *Caenorhabditis elegans* and other related nematodes. *Neurosci. Res*, https://doi.org/10.1016/j.neures.2019.04.006 (2019).
21. Liu, Z. *et al*. Predator-secreted sulfolipids induce defensive responses in *C. elegans*. *Nature Communications* **9** (2018).
22. Moreno, E. *et al*. DAF-19/RFX controls ciliogenesis and influences oxygen-induced social behaviors in *Pristionchus pacificus*. *Evol. Dev.* **20**, 233–243 (2018).
23. Markov, G. V. *et al*. Functional conservation and divergence of *daf-22* paralogs in *Pristionchus pacificus* dauer development. *Mol. Biol. Evol.* **33**, 2506–2514 (2016).
24. Markov, G. V., Baskaran, P. & Sommer, R. J. The same or not the same: lineage-specific gene expansions and homology relationships in multigene families in nematodes. *J. Mol. Evol.* **80**, 18–36 (2015).
25. Namdeo, S. *et al*. Two independent sulfation processes regulate mouth-form plasticity in the nematode. *Development* **145** (2018).
26. Rödelsperger, C. Comparative genomics of gene loss and gain in *Caenorhabditis* and Other Nematodes. In *Methods in Molecular Biology* 419–432 (2018).
27. Korf, I. Gene finding in novel genomes. *BMC Bioinformatics* **5**, 59 (2004).
28. Stanke, M. *et al*. AUGUSTUS: ab initio prediction of alternative transcripts. *Nucleic Acids Res.* **34**, W435–9 (2006).
29. Hoff, K. J., Lomsadze, A., Borodovsky, M. & Stanke, M. Whole-genome annotation with BRAKER. *Methods Mol. Biol.* **1962**, 65–95 (2019).
30. Cantarel, B. L. *et al*. MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Res.* **18**, 188–196 (2008).
31. Holt, C. & Yandell, M. MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. *BMC Bioinformatics* **12**, 491 (2011).
32. Lightfoot, J. W. *et al*. Small peptide-mediated self-recognition prevents cannibalism in predatory nematodes. *Science* **364**, 86–89 (2019).
33. Moreno, E. *et al*. Regulation of hyperoxia-induced social behaviour in *Pristionchus pacificus* nematodes requires a novel cilia-mediated environmental input. *Sci. Rep.* **7**, 17550 (2017).
34. Kieninger, M. R. *et al*. The nuclear hormone receptor NHR-40 acts downstream of the sulfatase EUD-1 as part of a developmental plasticity switch in *Pristionchus*. *Curr. Biol.* **26**, 2174–2179 (2016).
35. Baskaran, P. & Rödelsperger, C. Microevolution of duplications and deletions and their impact on gene expression in the Nematode *Pristionchus pacificus*. *PLoS One* **10**, e0131136 (2015).
36. Weller, A. M., Rödelsperger, C., Eberhardt, G., Molnar, R. I. & Sommer, R. J. Opposing forces of A/T-biased mutations and G/C-biased gene conversions shape the genome of the nematode *Pristionchus pacificus*. *Genetics* **196**, 1145–1152 (2014).
37. Prabh, N. & Rödelsperger, C. Divergence, and mixed origin contribute to the emergence of orphan genes in nematodes. *G3* **9**, 2277–2286 (2019).
38. Salzberg, S. L. Next-generation genome annotation: we still struggle to get it right. *Genome Biol.* **20**, 92 (2019).
39. Smythe, A. B., Holovachov, O. & Kocot, K. M. Improved phylogenomic sampling of free-living nematodes enhances resolution of higher-level nematode phylogeny. *BMC Evolutionary Biology* **19** (2019).
40. Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V. & Zdobnov, E. M. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**, 3210–3212 (2015).
41. McLean, F., Berger, D., Laetsch, D. R., Schwartz, H. T. & Blaxter, M. Improving the annotation of the *Heterorhabditis bacteriophora* genome. *Gigascience* **7** (2018).
42. Hiraki, H. *et al*. Genome analysis of Diploscapter coronatus: insights into molecular peculiarities of a nematode with parthenogenetic reproduction. *BMC Genomics* **18**, 478 (2017).
43. Rödelsperger, C. *et al*. Characterization of genetic diversity in the nematode *Pristionchus pacificus* from population-scale resequencing data. *Genetics* **196**, 1153–1165 (2014).

44. Werner, M. S. *et al*. Young genes have distinct gene structure, epigenetic profiles, and transcriptional regulation. *Genome Res.* **28**, 1675–1687 (2018).
45. Rödelsperger, C., Menden, K., Serobyan, V., Witte, H. & Baskaran, P. First insights into the nature and evolution of antisense transcription in nematodes. *BMC Evol. Biol.* **16**, 165 (2016).
46. Baskaran, P. *et al*. Ancient gene duplications have shaped developmental stage-specific expression in *Pristionchus pacificus*. *BMC Evol. Biol.* **15**, 185 (2015).
47. Domazet-Loso, T., Brajković, J. & Tautz, D. A phylostratigraphy approach to uncover the genomic history of major adaptations in metazoan lineages. *Trends Genet.* **23**, 533–539 (2007).
48. Ekseth, O. K., Kuiper, M. & Mironov, V. orthAgogue: an agile tool for the rapid prediction of orthology relations. *Bioinformatics* **30**, 734–736 (2014).
49. Remm, M., Storm, C. E. & Sonnhammer, E. L. Automatic clustering of orthologs and in-paralogs from pairwise species comparisons. *J. Mol. Biol.* **314**, 1041–1052 (2001).
50. Tatusov, R. L. A Genomic Perspective on Protein Families. *Science* **278**, 631–637 (1997).
51. Li, L., Stoeckert, C. J. Jr & Roos, D. S. OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res.* **13**, 2178–2189 (2003).
52. Verster, A. J., Ramani, A. K., McKay, S. J. & Fraser, A. G. Comparative RNAi Screens in *C. elegans* and *C. briggsae* Reveal the Impact of Developmental System Drift on Gene Function. *PLoS Genetics* **10**, e1004077 (2014).
53. Sieriebriennikov, B., Markov, G. V., Witte, H. & Sommer, R. J. The Role of DAF-21/Hsp90 in Mouth-Form Plasticity in *Pristionchus pacificus*. *Mol. Biol. Evol.* **34**, 1644–1653 (2017).
54. Moreno, E. & Sommer, R. J. A cilia-mediated environmental input induces solitary behaviour in *Caenorhabditis elegans* and *Pristionchus pacificus* nematodes. *Nematology* **20**, 201–209 (2018).
55. Prabh, N. & Rödelsperger, C. Are orphan genes protein-coding, prediction artifacts, or non-coding RNAs? *BMC Bioinformatics* **17**, 226 (2016).
56. Borchert, N. *et al*. Proteogenomics of *Pristionchus pacificus* reveals distinct proteome structure of nematode models. *Genome Res.* **20**, 837–846 (2010).
57. Sinha, A., Sommer, R. J. & Dieterich, C. Divergent gene expression in the conserved dauer stage of the nematodes *Pristionchus pacificus* and *Caenorhabditis elegans*. *BMC Genomics* **13**, 254 (2012).
58. Stoltzfus, J. D., Minot, S., Berriman, M., Nolan, T. J. & Lok, J. B. RNAseq analysis of the parasitic nematode *Strongyloides stercoralis* reveals divergent regulation of canonical dauer pathways. *PLoS Negl. Trop. Dis.* **6**, e1854 (2012).
59. Sanghvi, G. V. *et al*. Life history responses and gene expression profiles of the nematode *Pristionchus pacificus* cultured on *Cryptococcus* yeasts. *PLoS One* **11**, e0164881 (2016).
60. Serobyan, V. *et al*. Chromatin remodelling and antisense-mediated up-regulation of the developmental switch gene *eud-1* control predatory feeding plasticity. *Nat. Commun.* **7**, 12337 (2016).
61. Moreno, E., McGaughran, A., Rödelsperger, C., Zimmer, M. & Sommer, R. J. Oxygen-induced social behaviours in *Pristionchus pacificus* have a distinct evolutionary history and genetic regulation from *Caenorhabditis elegans*. *Proc. Biol. Sci.* **283**, 20152263 (2016).
62. Buels, R. *et al*. JBrowse: a dynamic web platform for genome visualization and analysis. *Genome Biol.* **17**, 66 (2016).
63. Slater, G. S. C. & Birney, E. Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics* **6**, 31 (2005).

Acknowledgements

The authors would like to thank the complete *Pristionchus* community for their long-term interest in studying *P. pacificus* and thus motivating this work. Further thanks to Bogdan Sieriebriennikov for providing additional manual curations and to all members of the Sommer lab for general discussions. Finally, special thanks to Michael Paulini for incorporating the updated gene models into WormBase. This work was funded by the Max Planck Society.

Author contributions

Conceptualization, C.R.; Investigation, C.R., M.A., M.L., T.T., S.S., M.D., S.W., W.H., D.R.S. and Z.H.; Writing – Original Draft, C.R.; Writing – Review & Editing, C.R., M.A., M.L., T.T., S.S., M.D., S.W., W.H., D.R.S. and Z.H.; Supervision, C.R.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41598-019-55359-5.

Correspondence and requests for materials should be addressed to C.R.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

nature protocols

PROTOCOL

https://doi.org/10.1038/s41596-020-0347-z

Check for updates

Geometric morphometrics of microscopic animals as exemplified by model nematodes

Tobias Theska[1], Bogdan Sieriebriennikov[1,2], Sara S. Wighard[1], Michael S. Werner[1]✉ and Ralf J. Sommer[1]✉

While a host of molecular techniques are utilized by evolutionary developmental (evo-devo) biologists, tools for quantitative evaluation of morphology are still largely underappreciated, especially in studies on microscopic animals. Here, we provide a standardized protocol for geometric morphometric analyses of 2D landmark data sets using a combination of the geomorph and Morpho R packages. Furthermore, we integrate clustering approaches to identify group structures within such datasets. We demonstrate our protocol by performing exemplary analyses on stomatal shapes in the model nematodes *Caenorhabditis* and *Pristionchus*. Image acquisition for 80 worms takes 3–4 d, while the entire data analysis requires 10–30 min. In theory, this approach is adaptable to all microscopic model organisms to facilitate a thorough quantification of shape differences within and across species, adding to the methodological toolkit of evo-devo studies on morphological evolution and novelty.

Introduction

More than 150 years after *The Origin of Species*, identifying the mechanisms of morphological evolution remains a major focus in evolutionary developmental biology (evo-devo)[1–6]. In the modern era, investigations of morphological novelty primarily rely on three components: (i) morphometric quantification of evolutionary changes in the structure of interest, (ii) the usage of comparative developmental approaches to identify candidate genes and/or pathways promoting such morphological change and (iii) functional experiments aiming to demonstrate causative relationships between the observed morphological change and the candidate developmental mechanisms[7–9]. Dissecting the gene regulatory networks that underlie a given morphological trait has been a standard practice in the field, especially in studies on classical model organisms. However, efforts to quantify experimentally induced mutant phenotypes or anatomical differences resulting from divergent evolution have been scarce[7,9]. Here, we provide a protocol that facilitates the evaluation of differences in biological shape by combining geometric morphometrics with *k*-medoid and model-based clustering. This streamlined protocol can be used as a blueprint by all members of the evo-devo research community working on morphological evolution—particularly in microscopic animal taxa—to improve the ease of use and comparability of results. We hope that lowering the barrier to quantitative morphological techniques will stimulate future studies on morphological evolution in small animals to integrate all three components of contemporary evo-devo research.

Development of the approach for nematodes

Several established approaches can be applied to nematodes to search for candidate genes underlying character development, including quantitative trait locus mapping, in situ hybridization and RNA-seq screens[10–12]. Functional experiments on candidate genes are now also feasible due to the recent establishment of precise genome editing tools like the CRISPR/Cas9 system[13–15]. In contrast, for the quantification of morphological changes—the first component of evo-devo research programs—most studies in nematodes still rely on traditional morphometry (i.e., linear measurements of structures[16]) and comparative qualitative descriptions of homologous anatomical elements. Only a handful of studies include more sophisticated methods of landmark-based geometric morphometrics to thoroughly quantify morphological data[17–20]. Thus, the power of geometric morphometrics has not

[1]Department for Integrative Evolutionary Biology, Max Planck Institute for Developmental Biology, Tübingen, Germany. [2]Present address: Department of Biology, New York University, New York, NY, USA. ✉e-mail: michael.werner@tuebingen.mpg.de; ralf.sommer@tuebingen.mpg.de

yet been fully realized in nematodes, despite it offering novel insights into the nature of morphological change and its implications for long-standing predictions of evolutionary theory.

One such prediction is that phenotypic plasticity, the ability of one genotype to produce different phenotypes based on environmental input, plays a significant role in the evolution of morphological novelty[1,2]. Recently, nematodes of the Diplogastridae family, specifically *Pristionchus pacificus* and its close relatives, have been developed as model systems to investigate the role of plasticity in development and evolution[21–24]. These species are characterized by the presence of prominent cuticular teeth as part of their stomata, which are absent in other nematodes like *Caenorhabditis elegans*[25–27]. In addition, *P. pacificus* displays a polyphenism in this morphological novelty, where adults irreversibly adopt either a wide eurystomatous (Eu) or narrow stenostomatous (St) mouth-form, based on a number of environmental cues. Mouth-form plasticity also has important consequences on feeding strategies; for example, in *P. pacificus*, only Eu animals are facultative predators on other nematodes and fungi, whereas St animals are strict bacterial feeders[28].

The gene regulatory network controlling mouth-form plasticity in *P. pacificus*[24] has been revealed through unbiased genetic approaches and the subsequent functional manipulation of these genes[29–32]. In contrast, there is to date no standard protocol for quantifying morphological changes that accompany genetic perturbation. Similarly, the quantification of morphological changes through development and evolution has not been standardized in nematodes. Presumably, factors like their small body size and the triradiate symmetry of their pharynx have prevented the previous establishment of sophisticated quantification methods, like geometric morphometrics, while such methods are regularly used in larger animals like vertebrates and arthropods[33–35]. Thus, while nematodes have traditionally been used as model systems in molecular cell biology, their vast diversity of form and the depth of taxonomic sampling renders them a largely untapped resource for fruitful evo-devo studies on morphological evolution. Here, we describe a protocol that can be applied to microscopic structures including nematode stomata to quantify morphological shape differences within and between species.

Geometric morphometrics using R

The identification of patterns of shape variation in anatomical traits, as well as the covariation of shapes with environmental parameters, is a central goal of contemporary studies on ecology and evolution[36–41]. The most widely applied approach to quantitatively analyze shapes is landmark-based geometric morphometrics (GM)[36,38]. It utilizes 2D or 3D landmark coordinates containing information on the relative positions of homologous anatomical traits and boundary curves and surfaces, all of which serve as the basis for shape quantification[36,38]. Recently, many of the geometric morphometric techniques have been incorporated into various packages in the computational platform R/RStudio[42–44]. This platform has some inherent advantages: it is freely available for all common operating systems, an abundance of compatible software packages reduces the number of different applications needed to perform GM analyses and it is supported by a large active community of researchers from different fields[43]. Furthermore, several books, online forums and manuals and the interactive RStudio interface make it accessible to beginners. Two of the most widely used packages for GM analyses are geomorph[45,46] and Morpho[47,48]. While both of them are able to perform all fundamental steps of shape data analysis, each also comes with some useful functions that are unique to them. Therefore, a combination of both of these packages can be used to perform comprehensive geometric morphometric analyses.

Overview of the procedure

The general procedure of this protocol follows the four steps of the Procrustes paradigm[39] plus an additional stage to identify group structures through unbiased clustering. The Procrustes paradigm is based on the acquisition of landmark configurations for all specimens under investigation (Steps 1–8). These data are then modified using general Procrustes analysis (GPA), so that all non-shape information is removed from the configurations (Step 20). The shape data obtained via GPA are subsequently used to (i) visualize and describe patterns of shape (co-)variation (Steps 21–25 and 28) and (ii) perform multivariate statistical testing of hypotheses regarding group differences (Step 26). These results can be complemented by unbiased clustering approaches to find and analyze potential group-structures in the shape data set (Step 27). The following sections explain each step in more detail, followed by a step-by-step protocol with exemplary data sets from *Caenorhabditis* and *Pristionchus* species. Nevertheless, the experimental design is applicable to all microscopic organisms or structures for which a set of reproducible landmark configurations can be designed.

Landmark placement

First, the coordinates of defined sets of landmarks are acquired for each specimen[36,39] (Steps 1–8). These sets are called landmark configurations, and they contain information about the shape, size, orientation and location of homologous anatomical traits of interest. While landmarks always describe fixed homologous points defined by local anatomy, it is also possible to geometrically describe and quantify homologous curves or surfaces on which landmarks are sparse using 'semi-landmarks'[49]. A defined number of semilandmarks (Step 13) is placed along any given curve or surface in roughly corresponding positions across all specimens. The number of semilandmarks is dependent on the complexity of the curve or surface. Generally, one should use more semilandmarks with increasing structural complexity[49]. Unlike fixed landmarks, semilandmarks are allowed to slide along tangent lines to the respective curve (or tangent planes to the respective surface) that they describe to align their position, and thus establish a geometric correspondence of homologous curves or surfaces between individuals[38,49]. For semilandmarks located on a curve, it is generally desirable to start and end every curve with fixed landmarks, as this facilitates a reliable computation of tangent lines for the first and the last semilandmark along the curve[49].

GPA

In the second stage, GPA is performed to extract the shape information from all landmark configurations by removing the parameters of non-shape variation (i.e., differences in size, location and orientation) from the dataset[36,38] (Step 20). Information about location is removed by shifting the landmark configurations so that they share the same origin in a common coordinate system. Size differences are subsequently eliminated by scaling landmark configurations to an identical centroid size of 1.0. The centroid size describes the spread of landmarks around the centroid of a given landmark configuration and is calculated as the square root of the sum of squared distances between each landmark and the corresponding centroid[36,38]. All landmark configurations are subsequently rotated so that the sum of squared distances between single landmarks and their corresponding counterparts in the common target configuration is minimized[36,38]. Thus, orientation differences are eliminated. The resulting aligned landmark configurations therefore contain only shape information and are accordingly called 'shapes', with the landmarks representing a set of shape variables called 'Procrustes coordinates'[39]. The GPA approach is iterative in nature and aligns all landmark configurations to a target configuration, which in the first iteration is one of the actual landmark configurations contained in the data set. After the first alignment, an average configuration is calculated, which becomes the target for the second iteration. With ongoing repetitions of this process, newly calculated average configurations become asymptotically more similar, and superimposition is halted after a threshold of similarity is reached (Step 20). The underlying convergence criterion is fixed for geomorph's gpagen() function, but it can be manually set for Morpho's procSym() function using the 'tol=' argument. Alternatively, it is possible to relax the convergence criterion in both functions by setting a maximum number of iterations (i.e., by adding the 'max.iter=' argument to gpagen() or 'iterations=' argument to procSym() in Step 20). Newly calculated average configurations are always rescaled to a centroid size of 1.0[50]. Thus, all shapes are optimally aligned to the calculated average and each other.

If landmark configurations contain semilandmarks, it is necessary to computationally slide them to establish a proper spacing between them and their bordering (semi-)landmarks[49] (Steps 13 and 20). Thus, their initial arbitrary placement on the curve or surface is aligned in reference to the average shape of the data set. In principle, semilandmark sliding can be achieved by minimizing either Procrustes distances (i.e., the same procedure that is used for landmark superimposition) or bending energy[49] (Step 20). The 'minimize bending energy' approach treats landmark configurations as 'infinitely thin and infinitely large metal plates'[38,50,51] and describes the transformation from one 2D landmark configuration into another (i.e., the difference in shapes) by vertically bending this metal plate in the z axis. The amount of 'work' that is needed to bend the metal plate (i.e., a landmark configuration) has been coined the bending energy[38]. However, although bending energy is an inverse measure of a spatial scale, the metaphor of 'bending metal planes' is not to be equated with the actual biological processes that produce the observed differences in shape[50].

Both available sliding approaches iteratively minimize the shape differences of curves or surfaces between each specimen and the average shape of the data set. In the bending energy approach, all semilandmarks slide together, and the sliding process is affected by fixed landmarks[49]. However, if Procrustes distances are minimized, each semilandmark slides separately, which can potentially result in semilandmarks passing each other or even sliding beyond one of the endpoints of a curve.

Therefore, minimizing bending energy is generally preferred for semilandmark sliding, even though it might consume more time for computation[49].

Once the superimposition procedure (including semilandmark sliding) is finished, all of the obtained Procrustes coordinates describe the position of individual shapes in a curved (non-Euclidean) space, which is related to Kendall's shape space[36,39,52]. The non-Euclidean distances between landmarks in this curved shape space are called Procrustes distances[36,46]. However, the curvature of the shape space complicates the application of standard statistical procedures like multivariate analysis of variance (MANOVA) to such data sets (see 'Multivariate statistics' below). To avoid this problem, the Procrustes coordinates are projected into a linear tangent space (Step 20) that is defined at the level of the mean shape[36,39,53]. The Euclidean distances in the tangent space can be considered as approximations of the Procrustes distances and used for subsequent statistical analysis.

Multivariate statistics

In the third stage of the Procrustes paradigm (Step 26), multivariate statistical analysis is performed on the aligned landmarks to identify biologically relevant differences in mean shape, with grouping factors usually including categorical variables like sex, age, culture conditions or species[38]. Traditional MANOVA is a well established parametric tool to compare the values of multiple variables between two or more experimental groups[54]. In geometric morphometrics, such variables are the X and Y (for 2D landmarks) or X, Y and Z (for 3D landmarks) coordinates of superimposed landmarks, which are projected into a linear tangent space. To facilitate statistical testing for differences in mean shapes using a MANOVA design, it is advisable to use the modified permutational MANOVA version 'PERMANOVA' (in geometric morphometrics 'Procrustes ANOVA'; Step 26)[54–56]. This test generates inferences based on permutations while no assumption of multivariate normality is made, which renders it more applicable to the type of biological data that are used in geometric morphometrics. Furthermore, the number of variables (i.e., Procrustes coordinates) can be higher than the number of specimens in the data set, as it is based on permutations rather than estimated degrees of freedom[54]. Thus, PERMANOVA has some inherent advantages that make it applicable to the Euclidean distances between corresponding landmarks in the linear tangent space[45,53]. Nonetheless, the use of permutation tests to obtain *P* values in Procrustes ANOVA designs does not circumvent the problem that this kind of statistical test is sensitive to heterogeneity of variances between groups[54,55].

Visualization

The last stage of the Procrustes paradigm involves visualization and description of shape (co-) variation patterns[39,50]. The first plots generated are typically scatterplots that show the position of superimposed shapes in a tangent space (Step 20; see Supplementary Figs. 3–5). These plots can be used to obtain a first informal impression of the data, but they do not show holistic patterns of shape variation, and therefore should not be relied upon for a thorough examination of the variation in the data set[50]. Instead, ordination methods like principal component (PC) analysis (PCA) should be utilized. PCAs generate scatterplots (Step 21) depicting the distribution of observations (i.e., specimens) along axes of major variation (i.e., PCs)[57]. Therefore, PCA can be used to visualize patterns of shape variation in tangent space and visually appreciate the importance of grouping factors on the axes themselves or on a combination of axes[38,39]. Extensive overlaps of groups in the PCA plot indicate that they share a common shape, while no overlap indicates morphological differences. To check for outliers, thin plate spline (TPS) interpolation[51] can be used, which warps shapes and computes deformation grids that depict the degree of stretching and compressing required to deform a starting into a target landmark configuration (Step 22). During this procedure, the TPS interpolation algorithm ensures that corresponding landmarks in the start and target configurations are located on exactly corresponding positions in the untransformed and transformed grids (or shapes)[50].

Further useful information about shape differences can be gathered from the loadings and the PC scores obtained during PCA. The loadings describe the correlation between a PC and a variable (i.e., landmark coordinate)[57], which can be used to estimate extreme shapes along PC axes (Step 24). Extreme shapes can be superimposed and then compared in deformation grids, lollipop plots or wireframes. These plots thus provide an impression of the shape variation that underlies the distribution of individuals in the morphospace. Wireframe plots are particularly useful to depict structural differences between extreme shapes because landmarks, which are placed on the same anatomical element, are connected by simple lines to give a simplified representation of the overall shape. Lollipop plots, on the other hand, are useful to visually emphasize the change in position that each landmark undergoes (relative to all other landmarks) when a shape is shifted along the

corresponding PC. However, if semilandmarks were slid by minimizing bending energy, their positional shifts cannot be interpreted easily, as only the shape of the curve they define has biological meaning. Thus, if bending energy was minimized, it is recommended to use wireframe plots that draw curve outlines.

It is also possible to estimate the relative contribution (%) of each landmark to the PCs[58] (Step 25). Landmarks that contribute more than expected can be explored and selectively displayed, for example in the context of a lollipop plot (Step 24), to emphasize which aspects of the anatomy most prominently influence the shape change along the respective PC. Taken together, PCA plots, deformation grids, wireframes and lollipop plots facilitate an evocative visualization and thorough interpretation of differences in mean shape.

Identification of groups through clustering methods

Suspected group structure(s) in a data set can then be identified by clustering. A plethora of clustering software is available for R[42], and we implement two commonly used methods, *k*-medoid clustering[59,60] (Step 27A) and model-based clustering[61,62] (Step 27B). In *k*-medoid clustering, a medoid is defined as the central-most data point in a cluster with the smallest sum of distances to all other observations within the cluster[59,60]. This is not to be confused with an alternative approach, called '*k*-means clustering', in which data are partitioned around the mean (i.e., not around an actual data point) of a cluster[59,60]. First, a set of *k* representative medoids is chosen, and *k* corresponding clusters are generated around them. The algorithm then iteratively tests for better representative medoids by analyzing all possible combinations of representative medoids and non-representative observations for *k* clusters. The goal is achieved once a set of *k* medoids is identified that minimizes the sum of the dissimilarities of all observations to their closest medoids[59,60]. A heuristic and empirical way to determine the best number of medoids (*k*) for a given data set is to plot the average dissimilarity within the clusters against the number of clusters in which the data set can be arbitrarily partitioned (Step 27A(i)).

Alternatively, model-based clustering can utilize finite sets of Gaussian mixture models to identify group structures, which might underlie a data set[61,62]. Here, the best Gaussian mixture model for clustering, and therefore the best number of clusters, is selected according to the Bayesian information criterion[61]. In contrast to *k*-medoid clustering, model-based clustering independently identifies the best number of clusters and assorts all observations accordingly[62] (Step 27B(ii)). Both clustering approaches can be used to identify the presence and the number of clusters in the morphospace. This might be of particular use if, for example, one wishes to identify the number of distinct morphs in polyphenic nematodes a priori (i.e., without assuming the number of morphs). Thus, the clustering approaches can be used to complement the results obtained from Procrustes ANOVA, as the latter tests for differences between groups that have been defined before the analysis. Alternatively, previously unknown group structures that were identified by clustering can be recursively tested with Procrustes ANOVA.

Advantages and limitations of the method

Geometric morphometrics is a powerful technique for two explicit reasons. First, unlike traditional morphometrics, it captures information on spatial relations among landmarks, and it preserves these data throughout the entire analysis (statistical testing, PCA visualization, etc.). Second, it allows quantitative visualization of shape changes in an immediate anatomical context[38,50]. Thus, even the most complex morphological changes can be visualized and communicated in an effective and intuitive manner[36,38,39,50]. The set of landmarks depends on the microscopic animal of choice and particular anatomical question. In our exemplary data set, only a basic understanding of the nematode stoma is necessary to apply our proposed landmark configurations. However, users might need to practice landmark annotation before robust and repeatable data sets can be obtained for comparisons between different species and strains. In addition, only an elementary knowledge of the R language is necessary to apply the present protocol, as the entire code is included in the paper, and the computational workflows for geometric morphometrics and cluster analysis, which are utilized here, are well established and supported by active online communities[45–48,59–62]. Therefore, no particular expertise is needed to apply this protocol, rendering it feasible for all academic levels from undergraduates to postdoctoral researchers.

A current limitation of this protocol is that it is conceived for the analysis of 2D landmark data sets only, and that depending on the research question at hand, new sets of landmarks may need to be

defined. Furthermore, we have described semilandmarks to reconstruct curves, but a slightly modified procedure for surfaces may be required depending on the anatomical structure of interest. Finally, updates to the software packages used here might result in changes to the commands or in removal of deprecated functions, which might require adjustments to the code presented in this protocol in future applications.

Applications and future developments

The protocol can be used and further developed in multiple circumstances in nematodes and other microscopic animals. As described here, it can be used for the quantification of both types of morphological variation, genetic and environmentally induced (phenotypic plasticity), in adult *Pristionchus* and *Caenorhabditis* mouth structures. This includes functional genetic experiments (i.e., CRISPR-induced mutation) and ecomorphological studies (i.e., multiple environmental conditions).

Second, although developed for *Caenorhabditis* or *Pristionchus*, the protocol can also be adapted to investigate the stomata of other nematode species. In such a phylogenetic context, the protocol can be applied to identify shape changes due to divergent evolution. In fact, recent studies demonstrated that geometric morphometrics can be useful in delimiting 'cryptic' centipede species[35]. Hence, this protocol could also be applied in future nematode species descriptions, where it might be of interest to visualize the shape differences between a reference dimorphism and any potential intermediate morphs, or additional stomatal morphs in highly plastic nematodes[18]. We also hope it will encourage members of the nematode research community to integrate geometric morphometrics into future studies focusing on morphological evolution and its underlying causal mechanisms.

Third, it can be adapted to quantify shape differences of other organs systems (e.g., spicules of male nematodes or imaginal wing discs in holometabolous insects like *Drosophila*). Fourth, one can define alternative landmarks sets, which can be applied to all larval stages to describe the ontogenetic trajectories along which shapes develop. Finally, our protocol can be easily adapted to other microscopic specimens, such as tardigrades, rotifers, copepods or mites (see 'Experimental design'). Therefore, this simple, ready-to-use protocol fills a significant gap in the standard toolkit of evo-devo approaches on microscopic animals.

Experimental design

Adapting the protocol for other microscopic animals

In the step-by-step protocol below, we demonstrate our approach by performing geometric morphometric analysis of nematode mouth structures. However, we anticipate that our protocol is broadly applicable to several other microscopic animals and structures. Our experimental design is based on image stacks acquired with differential interference contrast (DIC) on a standard upright light microscope. Thus, colleagues working on largely translucent animals such as tardigrades, rotifers, certain crustaceans (e.g., *Daphnia*), kinorhynchs or loriciferans will profit most from our protocol, as they will be able to perform GM analysis on external and internal organs. Other microscopic animals like certain taxa of mites or copepods might be too opaque (due to stronger sclerotization or pigmentation) to clearly identify internal structures in standard DIC image stacks; however, GM on external body structures will still be possible.

The positioning of the animal will always play an important role for GM analysis. The orientation in which the animal is placed has to be as consistent as possible across specimens. Achieving this can be a challenge in animals that have rounded cross-sections, like nematodes or loriciferans. In such cases, it might be useful to prepare a larger number of individuals for imaging to ensure that enough samples in the proper orientation can be found (i.e., two to three times the number of desired samples). In the example described herein for nematodes, we used agarose pads as a substrate on which to place our animals and applied an anesthetic to keep them in position. This procedure works well in nematodes, and we suspect can similarly be applied to other microscopic animals. However, the diameter of some animals may be too large to place directly between an agarose pad and a cover slip. Instead, one can place specimens in a droplet of buffer and deposit small silicone pads in each corner of the cover slip as distance-holders between the glasses. For handling nematodes, we use self-made platinum sticks (see 'Reagent setup'), which also allow handling of other small (but robust) animals such as mites or tardigrades. Very fragile specimens can be picked up with an eyelash glued to a tooth pick or directly placed in buffer and subsequently transferred to the microscope slide using a pipette. Apart from the specific requirements in maintaining and handling specimen cultures, the largest modification that will need to be made is to define a landmark configuration for the structure

of interest. The most important component to defining a configuration is to identify landmarks that can be reliably identified across all specimens. Thus, landmark configurations that were initially conceived to compare adults of a species might not be applicable to larval stages as well, since certain structures could still be absent (or rudimentary) in earlier developmental stages. These issues aside, only minimal changes need to be made to the code chunks given below to analyze and visualize the data, including: (i) defining sliding landmarks and curves on structures that lack options for fixed landmarks (Step 13), (ii) defining grouping factors of interest (e.g., species, culture conditions or larval stages) (Steps 14–19), (iii) defining the outlines for the wireframe plots (Step 24) and (iv) adjusting the model formula in the allometry assessment to visualize group-specific trends (Step 28).

Example nematode data sets

The stomatal polyphenism of *Pristionchus* nematodes has become one of the most promising models in plasticity and novelty research. Therefore, we used this character complex to generate example data sets for this protocol (available as Supplementary Data 1–3). We obtained landmark data sets from the stomata of both morphs of *P. pacificus*, the secondarily monomorphic species *Pristionchus bucculentus* and *Pristionchus elegans*, as well as the stomata of *C. elegans* and *Caenorhabditis briggsae*, and performed exemplary geometric morphometric analyses to quantify shape differences between young adult hermaphrodites (or females). In addition, we performed exemplary clustering analyses on the PCA results obtained for the intraspecific comparison of *P. pacificus* morphs and for the interspecific comparison of stomata in several *Pristionchus* species. Lastly, we visualized the relationship of stomatal sizes and shapes in our data sets, to see whether shape differences on the species level or the polyphenism level are manifestations of allometry.

We followed the standard protocol for establishing and maintaining laboratory cultures of diplogastrid and rhabditid nematodes[63]. All species mentioned in this protocol were cultured on nematode growth medium agar plates for multiple generations using the uracil auxotroph OP50 strain of *Escherichia coli* as food source. Agar plates were stored in plastic boxes and kept in an incubator at a constant temperature of 20 °C. For microscopy, all specimens were mounted on slides with agarose pads containing sodium azide (NaN_3) to sedate the animals (Steps 1–3). All specimens were examined using a Zeiss Axio Imager.Z1 microscope with a Zeiss Plan-Apochromat 100 × 1.4 DIC objective, together with immersion oil. All image stacks were taken in identical x-/y-dimensions (800 × 684 pixels) using a Zeiss Axiocam 506 mono charge-coupled device (CCD) camera (Steps 4 and 5). Landmark coordinates of 80 individuals per data set were acquired using Fiji's[64] multi-point measurement tool (Steps 6–8). All data analysis steps (GPA, PCA, multivariate statistics and clustering) were performed in RStudio, using the R statistical computing environment[42]. The geometric morphometric procedures were performed using the geomorph package[45,46].

Specific instructions for landmark configurations for model nematode stomata

The stomata of model nematodes, like *Pristionchus* and *Caenorhabditis*, can generally be divided into three cuticular compartments: cheilostom, gymnostom and stegostom (Fig. 1; Supplementary Fig. 1)[25]. The cheilostom is the anterior-most compartment of the stoma, which in *C. elegans* is secreted by hypodermal cells[27]. The gymnostom forms the second compartment of the stoma and is secreted by the arcade syncytium[25,27]. The stegostom is the posterior-most compartment of the stoma and is further subdivided into pro-, meso-, meta- and telostegostom[25]. The promesostegostom is secreted by pharyngeal epithelial cells, while meta- and telostegostom are secreted by pharyngeal muscle cells[25,27]. The metastegostom forms the characteristic teeth (i.e., the morphological novelty) in diplogastrids such as *P. pacificus*, whereas it forms simple triangular 'flaps' in other species, like *C. elegans* (Fig. 2a,b)[26].

The two-dimensional landmark configuration proposed in Fig. 1 (also see Supplementary Fig. 1 and Supplementary Table 1) can likely be used for most species of the *Pristionchus* genus and many other nematodes of the Diplogastridae family, as the landmarks are placed on homologous structures, which are easily identifiable under a standard DIC microscope with a 100× oil objective. It is important to note that the proposed landmark configurations are suitable to quantify shape differences only in adult diplogastrids or *Caenorhabditis* nematodes, in which these landmarks can be appropriately identified. Thus, depending on the research question at hand—nematodes of different taxa, different developmental stages or other microscopic organisms—alternative or modified landmark configurations might need to be defined.

It is important to make sure that the orientation of all specimens is as similar as possible (Steps 3 and 4) to not introduce systematic error into the GM analysis. Since most nematodes tend to be lying on a side (either right or left), we applied the following criteria to make sure animals are facing right

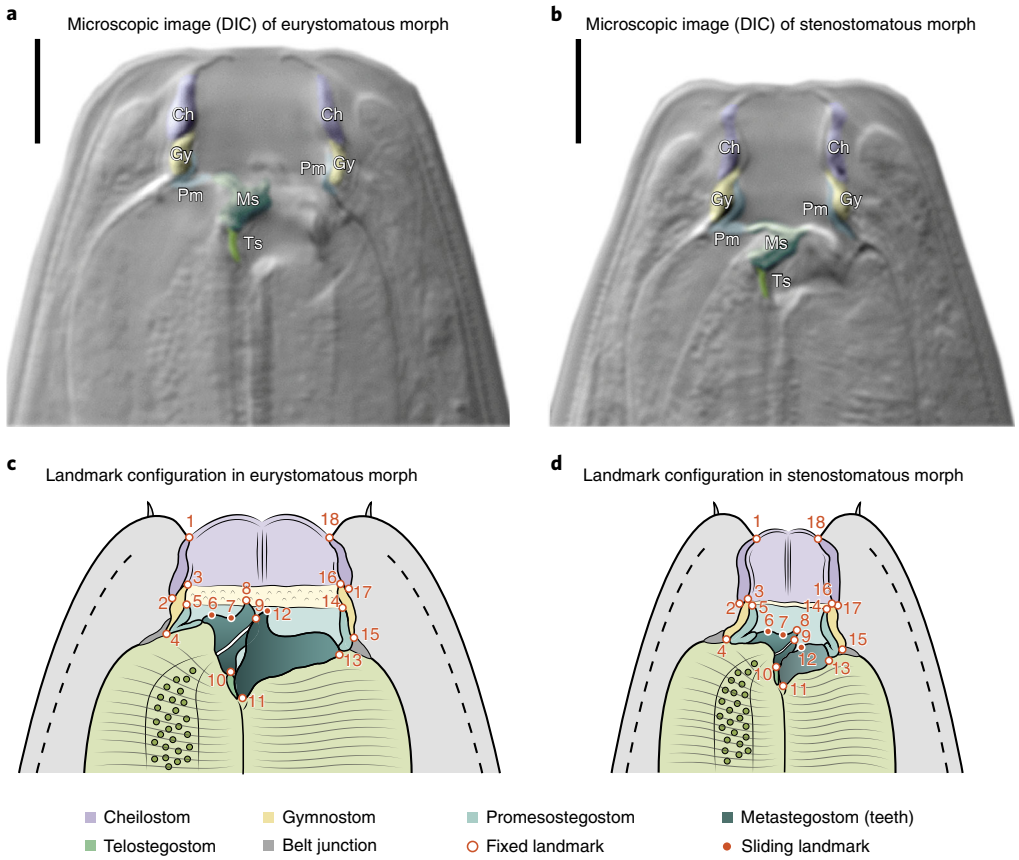


Fig. 1 | The stomatal polyphenism in *Pristionchus pacificus*. **a** and **b**, Overview of the stomatal cuticle elements in the eurystomatous (**a**) and stenostomatous (**b**) morph of *P. pacificus* (RS5205) under a 100× magnification using a DIC light microscope (only median plain; right lateral perspective). **c** and **d**, Schematic representation of the entire stomata of both morphs including the proposed landmark configurations (right lateral perspective; for a description of the landmarks, see Supplementary Table 1). Ch, cheilostom; Gy, gymnostom; Ms, metastegostom; Pm, promesostegostom; Ts, telostegostom. Scale bars in **a** and **b** are 5 µm.

side upwards (the right ventrosublateral tooth in diplogastrids is better visible in this position): (i) if the anterior end of the worm is on top of the field of view, and the posterior end of the worm is on the bottom field of view, the ventrally located vulva (or cloaca in males) must be found on the right side of the field of view, as the right lateral side of the animal is facing the viewer; (ii) confirm that the head of the worm is not rotated around its anterior-posterior axis by checking that the amphid opening is located close to the central axis of the animal, but usually with a slight offset toward the dorsal side of the body (i.e., opposite to the vulva or cloaca), and by checking that the right and left labial papillae align as close as possible with the central axis of the body; and (iii) confirm that the tips of the right and left labial papillae are approximately at the same level along the anterior-posterior axis to ensure that the head is not rotated around the dorsoventral body axis (i.e., tilted toward one of the lateral body sides). Note that rotation around the left-right lateral axis (i.e., the general orientation angle in which the worm appears in the image) can be ignored, as this kind of rotation is removed by GPA. Although not difficult, finding specimens that fit these criteria is the most time-consuming step in the protocol (Steps 1–5). In our experience, we typically found four to eight nicely oriented individuals per slide, which roughly equates to 10–20% of total specimens.

Sample size

It is important to have an appropriate sample size to perform geometric morphometric analyses and downstream multivariate statistical hypothesis testing. The number of specimens depends on the specific statistical test that is used for analysis of group differences. The permutation-based test designs, like PERMANOVA described above, do not assume any kind of distribution (e.g., Gaussian) of the data, and they are less sensitive to smaller sample sizes. Still, we recommend having at least two

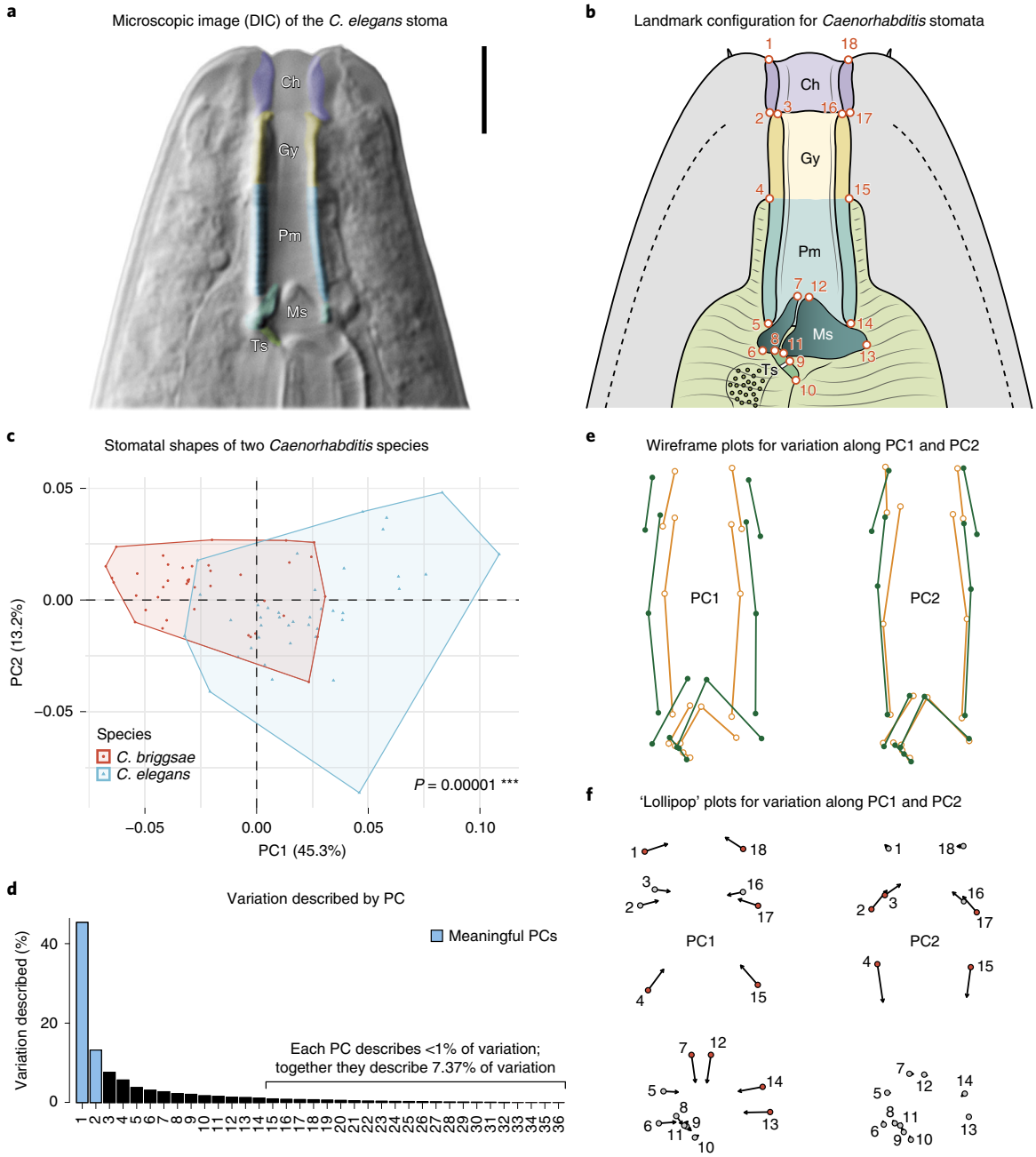


Fig. 2 | Analysis of shape differences between the stomata of *C. elegans* (N2) and *C. briggsae* (AF16). a, 100× DIC image of the *C. elegans* stoma (only median plain; right lateral perspective). **b**, Proposed landmark configuration for the *Caenorhabditis* stoma (for a description of the landmarks, see Supplementary Table 2). Ch, cheilostom; Gy, gymnostom; Ms, metastegostom; Pm, promesostegostom; Ts, telostegostom. **c**, PCA plot showing the distribution of specimens (i.e., individual shapes) in a tangent space. *P* value obtained by Procrustes ANOVA. F-statistic: 19.928; effect size (Z): 5.5305; d.f.: 1 for species, 78 for residuals (total d.f.: 79). **d**, Barplot showing the variation described by each PC. Meaningful PCs are estimated by the getMeaningfulPCs() function of the Morpho package. **e**, Wireframe representations of superimposed stomatal extreme shapes along PC1 and PC2. Dark green, maximum (+) extreme; dark orange, minimum (−) extreme. **f**, Lollipop representations of superimposed stomatal extreme shapes along PC1 and PC2. Circles, maximum extreme; arrowtips, minimum (−) extreme. Landmark numbers are indicated in Arabic numerals. Red, landmarks that contribute more than average to the respective PC. Note that, when semilandmarks are slid by minimizing bending energy, their exact position on the curve cannot be interpreted easily. Thus, we recommend generating wireframe plots, rather than lollipop plots, in such cases. Scale bar in **a** is 5 µm.

to four times more individuals than shape variables[36] to obtain robust results with Procrustes ANOVA. In our examples, we used 18 landmarks each for *Pristionchus* or *Caenorhabditis* species (Fig. 1c,d; Supplementary Fig. 1; Fig. 2b), which means that there are 36 initial variables (i.e., x and y landmark coordinates) in every analysis. During Procrustes superimposition, four dimensions are lost due to removal of differences in location (two dimensions), orientation and size (one dimension

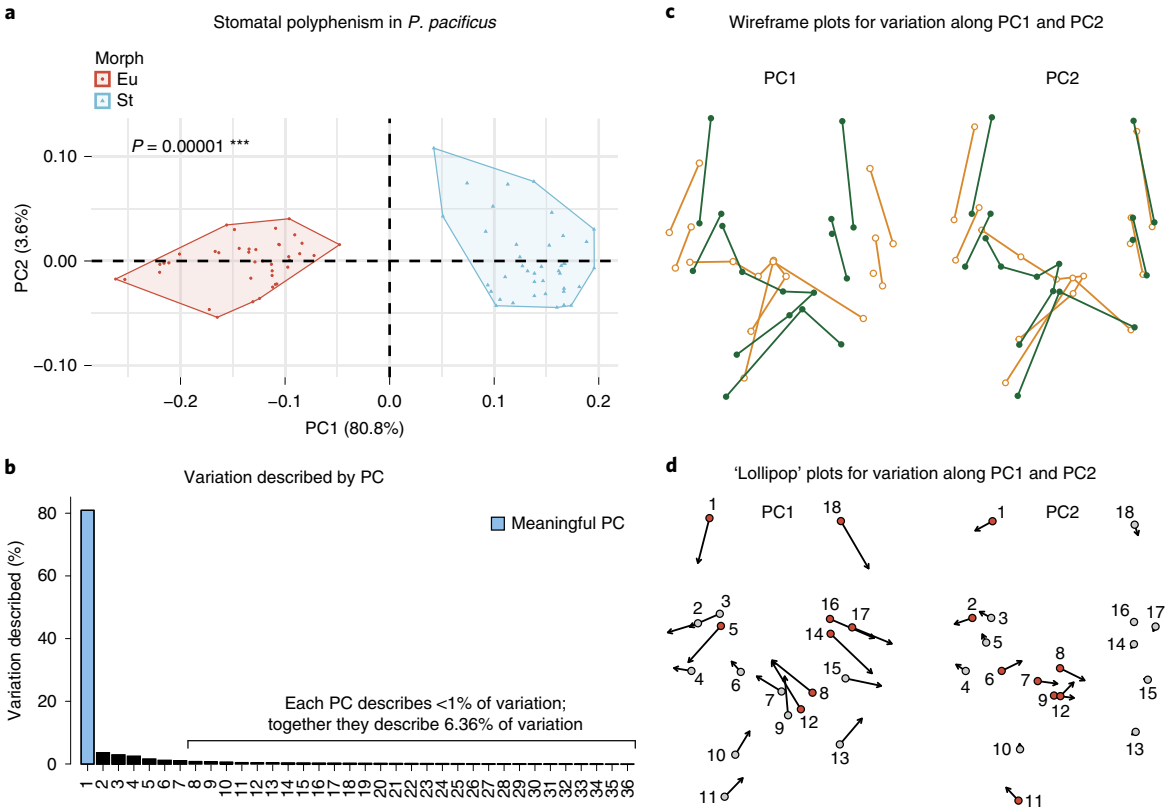


Fig. 3 | Analysis of shape differences between the two stomatal morphs of *P. pacificus* (RS5205). a, PCA plot showing the distribution of specimens (i.e., individual shapes) in a tangent space. *P* value obtained by Procrustes ANOVA. F-statistic: 218.31; effect size (Z): 6.0452; d.f.: 1 for morphs, 78 for residuals (total d.f.: 79). Eu, eurystomatous morph; St, stenostomatous morph. **b**, Barplot showing the variation described by each PC. Meaningful PCs are estimated by the getMeaningfulPCs() function of the Morpho package. **c**, Wireframe representations of superimposed stomatal extreme shapes along PC1 and PC2. Dark green, maximum (+) extreme; dark orange, minimum (−) extreme. **d**, Lollipop representations of superimposed stomatal extreme shapes along PC1 and PC2. Circles, maximum (+) extreme; arrowtips, minimum (−) extreme. Landmark numbers are indicated in Arabic numerals. Red, landmarks that contribute more than average to the respective PC. Note that, when semilandmarks are slid by minimizing bending energy, their exact position on the curve cannot be interpreted easily. Thus, we recommend generating wireframe plots, rather than lollipop plots, in such cases.

each)[44]. Therefore, we end up with 36 − 4 = 32 shape variables, and therefore ≥64–128 specimens should be contained in such datasets (here, we arbitrarily chose 80 animals for each example data set; Figs. 2c,3a and 4a). However, larger data sets might be necessary if the shape differences between groups are very small. In addition, we used equal numbers of individuals for each morph in the example analyses. However, it might be more useful to have different morphs represented in numbers that reflect morph frequency, especially if one aims to estimate the frequency from the results of the GM analysis. (Note that the *P. pacificus* strain RS5205 we use here shows a 1:1 morph ratio under standard laboratory conditions.) In summary, we recommend two to four times more individuals than shape variables with our pipeline for PERMANOVA statistical testing; yet, one has to make sure that the data set is large enough to address the specific question at hand.

Replicability and error

There are many sources of error in geometric morphometric analyses, including poor biological material or images, inherently variable anatomical structures and imprecise landmark labeling by the annotator. Here, it is crucial to assess whether a newly defined landmark configuration, such as the one proposed in Fig. 1 for model nematodes, can actually be used to produce clear and reproducible results. One can use several methods to assess the replicability of new landmark configurations for any microscopic animal. In this protocol, we provide two related tools to calculate replicability (Box 1): (i) estimation of the average repeatability (RPT)[65,66] of a shape by averaging the repeatability of Procrustes coordinates, and (ii) calculating the percent measurement error (%ME)[43,44,67] for shapes using PERMANOVA.

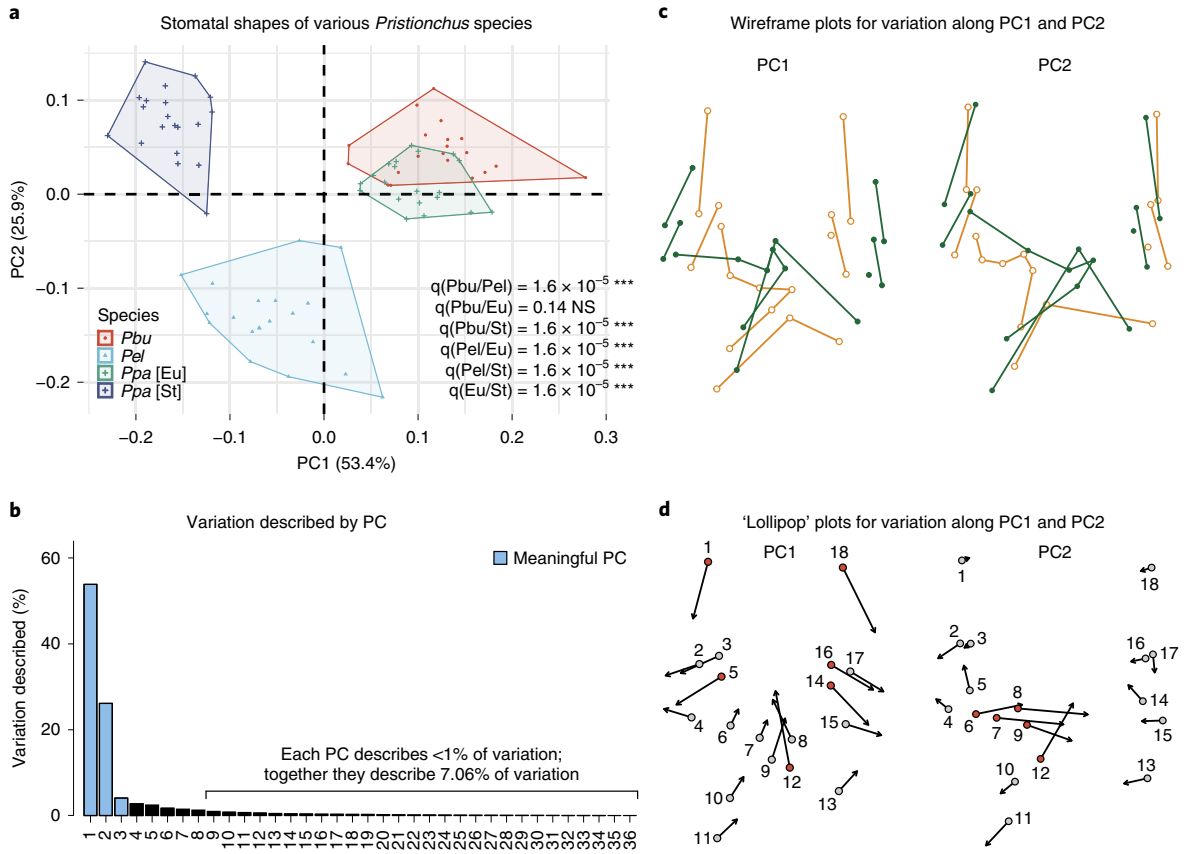


Fig. 4 | Analysis of shape differences between the stomatal morphs of *P. pacificus* (RS5205), *P. bucculentus* (RS5596) and *P. elegans* (RS5698). **a**, PCA plot showing the distribution of specimens (i.e., individual shapes) in a tangent space. *P* values obtained by Procrustes ANOVA and subsequently corrected for false discoveries (*q* values). F-statistic: 61.769; effect size (Z): 8.8245; d.f.: 3 for species, 76 for residuals (total d.f.: 79). Pairwise effect sizes (Z) for *Pbu:Pel* = 10.2778, *Pbu:Ppa[Eu]* = 1.3238, *Pbu:Ppa[St]* = 12.2343, *Pel:Ppa[Eu]* = 8.3855, *Pel:Ppa[St]* = 9.5263, *Ppa[Eu]:Ppa[St]* = 11.269. NS, not significant; Eu, eurystomatous morph; St, stenostomatous morph; *Pbu*, *P. bucculentus*; *Pel*, *P. elegans*; *Ppa*, *P. pacificus*. **b**, Barplot showing the variation described by each PC. Meaningful PCs are estimated by the getMeaningfulPCs() function of the Morpho package. **c**, Wireframe representations of superimposed stomatal extreme shapes along PC1 and PC2. Dark green, maximum (+) extreme; dark orange, minimum (−) extreme. **d**, Lollipop representations of superimposed stomatal extreme shapes along PC1 and PC2. Circles, minimum (−) extreme; arrowtips, maximum (+) extreme. Landmark numbers are indicated in Arabic numerals. Red, landmarks that contribute more than average to the respective PC. Note that, when semilandmarks are slid by minimizing bending energy, their exact position on the curve cannot be interpreted easily. Thus, we recommend generating wireframe plots, rather than lollipop plots, in such cases.

In this protocol, we replicated stomatal shapes for 20 individual *P. pacificus* nematodes (10 eurystomatous and 10 stenostomatous) using the configuration proposed in Fig. 1, and compared replicability for three annotators with varying levels of experience on stomatal morphology. Each of the annotators labelled all worms three times, with 20-min intervals between labelling specimens, on three different days spread over a 5-d period. Note that it was not possible to re-mount and re-image the same worm several times, as the animals do not survive such repeated stress. The 20 individuals were randomly selected from culture plates, and all specimen IDs had been encrypted during the replication sessions.

The RPT describes the fraction of the total phenotypic variance in a data set that can be attributed to variation among (biological replicates) or within (technical replicates) groups[65,66]. In our case, groups refer to the 20 individual worms. As an example, a repeatability of 80% would indicate that 80% of the total phenotypic variation is attributable to the variation that is naturally present between the 20 different specimens, while the remaining 20% of variation is attributable to differences between the three technical replicates obtained for each individual. In other words, 20% of the total phenotypic variance is due to the annotator's imprecision of placing landmarks in the same specimen over several replicates. Here, we used the rptR package (version 0.9.22)[66] in R to estimate average repeatabilities obtained for each shape coordinate of all replicates (technical and biological). In addition, we calculated %ME for shapes based on a PERMANOVA design, following Claude[43,44]. This approach, similar to RPT, attempts to decompose the total phenotypic variation into within-group

Box 1 | Testing the replicability of shapes by estimating RPT and %ME

Before starting the replicability analysis, perform Steps 1–20 to obtain shape coordinates for a landmark data set containing biological and technical replicates (for details, see 'Replicability and error' in 'Experimental design'). Expand the specimen ID ('ID=') so that it contains an additional part specifying the replicate number (e.g., 'RS5205-Eu-ind-1-2', where 1 indicates the number of the worm, and 2 indicates the technical replicate of that worm). Make sure to accordingly change the number of parts in Step 15 and to expand the code in Steps 16 and 17 so that you also extract the grouping factor information for 'specimen' and 'replicate' (e.g., we used 'BR.part' and 'BR' to get the information on the organism and 'TR.part' and 'TR' to extract the respective technical replicate) (see also Supplementary Data 4 and 5).

▲ **CRITICAL** When testing the code to estimate the RPT of your data, the number of parametric bootstraps and/or permutations should be changed from '1000' to, e.g., '10', as the full analysis will run for several minutes with both values set to '1000'.

▲ **CRITICAL** To calculate the %ME properly, the number of technical replicates (i.e., repeated measurements of the same animal) needs to be adjusted manually in the line that estimates 's2among'.

Procedure

To start the replicability analysis, run the following code:

```
```{r}
##Option-A: repeatability (RPT) estimation for shape coordinates
#load rptR package for repeatability estimation
library(rptR)
#generate input data frame for rptR package based on your superimposed landmarks:
#Input-Option-1 if superimposition was generated with *geomorph* use:
rptr_input <- sup.geo$data[grep("coord",colnames(sup.geo$data))]
#Input-Option-2 if superimposition was generated with *Morpho* use:
rptr_input <- data.frame(two.d.array(sup.mor$orpdata))
colnames(rptr_input) <- paste("coords",rep(seq(1,ncol(rptr_input)/2),each=2),rep(c("X","Y"),ncol
(rptr_input)/2),sep=".")
#add a column with the names of the individuals
rptr_input <- cbind(rptr_input,Individual=paste(SP,MF,BR,sep="-"))
#rptR - Repeatability estimation for individual coordinates
rptR_RPT <- c()
rptR_SE <- c()
for (i in 1:(ncol(rptr_input)-1)) {
formula.i <- formula(paste(colnames(rptr_input)[i]," ~
(1|Individual)",sep=""))
rptr_resul.i <- rptGaussian(formula = formula.i, grname =
"Individual", nboot = 1000, npermut = 1000, data = rptr_input)
rptR_RPT <- c(rptR_RPT,rptr_resul.i$R[1,1])
rptR_SE <- c(rptR_SE,rptr_resul.i$se[1,1])
}
# estimate average repeatability and standard error for whole shapes
mean(rptR_RPT)
mean(rptR_SE)
##Option-B: estimate percent measurement error (%ME)
#prepare a geomorph data frame for statistical analysis
ME_input <- geomorph.data.frame(shape = sup.geo$coords, individual = BR, replicate = TR)
#BR = biological replicate and TR = technical replicate
#Permutational Anova for replicate factor
replicate_factor <- summary(procD.lm(shape ~ TR, data = ME_input, iter = 1000, seed = 12345, RRPP = F, effect.
type = "F"))
#view influence of replicates on shape differences
replicate_factor
#Permutational Anova for specimen factor
individual_MS <- summary(procD.lm(shape ~ BR, data = ME_input, iter = 1000, seed = 12345, RRPP = F, effect.type
= "F"))
#view influence of individual differences on shape differences
individual_MS
# defining variation within individuals (across technical replicates)
MSwithin <- individual_MS[[1]] [2,3]
s2within <- MSwithin
# defining variation among individuals (across specimens)
MSamong <- individual_MS[[1]] [1,3]
s2among <- (MSamong - MSwithin)/3
#adjust number of technical replicates (here "3")
#calculate percentage of measurement error(ME) in the data set
ME <- s2within/(s2within + s2among) * 100
ME
```
```

and among-group variation components. However, %ME is based on the mean squares obtained via a PERMANOVA with individuals as sources of variation[43] and estimates the within-group variation (i.e., annotation imprecision)[67]. Meanwhile, RPT directly estimates the among-group variation (also in percentages)[65]. Thus, repeatability and %ME are complementary approaches for estimating replicability of landmark placement.

The estimated RPTs and %ME values in our exemplary data set clearly show differences in replicability of shapes (Supplementary Figs. 6 and 7), which related to the level of experience of the annotator. Interestingly, stenostomatous animals consistently show a smaller replicability across all users. Nevertheless, our replicability assessment passed two thresholds: (i) the amount of among-group variation was greater than the within-group variation (i.e., >50%, except for one case where 100 – %ME was 47.9%); and (ii) for %ME, the specimen factor was always significant for both *P. pacificus* morphs while the technical replicate factor was not (Supplementary Fig. 6 legend). This indicates that the shape variation across worms was stronger than the variation across replicates of the same individual[44]. Thus, our proposed landmark configuration (Fig. 1) can be used to produce robust shape data sets with high repeatability and a normal measurement error across different levels of experience. We advise checking your ability to replicate measurements using the same criteria before using a new landmark configuration for shape analysis. In our experience, practicing landmark placement for a few days before obtaining the first real data set will vastly increase replicability of results (Supplementary Figs. 6 and 7).

Another concern is that the anatomical character complex of interest contains moveable elements. Such structures can bias the entire GM analysis, since GPA aligns landmark configurations by translation, rotation and scaling without allowing for transformations like shearing. Thus, moveable elements oftentimes introduce systematic error, as their position relative to the rest of the landmark configuration can vary considerably. This caveat renders replicability assessments even more important. However, in addition to estimating shape repeatability or the percent measurement error of an annotator, we can also confirm the variance of each landmark across all specimens in a shape data set. This way, one can investigate whether landmarks that are placed on a moveable element collectively show increased variance in their position (after GPA), relative to the rest of the shape. This can be indicative of relative movements, which could lead the annotator to remove these specific landmarks from the analysis as they might introduce more systematic error than they provide morphometric information. For instance, we wondered if the metastegostomal teeth or flaps in our example nematode data set, which can be moved by muscular contractions, contribute excess variability compared to immovable elements. The NaN_3 contained within the agarose pads sedates the worms and causes their musculature to relax. Thus, we expected that the teeth and flaps should always be in the same (non-protruded) position relative to the rest of the buccal cavity. Nevertheless, to validate that these movable elements do not introduce systematic errors, we performed a variance analysis of landmark position as mentioned above. Specifically, we performed Procrustes alignment (GPA) on 40 'Eu' and 40 'St' specimens of *P. pacificus*, as well as 40 *C. elegans* specimens. Afterwards, we estimated the variance of each landmark coordinate across each of the 40 individuals of the three different groups mentioned above. Then, we summed up the variances of the x and y coordinates of corresponding landmarks to obtain the total variance of each landmark in each of the three groups. This logic can be applied to any potentially moveable element in a microscopic organism. From our data, it is apparent that landmarks that are jointly placed on movable elements (teeth and flaps) do not show increased variances as compared to landmarks placed on non-movable elements (Supplementary Fig. 2), confirming that all of the proposed landmarks can be treated identically within both model species.

Allometry

An important phenomenon one might want to consider when interpreting shape differences is allometry, the relationship of biological shape with size[68–70]. A common observation is that allometry is a major factor of shape variation in many morphometric data sets, as shapes of animals often change when they grow. To investigate ontogenetic allometry[68,69], one can establish trajectories of shape change throughout development and subsequently compare their slopes across several groups of interest (typically species). However, it is also possible to check for static allometry (i.e., the size-shape covariation in animals within the same developmental stage), for example, in adults of different species or strains. Thus, static allometry can provide important insight into the coevolution of size and shape across species and facilitate studies on functional adaptation[70].

Given its importance, several approaches have been applied to visualize allometry. A widely used class of plots that are based on linear models of shape change proved particularly informative when comparing allometries across different groups[39,46]. Here, we include one such approach (Step 28), in which a PCA is performed on predicted values obtained from a multivariate regression of shape on size, and the first PC of this PCA is plotted against the logarithm of the centroid size of each individual[39,68,69]. This way, one can obtain stylized graphs that show group-specific allometric trends in the form of straight lines. The slopes of these trend lines can then be compared to assess the presence of group allometry and to interpret differences in the degree of shape-size covariation between groups. In addition, statistical significance can be evaluated using, for example, the multivariate PERMANOVA approach we describe above.

Materials

Biological material

The approach described in this protocol can be applied to a broad range of microscopic animals (for details, see 'Adapting the protocol for other microscopic animals' in 'Experimental design'). In the example shown in this protocol, we used the African strain RS5205 of *P. pacificus* (morph ratio of 1:1), the RS5596 strain of *P. bucculentus*, the RS5720 strain of *P. elegans*, the N2 reference strain of *C. elegans* and the AF16 reference strain of *C. briggsae*. Upon request, the Sommer laboratory can provide a plethora of *Pristionchus* strains, including the following: the reference strain PS312, numerous naturally isolated lines from multiple locations around the world, various transgenic and CRISPR-edited strains and a broad spectrum of other diplogastrid species and genera **! CAUTION** All national laws and institutional directives on animal welfare and the contained use of genetically modified organisms must be followed when working with such strains. The nematode strains/species used in this study were not genetically modified and are not included in the animal welfare act of Germany. Therefore, no ethical approval or guidance was required to work with any of the selected strains/species.

Reagents

▲ CRITICAL The reagents specified in this section are specific for housing and sample preparation of nematodes. Different microscopic animals will require specific growth media and possibly different immobilization strategies (see 'Experimental design').

- Absolute ethanol (Merck, cat. no. 986) **! CAUTION** Ethanol is flammable and may cause serious eye irritation. Make sure to wear safety glasses while handling the substance and avoid using it close to open flames.
- Agarose (Sigma, cat. no. A9539)
- $CaCl_2$ (Sigma, cat. no. C5080)
- Cholesterol (Fluka, cat. no. 26740)
- KH_2PO_4 (Sigma-Aldrich, cat. no. P3786)
- $MgSO_4$ (Fluka, cat. no. 63135-1KG-F)
- NaCl (Roth, cat. no. 3957.1)
- Na_2HPO_4 (cat. no, S5136)
- NaN_3 (Sigma, cat. no. S8032) **! CAUTION** Sodium azide can be lethal if swallowed, absorbed through skin or inhaled; it may cause organ damage and is toxic to aquatic organisms. Make sure to protect yourself from skin contact by wearing Nitrile gloves.
- Tryptone (Fluka, cat. no. 61044)

Equipment

Equipment for nematode handling and housing

▲ CRITICAL The equipment in this section is necessary for housing and handling of nematodes. Different microscopic animals may require specific handling strategies or housing conditions (see 'Experimental design').

- Glass alcohol burner (DWK Life Sciences, cat. no. 04-245-1)
- Glass pasteur pipettes (BRAND GmbH, cat. no. 7477 15)
- Incubator (Thermo Scientific Heraeus BK 6160, model no. 10759151)
- Platinum strips, 0.3 mm (Häberle, cat. no. 9.160 703)
- Petri dishes (Greiner Bio-One, cat. no. 628102)

Imaging equipment

- Block heater (Techne, Dri-Block heater, model no. DB-3)
- CCD microscope camera (Zeiss Axiocam 506 mono, model no. 426557-0000-000)
- Coverslips, 18 × 18 mm (Carl Roth, cat. no. 0657)
- Flexible light guide (SCHOTT, cat. no. P/N 121.101)
- Gloves (e.g., KIMTECH Sterling nitrile, cat. no. 99212)
- Immersion oil (Zeiss Immersol 518F, cat. no. 444962-0000-000)
- Light source (SCHOTT KL 300 LED, cat. no. P/N 120 300)
- Microscope (Zeiss, model no. Axio Imager.Z1)
- Microscope lens swipes (Assistent, cat. no. 41019010)
- Microscope objective with DIC (Zeiss Plan-Apochromat 100 × 1.4 Oil, cat. no. 440780-9904)
- Microscope slides 76 × 26 mm (Carl Roth, cat. no. H868)
- Pipettes (2–20 µl, 20–200 µl, 100–1,000 ml) (Eppendorf, model Research @ plus, cat. nos. 3123000098, 3123000055 and 3123000063)
- Pipette tips (2–20 µl, 20–200 µl, 100–1,000 ml) (Greiner Bio-One, cat. nos. 765290, 739290 and 686295)
- Stereo microscope (Zeiss, Stemi 508, cat. no. 435064-9000-000)

Software for image acquisition and data analysis

- Fiji[64] (ImageJ) (version 2.0.0 or later; https://imagej.net/Fiji/Downloads)
- Microsoft Excel (version 15.15 or later) ▲ **CRITICAL** If you want to use Microsoft Excel, install version 15.15 or later. Alternatively, any other text editor of choice could be used to generate the landmark file.
- BBEdit (version 13 or later; https://www.barebones.com/products/bbedit/download.html)
- R (version 3.6.1; https://cran.r-project.org/mirrors.html) ▲ **CRITICAL** Make sure to install the most recent version of R (3.6.1 or later; https://cran.r-project.org/mirrors.html) before you start analyzing your data.
- RStudio (version 1.2.5001 or later; https://rstudio.com/products/rstudio/download/) ▲ **CRITICAL** Make sure to install the most recent version of RStudio (1.2.5001 or later; download available at https://www.rstudio.com) before you start analyzing your data.
- ZEN 2 Pro (Carl Zeiss Microscopy GmbH, 2011; version 2.0.14283.302; 64-bit)
- geomorph (R package version 3.2.1; https://cran.r-project.org/package=geomorph) ▲ **CRITICAL** Version 3.2.1 of geomorph was just released, and the inbuilt plotAllometry() function currently seems to be unable to plot group-specific allometric trends (see Step 28). However, a corrected version of this function is provided by the package authors under the following link: https://github.com/geomorphR/geomorph/blob/c027fbcf4c423a8e1b64b827966661d89194bcee/R/plotAllometry.r. Please copy the function code from there and run it once in the console of R to overwrite the inbuilt version before you run the code chunk in Step 28.
- Morpho (R package version 2.8; https://cran.r-project.org/web/packages/Morpho/index.html)
- cluster (R package version 2.1.0; https://cran.r-project.org/web/packages/cluster/index.html)
- mclust (R package version 5.4.6; https://cran.r-project.org/web/packages/mclust/index.html)
- factoextra (R package version 1.0.7; https://cran.r-project.org/web/packages/factoextra/index.html)

Reagent setup

▲ **CRITICAL** All items specified in Reagent setup are specific for housing and sample preparation of nematodes. Different microscopic animals likely require specific growth media and housing conditions (for details, see 'Adapting the protocol for other microscopic animals' in 'Experimental design').

- M9 buffer: Dissolve 3 g of KH_2PO_4, 6 g of Na_2HPO_4 and 5 g of NaCl into 1 l of double deionized water and mix thoroughly. Autoclave the solution. Add 1 ml of 1 M $MgSO_4$ to the autoclaved solution. Store indefinitely at 4 °C.
- Nematode growth medium: Dissolve 17 g of agar, 2.9 g of NaCl and 2.5 g of tryptone in 1 l of double deionized water and mix thoroughly. Autoclave the solution. Cool to 55 °C and add the following (while swirling): 1 ml of cholesterol (5 mg/ml in ethanol), 1 ml of 1 M $CaCl_2$, 1 ml of 1 M $MgSO_4$ and 25 ml of 1 M KH_2PO_4 (pH 6.0). Store indefinitely at 4 °C.
- Platinum picks: Insert an ~4-cm-long platinum wire into the slender opening of a glass Pasteur pipette and shortly melt the glass tip over a glass ethanol burner to permanently attach the wire to the glass pipette. Let the glass cool down for 1–2 min. Slightly flatten and/or curve the tip of the wire with a

hammer to get a device that can be slid under the microscopic animal to pick it up. Picks can be stored at the workbench and should be disinfected using glass ethanol burners before handling specimens.

Procedure

Digital microscopy ● Timing ~4 d for 80 animals (nematodes)

▲ CRITICAL The procedures for sample preparation (Steps 1–3) and microscopy data acquisition (Step 4) need to be optimized for different animals (for details, see 'Adapting the protocol for other microscopic animals' in 'Experimental design'). In this example, we specify how to generate a dataset for geometric morphometrics of nematode stomata.

1 Melt 4 ml of 5% (wt/vol) noble agarose in a microwave and add 40 µl of 1 M NaN_3 (10 mM final concentration). Use a block heater (set to ~70 °C) to keep the agarose solution liquid during the microscopy session.
! CAUTION NaN_3 may cause organ damage and can be lethal if swallowed, absorbed through skin or inhaled. Make sure to wear nitrile gloves to avoid skin contact with the substance while pipetting.

2 To get slides with agar pads, place three object slides next to each other (long sides contacting one another) and stick a line of common laboratory tape to the surface of the two outer slides to make them higher than the middle one. The strips of laboratory tape will act as spacers between the slides and ensure that the pad is of the right thickness. Use a pipette to place a drop of the agarose + NaN_3 solution on the object slide in the middle and flatten the drop into a pad by perpendicularly placing another object slide on top.

3 Place ~40 individual nematodes on each slide using 5–10 µl of M9 buffer as a medium between the agar pads and the coverslip.
▲ CRITICAL STEP Transfer nematodes (or other microscopic organisms) with self-made platinum picks and use glass burners filled with ethanol for sterilization of the pick.

4 Generate image stacks for the structure of interest in each specimen using a digital microscope with a CCD camera and a 100× DIC objective together with immersion oil.
▲ CRITICAL STEP All image stacks should be taken using the same settings. Keep the image dimensions constant for all image stacks that you obtain and use a roughly equal thickness (z-axis) for single images in all stacks. For nematode stomata, take ~40–50 images per z-stack. Note that image dimensions and the number of images per stack might need to be adjusted if the protocol is applied to anatomical structures other than stomata or to different microscopic animals.
? TROUBLESHOOTING

5 Save the image stacks as single files in either the native format for the microscopy software (e.g., '.czi' for Zeiss' 'Zen 2 Pro') or as single '.tif' files, to retain scale and size information.
▲ CRITICAL STEP Make sure to save all image stacks in the same file format, to avoid potential differences in scaling.

Obtaining landmark data for further analysis ● Timing ~6 h for 80 specimens

6 Open each image stack in Fiji[64] and place all (semi-)landmarks on their respective locations by using the multi-point tool (Toolbar → Multi-point tool).
▲ CRITICAL STEP Always use the same total number of (semi-)landmarks for all of your specimens and place them in their exact numerical sequence.

7 Apply the measurement tool of Fiji[64] (Analyze → Measure) to obtain the x and y coordinates of each landmark and copy these coordinates to a spreadsheet in Microsoft Excel. Repeat this for all of your specimens.
▲ CRITICAL STEP Make sure to copy/keep only the landmark coordinates in the spreadsheet and not any other information from Fiji's measurement panel.

8 Once all measurements are copied into a common spreadsheet, save this landmark file as 'lands.txt'.
▲ CRITICAL STEP To use this landmark file for geometric morphometric analysis in RStudio, it has to have a format that is readable for the respective functions. Box 2 provides a step-by-step guide on how to format the text file manually. Alternatively, we provide a short R routine that can be used for automated formatting in Supplementary Data 6. Note that the functionality of the R routine can potentially be affected by the version of the text editor that was used to generate the original .txt file. The correctly formatted landmark files for all our example data sets are provided in Supplementary Data 1–3 and can be used for comparison.

Box 2 | Manual formatting of the landmark file to fit the requirements of the readlands.tps() function

The landmark file can be formatted using any text editor, as long as the following rules are followed:

1 In Microsoft Excel, arrange all landmark coordinates for separate specimens in subsequent rows of the same two columns, with columns one and two representing the x and y coordinates, respectively. Note that each subsequent specimen has to be separated by an empty row.
2 Give the total number of landmarks per specimen as initial information, by putting 'LM=x' (no quotation marks; x = number of landmarks) in the first row of the first column of every specimen. Lastly, give an ID to each specimen. In our example, we use the following ID code: 'Strain-mouthform-individual-number'. For a eurystomatous specimen of the *P. pacificus* strain RS5205, the ID for the third individual in the analysis would look like this: 'RS5205-Eu-ind-3'. Put 'ID=yourID' (no quotation marks; yourID = the chosen ID code) into the first column (i.e., x-coordinate) of the last row of every individual specimen in the spreadsheet.
3 Copy and paste the landmarks obtained with the measurement tool of Fiji into the text editor of choice (e.g., BBEdit). Make sure to show hidden symbols (view tab→text display→show invisibles). Remove all tabs by selecting 'detab' from the 'text' tab and replace them with a single space. Remove 'space-next line' with just 'next-line' by finding ' \n' and replacing it with '\n' (without a preceding space). Finally, make sure there is a 'next line/return/' on the last line, and replace all commas in coordinates with periods.
4 Save the output as a text file (.txt) and name it 'lands.txt' (e.g., Supplementary Data 1–3).

Setting up R markdown for data analysis ● Timing 1–10 min

9 Install and load the R packages necessary for data analysis and visualization. These packages include the following: geomorph and Morpho[45–48], two alternative (but compatible) packages for GPA of points, curves and surfaces; cluster[60] for cluster analysis; mclust[61,62] for model-based clustering and factoextra[58] for extraction and visualization of the most important information generated by multivariate data analyses (e.g., PCA).

10 Open a new markdown file in RStudio and copy the code given below (including the code chunk delimiters ‴{r} and ‴) into the upper left pane of RStudio (markdown pane), or load one of the ready-to-use markdown files we provide in Supplementary Data 4 and 5. Run the code chunks in the markdown pane separately by clicking on the 'play' button in its upper right corner. Packages can be found and downloaded via the 'packages' tab in the lower right pane of RStudio.

```
```{r}
library(geomorph)
library(Morpho)
library(mclust)
library(cluster)
library(factoextra)
```
```

▲ **CRITICAL STEP** The code chunks given in the main text are for a main data analysis based on the geomorph package[45,46]. If not noted otherwise, data containing figures depicted in this study were generated based on this code (see Supplementary Data 4 for the corresponding R markdown file). When semilandmarks were present, sliding was achieved by minimizing bending energy. An alternative code, which utilizes native functions of the Morpho package[47,48] for GPA and multivariate statistical analysis, is available as another ready-to-use R markdown file in Supplementary Data 5.

11 Set your working directory by pasting its path between the quotation marks in the setwd() function.

```
```{r}
setwd("D:\my\directory\morphometrics_folder")
```
```

▲ **CRITICAL STEP** Note that the example directory path we give below is for Microsoft Windows users. For MacOS users, folder names should be separated using slashes (e.g., /Users/my/directory/morphometrics_folder).

12 Read the formatted landmark file 'lands.txt' using the readland.tps() function from geomorph and attribute it to a new object called 'lands'. Enter the path to your landmark file in between the quotation marks in the following code chunk.

```r
lands <- readland.tps(
"D:\my\directory\morphometrics_folder\lands.txt",
specID = "ID")
```

▲ **CRITICAL STEP** Note that the example directory path we give below is for PC users. For MacOS users, folder names should be separated using slashes (e.g., /Users/my/directory/morphometrics_folder/lands.txt).

? TROUBLESHOOTING

13 Define the number of semilandmarks and their respective curves by using the define.sliders() function from geomorph and attribute each of the returned matrices to a new object called 'curveslide_partX', with 'X' being the number of the curve. After defining semilandmarks and curves, combine all of your curve objects into one matrix using the rbind() function. Assign the combined matrix to a single object called 'curveslide'.

```r
curveslide_part1 <- define.sliders(c(5, 6, 7, 8), nsliders = 2)
curveslide_part2 <- define.sliders(c(11, 12, 13), nsliders = 1)
curveslide <- rbind(curveslide_part1, curveslide_part2)
curveslide
```

▲ **CRITICAL STEP** The define.sliders() function needs two inputs: (i) a vector containing a sequence of numbers that correspond to all (semi-)landmarks in the exact order in which they are placed along the curve and (ii) the number of sliding landmarks on that curve. Our proposed landmark configuration for *Pristionchus* nematodes contains two curves. The first curve runs between two fixed landmarks (landmarks 5 and 8), and two semilandmarks are allowed to slide along this curve (semilandmarks 6 and 7); the second curve runs between landmarks 11 and 13, with landmark 12 sliding between them (Fig. 1c, d). To define the sliding landmarks on a curve, create a vector with the c() function and list the curve landmarks in their correct sequence as an input (e.g., 5,6,7,8). Then, define the number of semilandmarks on the curve in the ‚nsliders = 2' argument. This way, the define.sliders() function treats the first and last values of the vector as fixed landmarks and all landmarks in between them as semilandmarks. Apply the same logic to all curves in your landmark configuration.

14 Extract specimen IDs from the landmark file using the dimnames() function and assign it to a new object called 'lands.names'. This object is a character that contains all IDs from all specimens in your dataset.

```r
lands.names <- dimnames(lands)[[3]]
```

15 Define the number of parts in which the specimen ID can be split and save it to an object called 'n.parts'.

```r
n.parts <- 4
```

▲ **CRITICAL STEP** In our case, the number of ID components is four (i.e., ID = 'RS5205-Eu-ind-01' can be split into 'RS5205', 'Eu', 'ind' and '01'). Adjust this number according to how you named your specimen.

16 Define which of the four parts of the ID specify your grouping factors of interest.

```r
SP.part <- 1
MF.part <- 2
```

▲ **CRITICAL STEP** In our example, these factors are 'strains of species (SP)' and 'mouth-form (MF)', which are specified in the first and second parts of the specimen ID, respectively. Save this information in respective objects. In our example, we call them 'SP.part' and 'MF.part'.

17 Extract the information on your grouping factors from the specimen ID and assign it to new objects.

```
```{r}
SP <- factor(matrix(unlist(strsplit(lands.names,"-")), length(lands.
names), n.parts,
byrow=T)[,SP.part])
MF <- factor(matrix(unlist(strsplit(lands.names,"-")), length(lands.
names), n.parts,
byrow=T)[,MF.part])
```
```

▲ **CRITICAL STEP** Here, we generate the two objects 'SP' and 'MF', which contain the information on species and mouth-form for each specimen, respectively. You may need to adjust the name of the object according to the grouping factor(s) of your own interest. Add the names of the objects you created in the previous step behind the comma in the square brackets at the end of both lines of code. This defines the part of the specimen ID that the factor() function has to use as an input. The strsplit() function splits the specimen IDs into parts, which are defined based on the hyphens in the ID name, and returns those parts as a list. The unlist() function will convert this list into a character (i.e., a simple vector containing all ID parts of all specimens). The matrix() function rearranges the output of the unlist() function into a matrix in which each column contains one part of the specimen ID and each row represents one specimen. The number of rows is extracted from the 'lands.names' object (see Step 14) and returned as an integer by the length() function. The number of columns is defined by the number of parts in the specimen ID ('n.parts'). The matrix will be converted into a simple factor that contains only the information on your grouping factor(s) of interest for each specimen, as well as all levels of these factors. This is achieved by applying the factor() function. All of the functions in this code chunk come with the base R package.

18 Define a color code designating the levels of your first grouping factor. Use, for example, 'darkgreen' for the eurystomatous and 'darkorange' for stenostomatous (Fig. 1; Supplementary Fig. 3).

```
```{r}
g_color <- rep("", length(lands.names))
g_color[which(MF=="Eu")] <- "darkgreen"
g_color[which(MF=="St")] <- "darkorange"
```
```

▲ **CRITICAL STEP** Adjust the color coding according to your preferences and the number of grouping factor levels. If necessary, replace our first grouping factor (i.e., 'MF') and its levels (i.e., the two morphs 'Eu' and 'St') with your grouping factor and its levels in the following code chunk. More lines of code to assign various colors can be added as per your needs, if you have more or less than two grouping factor levels (e.g., Supplementary Fig. 4). The rep() function from base R is a generic function that replicates the action of designating a specimen with a color for all individuals in the data set (in our example, based on mouth-form). The which() function defines the subsets of specimens based on the level of the grouping factor of interest (e.g., grouping factor = 'MF' and level = 'Eu').

19 After defining a color code for the first grouping factor, assign different symbols to designate a second grouping factor of interest. Here, we use symbols to differentiate between species.

```
```{r}
g_pch<-rep(0,length(lands.names))
g_pch[which(SP=="Pel")] <- 21 #this number gives a circle
g_pch[which(SP=="Pbu")] <- 22 #this number gives a square
g_pch[which(SP=="RS5205E")] <- 23 #this number gives a diamond
g_pch[which(SP=="RS5205S")] <- 24 #this number gives a triangle
```
```

▲ CRITICAL STEP If necessary, replace our second grouping factor (i.e., 'SP') and its levels (e.g., 'Pbu' or 'Pel' for *P. bucculentus* and *P. elegans*, respectively) with your grouping factor of interest and its corresponding levels in the following code chunk. More species can be added with the same line of code we give below. The plotting symbols (e.g., triangles or squares) can be adjusted according to your preferences, by exchanging the pch numbers.

Superimposition of landmark configurations ● Timing ~1 min

20 Perform the (semi-)landmark superimposition according to GPA. Use the landmark file 'lands' and, if present, the curveslide object as input data for the gpagen() function of geomorph[45,46]. Assign the GPA results to a new object called 'sup.geo' and visualize them with the R graphics function plot() (see Supplementary Figs. 3–5).

```
```{r}
sup.geo <- gpagen(lands, curves=curveslide, ProcD=F, Proj = T)
plot(sup.geo$coord[,1,], sup.geo$coord[,2,], Type="n",
xlim = c(-0.6,0.6), ylim = c(-0.6,0.6),
asp = 1, main = "geomorph / bending energy")
for (i in 1:dim(sup.geo$data)[1]) {
points(sup.geo$coord[,,i], col = g_color[i])
}
```
```

▲ CRITICAL STEP Semilandmarks can be slid along a curve by minimizing either Procrustes distance or bending energy. To use the Procrustes Distance approach, the argument 'ProcD=' has to be set to 'T' (or 'TRUE'), while setting it to 'F' (or 'FALSE') will make gpagen() use the bending energy approach. The last argument of the function 'Proj =' specifies whether the aligned Procrustes coordinates should be projected into a linear tangent space. Since this projection facilitates the downstream multivariate statistical analysis[53], set this argument to 'TRUE' (or 'T'). The arguments 'sup.geo$coord[,1,] ' and 'sup.geo$coord[,2,] ' represent the x- and y-coordinates of the superimposed landmarks of all specimens. Use 'xlim' and 'ylim' to set the limits of the x and y axes, respectively. Define the aspect ratio in the 'asp=1' argument and add a graph title with 'main' argument. The for() loop in the code chunk gives all of the superimposed landmarks a color based on the previously defined color code (see Step 18).

▲ CRITICAL STEP Before performing the final steps of shape analysis (Steps 21–28), it is recommended to test the replicability of shapes, especially if new landmark configurations are tested or if the protocol is adapted to microscopic animals other than nematodes. To test the replicability, a separate data set containing landmarks of technical replicates of your specimen has to be generated (for a detailed example, see 'Replicability and error' in 'Experimental design'). If you wish to check your replicability, perform Steps 1–20 on this separate replicability data set and run the code that is provided in Box 1 on the obtained shapes, to estimate their RPT or %ME. If the replicability of shapes is acceptable, one can obtain the actual landmark data set, perform Steps 1–20, and then proceed beyond this step to analyze shape differences (Steps 21–28).

? TROUBLESHOOTING

PCA ● Timing 1–15 min

21 Run a PCA on the Procrustes coordinates obtained from the GPA in Step 20. Use the prcomp() function that comes with the R stats package to perform the analysis and save the result to an object called 'pca.geo'. Visualize the results of the analysis by generating a ggplot2-based[71] PCA plot with the fviz_pca_ind() function of the the factoextra package[58] and create a bar plot depicting the amount of variation described by each PC (in percentages). Use convex hulls in the PCA plot to visualize the portions of the morphospace that are occupied by our groups of interest.

```
```{r}
#Principal component analysis (PCA) using a base R function
pca.geo <- prcomp(two.d.array(sup.geo$coord))
#ggplot2-based plot of the PCA results using a factoextra function
fviz_pca_ind(pca.geo,
```
```

```
geom.ind = "point", # show points without labels
col.ind = SP, #adjust to your group
palette = "npg", #color palette, adjust to preferences
addEllipses = T, #adds ellipses or convex hulls
ellipse.type = "convex", #convex hulls
mean.point = F, #don't show average point
legend.title = "SP" #adjust accordingly) + coord_fixed(ratio = 1)
#adjustable for variation-weighted plots
# Barplot of variation described (in %) by each principal component
barplot(rev(100*pca.geo$sdev^2/sum(pca.geo$sdev^2)),horiz=T,
yaxt="n",xlab="variation described, %",xlim=c(0,100))
```

▲ **CRITICAL STEP** The geomorph function two.d.array()[45,46] converts the 3D array of Procrustes landmark coordinates into a 2D matrix, which is a readable input class for the prcomp() function. Set the 'ellipse.type =' argument to 'convex' so that the groups will be enclosed by convex hulls. Alternatively, setting the argument to 'confidence' will add confidence ellipses for each group (default level = 0.95). If necessary, adjust grouping factors according to your input data. Keep the 'ratio=' argument in the coord_fixed() function set to 1 to obtain PCA plots with equal unit length along the x and y axes. You can adjust the ratio according to the relative difference in variation that is explained by each PC, to obtain variation-weighted PCA plots (see Supplementary Fig. 12). The amount of variation described by each PC is calculated by dividing the squared standard deviations of each PC ('pca.geo$sdev^2') by the sum of squared standard deviations of all PCs ('sum(pca.geo $sdev^2)') and multiplying it by 100. The base R function rev() produces a reversed sequence of this argument, ordering the PCs from highest 'variance explained' to lowest.

? TROUBLESHOOTING

22 Before proceeding with the analysis, check whether there are potential outliers in the data set by using the plotOutliers() function[45,46] from the geomorph package and plot deformation grids for potential outliers using geomorph's plotRefToTarget() function[45,46] to visually check for errors (e.g., incorrect landmark sequence).

```
```{r}
plotOutliers(sup.geo$coords) #generate an outlier plot
#if a potential outlier is found, compare it to the mean shape of the data
set
shape1 <- sup.geo$consensus #mean shape
shape2 <- sup.geo$coords[,,52]
#adjust the value so that it fits the number that corresponds to the
outlier in "lands.txt"
#generate deformation grid of mean shape and potential outlier to check
for errors
plotRefToTarget(shape1, shape2, method = "TPS", mag = 2)
```
```

▲ **CRITICAL STEP** This function plots the Procrustes distance of each specimen to the mean shape. Specimens that exceed the upper quantile of distances are marked in red as potential outliers. The 'method=' argument of the function is set to 'TPS'. The 'mag=' argument allows the user to define the degree of magnification that is desired when visualizing the shape differences. Here, we use a twofold magnification ('mag=2'). Note that plotOutliers() identifies only 'potential' outliers. These are not necessarily 'wrong'; nor do they need to be removed in every case[46]. If, however, erroneous specimens can be found, one can remove them manually from the landmark file ('lands.txt') and repeat the analysis. We intentionally did not remove potential outliers in our example data sets, so that readers can identify such cases by using the example landmark files provided in the supplements (Supplementary Data 1–3). If one uses, for example, the file containing landmarks of different *Pristionchus* species (Supplementary Data 2), one can identify a single potential outlier in specimen 'Pbu-Eu-ind-21'. Checking the deformation grid of this individual and the mean shape indicates a potentially incorrect sequence of landmarks, as the grids folds over in the middle of the

plain. Such outliers might need to be removed from the data set, and the analysis might need to be repeated, if the potential outlier drastically affects PCA or downstream statistical analysis (see Step 26). Note that in our example, the significance of the respective result does not change regardless of whether these potential outliers are removed.

23 Identify important or meaningful PCs from the PCA results. Calculate the variation described by each PC and arrange the returned values from highest to lowest. Use the eigenvalues obtained during the PCA (i.e., the squared standard deviations of the PCs obtained with prcomp() in Step 21) and the number of individuals (in our example, $n = 80$) to determine whether a given PC is qualified to be considered meaningful with the getMeaningfulPCs() function of the Morpho package[47,48].

```r
#view how much variation is described by each PC (ordered from PC1 to last PC)
100*pca.geo$sdev^2/sum(pca.geo$sdev^2)
#get meaningful PCs
getMeaningfulPCs(pca.geo$sdev^2, n = 80)
#adjust sample size (n=) accordingly
```

▲ **CRITICAL STEP** This method calculates a threshold between a PC and its successor based on a log-likelihood ratio and the sample size[72]. Both the meaningful PCs and the threshold value are returned as outputs of the function. In our comparison of different *Pristionchus* species (including 80 animals), a PC is considered meaningful if its value is >1.37 times the value of the subsequent PC. This criterion applied only to the first three PCs (Fig. 4b; compare to Figs. 3b and 2d).

24 Estimate extreme shapes for PC1 and PC2 and visualize underlying shape differences along these principal components by drawing superimposed wireframe and lollipop plots (Figs. 2e,f, 3c,d and 4c,d).

```r
#estimate extreme shapes along PC1 and PC2
meanshape <- as.vector(t(sup.geo$consensus))
min_PC1 <- t(matrix(meanshape+(min(pca.geo$x[,1])*pca.geo$rotation[,1]),2,18))
max_PC1 <- t(matrix(meanshape+(max(pca.geo$x[,1])*pca.geo$rotation[,1]),2,18))
min_PC2 <- t(matrix(meanshape+(min(pca.geo$x[,2])*pca.geo$rotation[,2]),2,18))
max_PC2 <- t(matrix(meanshape+(max(pca.geo$x[,2])*pca.geo$rotation[,2]),2,18))
#generate wireframe plot of superimposed PC1 extreme shapes
plotRefToTarget(max_PC1, min_PC1, method = "point", mag = 1)
jointline <- c(1:2, NA, 3:4, NA, 5:10, NA, 11:13, NA, 15:16, NA, 17:18)
lines(min_PC1[jointline,], lty = 1, col = "darkorange", lwd = 0.75)
lines(max_PC1[jointline,], lty = 1, col = "darkgreen", lwd = 0.75)
#generate wireframe plot of superimposed PC2 extreme shapes
plotRefToTarget(max_PC2, min_PC2, method = "point", mag = 1)
jointline <- c(1:2, NA, 3:4, NA, 5:10, NA, 11:13, NA, 15:16, NA, 17:18)
lines(min_PC2[jointline,], lty = 1, col = "darkorange", lwd = 0.75)
lines(max_PC2[jointline,], lty = 1, col = "darkgreen", lwd = 0.75)
#generate lollipop plots of superimposed extreme shapes for PC1 and PC2
plotRefToTarget(max_PC1, min_PC1, method = "vector", mag = 1)
plotRefToTarget(max_PC2, min_PC2, method = "vector", mag = 1)
```

▲ **CRITICAL STEP** Color coding of wireframe lines can be changed according to personal preferences. Depending on your landmark configuration, the wireframe line connections may need

to be redefined by changing the vectors c(), which give the input to the 'jointline' objects. Here, 'NA' indicates stops between lines, while numbers separated by colons (e.g., '5:10') defines lines connecting the corresponding landmarks 5–10. If necessary, change the line definitions so that the wireframes represent the structures in your landmark configurations appropriately. Here, the 'jointline' vectors define proper wireframes for the stomata of *Pristionchus* nematodes (Fig. 1).

? TROUBLESHOOTING

25 Use the fviz_contrib() function to estimate the contribution (percentage) of every landmark on PC1/PC2. Extract and investigate the landmarks that contribute more to these PCs than expected.

```r
#estimate contribution (%) of landmarks on PC1
PC1_contrib <- fviz_contrib(pca.geo, choice = "var", axes = 1)
LM_contrib_PC1 <- data.frame(LM = rep(NA, nrow(PC1_contrib$data)/2),
PC1_contrib = rep(NA, nrow(PC1_contrib$data)/2))
for (i in 1:(nrow(PC1_contrib$data)/2)) {
LM_contrib_PC1$PC1_contrib[i]    <-    sum(PC1_contrib$data[(i*2-1):
(i*2),]$contrib)
LM_contrib_PC1$LM[i] <- as.character(i)
}
#estimate contribution (%) of landmarks on PC2
PC2_contrib <- fviz_contrib(pca.geo, choice = "var", axes = 2)
LM_contrib_PC2 <- data.frame(LM = rep(NA, nrow(PC2_contrib$data)/2),
PC2_contrib = rep(NA, nrow(PC2_contrib$data)/2))
for (i in 1:(nrow(PC2_contrib$data)/2)) {
LM_contrib_PC2$PC2_contrib[i]    <-    sum(PC2_contrib$data[(i*2-1):
(i*2),]$contrib)
LM_contrib_PC2$LM[i] <- as.character(i)
}
#examine landmarks with major contribution (%) to PC1 and PC2
important_LMs_PC1 <- LM_contrib_PC1[LM_contrib_PC1$PC1_contrib > 100/
(nrow(PC1_contrib$data)/2),]
important_LMs_PC1[order(decreasing = T, important_LMs_PC1$PC1_con-
trib),]
important_LMs_PC2 <- LM_contrib_PC2[LM_contrib_PC2$PC2_contrib > 100/
(nrow(PC2_contrib$data)/2),]
important_LMs_PC2[order(decreasing = T, important_LMs_PC2$PC2_con-
trib),]
```

▲ **CRITICAL STEP** To estimate which landmarks contribute more than expected, a threshold is set as 100 divided by the number of landmarks. The value obtained is the contribution (percentage) the landmarks are expected to have if all of them contribute equally to the shape differences. Important landmarks are ordered decreasingly by the order() function of base R. These landmarks can subsequently be indicated in the lollipop plot (obtained in Step 24) to give an impression of which landmarks differ profoundly between groups in the context of superimposed extreme shapes.

Multivariate statistical testing using 'Procrustes ANOVA' ● Timing ~5 min

26 Perform a permutational MANOVA[54–56] (or 'Procrustes ANOVA' in geometric morphometrics[46,56]) and—if required—a pairwise permutational MANOVA (including FDR correction[73] of *P* values) to assess statistical hypotheses describing patterns of variation based on the Euclidean distances obtained from GPA in Step 20. Define a formula for the linear model in the procD.lm() function according to your needs.

```r
#prepare a geomorph data frame for your statistical analysis
```

```
#adjust grouping factor according to interest
gdf <- geomorph.data.frame(shape = sup.geo$coords, species = SP,
mouth_form = MF)
#Procrustes Anova to test for shape differences between groups
#adjust grouping factor according to interest
a <- procD.lm(shape ~ species, data = gdf, iter = 100000, seed = 12345,
RRPP = F, effect.type = "F")
summary(a)
#pairwise Procrustes Anova
pw <- pairwise(a, groups = SP) #adjust grouping factor according to
interest
pw_summary <- summary(pw)
pw_summary
pw_p.values <- pw_summary$pairwise.tables$P
#transform the pairwise table into a data frame
#make sure to adjust row/column names and numbers accordingly
results_pw <- as.data.frame(matrix(pw_p.values,
dimnames = list(c("Pbu", "Pel", "Ppa_Eu", "Ppa_St"),
c("Pbu", "Pel", "Ppa_Eu", "Ppa_St")),
ncol = 4, nrow = 4))
#FDR-correction of p-values obtained from pairwise Procrustes Anova
adj.pw <- p.adjust(pw_summary$pairwise.tables$P, method = "fdr")
#transform the pairwise table into a data frame
#make sure to adjust row/column names and numbers accordingly
results_FDR.adj.pw <- as.data.frame(matrix(adj.pw,
dimnames = list(c("Pbu", "Pel", "Ppa_Eu", "Ppa_St"),
c("Pbu (FDR-adj.)", "Pel (FDR-adj.)", "Ppa_Eu (FDR-adj.)",
"Ppa_St (FDR-adj.)")),
ncol = 4, nrow = 4))
#view results of pairwise Procrustes Anova and FDR-corrected pairwise
Procrustes Anova
results_pw
results_FDR.adj.pw
```

▲ **CRITICAL STEP** To facilitate statistical analysis of the shape data, the array containing the Procrustes coordinates ('sup.geo$coords') has to be converted into a list that can be used as a data frame. This is achieved by the geomorph.data.frame() function of geomorph[45,46], if the array and the grouping factors are provided as arguments (e.g., 'shape = sup.geo$coords' or 'species = SP'). Save the results to a new object called 'gdf'. This object serves as the input for geomorph's procD.lm () function[45,46], which performs the Procrustes ANOVA[46]. A formula for the linear model has to be provided as an argument (e.g., 'shape ~ species', which asks whether shapes differ between species), together with the geomorph.data.frame (i.e., 'data=gdf') and the number of permutations for significance testing (here: 100,000). The significance can be tested either by randomizing raw values (residuals of mean, 'RRPP=F') or by randomizing null model residuals ('RRPP=T')[74,75]. In our examples, we used raw value randomization. The results of the Procrustes ANOVA can be viewed by using the generic summary() function of base R. Subsequently, distributions of pairwise statistics for the linear model fit obtained from procD.lm() can be generated with the pairwise() function of the RRPP package[74,75], after defining the grouping factor for the statistical test as an argument (e.g., 'groups = SP'). The results of the pairwise analysis are saved to a new object, and the *P* values are subsequently converted into and displayed as a data frame. Finally, the *P* values obtained by pairwise comparisons are corrected for false discoveries[73] using the p.adjust() function that is part of the R stats package. The obtained FDR-corrected *P* values (or 'q-values') are also formatted into a simple data frame. Results can be shown inside the corresponding PCA plot of the analysis (Figs. 2c, 3a and 4a). Make sure to adjust the names and the number of all rows and columns in the data frames according to your input data set.

? TROUBLESHOOTING

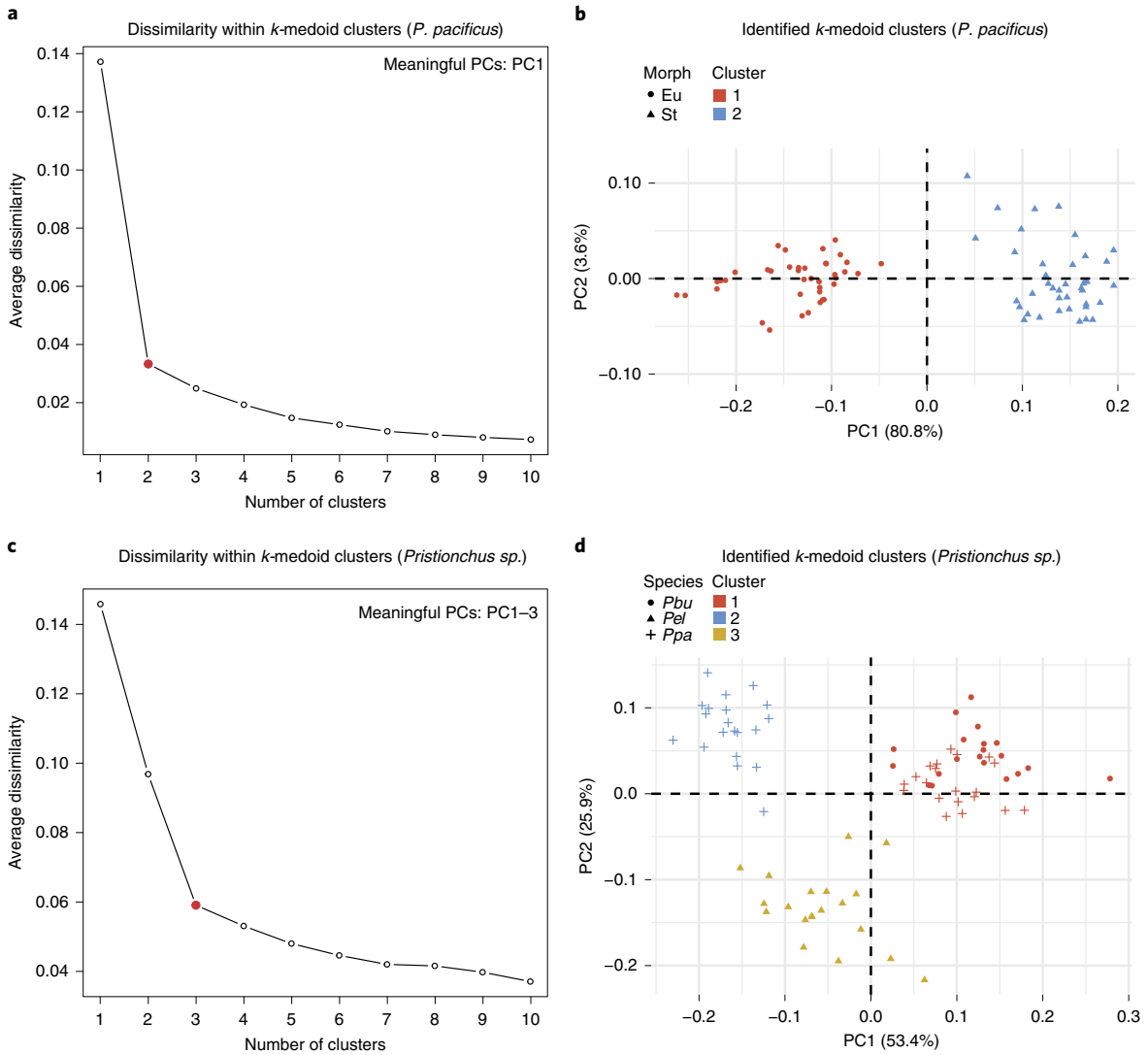


Fig. 5 | *k*-medoid clustering performed on the two *Pristionchus* datasets. a, Saturation plot showing the change of average dissimilarity within each cluster in relation to the number of clusters (*k*) for the *P. pacificus* dataset (including both morphs). The red dot indicates the best number of medoids for *k*-medoid clustering. **b**, Identified *k*-medoid clusters for the *P. pacificus* polyphenism dataset. Eu, eurystomatous morph; St, stenostomatous morph. **c**, Saturation plot showing the change of average dissimilarity within each cluster in relation to the number of clusters (*k*) for the *Pristionchus* dataset (including different species). The red dot indicates the best number of medoids for *k*-medoid clustering. **d**, Identified *k*-medoid clusters for the *Pristionchus* polyphenism and species dataset. *Pbu*, *P. bucculentus*; *Pel*, *P. elegans*; *Ppa*, *P. pacificus*.

Group identification by clustering

27 Perform group identification by clustering. For partitional clustering around *k*-medoids, follow option A. For model-based clustering, follow option B.

(A) **Partitional clustering around *k*-medoids** ● Timing ~2 min

(i) Identify the best number of medoids for the input data set by performing partitioning (i.e., clustering) around the medoids with the pam() function of the cluster package[60]. This calculates the average dissimilarity within the obtained clusters, depending on a number of clusters between 1 and 10. Generate a plot depicting the change in average dissimilarity of obtained clusters in relation to the number of medoids (*k*) (Fig. 5a,c; for simplicity, clustering was performed only for *Pristionchus* data sets).

```
```{r}
saturation <- numeric()
for (i in 1:10){
```
```

```
saturation[i] <- mean(pam(pca.geo$x[,1:(ncol(pca.geo$x)-33)],
i) [[6]][,3])
#remove non-meaningful PCs by adjusting the value behind the "-"
}
plot(1:10, saturation, xlab="number of clusters", ylab="average
dissimilarity", Type="b")
```

▲ **CRITICAL STEP** You may want to remove all non-meaningful PCs from your input matrix by adjusting the value behind the '-'. For example, we obtained only three meaningful PCs (i.e., PC1–3) from our stomatal shape comparison across several *Pristionchus* species (see Step 23). Since we had 36 PCs in total, we removed the 33 PCs that are considered non-meaningful (i.e., 'pca.geo$x[,1:(ncol(pca.geo$x)-33)]'). This procedure is optional, and one can lose information by PC removal. If no PCs are to be removed, just delete the term '-XX' with XX being the number of PCs.

(ii) Choose the number of medoids/clusters based on the dissimilarity plot obtained previously and save this input to a new object called 'best.n.kmedoid.clusters'. The best number of clusters corresponds to the point in the 'saturation plot' where the average dissimilarity curve starts to level out (i.e., the point where partitioning of the data set into more clusters barely changes the average dissimilarity within the clusters).

Afterwards, perform *k*-medoid clustering with the pam() function[60] for the chosen number of medoids.

```{r}
best.n.kmedoid.clusters <- 3 #set number of clusters here
pca.kmedoid.clusters <- pam(pca.geo$x[,1:(ncol(pca.geo$x)-33)],
best.n.kmedoid.clusters)
#adjust number of PCs you want to remove (exclude at least 4)
```

▲ **CRITICAL STEP** Again, non-meaningful PCs can be removed from your input matrix by adjusting the value behind the '-'. However, if all PCs should be taken into account, delete the term '-XX' (with XX being the number of PCs).

(iii) Generate a new PCA plot in which all observations (i.e., specimens) are colored according to the newly identified clusters (Fig. 5b,d).

```{r}
plot(pca.geo$x[,1:2], main = "k-medoid clustering",
pch = g_pch, bg = (1:best.n.kmedoid.clusters)[pca.kmedoid.
clusters$clustering], asp=1)
```

(B) **Model-based clustering** ● **Timing ~ 2 min**

(i) Perform model-based clustering on the PCA data by using the Mclust() function that comes with the mclust package[61,62].

```{r}
pca.model.clusters <- Mclust(pca.geo$x[,1:(ncol(pca.geo$x) -33)])
#adjust number of non-meaningful PCs to remove
```

▲ **CRITICAL STEP** A high number of variables could over-parameterize multivariate normal models. Thus, we recommend removing all non-meaningful PCs from your input matrix by adjusting the value behind the '-', depending on the outcome of Step 23. Always exclude the worst four PCs at the minimum, because four degrees of freedom are lost during GPA.

(ii) Extract the best number of clusters (i.e., the optimal number of mixture components) estimated in the previous step.

```
```{r}
best.n.model.clusters <- pca.model.clusters$parameters$variance$G
```
```

(iii) Finally, plot the PCA with observations (i.e., specimens) colored according to newly identified clusters (see Supplementary Fig. 8 for clustering results based on the *Pristionchus* example data sets and two different rules for utilization of PCs).

```
```{r}
plot(pca.geo$x[,1:2], main="model-based clustering", asp = 1,
pch = g_pch, bg = (1:best.n.model.clusters)[pca.model.clusters
$classification])
```
```

Visualize allometry ● Timing ~5 min

28 Visualize group-specific allometric trends in your data set. Prepare a list called 'allometry' that contains the original shape variables, your grouping factors and the centroid sizes of all specimens before GPA alignment with the help of the geomorph.data.frame() function. Use the procD.lm() function to fit a linear model for the correlation of shape with size (i.e., allometry) as well as other covariates (e.g., species) and assess the statistical significance with a distance-based PERMANOVA design including a residual randomization permutation procedure with ≥10,000 iterations. Generate a model-based allometry plot with geomorph's plotAllometry() function.

```
```{r}
#prepare geomorph dataframe for permutation test
#adjust according to your grouping factor and grouping factor levels
allometry <- geomorph.data.frame(shape = sup.geo$coords, species = SP,
mouth_form = MF, size = sup.geo$Csize)
#permutation test based on linear model
allometry_model <- procD.lm(data=allometry, shape ~ size * SP,
iter=100000) #adjust the linear model according to your grouping
factors and research question
summary(allometry_model)
#plot allometric trends (use "PredLine" or "RegScore")
#adjust symbols (pch=) and color (col=) according to preferences
plotAllometry(allometry_model, size = allometry$size, logsz = T,
method = "PredLine", pch = g_pch, col = g_color)
```
```

▲ **CRITICAL STEP** Please note that, in the most recent geomorph release (v. 3.2.1), the plotAllometry() function seems to be unable to visualize allometric trends of multiple groups in the same data sets as separate lines in one plot. Copy the overhauled source code of the function that is currently provided by the package authors (see 'Software for image acquisition and data analysis' in 'Materials') and run it once in the R console to overwrite the initial version before running the code chunk above. We recommend using 100,000 rather than 10,000 iterations in the permutation procedure. The 'size=' argument in plotAllometry() defines the input that contains the centroid sizes of all individuals (before GPA alignment), while setting the 'logsz=' argument to 'T' (or 'TRUE') will result in plotting the logarithm of the centroid size on the x axis. The 'method=' argument specifies the plotting procedure, which has to be set to 'PredLine' to plot PC1 on the predicted values (obtained from the linear regression of shape versus size) on the y axis. This argument can be changed to 'RegScore' if one wishes to use a different model-based visualization of allometry.

? TROUBLESHOOTING

Troubleshooting

For troubleshooting guidance, see Table 1.

Table 1 | Troubleshooting table

Step	Problem	Possible reason	Solution
4	Nematodes are still moving on the object slide	The NaN_3 concentration in agar pads is too low	1 Simply wait another 5–10 min 2 Kill nematodes by placing the object slide on a heat block for a maximum of 3–4 s at ~70 °C (note that the integrity of inner organs might be affected by this procedure)
12	The readlands.tps() function gives an error message and reports that curve points are detected but that numbers vary among specimens	The number of landmarks defined by 'LM=' is not equal for all specimens in the data set, and/or coordinates contain commas	Ensure that you have the same number of landmarks (in our example, 18) for all specimens in the landmark file and replace all commas in coordinate values with periods
20	The gpagen() function gives an error message and reports that 'x' must be an array of at least two dimensions	The number of semilandmarks or curves is too small	Include at least two semilandmarks on a single curve or have at least two curves with one semilandmark each (more curves and/or semilandmarks are always possible)
21	Semi-transparent filling of convex hulls is not included in the (exported) PCA plot generated by fviz_PCA_ind()	The image was exported in file formats like .png or .jpg	Export all plots as PDF and use a vector graphics editor like Adobe Illustrator for postediting
24	Parts of the wireframe or lollipop plots are outside the plot	The graphical window is not set appropriately	Change the sequence of the 'max_PC1/2' and 'min_PC1/2' arguments in the plotReftoTarget() functions that return the wireframe and lollipop plots
26	Pairwise comparisons yield results that clearly do not fit the observations in PCA	The data frame that displays the results of the pairwise comparison is set up incorrectly	Open the object 'pw_p.values' generated in the step that includes the pairwise() function and adopt the row and column names in their correct sequence into the functions that set up the data frames that contain *P* values from pairwise comparisons (normal and FDR adjusted)
28	Some members of a group are located far away from other members of the same group along the log(size) axis (see an example in Supplementary Fig. 9b)	Image stacks have been saved with an inconsistent scaling factor (e.g., some in micrometers versus some in inches)	If you save all image stacks with identical x-/y-dimensions, you can manually calculate a re-scaling factor based on these dimensions and the scale unit. Multiply affected centroid sizes (see object 'sup.geo$Czise') with that factor and create a new input dataframe that contains the corrected values
	Group-specific allometries are not separated as individual trend lines in the allometry plot	'bug' in the plotAllometry() function of the most recent geomorph release (v. 3.2.1)	Overwrite the inbuilt plotAllometry() function with the updated version (see 'Software for image acquisition and data analysis')
	The procD.lm() function gives an error message and reports 'in fits$full[[k]]: subscript out of bounds'	The procD.lm() function of the most recent geomorph release (v. 3.2.1) potentially assumes a fully crossed factorial design	Write out the model formula using only '+' between factors and their interactions (instead of using '*'); try to identify factor interactions that equate to zero and delete them from the formula

Timing

Most of the time that goes into data analysis depends on the amount of troubleshooting and the number of adjustments that are needed to fit the analyses to your grouping factors of interest. The timeframe for replicability assessments (Box 1; or see the end of the R code in Supplementary Data 4 and 5) depends on the design of the study. If one follows the approach used herein (20 nematode specimens in three replicates on different days), then the according landmark files are obtained over a course of 5 d. Their analysis takes ~20–30 min per annotator.

Steps 1–5, acquisition of raw image stacks for 80 animals: ~4 d (for nematodes)

Steps 6–8, obtaining landmark data for 80 animals: ~6 h

Steps 9–19, setting-up the data in an R markdown file: 1–10 min

Step 20, generalized Procrustes analysis (i.e., superimposition): ~1 min

Steps 21–25, PCA: 1–15 min

Step 26, multivariate statistics (Procrustes ANOVA): ~5 min

Step 27A, *k*-medoid clustering: ~2 min
Step 27B, model-based clustering: ~2 min
Step 28, visualization of allometric trends: ~5 min
Box 1, Replicability assessment: 5 d

Anticipated results

The combination of geometric morphometric analysis and clustering methods allows quantification and statistical testing of morphological differences between groups of interest as well as an estimation of the number of separate groups that are present in the morphospace. In an appropriate context, these differences can be investigated as morphological shape changes due to, for example, divergent evolution (phylogeny), embryonic development (ontogeny), environmental inputs (ecomorphology) or mutagenesis (functional genetics). Therefore, the expected results from the application of the presented protocol will depend on the biological system under investigation, as well as on the context of the study.

Researchers who acquire landmark data (Steps 1–8) using one of the landmark configurations we proposed above (Figs. 1 and 2a,b) will be able to robustly quantify shape differences between adults of many species or strains within the nematode genera *Pristionchus* or *Caenorhabditis* in any of the respective study contexts mentioned above. With minor adjustment and a suitable landmark configuration, this protocol should allow similar shape quantification in any microscopic animal or structure of choice. However, it might be necessary to practice the correct placement of landmarks over a few days before obtaining the final landmark data set. Even when RPT and %ME are acceptable to obtain reproducible and significant results, it is clear that both estimates can be substantially affected by the level of experience of the annotator (Supplementary Figs. 6 and 7; compare annotators 1 and 2 with 3). Importantly, if the variation within technical replicates is higher than the variation across different individuals, it might be difficult to obtain meaningful results from geometric morphometric analyses. Besides practicing landmark placements, one could also label each specimen multiple times and perform GPA on the average landmark configurations of each individual, in order to reduce the measurement error in a data set[67].

In the following sections, we provide a more detailed description of exemplary results in nematode mouth structures. First, we describe stomatal shape differences that were quantified for three species of *Pristionchus* nematodes using GPA, PCA, multivariate statistics and *k*-medoid clustering. We then proceed to describe the patterns of stomatal shape variation we identified for two species of *Caenorhabditis* nematodes. These conclusions are based on the lollipop and wireframe plots generated for (estimated) extreme shapes along axes of major variation. In addition, we explore which landmarks in the *Caenorhabditis* data set have a major contribution to PC1 and PC2 by assessing variable loadings obtained from PCA. We close by shortly discussing stomatal shape differences between the two adult morphs of *P. pacificus* in relation to changes in size (i.e., allometry).

Quantification of shape differences in nematode stomata

As can be seen from our example data sets, our protocol can be used to quantify stomatal shape differences on the intraspecific level (Fig. 3) and the interspecific level (Figs. 2 and 4). The PCA analysis, together with the multivariate Procrustes ANOVA approach, allows statistical assessment of shape differences. Our analysis of different *Pristionchus* species, for example, reveals that the eurystomatous morph of *P. pacificus* overlaps with the monomorphic species *P. bucculentus* in the shape space (Fig. 4a). On the other hand, these two groups do not overlap either with the stenostomatous morph of *P. pacificus* or with *P. elegans*. This pattern of shape variation indicates that the shapes of *P. bucculentus* and eurystomatous *P. pacificus* are similar (if not identical), while stenostomatous *P. pacificus* and *P. elegans* occupy very different parts of the morphospace. This is particularly interesting in the context of canalization and morphological evolution, as *P. bucculentus* and *P. elegans* are described as monomorphic eurystomatous or monomorphic stenostomatous species, respectively[76,77]. The results of our exemplary analysis thus indicate that eurystomatous stomatal shapes can be similar across species, while stenostomatous morphs can indeed differ in terms of shape. This is also supported by the results of the Procrustes ANOVA on shape across morphs, as it estimates high effect sizes and low *P* values for all pairwise comparisons, except for the comparison of eurystomatous *P. pacificus* and *P. bucculentus* (Fig. 4a), which indicates that there is no significant shape difference between them. On top of that, *k*-medoid clustering provides supporting evidence for this assumption as well. We estimated the optimal number of clusters in this data set to be three

(Fig. 5c), and subsequent partitioning of data points around three medoids reveals a group structure in which all eurystomatous *P. pacificus* and *P. bucculentus* specimens form a common cluster (Fig. 5d). Furthermore, specimens of stenostomatous *P. pacificus* and those of *P. elegans* each form a distinct, non-overlapping cluster. Thus, the group structure revealed by *k*-medoid clustering (without an a priori hypothesis) perfectly complements the results obtained via multivariate Procrustes ANOVA performed on groups that were defined a priori (compare Fig. 4a with Fig. 5d).

Visualizing and interpreting the quantified stomatal shape differences

Morphological variation can be described by depicting wireframe and lollipop plots generated for superimposed extreme shapes along PC1 and PC2 (Fig. 3c,d; Fig. 4c,d; and Fig. 2e,f). For example, if one considers the wireframe plots obtained for the stomatal shape differences between *C. elegans* and *C. briggsae* (Fig. 2e), it appears that along PC1 the entire buccal cavity is wider (on average) in *C. elegans* than in *C. briggsae*, while the relative length of the cheilo- and gymno/promesostegostom remains similar across species. Another surprising observation is that along PC1 the metastegostomal flaps are (on average) larger in *C. elegans* than in *C. briggsae*. The corresponding lollipop plots (Fig. 2f) can then be used to describe the change in landmark position (relative to all other landmarks) along PC1/2. In these plots, landmarks 2, 3, 16 and 17 move interiorly in a nearly perpendicular manner to the anteroposterior body axis in *C. briggsae*. Therefore, while the overall length of the buccal cavity is maintained, its lumen is constricted in *C. briggsae* as compared to *C. elegans* (Fig. 2f). It is conceivable that these differences have functional consequences related to bacterial diet, thus demonstrating the power of GM and our simplified protocol to provide new testable hypotheses for morphological evolution of microscopic organisms.

When interpreting such data, one might be tempted to deduce where the shape changes originate by looking at individual landmarks. However, it is important to properly interpret these results as follows. If we look at the loading-based estimation of relative contribution, we can see that landmarks 4 and 15, which are the main predictors of stomatal width, have a major influence on PC1 (Fig. 2f). This correlates with the separation of the two species along PC1 (Fig. 2c). On the other hand, the shape of the telostegostom (landmarks 8–10) is very similar across species along both PC1 and PC2 (Fig. 2e). Thus, landmarks 8–10 do not contribute substantially to the first two components of shape variation (Fig. 2f). However, we note that (due to the nature of GPA) differences in shapes always result from shifts of a given landmark relative to all other landmarks in the configuration. Furthermore, shape differences do not originate from landmarks themselves, but within the tissues on which these landmarks are placed[50]. Thus, landmarks that contribute more to the distribution of specimens in the tangent space than expected exclusively point to (i) the portion of the overall shape that differs most prominently between groups and (ii) the tissue(s) that might cause these morphological differences. They do not indicate contributions independent of other landmarks. Here, it appears that essentially all tissues that secrete the stoma are involved in causing the shape change, as shape differences are obvious in all stomatal structures (i.e., cheilostom, gymnostom and stegostom) of *C. briggsae* and *C. elegans*.

Identification of size-related changes in stomatal shape

Lastly, we can visualize and describe group-specific allometries in our example data sets. By plotting predicted shapes against the logarithm of the centroid size, we can obtain stylized allometric trend lines for different groups. Allometry strongly influences shape differences in many morphometric data sets, and it often closely aligns with PC1 of the tangent space[70]. Thus, we might want to know whether the large shape differences between the two morphs in *P. pacificus* (Fig. 3a) are mere manifestations of allometry. In our case, allometry might be a good null hypothesis, especially since the respective data set is nearly one dimensional, with PC1 describing 80.8% variation (Fig. 3a,b). However, the allometric trend lines obtained for the two morphs indicate a contrasting situation. While there is slight allometry within each morph, they do not share a common allometric trajectory (Supplementary Fig. 9a). Simply put: it is impossible to obtain a eurystomatous mouth by just increasing the size of a stenostomatous one. Thus, the shape differences between them are not just allometric (Supplementary Fig. 9a). This observation has important implications regarding the genetic mechanisms regulating adult stomatal polyphenism. Assessing allometry might also be important when comparing wildtype and mutant phenotypes, because it could reveal whether mutant phenotypes represent different shapes or just mutation-induced allometric scaling effects.

Reporting Summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

All data generated and analyzed in this study are included in the paper or its Supplementary Data 1–3. Upon request, the raw data that were used to generate the example results are available from the corresponding authors.

Code availability

The entire code for an analysis based on geomorph can be copied from the different steps of the procedure. In addition, two ready-to-use R markdown files that contain the code for both geometric morphometrics packages (geomorph and Morpho) and a R routine that allows the generation of landmark files can be downloaded from the supplementary material of this paper (Supplementary Data 4–6).

References

1. Pigliucci, M. *Phenotypic Plasticity: Beyond Nature and Nurture* (Johns Hopkins University Press, 2001).
2. West-Eberhard, M. J. *Developmental Plasticity and Evolution* (Oxford University Press, 2003).
3. Shubin, N., Tabin, C. & Carroll, S. Deep homology and the origins of evolutionary novelty. *Nature* **457**, 818 (2009).
4. Wagner, G. P. *Homology, Genes, and Evolutionary Innovation* (Princeton University Press, 2014).
5. Moczek, A. P. et al. The significance and scope of evolutionary developmental biology: a vision for the 21st century. *Evol. Dev.* **17**, 198–219 (2015).
6. Minelli, A. Grand challenges in evolutionary developmental biology. *Front. Eco. Evol.* **2**, 1–11 (2015).
7. Mallarino, R. & Abzhanov, A. Paths less traveled: evo-devo approaches to investigating animal morphological evolution. *Annu. Rev. Cell Dev. Biol.* **28**, 743–763 (2012).
8. Klingenberg, C. P. Evolution and development of shape: integrating quantitative approaches. *Nat. Rev. Genet.* **11**, 623–635 (2010).
9. Parsons, K. J. & Albertson, R. C. Unifying and generalizing the two strands of evo-devo. *Trends Ecol. Evol.* **28**, 584–591 (2013).
10. Andersen, E. C. et al. A powerful new quantitative genetics platform, combining *Caenorhabditis elegans* high-throughput fitness assays with a large collection of recombinant strains. *G3 (Bethesda)* **5**, 911–920 (2015).
11. Seydoux, G. & Fire, A. Whole-mount in situ hybridization for the detection of RNA in *Caenorhabditis elegans* embryos. *Methods Cell Biol.* **48**, 323–337 (2015).
12. Pertea, M., Kim, D., Pertea, G. M., Leek, J. T. & Salzberg, S. L. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat. Protoc.* **11**, 1650–1667 (2016).
13. Friedland, A. E. et al. Heritable genome editing in *C. elegans* via a CRISPR-Cas9 system. *Nat. Methods* **10**, 741–743 (2013).
14. Witte, H. et al. Gene inactivation using the CRISPR/Cas9 system in the nematode *Pristionchus pacificus*. *Dev. Genes Evol.* **225**, 55–62 (2015).
15. Au, V. et al. CRISPR/Cas9 methodology for the generation of knockout deletions in *Caenorhabditis elegans*. *G3 (Bethesda)* **9**, 135–144 (2019).
16. Yoshida, K. et al. Two new species of *Pristionchus* (Nematoda: Diplogastridae) from Taiwan and the definition of the *pacificus* species-complex sensu stricto. *J. Nematol.* **50**, 355–368 (2018).
17. Susoy, V., Ragsdale, E. J., Kanzaki, N. & Sommer, R. J. Rapid diversification associated with a macro-evolutionary pulse of developmental plasticity. *Elife* **4**, e05463 (2015).
18. Susoy, V. et al. Large-scale diversification without genetic isolation in nematode symbionts of figs. *Sci. Adv.* **2**, e1501031 (2016).
19. Sieriebriennikov, B., Markov, G. V., Witte, H. & Sommer, R. J. The role of DAF-21/Hsp90 in mouth-form plasticity in *Pristionchus pacificus*. *Mol. Biol. Evol.* **34**, 1644–1653 (2017).
20. Sieriebriennikov, B. et al. Conserved nuclear hormone receptors controlling a novel trait target fast-evolving genes expressed in a single cell. *PLoS Genet* **16**, e1008687 (2020).
21. Hong, R. L. & Sommer, R. J. *Pristionchus pacificus*: a well-rounded nematode. *Bioessays* **28**, 651–659 (2006).
22. Sommer, R. J. & McGaughran, A. The nematode *Pristionchus pacificus* as a model system for integrative studies in evolutionary biology. *Mol. Ecol.* **22**, 2380–2393 (2013).
23. Sommer, R. J. *Pristionchus pacificus: A Nematode Model for Comparative and Evolutionary Biology* (Brill, 2015).
24. Sommer, R. J. et al. The genetics of phenotypic plasticity in nematode feeding structures. *Open Biol.* **7**, 160332 (2017).
25. De Ley, P., Van De Velde, M. C., Mounport, D., Baujard, P. & Coomans, A. Ultrastructure of the stoma in Cephalobidae, Panagrolaimidae and Rhabditidae, with a proposal for a revised stoma terminology in Rhabditida (Nematoda). *Nematologica* **41**, 153–182 (1995).

26. Von Lieven, A. F. & Sudhaus, W. Comparative and functional morphology of the buccal cavity of Diplogastrina (Nematoda) and a first outline of the phylogeny of this taxon. *J. Zool. Syst. Evol. Res.* **38**, 37–63 (2000).
27. Jay Burr, A. & Baldwin, J. G. The nematode stoma: homology of cell architecture with improved understanding by confocal microscopy of labeled cell boundaries. *J. Morphol.* **277**, 1168–1186 (2016).
28. Wilecki, M., Lightfoot, J. W., Susoy, V. & Sommer, R. J. Predatory feeding behaviour in *Pristionchus* nematodes is dependent on phenotypic plasticity and induced by serotonin. *J. Exp. Biol.* **218**, 1306–1313 (2015).
29. Ragsdale, E. J., Müller, M. R., Rödelsperger, C. & Sommer, R. J. A developmental switch coupled to the evolution of plasticity acts through a sulfatase. *Cell* **155**, 922–933 (2013).
30. Kieninger, M. R. et al. The nuclear hormone receptor NHR-40 acts downstream of the sulfatase EUD-1 as part of a developmental plasticity switch in *Pristionchus*. *Curr. Biol.* **26**, 2174–2179 (2016).
31. Namdeo, S. et al. Two independent sulfation processes regulate mouth-form plasticity in the nematode *Pristionchus pacificus*. *Development* **145**, dev166272 (2018).
32. Moreno, E., Lightfoot, J. W., Lenuzzi, M. & Sommer, R. J. Cilia drive developmental plasticity and are essential for efficient prey detection in predatory nematodes. *Proc. Biol. Sci.* **286**, 20191089 (2019).
33. Bardua, C., Wilkinson, M., Gower, D. J., Sherratt, E. & Goswami, A. Morphological evolution and modularity of the caecilian skull. *BMC Evol. Biol.* **19**, 30 (2019).
34. Tatsuta, H., Takahashi, K. H. & Sakamaki, Y. Geometric morphometrics in entomology: basics and applications. *Entomol. Sci.* **21**, 164–184 (2018).
35. Siriwut, W., Edgecombe, G. D., Sutcharit, C. & Panha, S. The centipede genus *Scolopendra* in mainland Southeast Asia: molecular phylogenetics, geometric morphometrics and external morphology as tools for species delimitation. *PLoS One* **10**, e0135355 (2015).
36. Zelditch, M. L., Swiderski, D. L. & Sheets, H. D. *Geometric Morphometrics for Biologists: A Primer* (Academic Press, 2004).
37. Adams, D. C., Rohlf, F. J. & Slice, D. E. Geometric morphometrics: ten years of progress following the 'revolution'. *Ital. J. Zool.* **71**, 5–16 (2004).
38. Webster, M. & Sheets, H. D. A practical introduction to landmark-based geometric morphometrics. *Paleontological Soc. Pap.* **16**, 163–188 (2010).
39. Adams, D. C., Rohlf, F. J. & Slice, D. E. A field comes of age: geometric morphometrics in the 21st century. *Hystrix It. J. Mamm.* **24**, 7–14 (2013).
40. Collyer, M. L. & Adams, D. C. Phenotypic trajectory analysis: comparison of shape change patterns in evolution and ecology. *Hystrix It. J. Mamm.* **24**, 75–83 (2013).
41. Feilich, K. L. & López-Fernández, H. When does form reflect function? Acknowledging and supporting ecomorphological assumptions. *Integr. Comp. Biol.* **59**, 358–370 (2019).
42. R Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2020).
43. Claude, J. *Morphometrics with R* (Springer, 2008).
44. Claude, J. Log-shape ratios, Procrustes superimposition, elliptic Fourier analysis: three worked examples in R. *Hystrix It. J. Mamm.* **24**, 94–102 (2013).
45. Adams, D. C. & Otárola-Castillo, E. geomorph: an R package for the collection and analysis of geometric morphometric shape data. *Methods Ecol. Evol.* **4**, 393–399 (2013).
46. Adams, D. C., Collyer, M. & Kaliontzopoulou, A. Geomorph: Software for geometric morphometric analyses. R package version 3.2.1. https://cran.r-project.org/package=geomorph (2020).
47. Schlager, S. Morpho and Rvcg–Shape Analysis in R: R-Packages for geometric morphometrics, shape analysis and surface manipulations. in *Statistical Shape and Deformation Analysis* (eds. Zheng, G., Li, S. & Szekely, G.) 217–256 (Academic Press, 2017).
48. Schlager, S. Morpho: calculations and visualisations related to geometric morphometrics. R package version 2.8. https://rdrr.io/cran/Morpho/ (2020).
49. Gunz, P. & Mitteroecker, P. Semilandmarks: a method for quantifying curves and surfaces. *Hystrix It. J. Mamm.* **24**, 103–109 (2013).
50. Klingenberg, C. P. Visualizations in geometric morphometrics: how to read and how to make graphs showing shape changes. *Hystrix It. J. Mamm.* **24**, 15–24 (2013).
51. Bookstein, F. L. Principal warps: thin-plate splines and the decomposition of deformations. *IEEE Trans. Pattern Anal. Mach. Intell.* **11**, 567–585 (1989).
52. Rohlf, F. J. Shape statistics: Procrustes superimpositions and tangent spaces. *J. Classif.* **16**, 197–223 (1999).
53. Dryden, I. L. & Mardia, K. V. Multivariate shape analysis. *Sankhya* **55**, 460–480 (1993).
54. Anderson, M. J. A new method for non-parametric multivariate analysis of variance. *Austral Ecol.* **26**, 32–46 (2001).
55. Anderson, M. J. Permutational multivariate analysis of variance (PERMANOVA). in *Wiley StatsRef: Statistics Reference Online* (eds. Balakrishnan, N. et al.) 1–15. https://onlinelibrary.wiley.com/doi/full/10.1002/9781118445112.stat07841 (2014).
56. Goodall, C. Procrustes methods in the statistical analysis of shape. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **53**, 285–321 (1991).
57. Abdi, H. & Williams, L. J. Principal component analysis. *Wiley Interdiscip. Rev. Comput. Stat.* **2**, 433–459 (2010).

58. Kassambara, A. & Mundt, F. Factoextra: extract and visualize the results of multivariate data analyses. R package version 1.0.7. https://cran.r-project.org/web/packages/factoextra/index.html (2020).
59. Jin, X. & Han, J. K-medoids clustering. in *Encyclopedia of Machine Learning and Data Mining* (eds. Sammut, C. & Webb, G. I.) 697–700 (Springer, 2017).
60. Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M. & Hornik, K. Cluster: cluster analysis basics and extensions. R package version 2.1.0. https://cran.r-project.org/web/packages/cluster/index.html (2019).
61. Scrucca, L., Fop, M., Murphy, T. B. & Raftery, A. E. mclust 5: clustering, classification and density estimation using Gaussian finite mixture models. *R. J.* **8**, 289–317 (2016).
62. Fraley, C., Raftery, A. & Scrucca, L. mclust: normal mixture modeling for model-based clustering, classification, and density estimation. R package version 5.4.6. https://cran.r-project.org/web/packages/mclust/index.html (2020).
63. Stiernagle, T. Maintenance of *C. elegans*. In WormBook (eds. The *C. elegans* research community, WormBook). http://www.wormbook.org (2006).
64. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676 (2012).
65. Stoffel, M. A., Nakagawa, S. & Schielzeth, H. rptR: repeatability estimation and variance decomposition by generalized linear mixed-effects models. *Methods Ecol. Evol.* **8**, 1639–1644 (2017).
66. Schielzeth, H. & Nakagawa, S. rptR: Repeatability for Gaussian and non-Gaussian data. R package version 0.9.22. https://cran.r-project.org/web/packages/rptR/index.html (2019).
67. Yezerinac, S. M., Lougheed, S. C. & Handford, P. Measurement error and morphometric studies: statistical power and observer experience. *Syst. Biol.* **41**, 471–482 (1992).
68. Adams, D. C. & Nistri, A. Ontogenetic convergence and evolution of foot morphology in European cave salamanders (Family: Plethodontidae). *BMC Evol. Biol.* **10**, 216 (2010).
69. Esquerré, D., Sherratt, E. & Keogh, J. S. Evolution of extreme ontogenetic allometric diversity and heterochrony in pythons, a clade of giant and dwarf snakes. *Evolution* **71**, 2829–2844 (2017).
70. Mitteroecker, P. et al. A brief review of shape, form, and allometry in geometric morphometrics, with applications to human facial morphology. *Hystrix It. J. Mamm.* **24**, 59–66 (2013).
71. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis* (Springer, 2016).
72. Bookstein, F. L. *Measuring and Reasoning: Numerical Inference in the Sciences* (Cambridge University Press, 2014).
73. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **57**, 289–300 (1995).
74. Collyer, M. L. & Adams, D. C. RRPP: An r package for fitting linear models to high-dimensional data using residual randomization. *Methods Ecol. Evol.* **9**, 1772–1779 (2018).
75. Collyer, M. & Adams, D. RRPP: linear model evaluation with randomized residuals in a permutation procedure. R package version 0.5.0. https://cran.r-project.org/web/packages/RRPP/index.html (2019).
76. Kanzaki, N. et al. *Pristionchus bucculentus n. sp.* (Rhabditida: Diplogastridae) isolated from a shining mushroom beetle (Coleoptera: Scaphidiidae) in Hokkaido, Japan. *J. Nematol.* **45**, 78–86 (2013).
77. Kanzaki, N. et al. Two new species of *Pristionchus* (Rhabditida: Diplogastridae): *P. fissidentatus n. sp.* from Nepal and La Réunion Island and *P. elegans n. sp.* from Japan. *J. Nematol.* **44**, 80–91 (2012).

Acknowledgements

J. Claude (Institut des Sciences de l'Évolution-Montpellier, Montpellier, France) is gratefully acknowledged for his advice on geometric morphometrics using geomorph and Morpho in RStudio during a workshop given in Calgary (attended by B.S. in 2017). Earlier drafts of this manuscript were improved by advice on many steps of data analysis by F. Chan (Friedrich Miescher Laboratory, Tübingen, Germany) and critical comments on the general methodology of geometric morphometrics by P. Arnold (University of Potsdam, Potsdam, Germany). Furthermore, we would like to thank A. Rohrlach (Max Planck Institute for the Science of Human History, Jena, Germany) for his comments on statistical procedures and multivariate hypothesis testing using R. N. Kanzaki (Forestry and Forest Products Research Institute, Tsukuba, Japan) provided expertise on nematode stomatal morphology in an early phase of this project. Further thanks go to M. Riebesell for providing the TEM image of the *P. pacificus* stoma (Supplementary Fig. 1a). D. Sharma and J. de la Cuesta are to be thanked for general discussion. This work was funded by the Max Planck Society. T.T. was supported by the IMPRS 'From Molecules to Organisms'.

Author contributions

Conceptualization: T.T., B.S. and M.S.W. Coding: B.S. and T.T. Data acquisition: T.T., S.S.W. and M.S.W. Data analysis: T.T. Writing the original draft: T.T. and S.S.W. Reviewing and editing the initial draft: B.S., M.S.W. and R.J.S. Revision of the original manuscript: T.T., B.S., S.S.W., M.S.W. and R.J.S. Supervision: M.S.W. and R.J.S.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41596-020-0347-z.

Correspondence and requests for materials should be addressed to M.S.W. or R.J.S.

Peer review information *Nature Protocols* thanks Ming Bai and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Received: 24 January 2020; Accepted: 27 April 2020;
Published online: 6 July 2020

Related links

Key references using earlier versions of this protocol

Sieriebriennikov, B. et al. *Mol. Biol. Evol.* **34**, 1644–1653 (2017): https://academic.oup.com/mbe/article/34/7/1644/3067497

Sieriebriennikov, B. et al. *PLoS Genetics.* **16**, e1008687 (2020): https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1008687

Early paper from our lab using a similar approach but a different protocol

Susoy, V. et al. *Elife* **4**, e05463 (2015): https://elifesciences.org/articles/05463

nature protocols

SUPPLEMENTARY INFORMATION

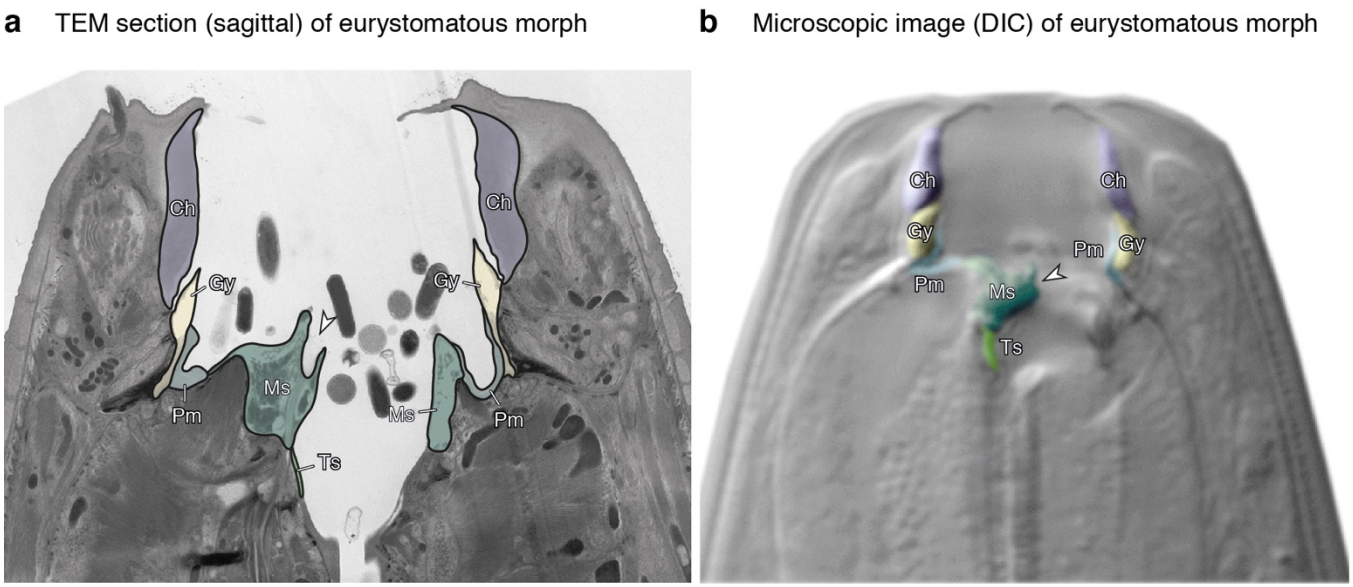
https://doi.org/10.1038/s41596-020-0347-z

In the format provided by the authors and unedited.

Geometric morphometrics of microscopic animals as exemplified by model nematodes

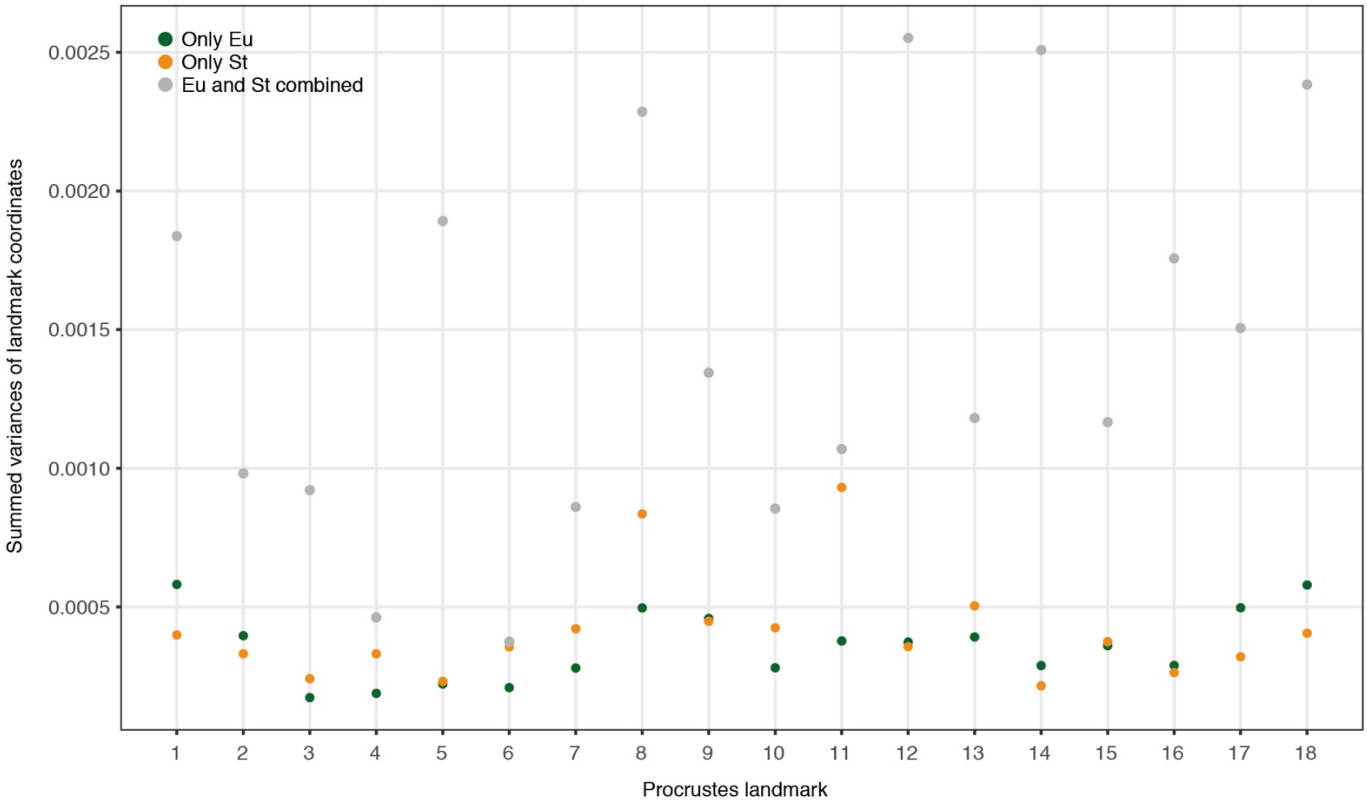
Tobias Theska[1], Bogdan Sieriebriennikov[1,2], Sara S. Wighard[1], Michael S. Werner[1]✉ and Ralf J. Sommer[1]✉

[1]Department for Integrative Evolutionary Biology, Max Planck Institute for Developmental Biology, Tübingen, Germany. [2]Present address: Department of Biology, New York University, New York, NY, USA. ✉e-mail: michael.werner@tuebingen.mpg.de; ralf.sommer@tuebingen.mpg.de

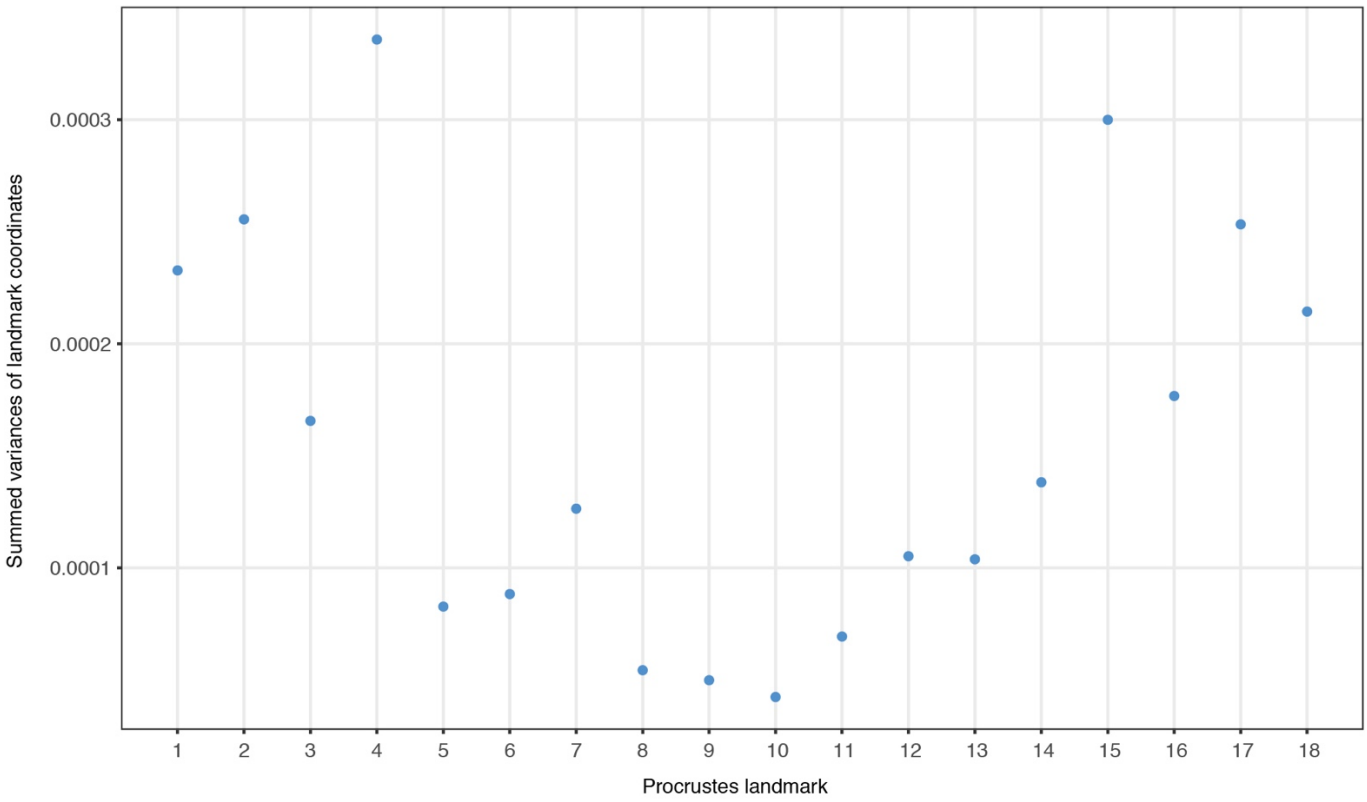


Supplementary figure 1. Comparison of the eurystomatous morph of *P. pacificus* under TEM and DIC microscope. **a,** TEM image of the stoma. **b,** Light microscope (DIC) image of the stoma (100x). Both images show the sagittal plane in right lateral view. Arrowheads indicate the position of the dorsal gland opening. ch, cheilostom; gy, gymnostom; ms, metastegostom; pm, promesostegostom; ts, telostegostom.

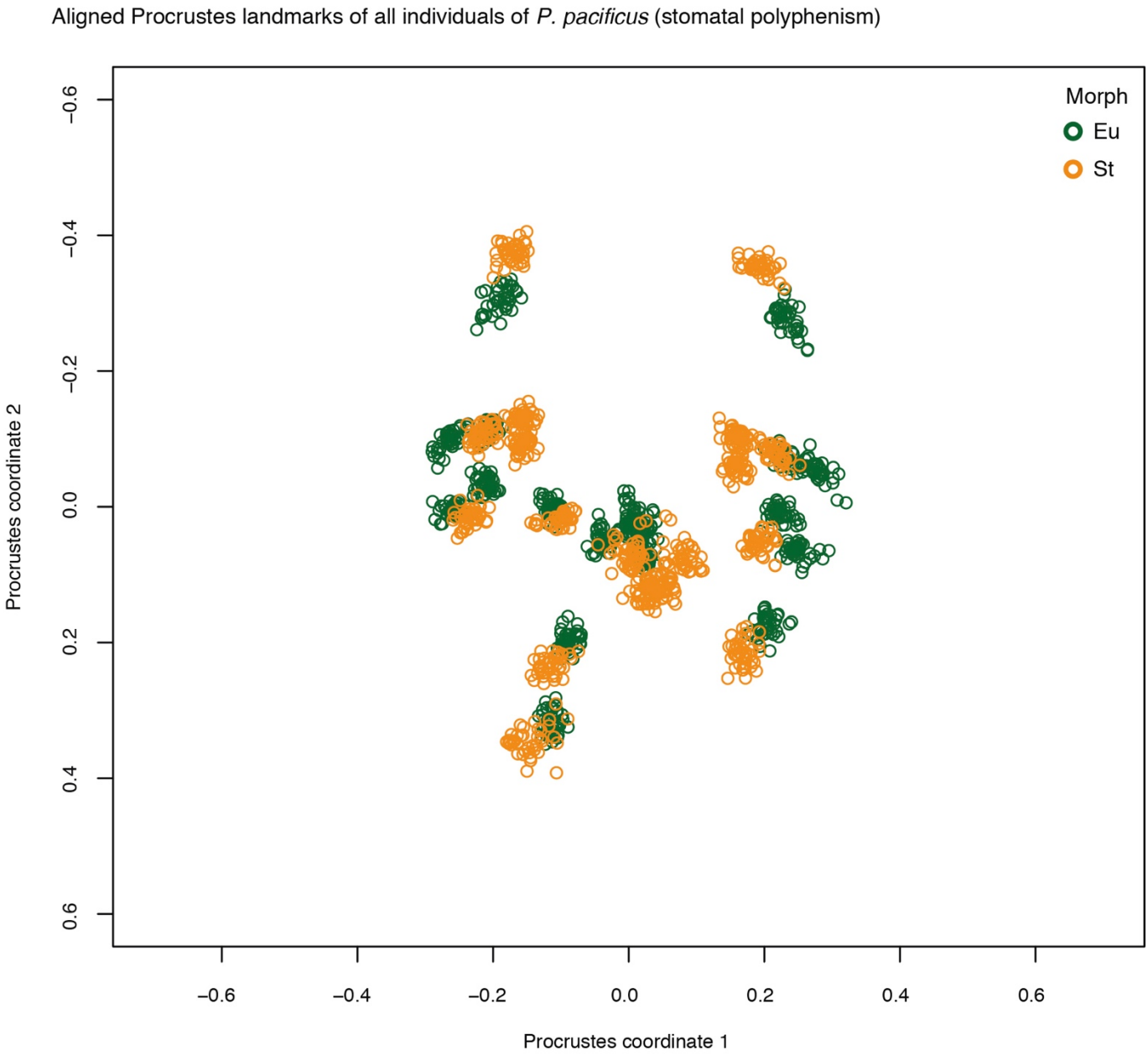
a Summed variances for each landmark in *P. pacificus*



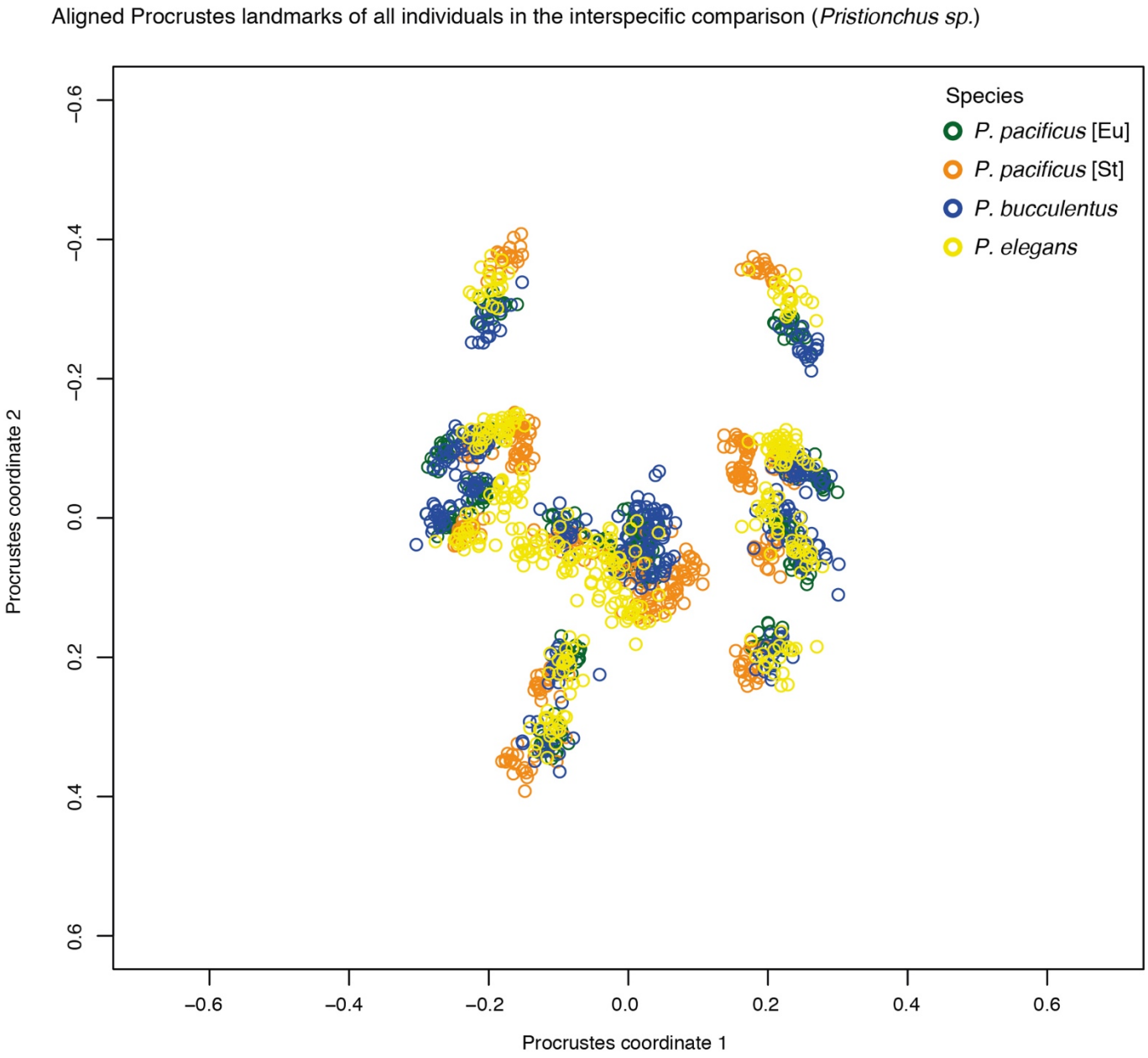
b Summed variances for each landmark in *C. elegans*



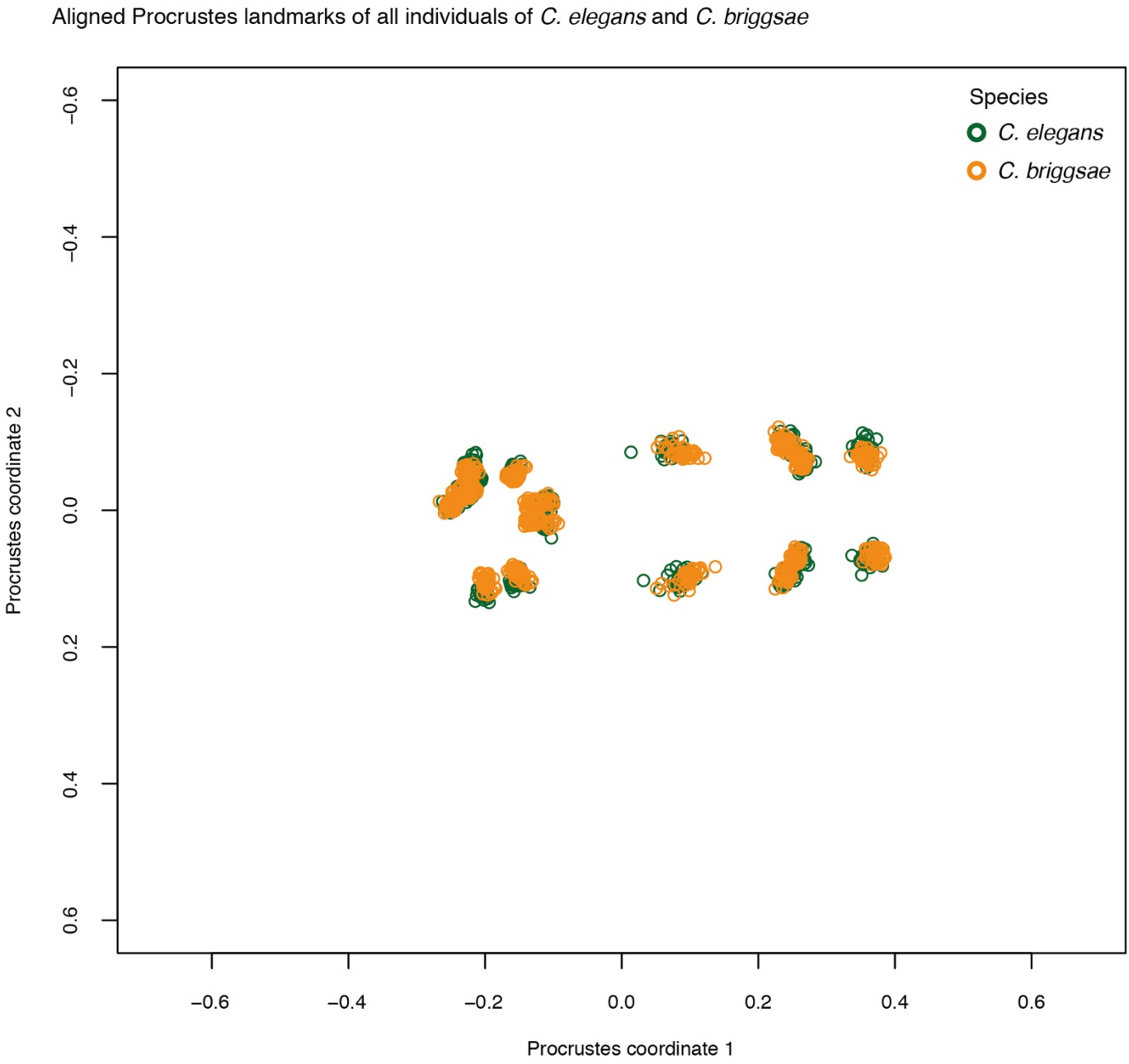
Supplementary figure 2. Variances of single landmarks for *P. pacificus* and *C. elegans*. Total variances for each landmark were calculated by summing up the variance in both coordinates of a respective landmark. Single coordinate variances were calculated using the **var()** function of the R statistics package. Note that the *Procrustes coordinates*, and thus also their variances, are dimensionless. **a,** Total variance of single landmarks for each morph (Eu and St) of *P. pacificus*, as well as total variance of single landmarks across all specimens regardless of the morph. **b,** Total variance of single landmarks for all *C. elegans* specimens. **a, b,** Smaller variances indicate a robust landmark positioning. Larger variances can indicate either inherently noisy landmarks (i.e. a high degree of naturally occurring variation in the structure), or biologically meaningful differences between groups. For example: Landmark "5" shows very small variances within each of the two morphs of *P. pacificus*, which indicates that it can be found reliably within each morph. However, the variance in landmark position is very high if all individuals (i.e. Eu and St combined) are used for estimation. This indicates that the position of landmark "5", relative to all other landmarks of the configuration, is very different between the morphs and thus potentially relevant in explaining biological shape differences. On the other hand, the high variance of landmark "11" across all *P. pacificus* specimens is due to the slightly increased variance of this landmark in stenostomatous animals (St) as compared to the low variance observed for eurystomatous animals (Eu). Thus, this landmark might be less reliable for the identification of shape differences between morphs. Note that landmarks which are jointly placed on movable elements (e.g. landmark 6, 7, 8, 9 &10 in *P. pacificus*) do not show systematically increased variances as opposed to landmarks which are placed on non-movable elements (e.g. 1-4 and 15-18 in *P. pacificus*). This indicates that landmarks on movable and non-movable elements can be treated equally for downstream analysis.



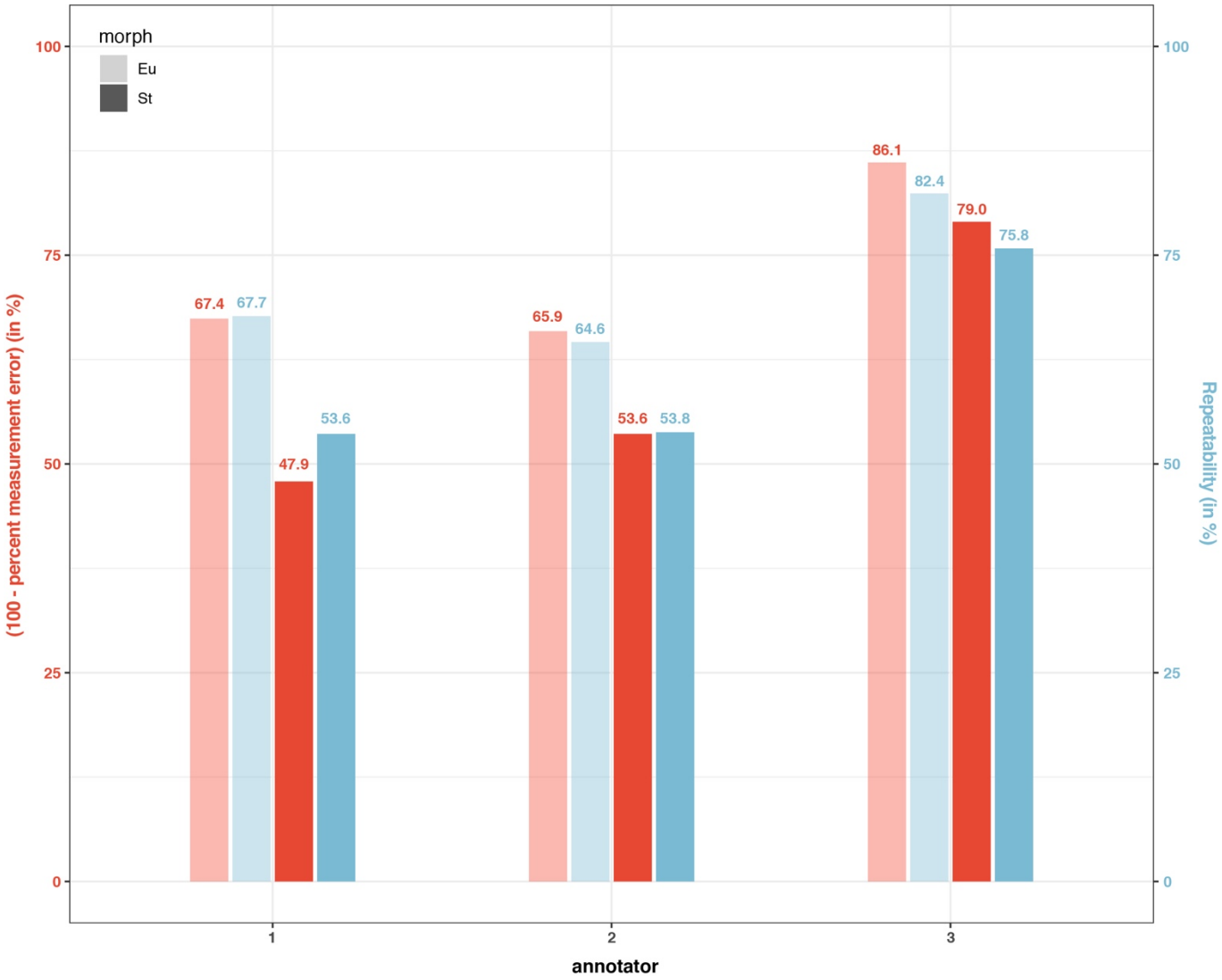
Supplementary figure 3. Scatterplot of Procrustes aligned coordinates obtained for the *P. pacificus* dataset (including both morphs). Generalized Procrustes analysis (GPA) was carried out using the **gpagen()** function of the *geomorph* package (see step 20 of the Procedure). *Procrustes coordinates* are projected into a linear (euclidean) tangent space. Semilandmarks were slid by minimizing bending energy. Output image was flipped horizontally, to depict aligned shapes in the original (right lateral) orientation.



Supplementary figure 4. Scatterplot of Procrustes aligned coordinates obtained for the *Pristionchus* dataset (including *P. pacificus, P. bucculentus* and *P. elegans*). Generalized Procrustes analysis (GPA) was carried out using the **gpagen()** function of the *geomorph* package (see step 20 of the Procedure). *Procrustes coordinates* are projected into a linear (euclidean) tangent space. Semilandmarks were slid by minimizing bending energy. Output image was flipped horizontally, to depict aligned shapes in the original (right lateral) orientation.

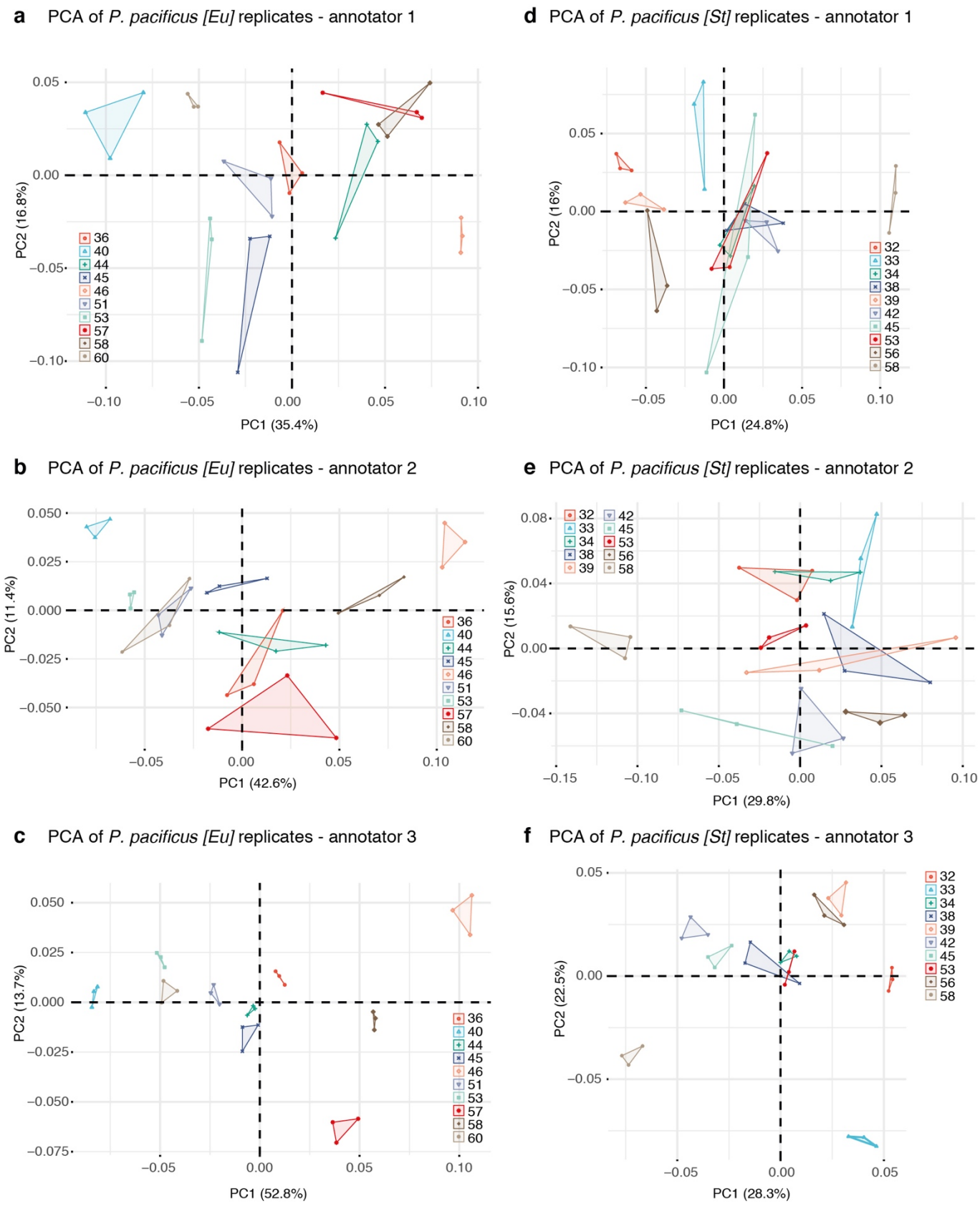


Supplementary figure 5. Scatterplot of Procrustes aligned coordinates obtained for the *Caenorhabditis* dataset (including *C. elegans* and *C. briggsae*). Generalized Procrustes analysis (GPA) was carried out using the **gpagen()** function of the *geomorph* package (see step 20 of the Procedure). *Procrustes coordinates* are projected into a linear (euclidean) tangent space. Output image was flipped horizontally, to depict aligned shapes in the original (right lateral) orientation.



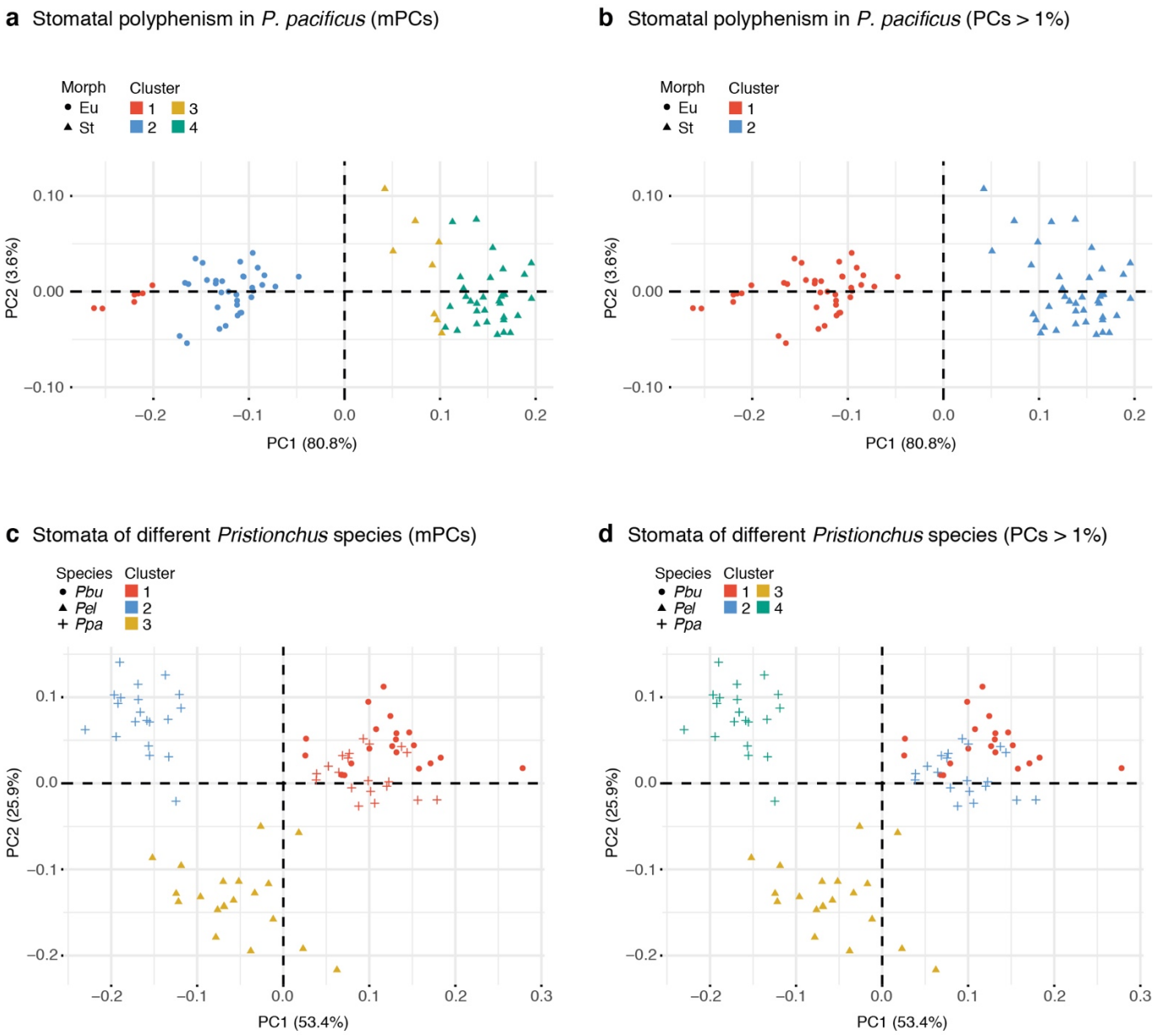
Supplementary figure 6. Replicability performance of annotators with different levels of training. Repeatability (i.e. the amount of variation between individuals) and percent measurement error (i.e. the amount of variation within individuals/among replicates) were calculated for shapes using the code in Box 2 of the main text. Annotators 1 and 2 are less trained in the procedure of landmark placement, while annotator 3 is familiar with the procedure. Color coding in y-axis corresponds to color coding of bars. Red, percent measurement error (%ME); blue, repeatability (RPT). Light bars indicate the values obtained for Eu, and dark bars indicate those obtained for St animals. Abbreviations: Eu, eurystomatous and St, stenostomatous morph of *P. pacificus*. Superimposition was performed using the **gpagen()** function of *geomorph* (see step 20 of the Procedure). Semilandmarks were slid by minimizing bending energy. For repeatability estimation, the number of parametric bootstraps and permutations was set to 1,000. In all cases (across morphs and annotator), when estimating the %ME of shapes, the specimen factor was significant while the replicate factor was not. This means that the variation between worms was always

stronger than the variation across technical replicates of the same individual. Statistics (PERMANOVA on shape) for annotator-1: ME[Eu] = 32.6%, ME[St] = 52.1%. Specimen factor [Eu], F = 7.2156, effect size (Z) = 8.7559, $P < 0.001$. Replicate factor [Eu], F = 0.4828, effect size (Z) = -1.6918, $P = 0.96$. Specimen factor [St], F = 3.7629, effect size (Z) = 7.5127, $P < 0.001$. Replicate factor [St], F =0.7633, effect size (Z) = -0.7179, $P = 0.76$. Statistics (PERMANOVA on shape) for annotator-2: ME[Eu] = 34.1%, ME[St] = 46.4%. Specimen factor [Eu], F =6.8098, effect size (Z) = 7.4885, $P < 0.001$. Replicate factor [Eu], F = 0.7233, effect size (Z) = -0.63636, $P = 0.72$. Specimen factor [St], F = 4.465, effect size (Z) = 7.4693, $P < 0.001$; Replicate factor [St], F = 0.6583, effect size (Z) = -1.0927, $P = 0.88$. Statistics (PERMANOVA on shape) for annotator-3: ME[Eu] = 13.9%, ME[St] = 21.0%. Specimen factor [Eu], F = 19.647, effect size (Z) = 9.2233, $P < 0.001$; Replicate factor [Eu], F = 0.1969, effect size (Z) = -2.9198, $P = 0.999$. Specimen factor [St], F = 12.286, effect size (Z) = 11.612, $P < 0.001$; Replicate factor [St], F = 0.3595, effect size (Z) = -2.5964, $P = 0.995$.



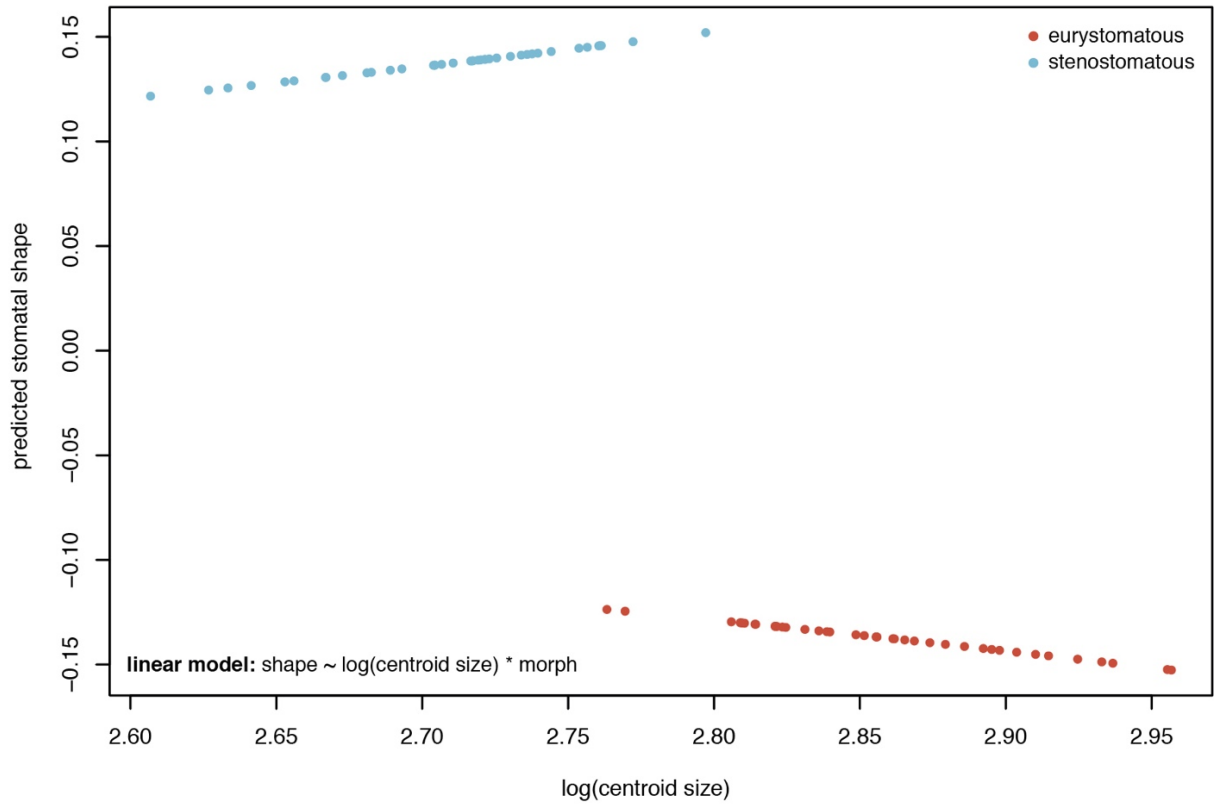
Supplementary figure 7. Annotator-specific distribution of technical replicates (stomatal shapes) in the shape tangent space. Each individual was labeled three times independently by three annotators of

different levels of training in landmark placement. Numbers in the legends refer to identifiers of the ten eurystomatous (a-c) or ten stenostomatous (d-f) specimens (i.e. the biological replicates of each morph). Generalized Procrustes analysis (GPA) was carried out using the **gpagen()** function of the *geomorph* package (see step 20 of the Procedure). *Procrustes coordinates* are projected into a linear (euclidean) tangent space. Semilandmarks were slid by minimizing bending energy. PCA was performed according to step 21 of the Procedure. Note that specimen ID numbers are *not* identical between eurystomatous (a-c) and stenostomatous (d-f) individuals (e.g. '58' refers to two different animals in a-c versus d-f).

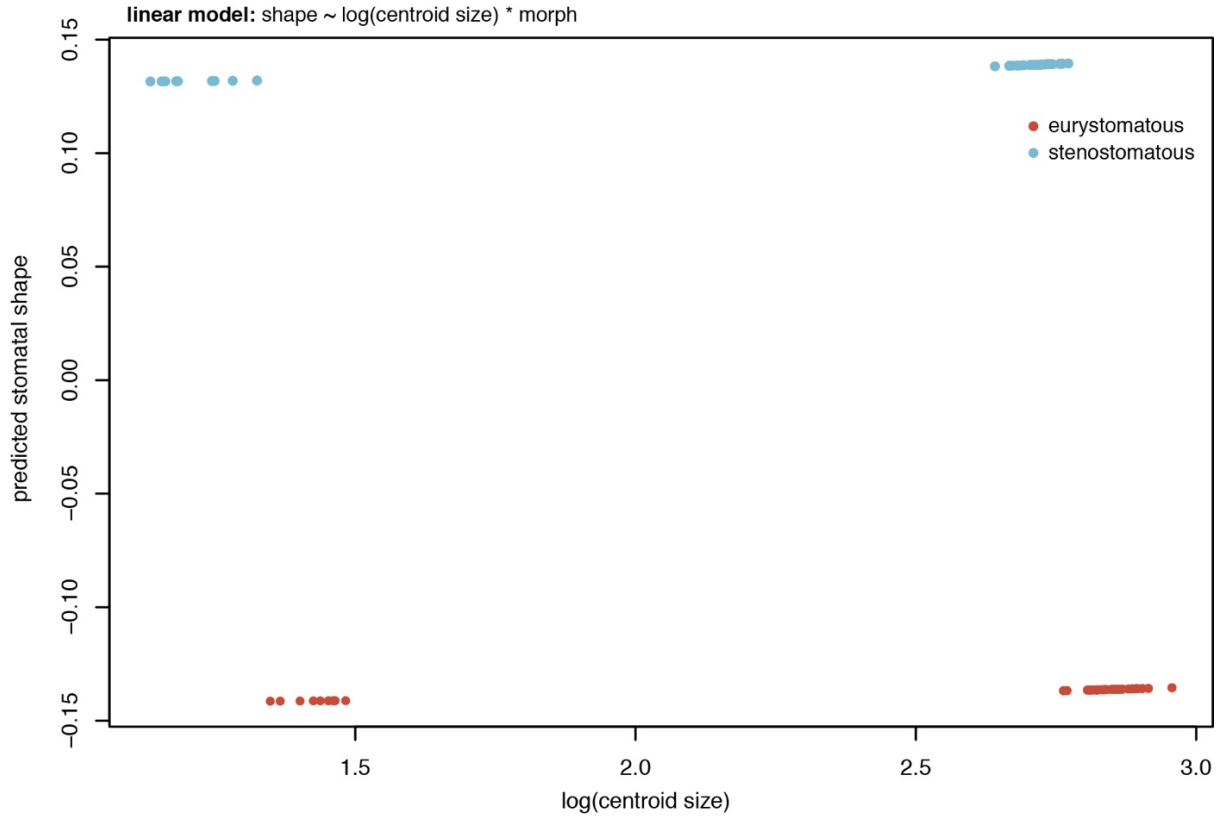


Supplementary figure 8. model-based clustering performed on the two *Pristionchus* datasets using two different approaches for PC inclusion. **a,** Clusters identified for the *P. pacificus* dataset by using only meaningful PCs (mPCs) as an input. **b,** Clusters identified for the *P. pacificus* dataset by including all PCs which describe at least 1% variation (PCs>1%) as an input. **c,** Clusters identified for the *Pristionchus* species dataset by using only meaningful PCs (mPCs) as an input. **d,** Clusters identified for the *Pristionchus* species dataset by including all PCs which describe at least 1% variation (PCs>1%) as an input. Abbreviations: Eu, eurystomatous morph; St, stenostomatous morph; *Pbu*, *P. bucculentus*; *Pel*, *P. elegans*; *Ppa*, *P. pacificus*. Meaningful PCs are estimated by the getMeaningfulPCs() function of the *Morpho* package.

a Allometric trajectories of the two different stomatal morphs in *P. pacificus*

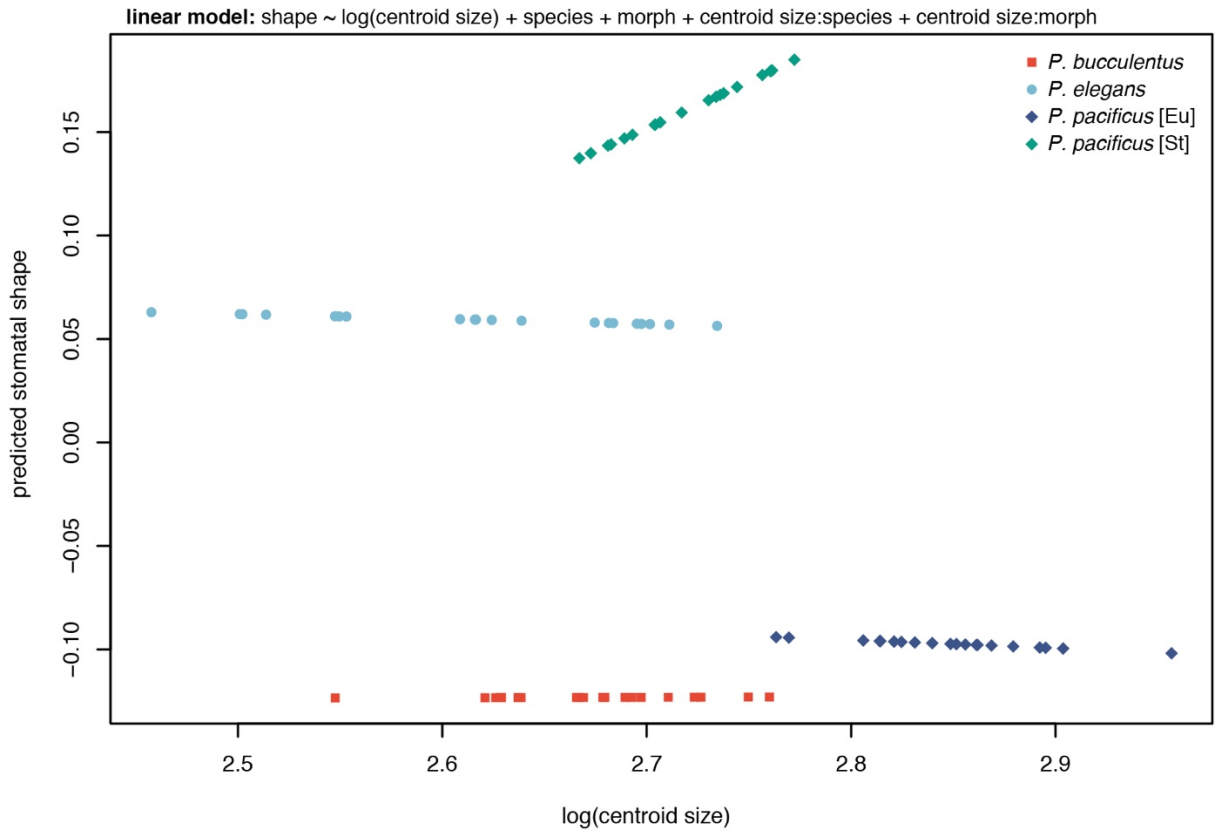


b Allometric trajectories for the two morphs of *P. pacificus* based on uncorrected centroid sizes

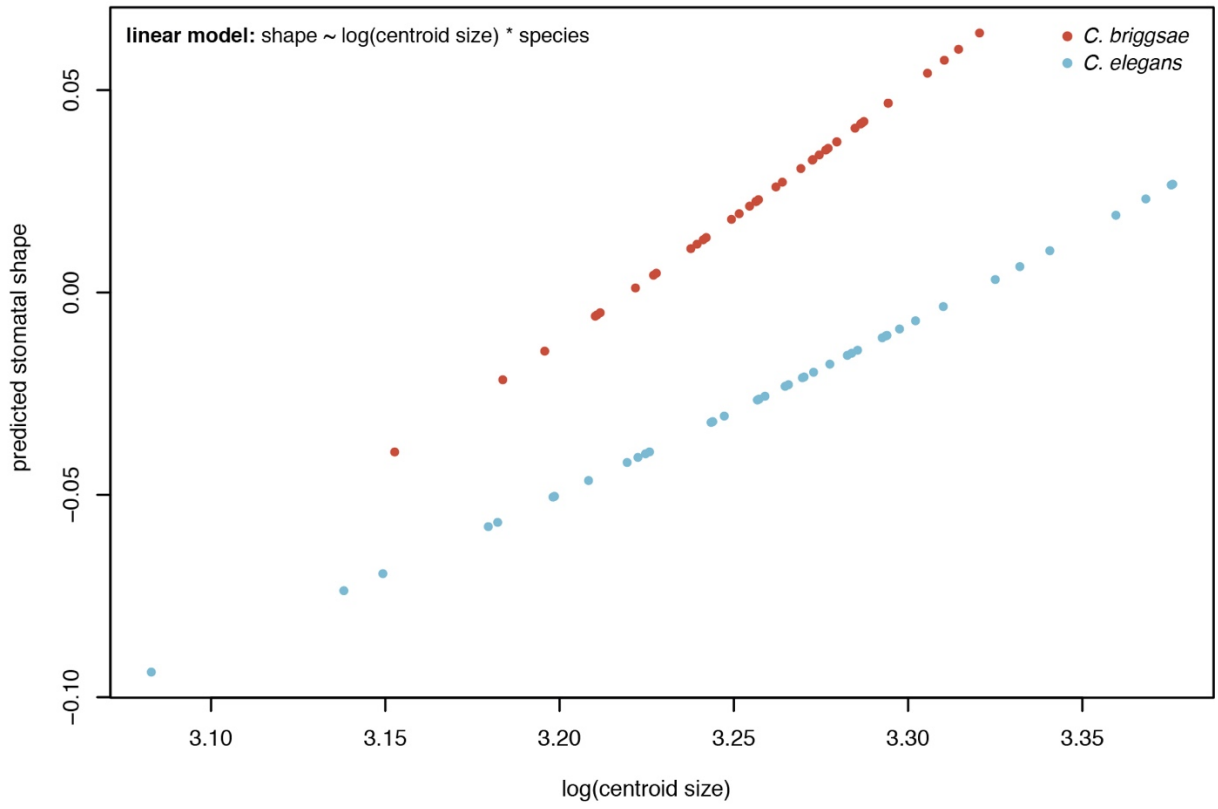


Supplementary figure 9. Allometry plots for young adults of the two *P. pacificus* morphs. **a,** Morph-specific (static) allometric trajectories after troubleshooting (see main text) differences in centroid size due to inconsistent image stack scaling (inches versus microns). **b,** Morph-specific (static) allometric trajectories before troubleshooting centroid sizes (note that individuals whose centroid size was in inches separated from their kin towards the left of the x-axis). Predicted stomatal shapes on the y-axis represent the PC1 scores of the PCA that was performed on the predicted values from multivariate regression on shape versus log(centroid size). Superimposition was performed using the **gpagen()** function of *geomorph* (see step 20 of the Procedure). Note that, while each morph shows static allometry, the differences between morphs are not manifestations of allometric scaling as they do not share a common trajectory. Linear models are indicated in the panels. Statistics on shape (PERMANCOVA): size, F = 165.8679, effect size (Z) =5.7801, $P < 0.0001$. morph, F = 59.6372, effect size (Z) = 6.1827, $P < 0.0001$. size x morph, F = 1.7444, effect size (Z) = 1.356, $P = 0.097$.

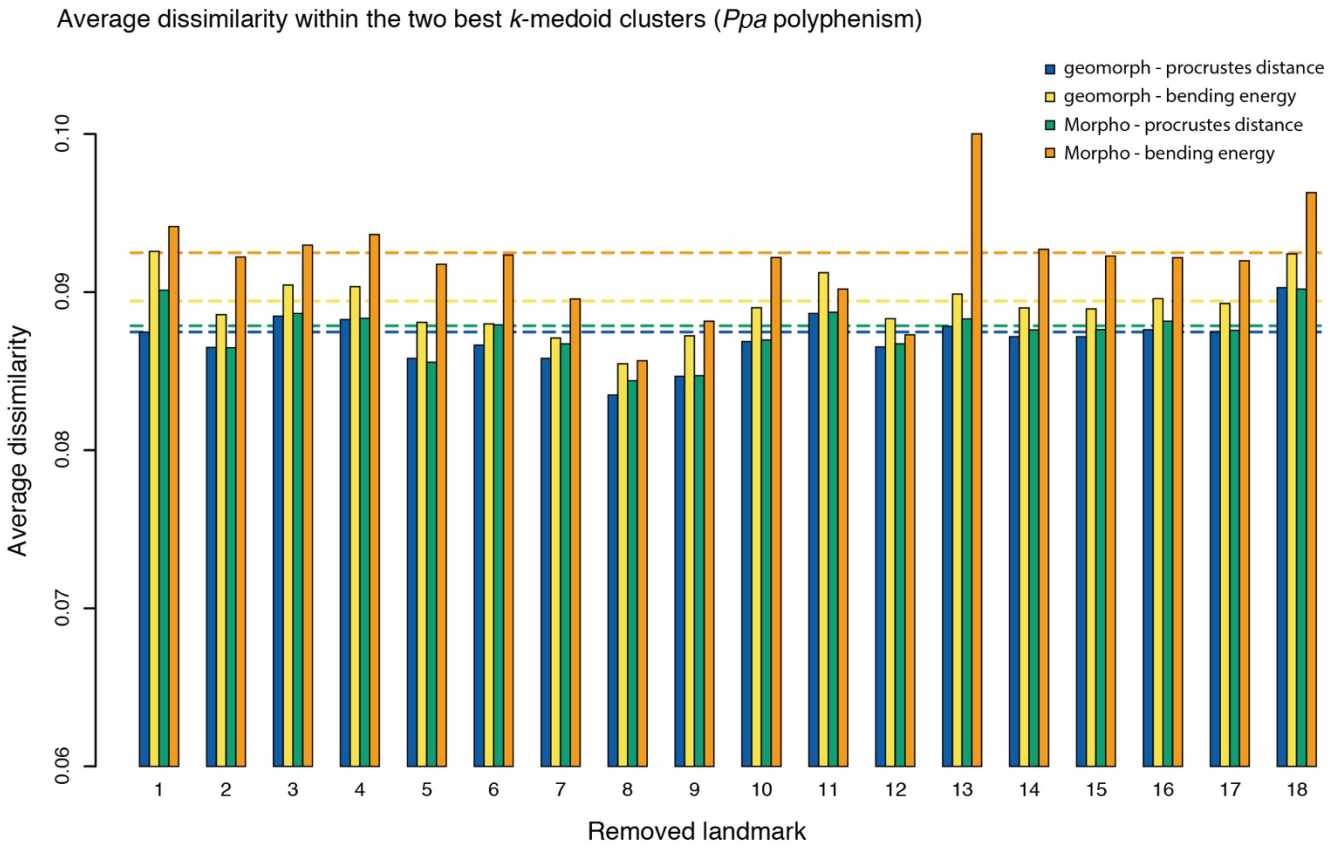
a Allometric trajectories of the stomata in various *Pristionchus* species



b Allometric trajectories of the stomata in *C. briggsae* and *C. elegans*

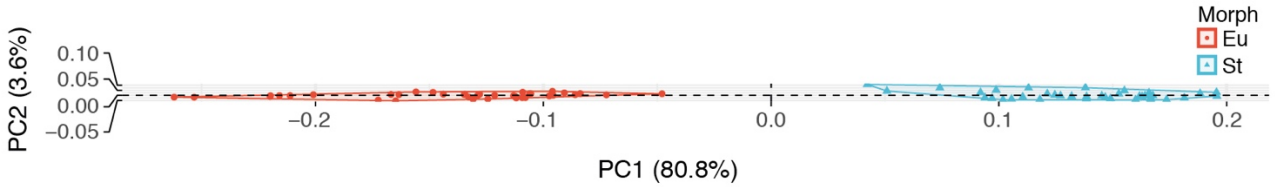


Supplementary figure 10. Static allometry trajectories estimated for young adults of (**a)** three *Pristionchus* species and **(b)** two *Caenorhabditis* species. Note that differences in centroid size due to inconsistent image stack scaling (inches versus microns) were corrected before plotting panel **a** (see 'troubleshooting' in the main text). Predicted stomatal shapes on the y-axis represent the PC1 scores of the PCA that was performed on the predicted values from multivariate regression on shape versus log(centroid size). Superimposition was performed using the **gpagen()** function of *geomorph* (see step 20 of the Procedure). Note that *P. elegans*, eurystomatous *P. pacificus* and *P. bucculentus* (a) show essentially isometric trendlines (i.e. shapes do not change with increasing body size), while the stenostomatous morph of *P. pacificus* shows a specific allometric trajectory. Both *Caenorhabditis* species **(b)** show strong allometric relationships in stomatal shape, indicative of increasing morphological divergence between species with growth. Abbreviations: Eu, eurystomatous and St, stenostomatous morph. Linear models are indicated in the panels. Statistics on shape (PERMANCOVA) for the *Pristionchus* data set **(a)**: size, F = 20.2303, effect size (Z) = 4.3164, $P < 0.0001$. species, F = 62.7531, effect size (Z) = 7.4549, $P < 0.0001$. morph, F = 48.0781, effect size (Z) = 6.2528, $P < 0.0001$. size x species, F = 1.2129, effect size (Z) = 0.7009, $P = 0.23831$. size x morph, F = 2.1124, effect size (Z) = 1.758, $P < 0.05$. Statistics on shape (PERMANCOVA) for the *Caenorhabditis* data set **(b)**: size, F = 27.1836, effect size (Z) = 5.1015, $P < 0.0001$. species, F = 30.6839, effect size (Z) = 6.4582, $P < 0.0001$. size x species, F = 4.3322, effect size (Z) = 3.7579, $P < 0.001$.

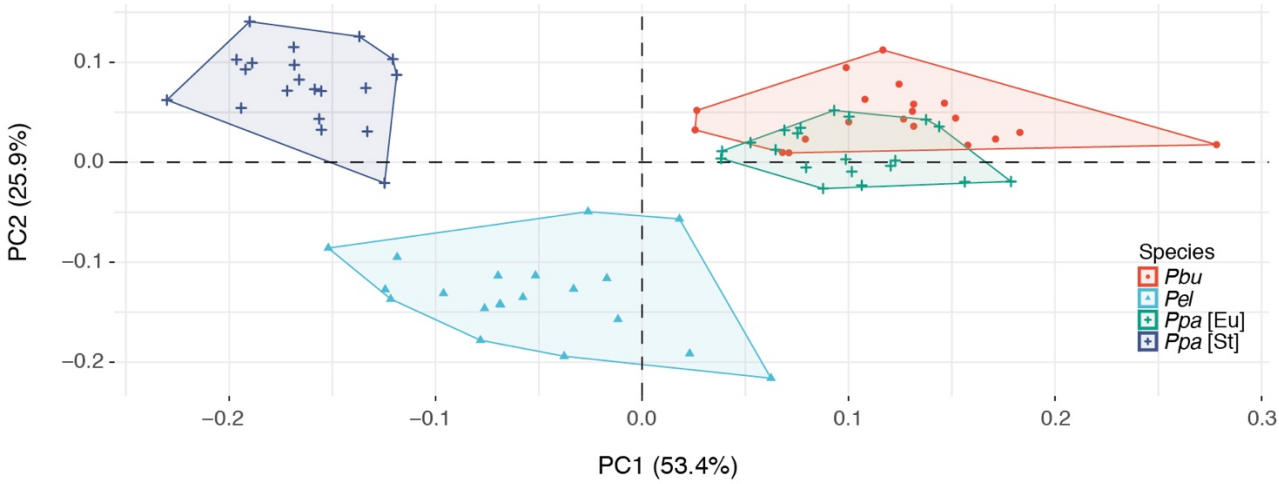


Supplementary figure 11. Change of average dissimilarity within the two-best *k*-medoid clusters upon single landmark removal in the *P. pacificus* dataset (including both morphs). Dotted lines represent the average dissimilarity of the best two *k*-medoid clusters in the data set, when no landmarks are removed (colors correspond to the package and sliding approach used to perform GPA). Note that the detected change for each landmark removal shows a similar direction, regardless of the package which was used for GPA (*geomorph* and *Morpho*) and regardless of the sliding approach that was chosen for GPA (minimization of Procrustes distances or bending energy). Changes of dissimilarity are generally subtle, indicating that the proposed landmark configuration is robustly able to quantify shape differences, even after single landmarks are removed.

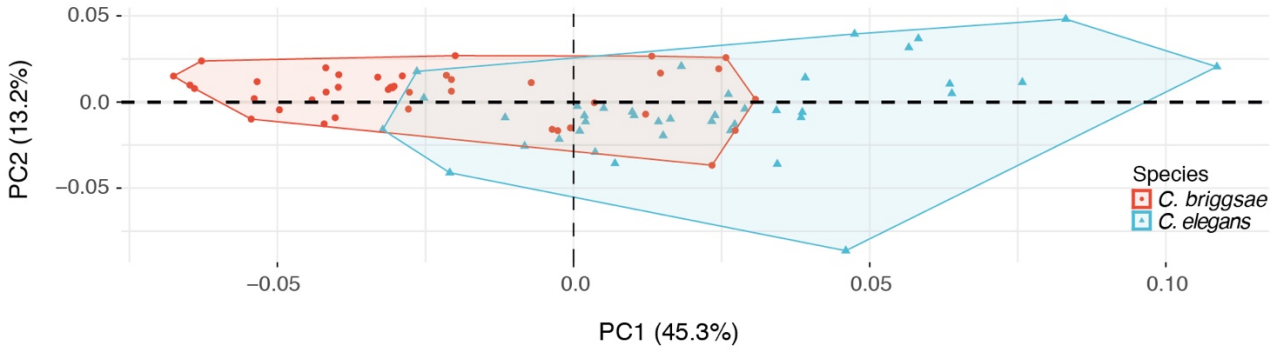
a Variation-weighted PCA plot showing the two different mouth-forms of *P. pacificus*



b Variation-weighted PCA plot showing stomatal shapes of various *Pristionchus* species



c Variation-weighted PCA plot showing stomatal shapes of two *Caenorhabditis* species



Supplementary figure 12. Variation-weighted PCA plots. Length of the y-axis is scaled relative to the x-axis, based on the variation described by PC1 and PC2. **a,** Comparison of two stomatal morphs in *Pristionchus pacificus*; **b,** Comparison of stomatal shapes across various *Pristionchus* species; **c,** Comparison of two *Caenorhabditis* species. Abbreviations: Eu, eurystomatous; St, stenostomatous; *Pbu*, *Pristionchus bucculentus*; *Pel*, *Pristionchus elegans*; *Ppa*, *Pristionchus pacificus*.

landmark	type	description
1	fixed	anteriormost point of cheilostom on the dorsal side
2	fixed	posteriormost point of cheilostom on the dorsal side
3	fixed	anteriormost point of gymnostom on the dorsal side
4	fixed	posteriormost point of gymnostom on the dorsal side
5	fixed	anteriormost point of promesostegostom on the dorsal side
6	sliding	most convex point on the dorsal outline of the dorsal tooth
7	sliding	most concave point on the dorsal outline of the dorsal tooth
8	fixed	anterior tip of the dorsal tooth
9	fixed	opening of the dorsal pharyngeal gland
10	fixed	posteriormost point of the metastegostom on dorsal side (i.e. base of dorsal tooth)
11	fixed	lateral base point of the right ventrosublateral tooth or ridge
12	sliding	anterior tip of the right ventrosublateral tooth or ridge
13	fixed	ventral base point of the right ventrosublateral tooth or ridge
14	fixed	anteriormost point of promesostegostom on the ventral side
15	fixed	posteriormost point of cheilostom on the ventral side
16	fixed	anteriormost point of gymnostom on the ventral side
17	fixed	posteriormost point of cheilostom on the ventral side
18	fixed	anteriormost point of cheilostom on the ventral side

Supplementary table 1. Description of landmarks for *Pristionchus*.

landmark	type	description
1	fixed	anteriormost point of cheilostom on the dorsal side
2	fixed	posteriormost point of cheilostom on the dorsal side (deepest point within the cheilostomal ridge)
3	fixed	anteriormost point of gymnostom on the dorsal side
4	fixed	border point of cheilo- and promesostegostom on the dorsal side (at the level of the pharyngeal sleeve tip)
5	fixed	posteriormost point of promesostegostom on the dorsal side
6	fixed	posterior base point of the dorsal flap
7	fixed	anterior tip of the dorsal flap
8	fixed	anteriormost point of the telostegostom on the dorsal side
9	fixed	opening of the dorsal pharyngeal gland
10	fixed	posteriormost point of the telostegostom on the dorsal side
11	fixed	lateral base point of the right ventrosublateral flap
12	fixed	anterior tip of the right ventrosublateral flap
13	fixed	ventral base point of the right ventrosublateral flap
14	fixed	posteriormost point of promesostegostom on the ventral side
15	fixed	border point of cheilo- and promesostegostom on the ventral side (at the level of the pharyngeal sleeve tip)
16	fixed	anteriormost point of gymnostom on the ventral side
17	fixed	posteriormost point of cheilostom on the ventral side (deepest point within the cheilostomal ridge)
18	fixed	anteriormost point of cheilostom on the ventral side

Supplementary table 2. Description of landmarks for *Caenorhabditis*.

Zootaxa 4943 (1): 001–066
https://www.mapress.com/j/zt/

Monograph

ISSN 1175-5326 (print edition)
ZOOTAXA
ISSN 1175-5334 (online edition)

https://doi.org/10.11646/zootaxa.4943.1.1
http://zoobank.org/urn:lsid:zoobank.org:pub:A603FBF3-FB8D-4BB0-A738-1BD18B0FADAD

ZOOTAXA

4943

Nine new *Pristionchus* (Nematoda: Diplogastridae) species from China

NATSUMI KANZAKI[1*], MATTHIAS HERRMANN[2,3*§], CHRISTIAN WEILER[2], WALTRAUD RÖSELER[2], TOBIAS THESKA[2,4], JÜRGEN BERGER[7], CHRISTIAN RÖDELSPERGER[2,5] & RALF J. SOMMER[2,6§]

[1]*Kansai Research Center, Forestry and Forest Products Research Institute, Kyoto 612-0855, Japan.*
https://orcid.org/0000-0001-8752-1674
[2]*Max Planck Institute for Developmental Biology, Department Evolutionary Biology, Max Planck Ring 9, 72076 Tübingen, Germany*
[3] *matthias.herrmann@tuebingen.mpg.de; https://orcid.org/0000-0002-4801-2524*
[4] *https://orcid.org/0000-0002-0920-4427*
[5] *https://orcid.org/0000-0002-7905-9675*
[6] *ralf.sommer@tuebingen.mpg.de; https://orcid.org/0000-0003-1503-7749*
[7]*Max Planck Institute for Developmental Biology, EM facility, Max Planck Ring 5 Tübingen, Germany*
[*]*these authors contributed equally*
[§]*Corresponding authors.* *matthias.herrmann@tuebingen.mpg.de; ralf.sommer@tuebingen.mpg.de*

Magnolia Press
Auckland, New Zealand

Accepted by M. Hodda: 12 Oct.. 2020; published: 17 Mar. 2021

NATSUMI KANZAKI, MATTHIAS HERRMANN, CHRISTIAN WEILER, WALTRAUD RÖSELER, TOBIAS THESKA, JÜRGEN BERGER, CHRISTIAN RÖDELSPERGER & RALF J. SOMMER
Nine new *Pristionchus* (Nematoda: Diplogastridae) species from China
(*Zootaxa* 4943)

66 pp.; 30 cm.

17 Mar. 2021

ISBN 978-1-77688-212-0 (paperback)

ISBN 978-1-77688-213-7 (Online edition)

FIRST PUBLISHED IN 2021 BY

Magnolia Press

P.O. Box 41-383

Auckland 1041

New Zealand

e-mail: magnolia@mapress.com

https://www.mapress.com/j/zt

ISSN 1175-5326 (Print edition)

ISSN 1175-5334 (Online edition)

Table of Contents

名不正，则言不顺

"If names be not correct, language is not in accordance with the truth of things."
(paraphrased as a chinese proverb stating "The beginning of wisdom is to call things by their proper name."). Confucius.

Abstract

The model organism *Pristionchus pacificus* and the genus *Pristionchus,* Kreis, 1932 have been intensively studied in the last decade with contemporary work focusing on the development, evolution, ecology, behavior, neurobiology, and genomics of this group of organisms. In particular, mechanistic studies on the development and evolution of mouth-form plasticity, predation and associated self-recognition processes enabled unique insight into life history strategies and the evolution of novelty. These studies include a comparative research agenda making use of the 39 available species of *Pristionchus*, all of which can be studied in living cultures. Sampling efforts revealed that Asia represents a biodiversity hotspot for *Pristionchus* worms. However, previous samplings have a bias towards northern and island areas, largely for logistic reasons. Here, we report on two extensive sampling trips to the Yunnan and Shaanxi provinces in Mainland China. We report the isolation of nine new *Pristionchus* species by morphology, morphometrics, mating experiments and genome-wide sequence analysis.

Key words: *Pristionchus pacificus*, Diplogastridae, scarab beetles, phenotypic plasticity, evolution of hermaphroditism

Introduction

The nematode *Pristionchus pacificus* Sommer, Carta, Kim & Sternberg, 1996 of the Diplogastridae family is a well-established model system in evolutionary developmental biology (Sommer 2009; Sommer 2015). A suite of analytical tools available for *P. pacificus* enable comparative and mechanistic studies building on a completely sequenced genome (Dieterich *et al.* 2008; Rödelsperger *et al.* 2017), comparative transcriptomics (Rödelsperger *et al.* 2018), DNA-mediated transformation (Schlager *et al.* 2009), and forward and reverse genetic tools, including the CRISPR/Cas9 technology (Sommer *et al.* 1996; Tian *et al.* 2008; Witte *et al.* 2015). The isolation of hundreds of worldwide strains of *P. pacificus* (Morgan *et al.* 2012) and the support of a robust phylogenetic framework, consisting of more than 20 diplogastrid genera (Susoy *et al.* 2015, 2016), offer the potential to study the evolution of morphological, behavioral and life history traits within and among *Pristionchus* species.

In order to link processes of micro- and macroevolution, a precise knowledge of closely related species is a prerequirement, which is however, hard to achieve in many animal taxa. In the genus *Pristionchus,* the isolation of new species was fostered by the discovery that these soil nematodes are often associated with scarab and related beetles (Herrmann *et al.* 2006a, b, 2007). Therefore, directly targeting beetles for nematode sampling has led to a boom in

the number of *Pristionchus* strains and species recovered (Kanzaki *et al.* 2012a, b, 2013a, b, c, 2014; Ragsdale *et al.* 2013). Currently, 39 *Pristionchus* species have been described, all of which are cultured in the laboratory and are thus available for biological research (Ragsdale *et al.* 2015). This includes *P. exspectatus* Kanzaki, Ragsdale, Herrmann, Mayer & Sommer, 2012, a species close enough to be considered the true sister species of *P. pacificus* isolated in the South-Western part of Japan in 2010 (Kanzaki *et al.* 2012a). Both of these species are part of the *pacificus*-group *sensu stricto*, which currently consists of 10 species. In general, the genus *Pristionchus* is typologically and phylogenetically separated into six major radiations that have sometimes been refered to as "species-groups", "species-complex", or "clades" (Ragsdale *et al.* 2015; Herrmann *et al.*, 2019). Specifically, there are i) the fig-associated clade that branches basal from all other beetle and soil-derived species, ii) another, beetle-associated basal radiation consisting currently of at least two small species groups, the *fissidentatus* and *elegans*-groups, respectively, iii) the *triformis*-group, iv) the *lheritieri*-group with species largely from Europe, v) the *maupasi*-group with species mainly from North-America, and vi) the *pacificus*-group. Within these groups, fig-associates have highly derived morphological and biological characters, *e.g.*, the species have stomal polyphenism showing five different morphotypes, inhabit fresh syconia of figs in the subgenus *Sycomorus* and are vectored by fig wasps (Susoy *et al.*, 2016). In contrast, all other species are morphologically very similar and only the use of sophisticated transcriptomic data allows proper phylogenetic groupings.

Reproductive biology within the genus *Pristionchus* is diverse. While the majority of *Pristionchus* species are gonochorists with equal numbers of males and females, hermaphroditism has arisen independently at least six times (Ragsdale *et al.* 2015; Weadick & Sommer 2016; Rödelsperger *et al.* 2018). Interestingly, only a few hermaphroditic species still form functional males, such as in the case of *P. pacificus*, representing a pre-requisite for genetic analysis. In contrast, other prominent hermaphrodites, such as the nearly cosmopolitan *P. maupasi* (Potts, 1910) Paramonov, 1952 and *P. entomophagus* (Steiner, 1929) Sudhaus & Fürst von Lieven, 2003*,* form only non-functional males, which prevents these species from being studied by forward genetic tools. Mating system transitions as seen between *P. pacificus* and *P. exspectatus* provide a powerful tool for comparative studies, as recently shown in the context of the evolution of life span and aging (Weadick & Sommer 2016). Therefore, the isolation and characterization of additional *Pristionchus* species is not only of taxonomic and systematic value, but offers unique opportunities for research in various areas of biology.

Asia is a hotspot of *Pristionchus* biodiversity and has recently been a rich source for novel *Pristionchus* species for two major reasons. First, most of the basal species of this genus and the outgroup *Parapristionchus* are from Asia (Kanzaki *et al.* 2012c, 2013a; Ragsdale *et al.* 2015). Second, all nine gonochoristic species of the *pacificus*-group *sensu stricto,* including *P. exspectatus* are from Asia, whereas the hermaphrodite *P. pacificus* is nearly cosmopolitan. Importantly, however, biogeographic studies of *Pristionchus* in Asia exhibit a sampling bias towards northern and island areas, and likely result in the current knowledge being incomplete (Kanzaki *et al.* 2012a, 2013a; Ragsdale *et al.* 2015). Mainland China was a *terra incognita* for diplogastrid nematodes, including the genus *Pristionchus,* until Li *et al.* (2015) described the first species. Subsequently, Yoshida *et al.* (2018), Kanzaki *et al.* (2018) and Herrmann *et al.* (2019) intensified the search for new *Pristionchus* species and described several species new to science from China. Realizing that China could be a treasure chest for new *Pristionchus* species, we conducted another two collecting trips with a focus on Yunnan and Shaanxi provinces. Here, we describe the results of these scarab beetle and substrate samplings and the resulting isolation of novel *Pristionchus* species from Mainland China in the years 2018 and 2019. Together, we describe nine new *Pristionchus* species by morphology, morphometrics, mating experiments and genome-wide sequence analysis.

Material and methods

In July 2018, nine locations in Yunnan province were visited and 250 beetles from more than 16 genera, mostly from the Superfamily Scarabaeoidea were collected. Beetles were brought to the laboratory of the Max-Planck Institute (MPI) for Developmental Biology in Tübingen, Germany alive, where they were dissected and put on petri dishes with agar. These dishes were screened for two weeks for emerging nematodes. In July 2019, six locations in the Shaanxi province were sampled and yielded 585 beetles from around 20 genera. These beetles were treated the same way and emerging nematodes could be isolated from several individual beetles. Beetle identification was carried out by M.H. using morphological criteria.

Isolation of isogenic nematode lines. Nematodes were isolated from substrate or insects as single gravid females and transferred to Petri dishes containing NGM agar and *E. coli* (strain OP50) to generate isogenic lines.

Morphological observation and type material preparation. Light microscopic observations for drawings and morphometrics were conducted using live nematode material, which was handpicked from culture plates (Kanzaki 2013). Nomarski micrographs were taken using a Zeiss Axio Imager Z.1 microscope and a Spot RT-SE camera supported by the program MetaMorph v.7.1.3 (Molecular Devices, Sunnyvale, CA, USA). For scanning electron microscopy (SEM) observations, nematodes were prepared by fixation in 2.5% glutaraldehyde in M9 buffer and then post-fixation with 1% osmium tetroxide. After several rinses with water, samples were dehydrated through a graded ethanol series, followed by critical-point drying from carbon dioxide. Specimens were mounted on poly-lysine-coated cover-slips, sputter-coated with 20 nm gold/palladium, and then imaged with a Hitachi S-800 field emission scanning electron microscope operating at 20 kV.

To prepare type material, nematodes were isolated from cultures, rinsed in distilled water to remove bacteria, heat killed at 65°C, fixed in 5% formalin, and processed through a glycerol and ethanol series using Seinhorst's method (Hooper 1986).

Molecular characterisation and phylogenetic analysis. A species phylogeny of the complete *Pristionchus* genus was reconstructed and described previously (Rödelsperger *et al.* 2018). For the nine novel species described here, a similar molecular and transciptomic characterization was performed. In brief, worms were grown on NGM plates seeded with *E. coli* OP50 at 20°C and total RNA was isolated from 2–3 mixed-stage plates per species using standard Trizol extraction following the manufacturers' instructions (Zymo Research, CA, USA). RNA-seq libraries were prepared using TruSeq RNA library preparation kit v2 (Illumina, Inc., CA, USA), according to the manufacturer's instructions, from 1 µg of total RNA in each sample and sequenced on the Illumina HiSeq3000 platform, yielding a median of 14 million paired reads (2 × 150 bp) per species. Raw reads were submitted to the European nucleotide archive under the study accession PRJEB20959. RNA-seq reads were assembled into transcriptomes using Trinity (version v2.2.0) (Grabherr *et al.* 2011). For further analysis, only the first reported isoform per gene was selected and the longest complete or partial ORF (>60 amino acids) was called. Orthologous clusters were generated by orthAgogue (Ekseth *et al.* 2014) and protein sequences were aligned using the MUSCLE software (version 3.8.31) (Edgar 2004). 2,092 high quality alignments containing at least 14 species (without any duplication), with at least 50 amino acid positions with coverage in all represented species, were concatenated into a supermatrix spanning 350,000 amino acids. On the basis of the previous analysis of dozens of gene families (Baskaran *et al.* 2015), we chose the LG substitution model to reconstruct a maximum-likelihood tree using RA×ML (version 8.2.9, options –m PROTGAMMAILG –f a –N 100) (Stamatakis 2014).

Intrageneric grouping of the genus *Pristionchus*. The genus *Pristionchus* Kreis, 1932 is typologically redefined by Sudhaus and Fürst von Lieven (2003) and Kanzaki & Giblin-Davis (2015) as described below in more detail.

1) Buccal cavity often dimorphic (steno- and eurystomatous morphs)
2) Cheilostom with six per- and interradial flaps and separated into six per- and interradial plates, with the plates appearing stippled in lateral view (*i.e.*, right and left subventral, light and left lateral and right and left dorsal, sectors - Cheilostomal plates of *Pristionchus* spp. are located between each sector, and the arrangement is regarded as per and interradial)
3) Stegostom of eurystomatous morph with dorsal claw-like tooth, right subventral tooth and left subventral serrated plate
4) Stegostom of stenostomatous morph with dorsal tooth, right subventral denticle and three left subventral denticles
5) Pharyngeal tubes of procorpus conspicuous
6) Female gonad amphidelphic
7) Gubernaculum with a paired thorn-like and an unpaired rounded anteriorly directed process

Recent dense taxon samplings followed by new species descriptions yielded several exceptional species. For example, three nominal and several unnamed fig-associated species have stomal polyphenism of up to five distinct morphs; two species have monomorphic stoma; and *triformis* group species have six to 12 cheilostomatal flaps and stenostomatous, eurystomatous and megastomatous morphs (Susoy *et al.* 2016; Kanzaki *et al.* 2018).

Results

Of the nine novel species to be described here, *P. nudus* **n. sp.** is the only one with several highly-derived, species-specific apomorphies, *e.g.* a relatively smooth body surface with a conspicuous lateral field and a less-developed median pharyngeal bulb. These typological characters do not fit the common characteristics of the genus *Pristionchus* under its current definition. In addition, the other two basal species, *P. paranudus* **n. sp.** and *P. chinensis* **n. sp.** also have some apomorphic characters. In the following, we first provide the description of the three basal species *P. nudus* **n. sp.**, *P. paranudus* **n. sp.** and *P. chinensis* **n. sp.** Next, we describe the common characters of the *triformis*-group, followed by the description of four new species of this group. Finally, we describe the common characters of the *maupasi*-group and two new species belonging to this group.

Description of three basal species

Pristionchus nudus n. sp.

urn:lsid:zoobank.org:act:6A84DD91-577A-4CDD-A53D-072C37D5F7EE

Etymology. The species name is derived from the absence of prominent striation, common in other *Pristionchus* species (*nudus* lat.: naked)

Measurements. See Table 1

Adult. Body cylindrical, stout, *i.e.*, body length-maximum body diam. ratio (a value) is ranging from 12-18, depending on culture condition; cuticle moderate in thickness, smooth with fine annulations, weak but conspicuous longitudinal striations, conspicuous lateral field with two bands, no clear annulation between bands. Head without apparent lips, six mound-like anteriorly directed expansions, one on each sector. Six labial sensilla one on each lip sector, long, bristle-like; two lateral sensilla clearly longer than other four. Cephalic papillae vestigial, probably degenerate and embedded in cephalic tissue. Amphidial apertures located on lateral sector, slightly dorsally shifted, at level of margin of cheilo- and gymnostom. Stomal dimorphism not observed; all individuals examined of stenostomatous form. Dorsal pharyngeal gland penetrating dorsal tooth to gland opening. Anterior pharynx (=pro and metacorpus) twice as long as posterior (=isthmus and basal bulb), muscular composed of tube-shaped procorpus and moderately developed metacorpus (median bulb) with roundish rectangular shape in lateral view; posterior pharynx glandular with nerve ring surrounding anterior end of isthmus. Pharyngo-intestinal junction (cardia) well-developed. Intestine simple tube, not forming pre-rectum, extended posteriorly from cardia to rectum; three (two subventral and one dorsal) rectal gland cells at distal end of intestine (margin between intestine and rectum). Secretory-excretory pore not conspicuous, ventrally located at level of isthmus to pharyngo-intestinal junction, excretory duct extending anteriad and reflexed back to position of pore; two large secretory-excretory cells around excretory duct. Deirid observed laterally on lateral field, located around posterior end of basal bulb to pharyngo-intestinal junction to a half body diameter posterior to junction, *ca* 0.5 body diam. posterior to secretory-excretory pore. Hemizonid not observed. Lateral glands (small pores connected to secretory cell) on lateral body surface, with positions inconsistent among individuals, numbering 5 to 8 for males and 9 to 13 for females. Postdeirid at anterior part of *vas deferens* in male and posterior end of posterior gonad in female, on the same striation with deirid (= lateral field) or the adjacent striation or the second dorsally neighboring striation to lateral field.

Stenostomatous form. Cheilostom consisting of six per- and interradial plates, each forming small flap at anterior end, posterior part of each plate broad, consisting of translucent and non-sclerotized tissue. Gymnostom short, cuticular ring-like anterior end overlapping cheilostom internally; two layers of metastegostomatal mounds consisting of weakly sclerotized tissue present on dorsal side of inner wall. Pro-meso stegostom forming a weakly cuticularized ring internally overlapping with gymnostom to connect gymnostom and metastegostom; broad and conspicuous in dorsal view. Metastegostom bearing dorsal mounds, small, conspicuous, triangular, movable tooth with clearly observed pharyngeal tube. Left subventral ridge with three minute, rounded adventitious denticles on a plate, and the most ventral denticle masked by remaining two in lateral view. Right subventral ridge with two rounded distal adventitious denticles. Telostegostom weakly sclerotized cup-like cavity connecting stoma and pharynx.

Male. Whole body ventrally arcuate, strongly ventrally curved at tail region when killed by heat. Testis single, ventrally located, anterior part reflexed to right or left side; spermatogonia arranged in three to five rows in reflexed

part, well-developed spermatocytes arranged as three to four rows in anterior two-thirds of main branch, mature amoeboid spermatids arranged in multiple rows in proximal part of gonad. *Vas deferens* not clearly separated from other parts of gonad. Posterior end of *vas deferens* and rectum fused to form a cloacal tube. Spicules paired, separate; spicules smoothly curved in ventral view, adjacent to each other for distal third of their length, each smoothly tapering to pointed distal end; spicule in lateral view smoothly ventrally arcuate, giving spicule about 100° curvature, oval manubrium at anterior end, lamina/calomus complex smoothly tapering to pointed distal end. Gubernaculum conspicuous, about one-third of spicule length, broad anteriorly such that dorsal wall is slightly recurved with dorsal and ventral walls separate at ca 30° angle at posterior end; dorsal side of gubernaculum possessing single,

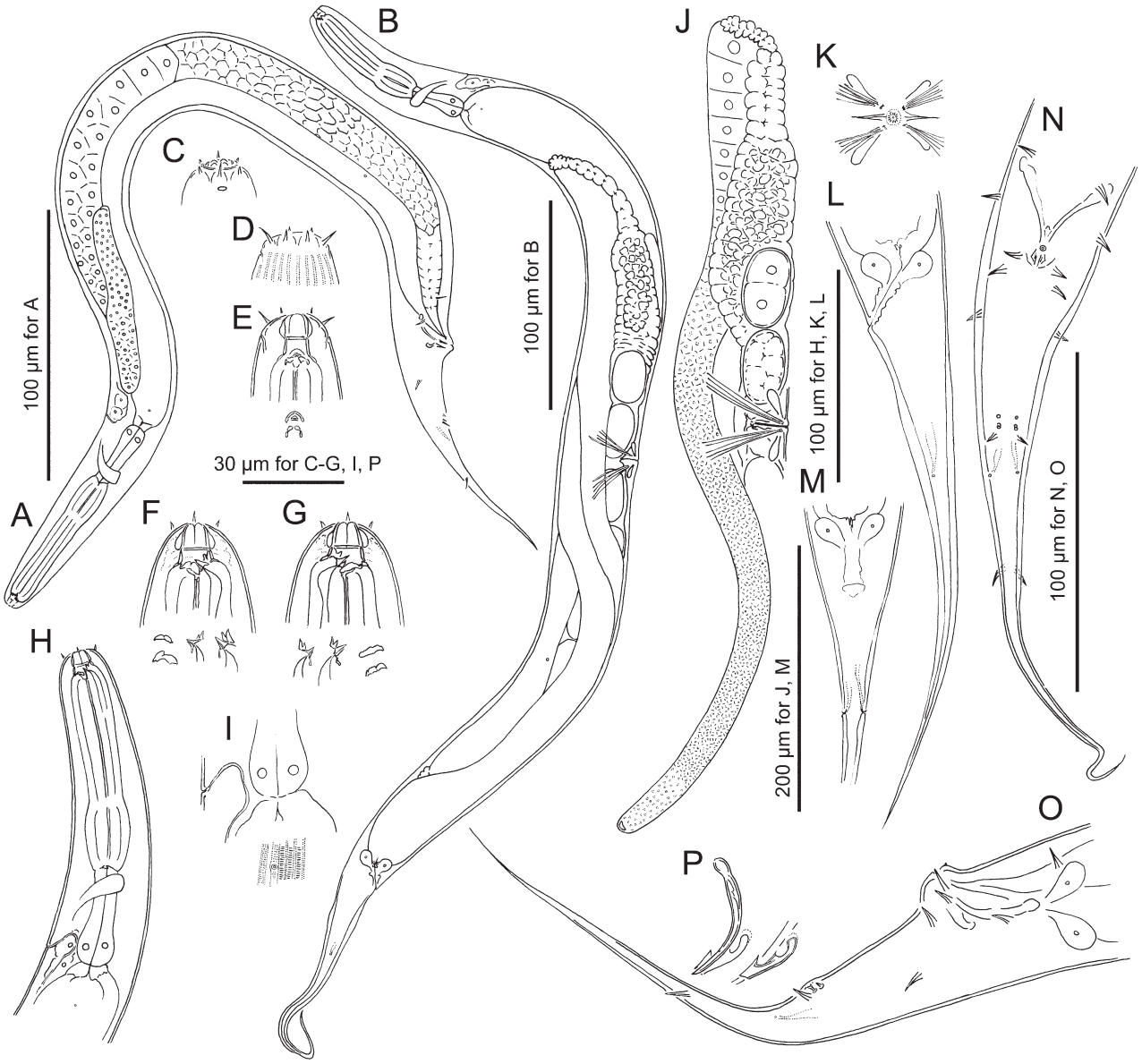


FIGURE 1. *Pristionchus nudus* **n. sp.** A: Right lateral view of adult male. B: Right lateral view of adult female. C: Left lateral view of male head region showing labial sensilla, cephalic papillae and amphid. D: Ventral view of male head region showing labial sensilla, cephalic papillae, amphid and the anterior end of longitudinal striations. E-G: Stomal region of adult female in ventral (E), left lateral (F) and right lateral (G) view. Morphological variations of tooth and ridges are separately drawn in each subfigure as dorsal tooth (upper) and right and left subventral ridges (lower) (E), variation in left subventral ridge (left) and dorsal tooth (middle and right) (F) and variation of dorsal (left and middle) tooth and right subventral (right) ridge (G). H: Anterior part of adult female in left lateral view. I: Body surface structure showing lateral field and the relative position of deirid. J: Anterior female gonad in right lateral view. K: Vulval region in ventral view. L, M: Female tail in left lateral (L) and ventral (M) view. N, O: Male tail in ventral (N) and left lateral (O) view. P: Spicule and gubernaculum in left lateral view.

FIGURE 2. *Pristionchus nudus* **n. sp.** A: Left lateral view of stenostomatous form (only form observed) in four different focal planes. B: Right lateral view of stenostomatous form in two different focal planes. C: posterior pharynx region in three different focal planes in left lateral view where deirid (d), nerve ring (nr) secretory-excretory pore (ep) median bulb (mb) and basal bulb (bb) are indicated.

membranous, anteriorly directed process and lateral pair of more sclerotized, anteriorly and obliquely ventrally directed processes. In lateral view, anterior half of gubernaculum with two successive curves separated by anteriorly and obliquely ventrally directed process, with anterior terminal curvature highly concave and almost closed, with deep posterior curvature being one-third of gubernaculum length; posterior half forming tube-like process enveloping spicules. Cloacal opening (co) slit-like in ventral view; one small, ventral, single genital papilla (vs) on anterior cloacal lip. All nine paired genital papillae relatively long, bristle-like. Tail elongate conoid to sharply pointed tip, *i.e.*, tail does not form a distinct spike. Paired papillae and phasmid are arranged as <v1, v2d, v3, co, v4, ad, (v5, v6, v7), ph, pd>, where v1 located about 1 cloacal body diameter (CBD) anterior to co; v2d a little less than 0.5 CBD anterior to co; v3 adcloacal; v4 at less than 1/10 CBD posterior to co, *i.e.*, v3, co and v4 are very close to each other; ad about 1 CBD posterior to co; v5–v7 forming triplet, but clearly separated from each other, about 1 CBD posterior to ad, *i.e.*, about 2 CBD posterior to co; ph at about 1/4 CBD posterior to v7; and pd about 1 CBD posterior to v7. v1, v4 and ph subventral, v2d and ad lateral, v3, v5–7 ventral, pd subdorsal in male tail. Bursa or bursal flap absent.

Female. Body relaxed or weakly ventrally arcuate when killed by heat. Gonad didelphic, amphidelphic; each gonadal system arranged from vulva/vagina as uterus, oviduct, and ovary; anterior gonad right of intestine, with uterus and oviduct extending ventrally and anteriorly on right of intestine and with totally reflexed (=antidromous

reflexion) ovary extending dorsally on left of intestine; oocytes mostly arranged in three to four or more rows in distal two-thirds of ovary and in double or single row in rest of ovary, distal tips of each ovary reaching oviduct of opposite gonad branch; anterior end of oviduct (=junction tissue between ovary and oviduct) consists of rounded cells; anterior part of oviduct consists of rounded cells, forming a simple tube; middle part of oviduct serving as spermatheca, consists of roundish and relatively large cells. Eggs in single to multiple-cell stage or even further developed at posterior part of oviduct (=uterus), in young females being composed of squared or angular cells, long enough to contain one well-developed oocyte. *Receptaculum seminis* not observed, *i.e.*, the organ is not independent, and a part of oviduct/uterus works as the organ; vaginal glands present but obscure; vagina perpendicular to body surface, surrounded by sclerotized tissue; vulva slightly protuberant in lateral view, pore-like in ventral view; rectum about one anal body diameter (ABD) long, intestine/rectum junction surrounded by well-developed sphincter muscle. Anus in form of dome-shaped slit, posterior anal lip slightly protuberant; phasmid about 2 ABDs posterior to anus. Tail elongate, conoid, with sharply pointed tip.

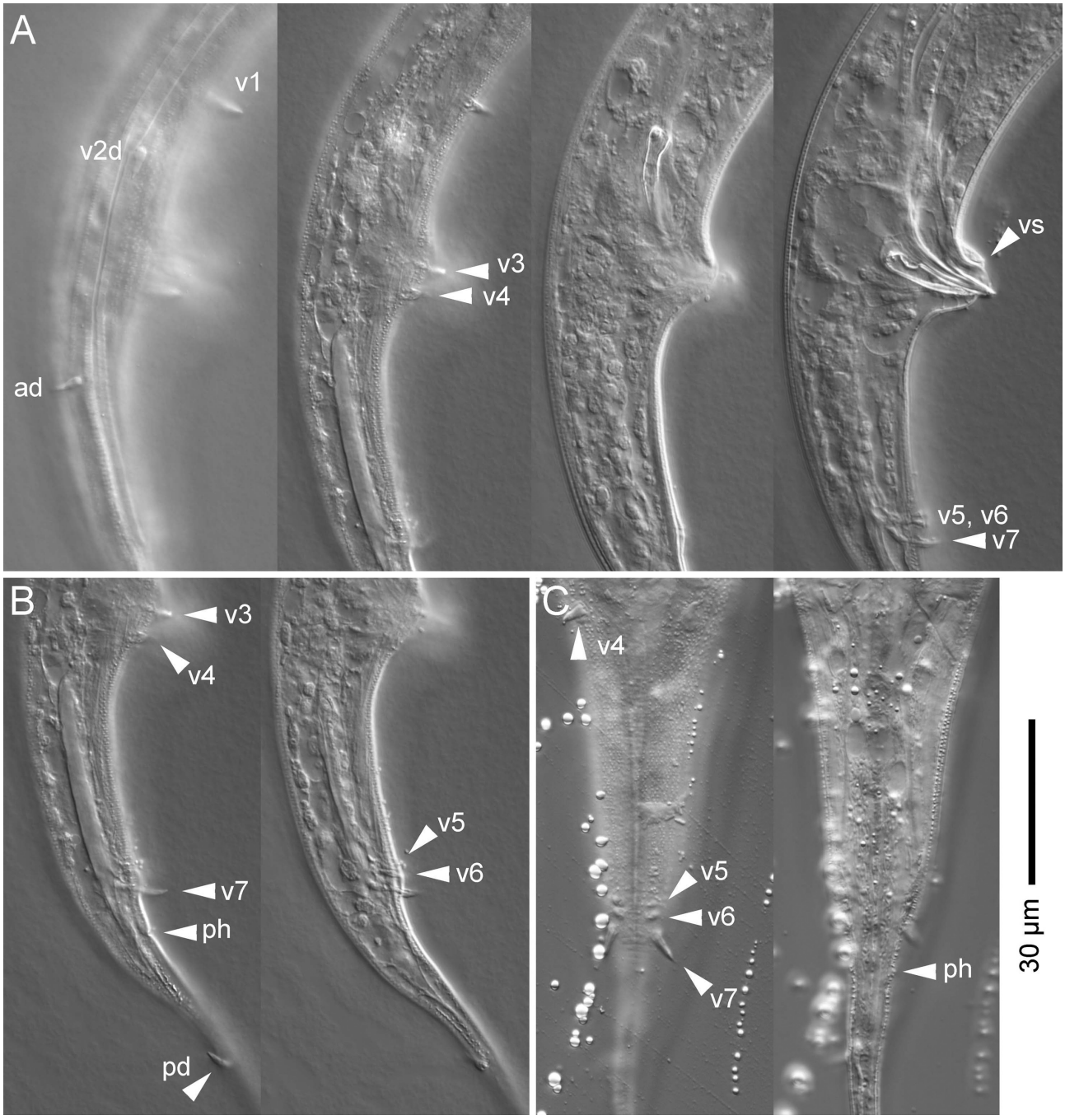


FIGURE 3. *Pristionchus nudus* **n. sp.** male tail with the position of genital papillae labelled according to the terminology of Sudhaus and Fürst von Lieven (2003). A: lateral right view in six focal planes. B: ventral view in two focal planes.

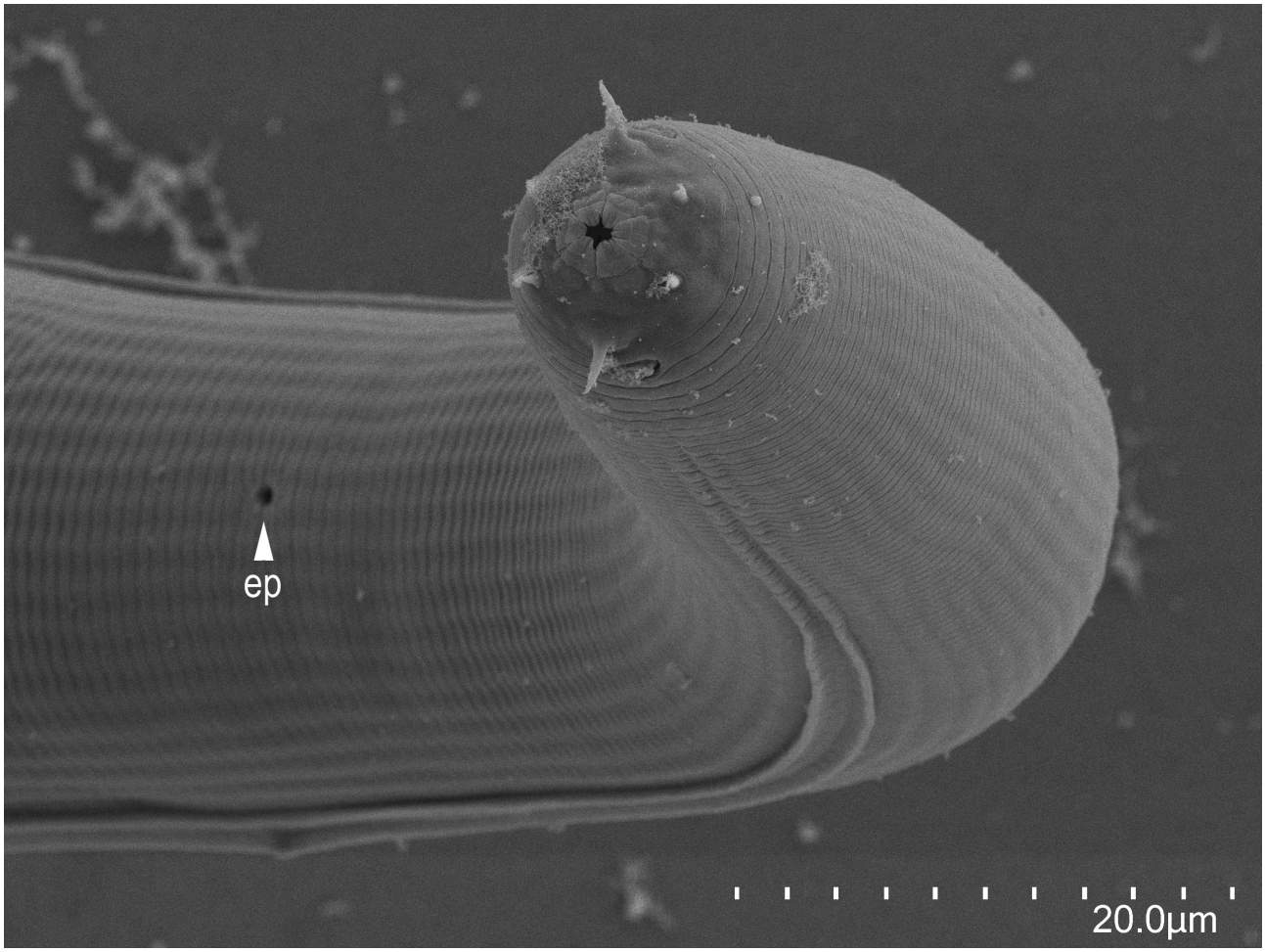


FIGURE 4. Scanning electron photomicrograph of *Pristionchus nudus* **n. sp.** female anterior end with secretory-excretory pore (ep) being indicated.

Diagnosis and relationships. *Pristionchus nudus* **n. sp.** has several characteristic typological traits, which have not been found in other species of the genus. The new species is characterized by its long and bristle-like labial sensilla where two lateral sensilla are much longer than the other four, lack of male cephalic sensilla, monomorphic stoma morphology (stenostomatous form only), where thick posterior half of cheilostomal plate is composed by translucent tissue, two-layered dorsal gymno- and stegostomal mound, small and triangular dorsal tooth, roundish rectangular-shaped metacorpus, position of nerve ring surrounding the anterior end of isthmus, long male tail without clear spike and the arrangement of male genital papillae, *i.e.*, <v1, v2d, v3, co, v4, ad, (v5, v6, v7), ph, pd>. The new species can be readily distinguished from all other *Pristionchus* species with the above species-specific characters, *i.e.*, none of the other species in the genus have these characters.

Phylogenetically, *P. nudus* **n. sp.** belongs to the basal species group of the genus, where two monomorphic species, *P. elegans* Kanzaki, Ragsdale, Herrmann, & Sommer, 2012 and *P. bucculentus* Kanzaki, Ragsdale, Herrmann, Röseler & Sommer, 2013 (*elegans* group), and two dimorphic species, *P. fissidentatus* Kanzaki, Ragsdale, Herrmann & Sommer, 2012 and *P. paulseni* Herrmann, Kanzaki, Weiler, Yoshida, Rödelsperger & Sommer, 2019 (*fissidentatus* group) have been described. *P. nudus* **n. sp.** shares the modification of the cheilostomatal plates with *P. elegans* and *P. bucculentus,* which have thin membrane-like cheilostomatal plates. Within these two species, *P. nudus* **n. sp.** shares the stoma pattern, monomorphic stenostomatous form with *P. elegans*. However, the new species is distinguished from *P. elegans* with its species-specific characters described above. Further, the new species is distinguished from all other species by mating experiments and also characterized by a *ca.* 1,600-bp fragment of the SSU rRNA gene (GenBank accession number MW017221), the sequence of which is distinct from that of all other *Pristionchus* species.

Type host and locality. Isolated from a longhorn beetle (Coleoptera: Cerambycidae) collected at Xishuangbanna Tropical Botanical Garden - Green stone forest, Yunnan province, PRC.

Type material and type strain. Type strain RS6026 frozen at the nematode collection of the MPI Tübingen and available as living culture upon request. Voucher specimens sent to the following museums: Holotype male, Paratype male and female: Museum für Naturkunde Karlsruhe, Germany; Paratype male and female: Swedish Natural History Museum, Stockholm, Sweden; Paratype male and female: University of California in Riverside Nematode Collection (UCRNC), Riverside, CA, USA.

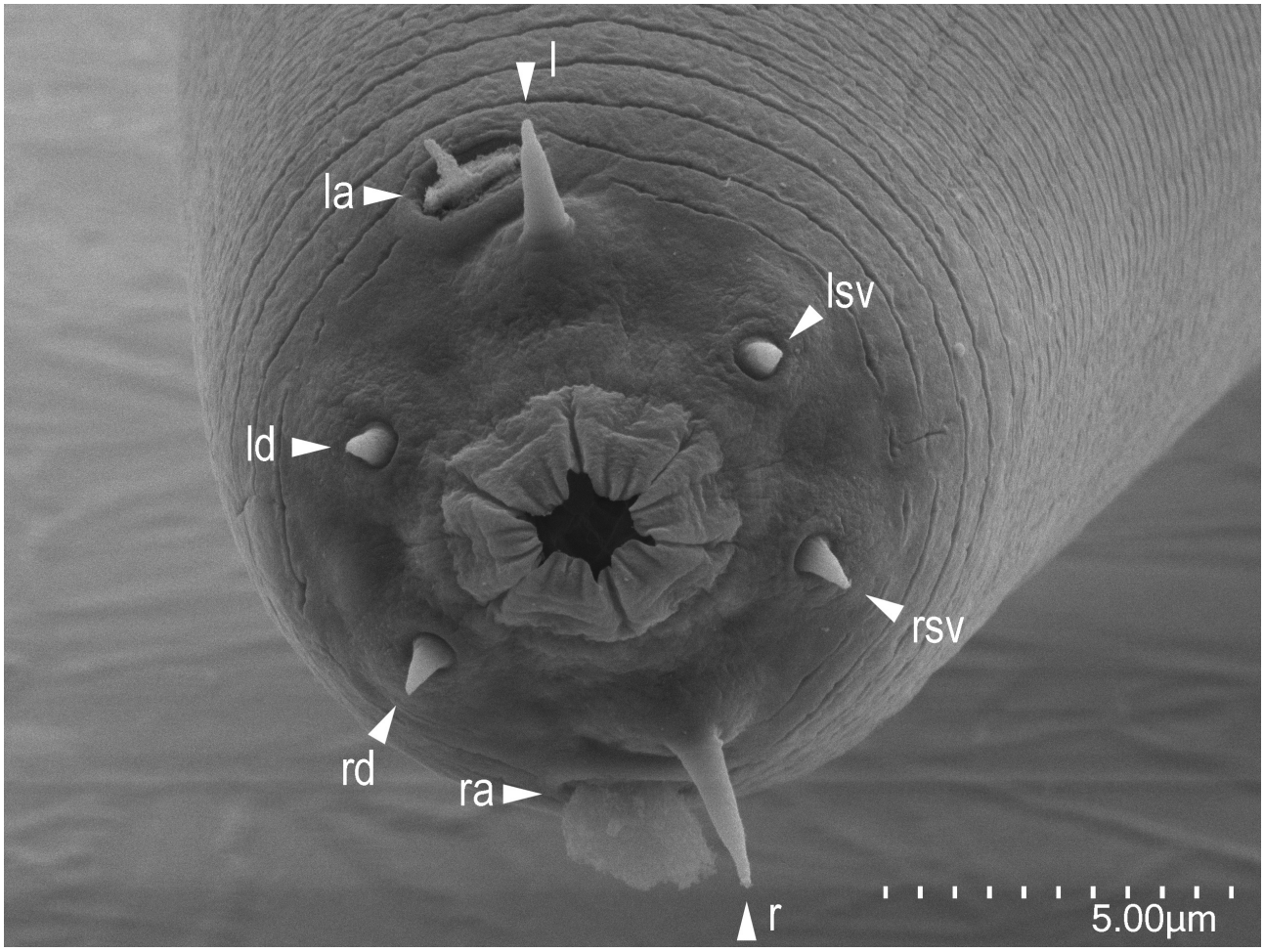


FIGURE 5. Scanning electron photomicrograph of *Pristionchus nudus* **n. sp.** male *en face* with left and right amphids (la and ra) and left and right subventral (lsv and rsv), lateral (l and r) and left and right dorsal (ld and rd) labial sensilla being indicated.

***Pristionchus paranudus* n. sp.**

urn:lsid:zoobank.org:act:23065D78-DFEB-4739-9C5E-B023DE10F678

Etymology. The species name is derived from its phylogenetic position as tentative sister to *P. nudus*.

Measurements. See Table 1

Adult. Body cylindrical, moderate to stout, *i.e.*, body length-maximum body diam. ratio (a value) is ranging from 15–22, depending on culture condition; cuticle moderate in thickness, smooth with fine annulations, weak longitudinal striations, conspicuous lateral field with two bands. No clear annulation between bands. Head without apparent lips, six mound-like anteriorly directed expansions, one on each sector. Six labial sensilla one on each lip sector, short, papilliform. Four small papilliform cephalic papillae in males on right and left subventral and right and left dorsal sectors. Amphidial apertures located on the lateral sector, slightly dorsally shifted, at level of margin of cheilo- and gymnostom. Stomal dimorphism present, and details are described below. Dorsal pharyngeal gland clearly observed, penetrating dorsal tooth to gland opening. Anterior pharynx 1.5 times as long as posterior pharynx, muscular, composed of tube-shaped procorpus and well-developed oval-shaped metacorpus; posterior pharynx glandular with nerve ring surrounding the middle of isthmus. Pharyngo-intestinal junction (cardia) well-developed.

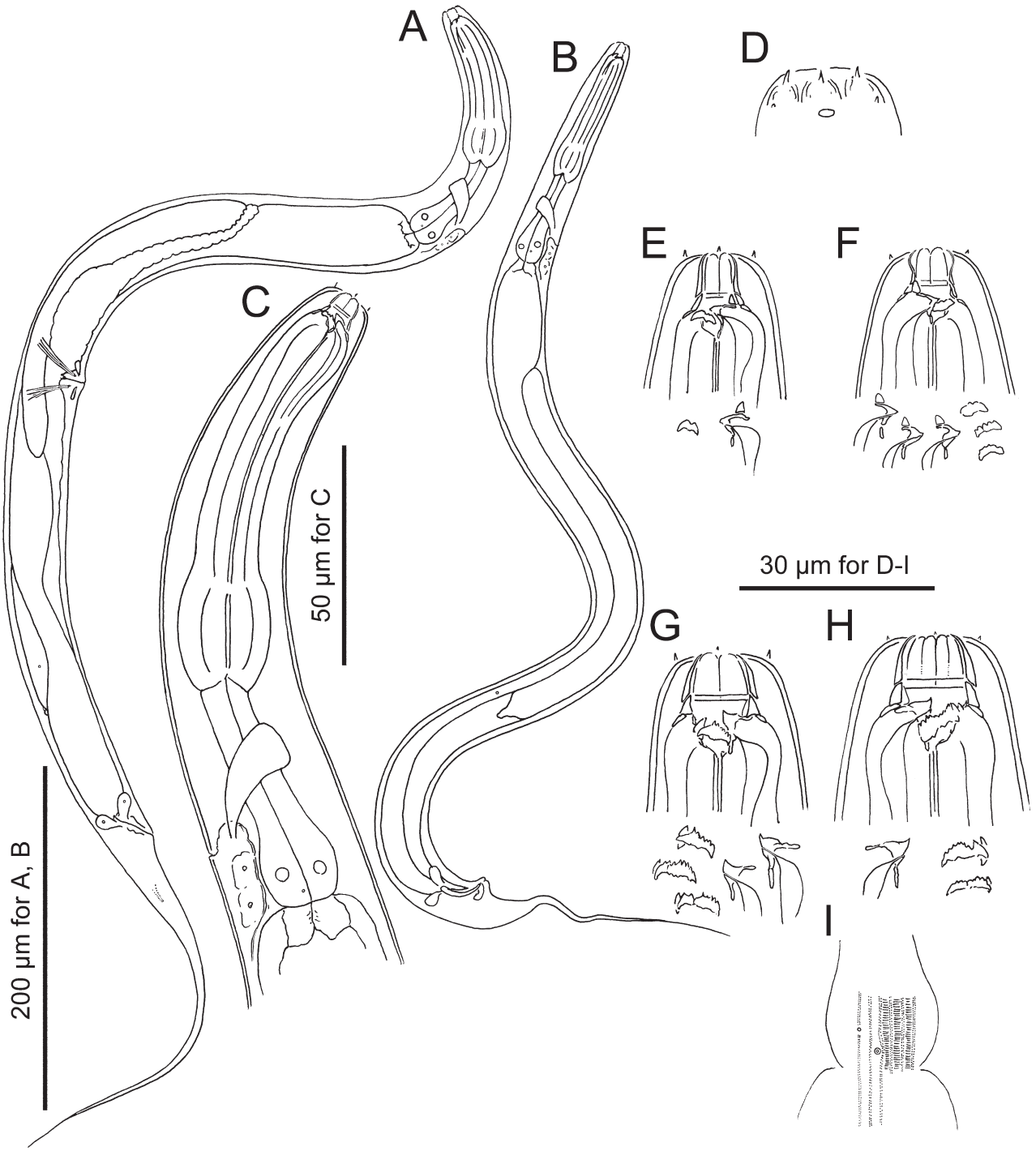


FIGURE 6. *Pristionchus paranudus* **n. sp.** A: Right lateral view of adult female. B: Right lateral view of adult male. C: Anterior part of adult female in left lateral view. D: Left lateral view of male head region showing labial sensilla, cephalic papillae and amphid. E, F: Stomal region of stenostomatous female in left (E) and right (F) lateral view. G, H: Stomal region of eurystomatous female in left (G) and right (H) lateral view. Morphological variations of tooth, ridges and serrated plate are separately drawn in each subfigure as left subventral ridge (left) and dorsal tooth (right) (E), dorsal tooth (left and middle) and right subventral ridge (right) (F), left subventral plate (left) and dorsal tooth (middle and right) (G), dorsal tooth (left) and right subventral plate (right) (H). I: Body surface structure showing lateral field and the relative position of deirid.

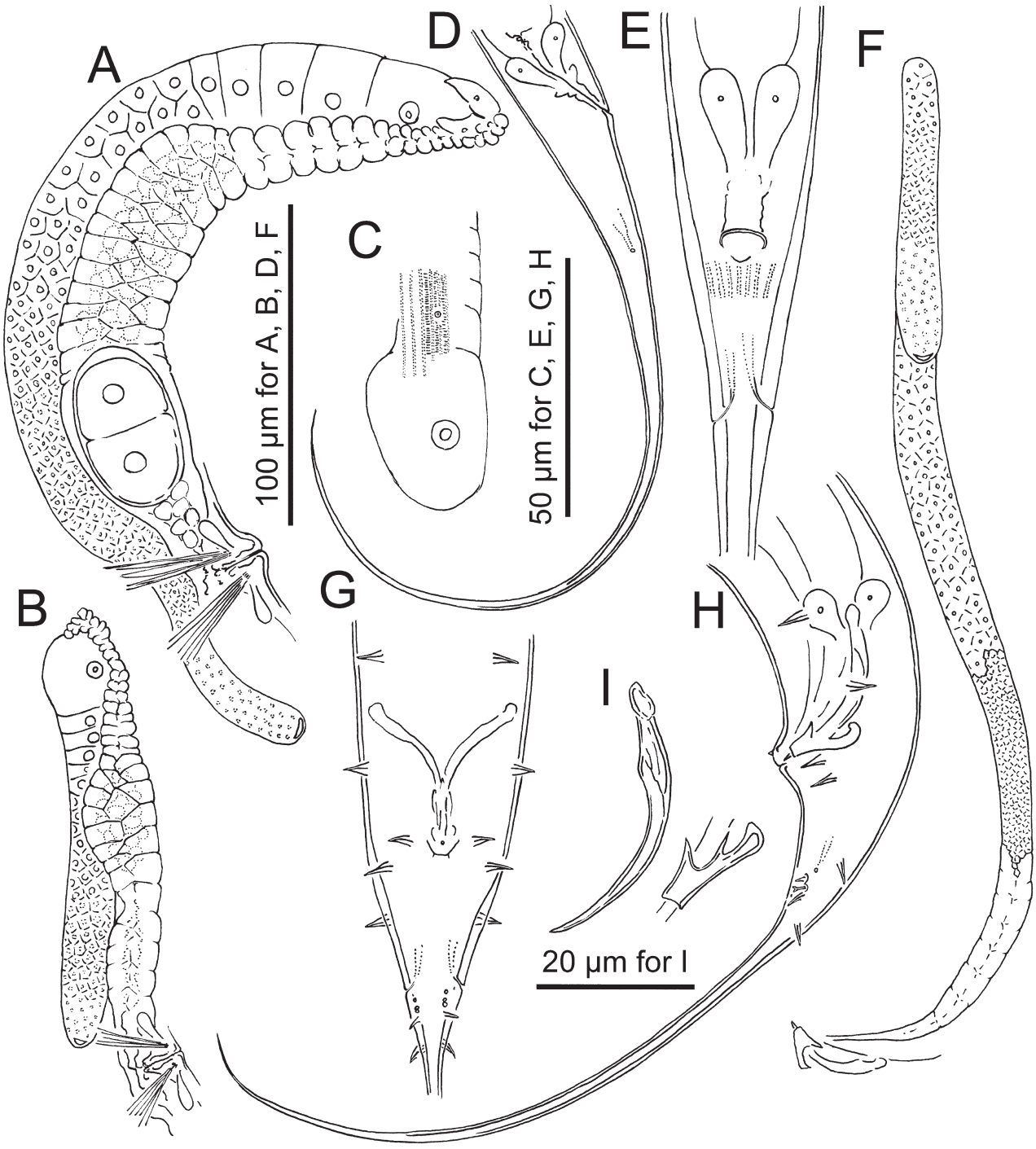


FIGURE 7. *Pristionchus paranudus* **n. sp.** A, B: Anterior female gonad of well-developed (A) and young (B) individual in right lateral view. C: Body surface structure at the postdeirid region in left lateral view showing lateral field and relative position of postdeirid. D, E: Female tail in right lateral (D) and ventral (E) view. F: Male gonad in left lateral view. G, H: Male tail in ventral (G) and left lateral (H) view. I: Spicule and gubernaculum in left lateral view.

Intestine simple tube, not forming pre-rectum, extended posteriorly from cardia to rectum; three (two subventral and one dorsal) rectal gland cells observed at distal end of intestine (margin between intestine and rectum). Secretory-excretory pore not conspicuous, ventrally located at level of isthmus to pharyngo-intestinal junction, excretory duct extending anteriad and reflexed back to position of pore; two large secretory-excretory cells around the excretory duct. Deirid observed laterally on lateral field, located at the level around the posterior end of basal bulb to pharyngo-intestinal junction to a half body diameter posterior to the junction, *ca* 0.5–1 body diam. posterior to secretory-excretory pore. Hemizonid not observed. Lateral glands (small pores connected to secretory cell) on lat-

eral body surface, with positions inconsistent among individuals, numbering 5 to 8 for males and 9 to 13 for females. Postdeirid at anterior part of *vas deferens* in male and the posterior end of posterior gonad in female, on the same striation with deirid (= lateral field) or on the adjacent striation or the second dorsally neighboring striation to lateral field.

Stenostomatous form. Cheilostom with six per- and interradial cuticular plates. Anterior end of each plate rounded and elongated to project from stomal opening and form a small flap. Gymnostom short, cuticular ring-like anterior end overlapping cheilostom internally with, metastegostomal mound consisting of weakly sclerotized tissue present on dorsal side of inner wall. Pro-meso stegostom forming a weakly cuticularized ring internally overlapping with gymnostoma to connect gymnostom and metastegostom, with dorsal side relatively thick and conspicuous, ventral side thin and inconspicuous. Metastegostom bearing dorsal mound, conspicuous, movable triangular or flint-shaped dorsal tooth with strongly sclerotized surface with an inverted V-shape in lateral view. Left subventral ridge with three minute, rounded adventitious denticles on a plate, most ventral denticle is masked by the remaining two in lateral view. Right subventral ridge with two rounded distal adventitious denticles. Telostegostom weakly sclerotized cup-like cavity connecting stoma and pharynx.

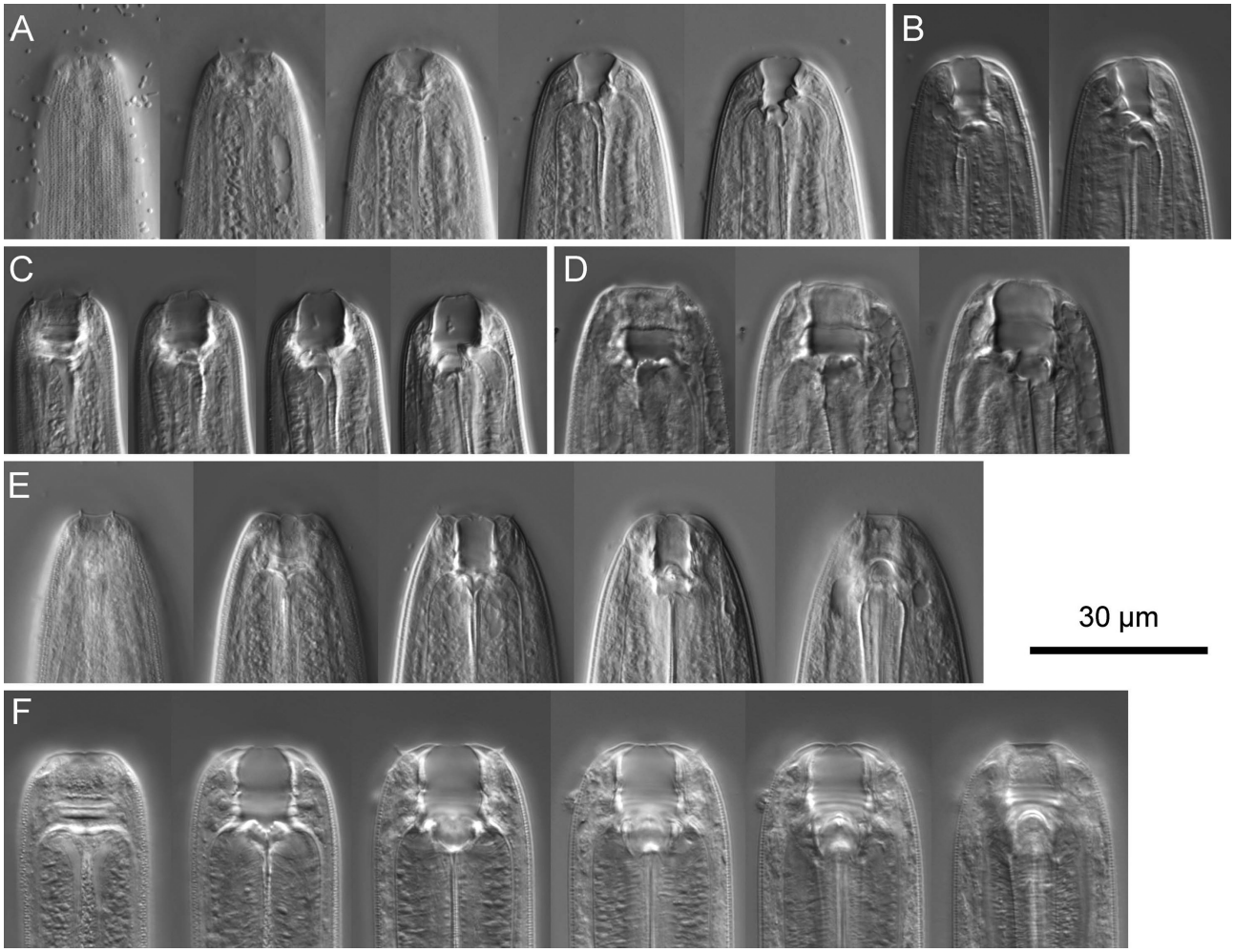


FIGURE 8. Stomal region of *Pristionchus paranudus* **n. sp.** A: Left lateral view of stenostomatous form in five different focal planes. B: Right lateral view of stenostomatous form in five different focal planes. C: Left lateral view of eurystomatous form in four different focal planes. D: Right lateral view of eurystomotous form in three different focal planes. E: Ventral view of stenostomatous form in five different focal planes. F: Ventral view of eurystomatous form in six different focal planes.

Eurystomatous form. Cheilostom divided into six per- and interradial plates. Anterior end of each plate rounded and elongated to project from stomal opening, forming a small flap. Tip of each cheilostomal plate sometimes split into two small flaps. Each cheilostomal plate inclined inwardly, *i.e.*, whole stoma appears to narrow anteriorly. Gymnostom with thick cuticle, forming short, ring-like tube with more heavily sclerotized wall in the posterior; anterior end of gymnostom internally overlapping posterior end of cheilostomatal plates. Pro-mesostegostom form-

ing a weakly cuticularized ring internally overlapping with gymnostoma to connect gymnostom and metastegostom with dorsal side relatively thick and conspicuous, ventral side thin and inconspicuous; metastegostom bearing large claw-like or triangular dorsal tooth; two right subventral serrated plates; three left subventral serrated plates. Separation between serrated plates often inconspicuous. Telostegostom forming weakly sclerotized cup-like cavity connecting stoma and pharynx.

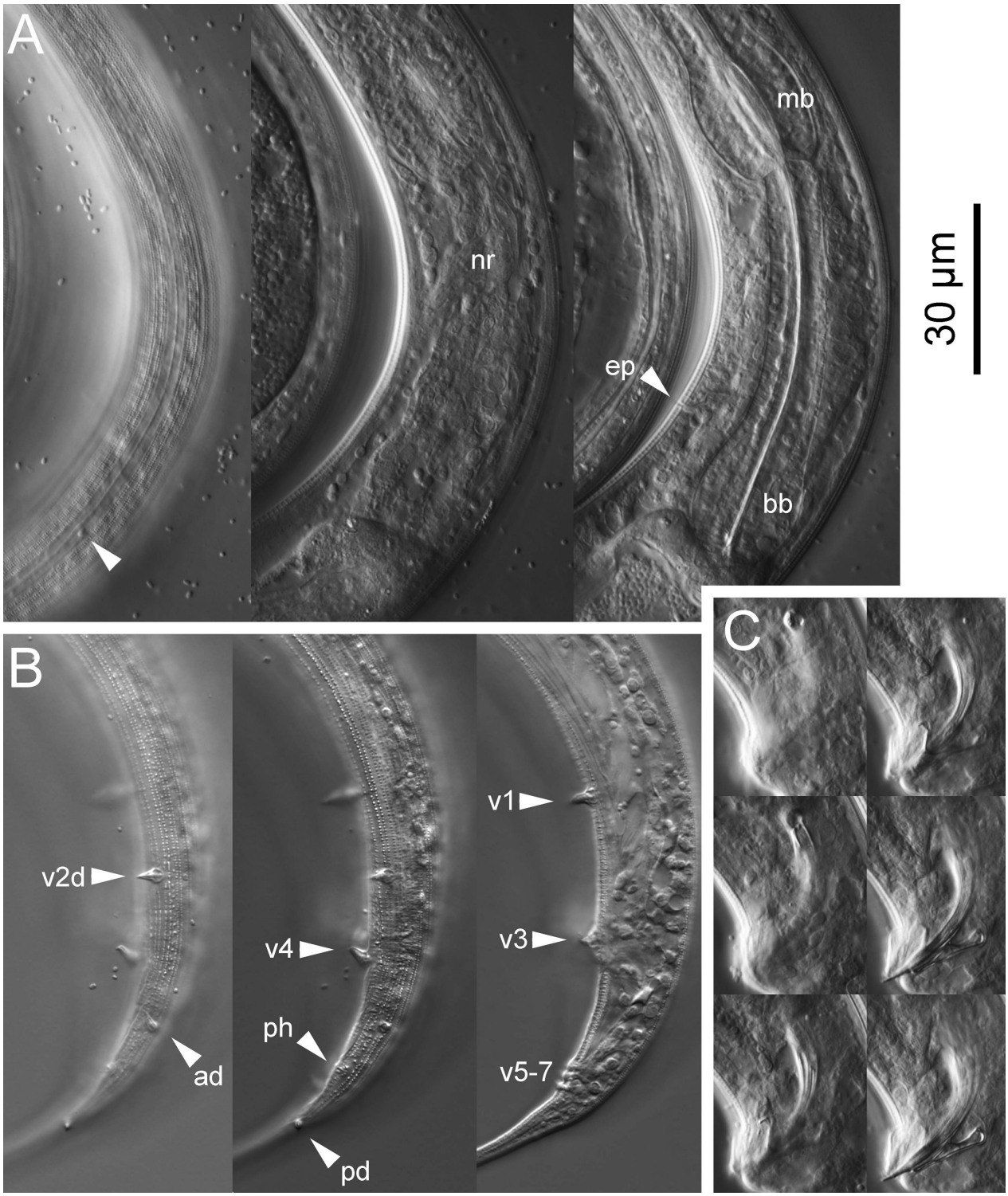


FIGURE 9. *Pristionchus paranudus* **n. sp.** A: posterior pharynx region in three different focal planes in left lateral view where deirid (arrowhead), nerve ring (nr) secretory-excretory pore (ep), median bulb (mb) and basal bulb (bb) are indicated. B: Left lateral view of male tail in three different focal planes, where genital papillae are labelled according to the terminology by Sudhaus & Fürst von Lieven (2003), and laterally located papillae are indicated with "d". C: Left lateral view of spicule and gubernaculum in six different focal planes.

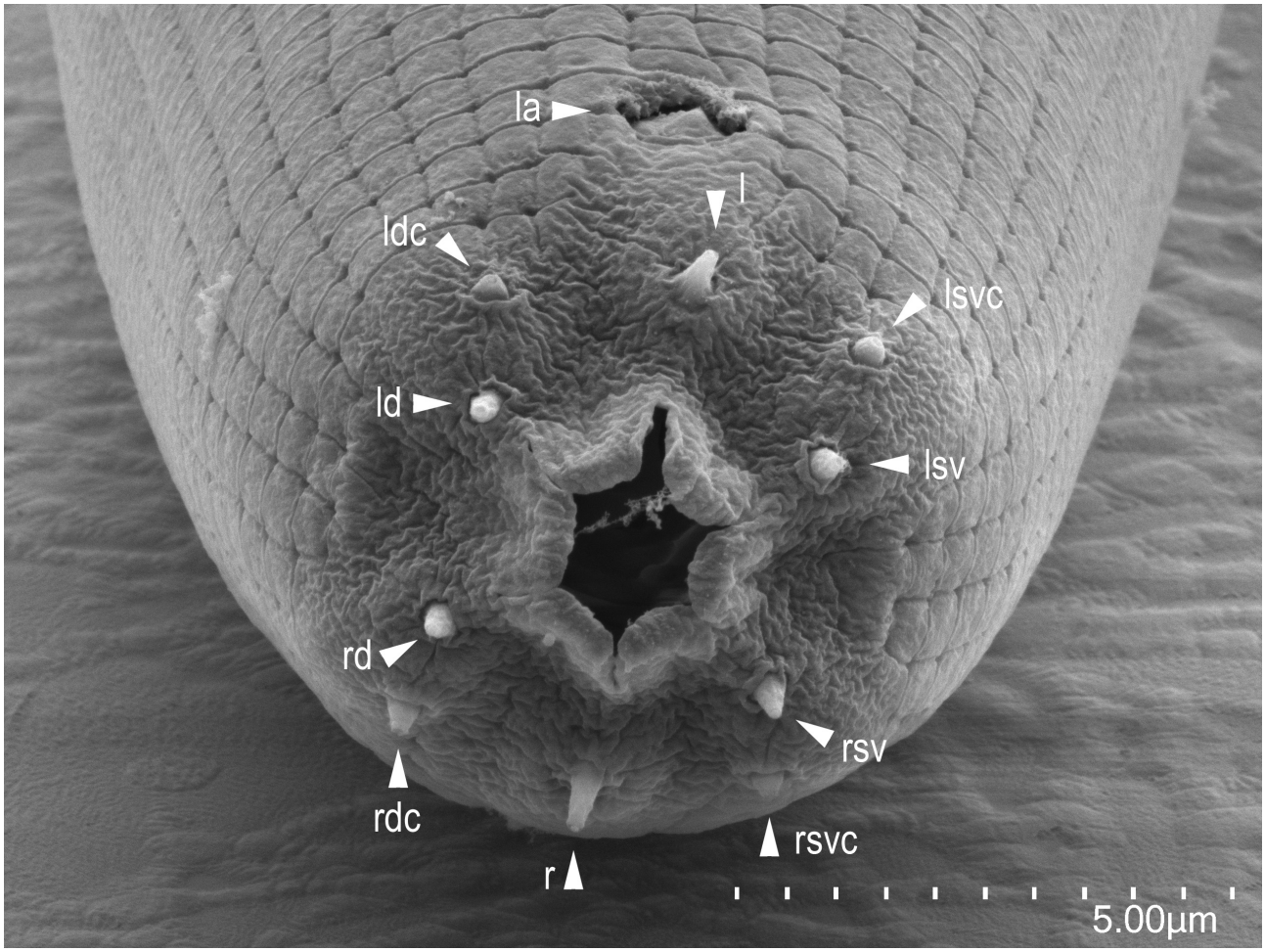


FIGURE 10. Scanning electron photomicrograph of *Pristionchus paranudus* **n. sp.** stenostomatous male *en face* view where left amphid (la), left and right subventral (lsv and rsv), left and right lateral (l and r) and left and right dorsal (ld and rd) labial sensilla, and left and right subventral (lsvc and rsvc) and dorsal (ldc and rdc) cephalic sensilla are indicated.

Male. Whole body ventrally arcuate, strongly ventrally curved at tail region when killed by heat. Testis single, ventrally located, anterior part reflexed to right or left side; spermatogonia arranged in three to five rows in reflexed part, well-developed spermatocytes arranged as three to four rows in anterior two-thirds of main branch, mature amoeboid spermatids arranged in multiple rows in proximal part of gonad. *Vas deferens* not clearly separated from other parts of gonad. Posterior end of *vas deferens* and rectum fused to form a cloacal tube. Spicules paired, separate; spicules smoothly curved in ventral view, adjacent to each other for distal third of their length, each smoothly tapering to pointed distal end; spicule in lateral view smoothly ventrally arcuate, giving spicule about 100° curvature, oval manubrium at anterior end, lamina/calomus complex expanded slightly (*ca* 1/4 of blade length) posterior to manubrium, then smoothly tapering to pointed distal end. Gubernaculum conspicuous, about one-third of spicule length, broad anteriorly such that dorsal wall is slightly recurved with dorsal and ventral walls separate at 50 to 60° angle at posterior end; dorsal side of gubernaculum possessing single, membranous, anteriorly directed process and lateral pair of more sclerotized, anteriorly and obliquely ventrally directed processes. In lateral view, anterior half of gubernaculum with two successive curves separated by anteriorly and obliquely ventrally directed process, with anterior terminal curvature highly concave and almost closed, with deep posterior curvature being one-third of gubernaculum length; posterior half forming tube-like process enveloping spicules. Cloacal opening (co) slit-like in ventral view; one small, ventral, single genital papilla (vs) on anterior cloacal lip. All nine paired genital papillae relatively long, bristle-like. Tail conoid with long spike occupying more than 2/3 of tail length, possessing filiform terminus. Paired papillae and phasmid arranged as <v1, v2d, v3, co, v4, ad, ph, (v5, v6, v7), pd>, where v1 located about 1.5 CBD anterior to co; v2d midway between v1 and co; v3 adcloacal; v4 at less than 1/10 CBD posterior to co, *i.e.*, v3, co and v4 situated close to each other; ad less than 1 CBD posterior to co; ph at 1 CBD posterior to co;

v5-v7 forming triplet, but clearly separated from each other, about 1 CBD posterior to ad, *i.e.*, less than 1/10 CBD posterior to ph; pd about 1/4 CBD posterior to v7. v1, v4 and ph subventral, v2d and ad lateral, v3, v5–7 ventral and pd subdorsal in male tail. Bursa or bursal flap absent.

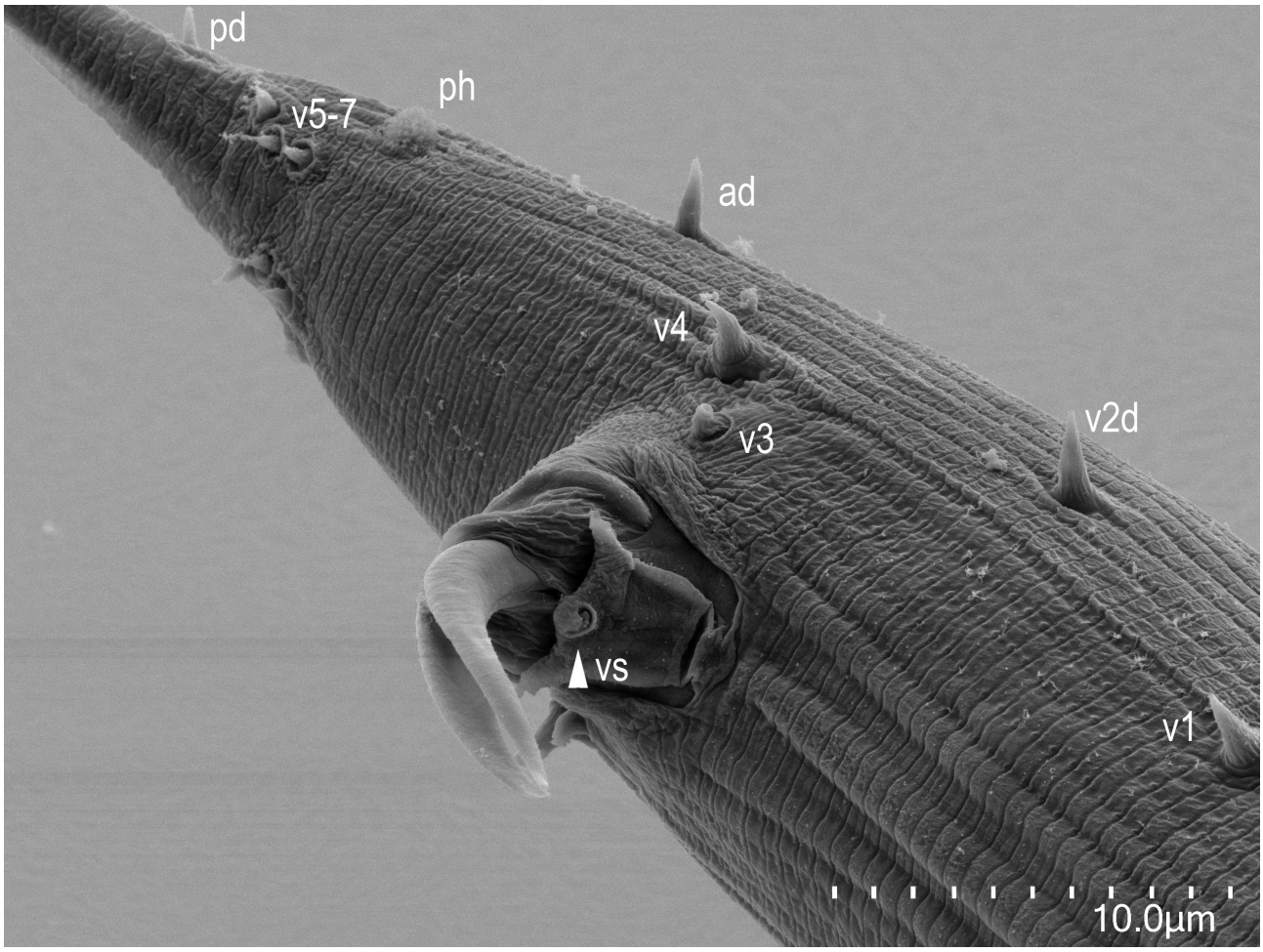


FIGURE 11. Scanning electron photomicrograph of *Pristionchus paranudus* **n. sp.** male tail with everted spicula with the position of genital papillae according to the terminology of Sudhaus and Fürst von Lieven (2003).

Female. Body relaxed or weakly ventrally arcuate when killed by heat. Gonad didelphic, amphidelphic; each gonadal system arranged from vulva/vagina as uterus, oviduct, and ovary; anterior gonad right of intestine, with uterus and oviduct extending ventrally and anteriorly on right of intestine and with totally reflexed (=antidromous reflexion) ovary extending dorsally on left of intestine; oocytes mostly arranged in three to four or more rows in distal two-thirds of ovary and in double or single row in rest of ovary, distal tips of each ovary reaching oviduct of opposite gonad branch; anterior end of oviduct (=junction tissue between ovary and oviduct) consists of rounded cells; anterior part of oviduct consists of rounded cells, forming a simple tube; middle part of oviduct serving as spermatheca, consists of roundish and relatively large cells. Eggs in single to multiple-cell stage or even further developed at posterior part of oviduct (=uterus), in young females being composed of squared or angular cells, long enough to contain one well-developed oocyte. *Receptaculum seminis* not observed, *i.e.*, the organ is not independent, and a part of oviduct/uterus works as the organ; vaginal glands present but obscure; vagina perpendicular to body surface, surrounded by sclerotized tissue; vulva slightly protuberant in lateral view, pore-like in ventral view; rectum about one anal body diameter (ABD) long, intestine/rectum junction surrounded by well-developed sphincter muscle. Anus in form of dome-shaped slit, posterior anal lip slightly protuberant; phasmid about 2–2.5 ABDs posterior to anus. Tail long, smoothly elongate, conoid, with filiform terminus.

Diagnosis and relationship. *Pristionchus paranudus* **n. sp.** is characterized by its cuticle structure with weak longitudinal striation and distinctive lateral field, presence of partially split cheilostomatal plates and right and left subventral metastegostomatal serrated plates of eurystomatous form, long spike of male tail and the arrangement of male genital papillae, *i.e.*, <v1, v2d, v3, co, v4, ad, ph, (v5, v6, v7), pd>. The new species shares partially split

cheilostomatal plates with *P. paulseni* and some of the *triformis* group species (*P. hoplostomus* Ragsdale, Kanzaki, Röseler, Herrmann, Sommer, 2013; *P. fukushimae* Ragsdale, Kanzaki, Röseler, Herrmann, Sommer, 2013 and *P. yamagatae* Herrmann, Kanzaki, Weiler, Yoshida, Rödelsperger & Sommer, 2019), and its surface cuticle structure (moderate in thickness with weak striations and distinctive lateral field) and the structure and arrangement of male genital papillae (long bristle-like and ventrally located v3, co and subventral v4 are very close to each other) with *P. nudus* **n. sp.** However, the new species is distinguished from *P. paulseni* and *triformis* group by its left subventral stegostomal plate of the eurystomatous form, cuticle structure and the structure and the arrangement of genital papillae, *i.e.*, *P. paulseni* and *triformis* group species have claw-like stegostomatous tooth on the left subventral sector of the eurystomatous form, thick cuticle with distinctive longitudinal striations and indistinctive lateral field, and v3 and v4 are clearly separated. The new species is also distinguished from *P. nudus* **n. sp.** by its median bulb, which is oval *vs.* somewhat rectangular in lateral view, the position of the nerve ring, middle *vs.* anterior end of isthmus, labial sensilla, papilliform *vs.* bristle-like, stomal morphology in stenostomatous form, with sclerotized chilostomatal plates, flint-shaped dorsal tooth with a metastegostomal mound on the dorsal side *vs.* thick and seemingly soft cheilostomatal plates, triangular dorsal tooth and two layers of dorsal metastegostomatal mounds, arrangement of genital papillae and phasmid, <v1, v2d, v3, co, v4, ad, ph, (v5, v6, v7), pd> *vs.* <v1, v2d, v3, co, v4, ad, (v5, v6, v7), ph, pd>, and the male tail shape with *vs.* without spike. Further, the new species is distinguished from all other species by mating experiments and also characterized by a *ca.* 1,600-bp fragment of the SSU rRNA gene (GenBank accession number MW017222), the sequence of which is distinct from that of all other *Pristionchus* species.

Type host and locality. Isolated from rotting water hyacinth bulbs (*Eichhornia*) on terraces below Dayutang village, Yuanyang, Yunnan province, PRC.

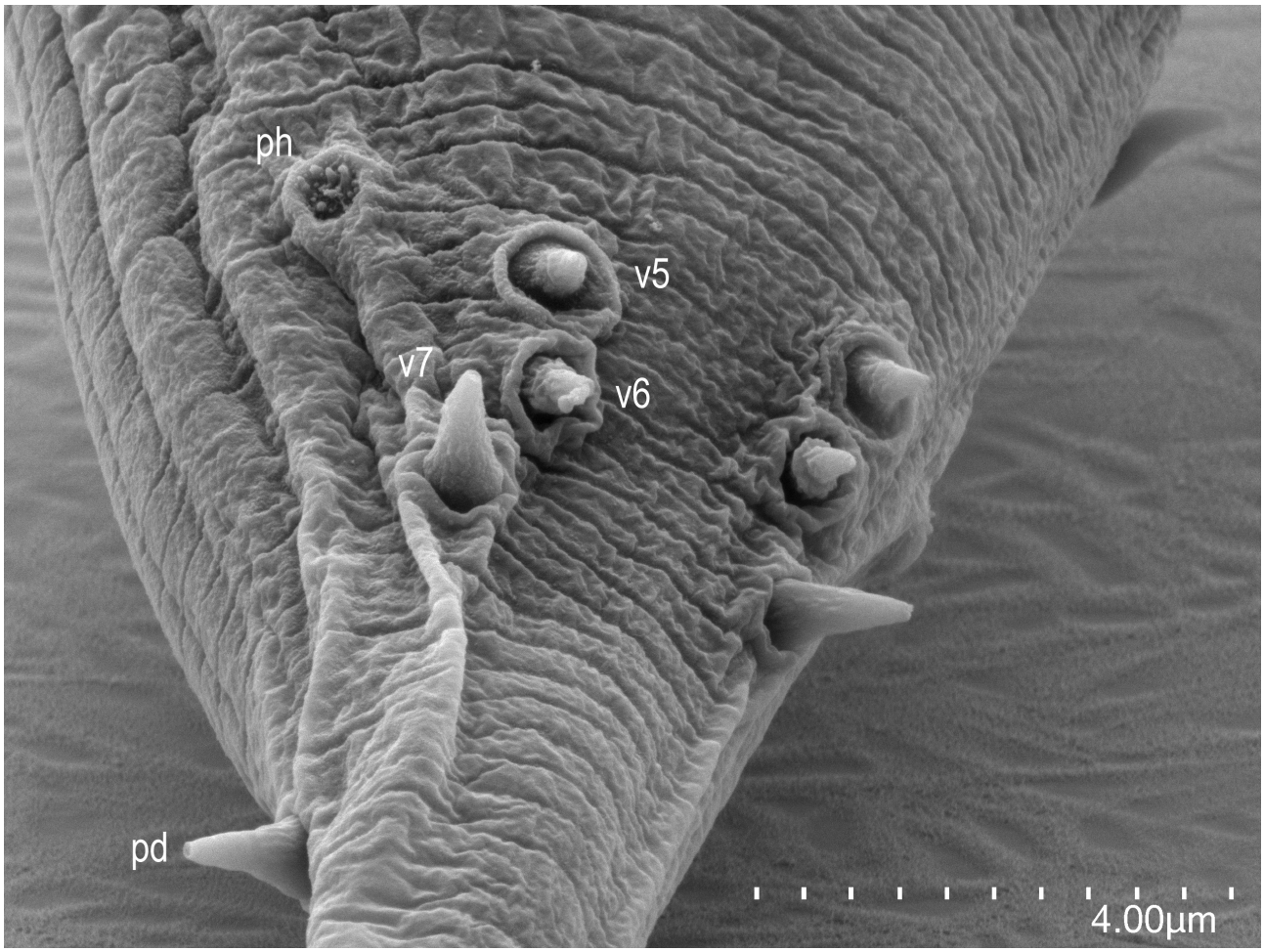


FIGURE 12. Scanning electron photomicrograph of *Pristionchus paranudus* **n. sp.** posterior part of male tail and papillae with phasmid (ph) and paired papillae (v5–7, pd) indicated.

TABLE 1.

	P. nudus (RS6026)		*P. paranudus* (RS5988)	
character	stenostomatous male	stenostomatous female	stenostomatous male	stenostomatous female
n	5	5	5	5
L	911 ± 128.7 (704–1029)	1174 ± 166.9 (911–1322)	761 ± 79.6 (660–864)	990 ± 89.7 (891–1106)
L’	748 ± 117.7 (572–872)	956 ± 101.1 (811–1079)	614 ± 60.3 (553–696)	733 ± 72.2 (674–825)
a	16 ± 1.1 (15–18)	14 ± 1.2 (12–15)	18 ± 3.0 (15–22)	18 ± 1.5 (16–20)
b	5.8 ± 0.6 (4.8–6.4)	6.9 ± 0.8 (5.7–7.9)	5.3 ± 0.4 (4.7–5.7)	6.6 ± 0.2 (6.3–6.9)
c	5.6 ± 0.7 (4.6–6.5)	5.9 ± 1.9 (4.4–9.1)	5.3 ± 0.6 (4.6–6.2)	3.9 ± 0.2 (3.6–4.1)
c’	3.7 ± 0.4 (3.3–4.1)	4.6 ± 1.1 (2.9–5.6)	4.8 ± 1.1 (3.3–6.1)	7.3 ± 0.7 (6.2–7.8)
ant. stoma length (cheilo- + gymnostom)	9.3 ± 1.4 (7.4–10.4)	11.1 ± 1.0 (10.2–12.7)	6.8 ± 0.9 (5.9–8.0)	8.2 ± 0.9 (6.9–9.1)
total stoma length	13.3 ± 1.2 (12.2–15.0)	15.5 ± 0.9 (14.5–17.0)	10.5 ± 0.6 (9.8–11.2)	13.2 ± 1.0 (11.6–14.4)
stoma width	7.3 ± 0.8 (6.7–8.6)	7.9 ± 0.5 (7.3–8.7)	5.0 ± 0.9 (4.0–5.9)	6.7 ± 0.9 (5.5–7.5)
ant. pharynx length (pro- + metacorpus)	93 ± 5.1 (87–100)	106 ± 5.2 (100–113)	78 ± 3.1 (73–82)	86 ± 10.0 (78–103)
post. pharynx length (isthmus + basal bulb)	49 ± 1.5 (47–50)	52 ± 4.7 (48–59)	56 ± 5.4 (49–62)	56 ± 2.2 (53–59)
total length pharynx	142 ± 5.4 (134–149)	158 ± 5.2 (150–163)	133 ± 8.0 (122–142)	142 ± 11.6 (133–162)
ant./total pharynx %	66 ± 1.3 (64–68)	67 ± 2.7 (64–70)	58 ± 1.8 (56–60)	60 ± 2.2 (58–63)
median bulb diameter	21 ± 1.8 (20–24)	28 ± 3.1 (25–31)	21 ± 2.4 (19–24)	27 ± 3.0 (22–29)
terminal bulb diameter	17 ± 1.9 (16–20)	24 ± 1.8 (22–27)	17 ± 1.7 (15–19)	22 ± 1.9 (18–23)
neck length (incl. stoma)	155 ± 5.9 (146–162)	169 ± 5.5 (161–174)	144 ± 8.4 (132–152)	151 ± 11.6 (141–169)
excretion pore	150 ± 12.6 (131–162)	168 ± 14.7 (154–186)	123 ± 8.4 (115–132)	133 ± 11.3 (119–145)
nerve ring	119 ± 9.3 (105–131)	133 ± 7.2 (122–141)	94 ± 6.0 (85–101)	105 ± 6.3 (98–112)
testis length	513 ± 136.9 (325–656)	-	376 ± 18.2 (352–402)	-
anterior gonad length	-	273 ± 40.8 (223–314)	-	186 ± 29.1 (161–234)
posterior gonad length	-	239 ± 21.7 (211–260)	-	164 ± 34.1 (136–207)
ant. end to vulva distance	-	566 ± 63.3 (465–633)	-	436 ± 38.5 (396–476)
vulva to anus distance	-	375 ± 64.5 (281–445)	-	303 ± 38.9 (268–357)
T or V	56 ± 7.9 (46–64)	48 ± 2.0 (46–51)	50 ± 3.4 (46–53)	44 ± 1.0 (43–45)
max. body diameter	56 ± 5.3 (48–63)	85 ± 12.6 (64–95)	43 ± 3.6 (39–48)	57 ± 6.1 (48–63)
cloacal or anal body diameter	44 ± 7.9 (37–56)	46 ± 9.0 (35–60)	31 ± 2.9 (28–34)	35 ± 4.4 (31–43)
tail length	162 ± 23.8 (132–193)	218 ± 72.7 (100–291)	147 ± 24.8 (107–168)	256 ± 23.4 (218–281)
spicule length (curve)	36 ± 3.1 (31–40)	-	39 ± 0.8 (38–40)	-
spicule length (chord)	31 ± 3.6 (25–34)	-	34 ± 1.2 (33–35)	-
gubernaculum length	17 ± 2.7 (13–20)	-	15 ± 1.1 (13–16)	-

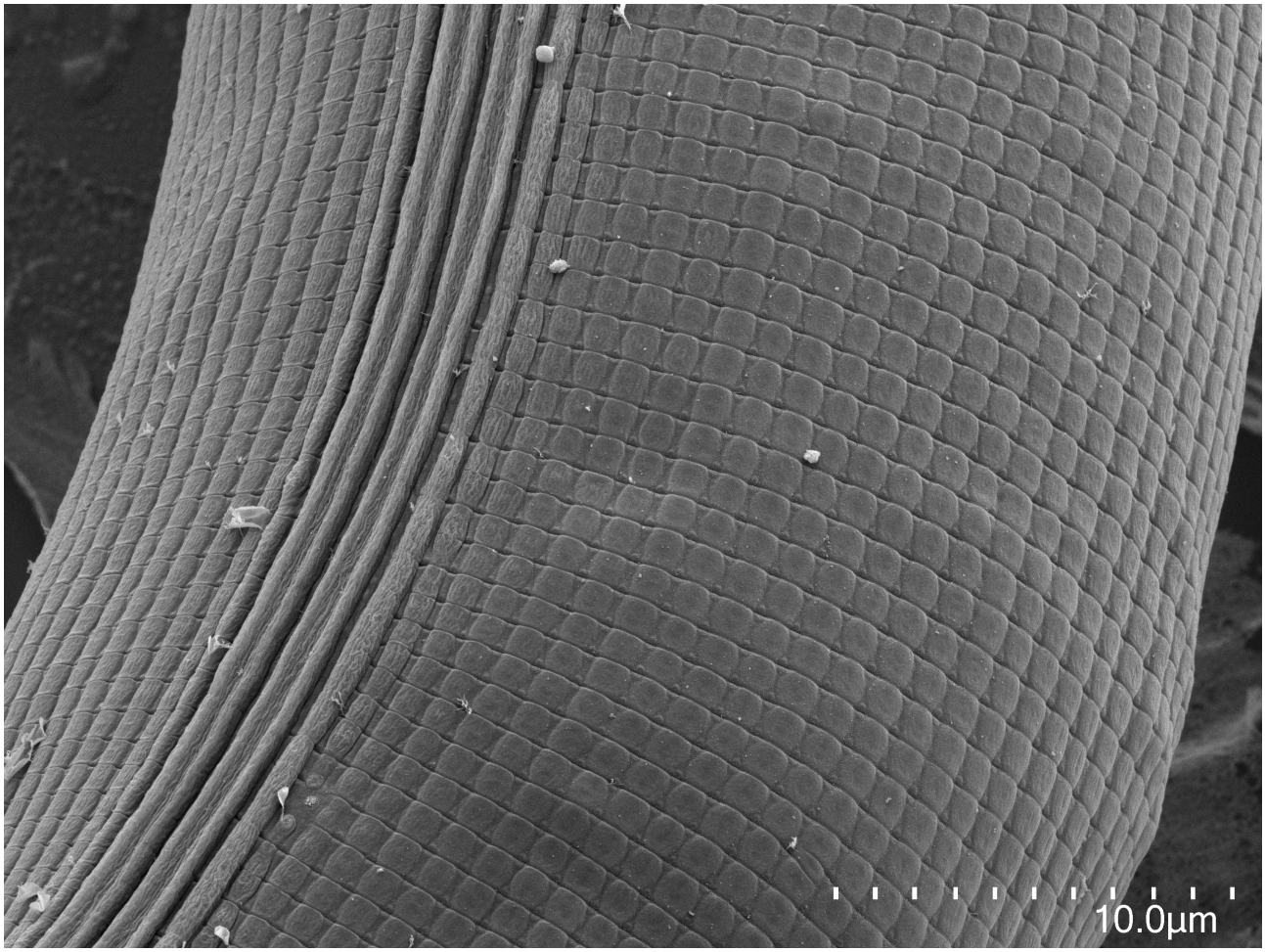


FIGURE 13. Scanning electron photomicrograph of *Pristionchus paranudus* **n. sp.** surface structure with lateral line.

Type material and type strain. Type strain RS5988 frozen at the nematode collection of the MPI Tübingen and available as living culture upon request. Voucher specimen sent to the following museums: Holotype male, Paratype male and female: Museum für Naturkunde Karlsruhe, Germany; Paratype male and female: Swedish Natural History Museum, Stockholm, Sweden; Paratype male and female: University of California in Riverside Nematode Collection (UCRNC), Riverside, CA, USA.

***Pristionchus chinensis* n. sp.**
urn:lsid:zoobank.org:act:2ECAEE4D-2089-48F7-83E0-86AA7F9591A5

Etymology. The species name is derived from China, the country of origin.

Measurements. See Table 2

Adult. Androdioecious species. Body cylindrical, stout, *i.e.*, body length-maximum body diam. ratio (a value) is usually ranging from 11–17, depending on culture condition; cuticle thick, with fine annulations, conspicuous longitudinal striations. Lateral field consisting of two lines, weakly separated from body striations by presence of deirid. Head without apparent lips, six mound-like anteriorly directed expansions, one on each sector. Six labial sensilla, one on each lip sector (on the mound); four small papilliform cephalic papillae in males on right and left subventral and right and left dorsal sectors; amphidial apertures located on lateral sector, slightly dorsally shifted, at level of margin of cheilo- and gymnostom. Stomal dimorphism present, and details are described below. Dorsal pharyngeal gland clearly observed, penetrating dorsal tooth to gland opening. Anterior pharynx 1.5 times as long as posterior pharynx, muscular, composed of tube-shaped procorpus, well-developed oval-shaped metacorpus (median bulb); posterior pharynx glandular with nerve ring surrounding middle of isthmus. Pharyngo-intestinal junction

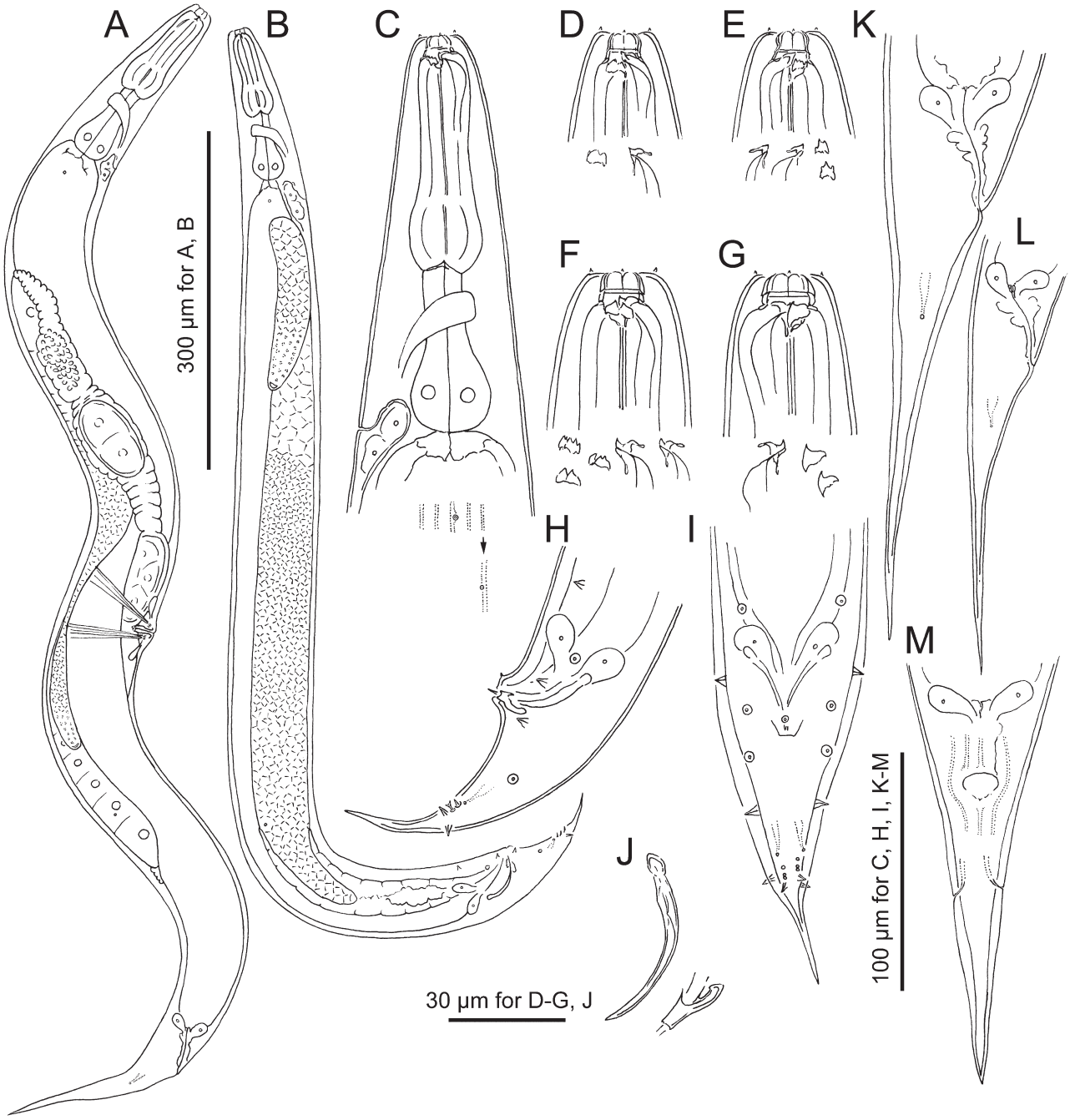


FIGURE 14. *Pristionchus chinensis* **n. sp.** A: Right lateral view of adult female. B: Right lateral view of adult male. C: Anterior part of adult female in left lateral view including the body surface structure and relative position of deirid and postdeirid. D, E: Stomal region of stenostomatous female in left (D) and right (E) lateral view. F, G: Stomal region of eurystomatous female in left (F) and right (G) lateral view. Morphological variations of teeth, ridges and denticle are separately drawn in each subfigure as left subventral ridge (left) and dorsal tooth (right) (D), dorsal tooth (left and middle) and right subventral ridge (right) (E), left subventral denticles (left) and dorsal tooth (middle and right) (F), dorsal (left) and right subventral (right) teeth (G). H, I: Male tail in left lateral (H) and ventral (I) view. J: Spicule and gubernaculum in left lateral view. K–M: Female tail of large (K) and small (L) individuals in right lateral view, and ventral view (M).

(cardia) well-developed. Intestine simple tube, not forming pre-rectum, extended posteriorly from cardia to rectum; three (two subventral and one dorsal) rectal gland cells observed at distal end of intestine (margin between intestine and rectum). Secretory-excretory pore not conspicuous, ventrally located at level of isthmus to pharyngo-intestinal junction, excretory duct extending anteriad and reflexed back to position of pore; two large secretory-excretory cells around the excretory duct. Deirid observed laterally on lateral field, located at the level around the posterior

end of basal bulb to pharyngo-intestinal junction to a half body diameter posterior to the junction, ca 0.5 body diam. posterior to secretory-excretory pore. Hemizonid not observed. Lateral glands (small pores connected to secretory cell) on lateral body surface, with positions inconsistent among individuals, numbering 5 to 8 for males and 9 to 13 for females. Postdeirid at anterior part of *vas deferens* in male and the posterior end of posterior gonad in female, on the same striation with deirid (= lateral field) or on the adjacent striation or the second dorsally neighboring striation to lateral field.

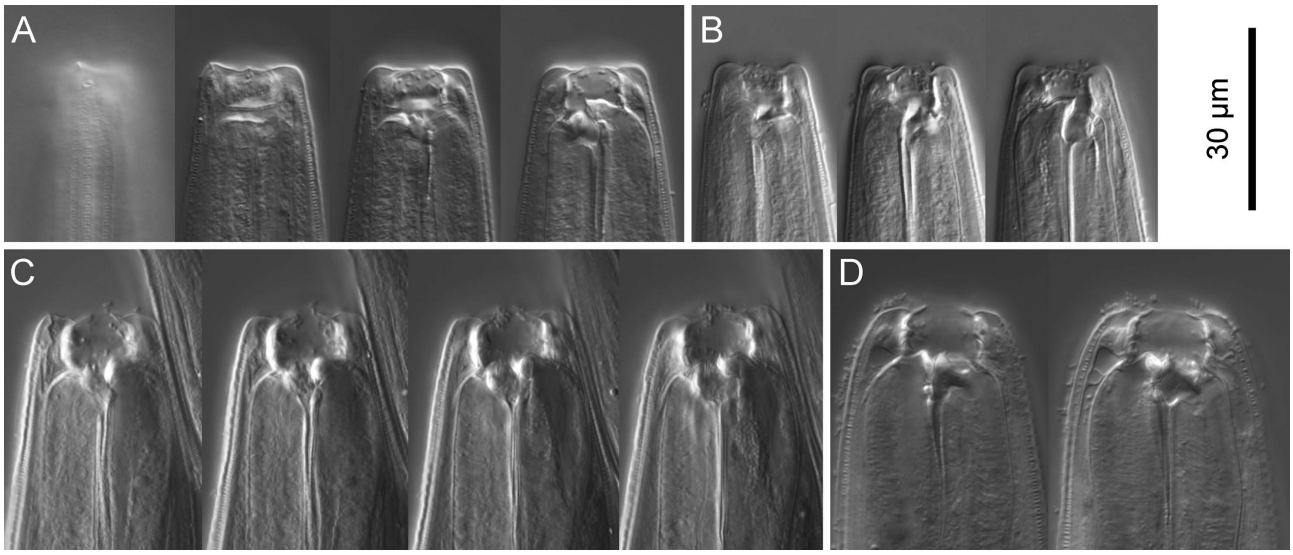


FIGURE 15. Stomal region of *Pristionchus chinensis* **n. sp.** A: Left lateral view of stenostomatous form in four different focal planes. B: Right lateral view of stenostomatous form in three different focal planes. C: Left lateral view of eurystomatous form in four different focal planes. D: Right lateral view of eurystomatous form in two different focal planes.

Stenostomatous form. Cheilostom with six per- and interradial cuticular plates. Incision between plates not always distinguished. Anterior end of each plate rounded and elongated to project from stomal opening to form small flap. Gymnostom short, cuticular ring-like anterior end overlapping cheilostom internally; dorsal gymnostomal wall with mound-like expansion probably derived from metastegostom which gives an appearance of thickened dorsal gymnostomal wall compared to ventral side, and a metastegostomal mound consisting of weakly sclerotized tissue present on the dorsal side of the inner wall. Pro-meso stegostom forming a weakly cuticularized ring internally overlapping with gymnostoma to connect gymnostom and metastegostom. Metastegostom bearing dorsal mound and conspicuous and movable triangular or flint-shaped dorsal tooth with strongly sclerotized surface giving an appearance of an inverted V-shape in light microscopy in lateral view. Left subventral ridge with three minute, rounded adventitious denticles on a plate, and the most ventral denticle is masked by the other two in the lateral view. Right subventral ridge with three bluntly pointed distal adventitious denticles. Telostegostom forming weakly sclerotized cup-like cavity connecting stoma and pharynx.

Eurystomatous form. Cheilostom divided into six distinctive per- and interradial plates. Anterior end of each plate rounded and elongated to project from stomal opening, forming a small flap. Each cheilostomal plate inclined inwardly, giving an appearance that whole stoma is narrowing anteriorly. Gymnostom with thick cuticle, forming short, ring-like tube with more heavily sclerotized wall in the posterior. Pro-mesostegostom forming a weakly cuticularized ring connecting gymnostom and metastegostom. Metastegostom bearing large claw-like dorsal tooth; claw-like right subventral tooth; and three left subventral denticles, where the tip of each denticle sometimes splits into two or three small ridges. Telostegostom forming weakly sclerotized cup-like cavity connecting stoma and pharynx.

Male. Whole body ventrally arcuate, strongly ventrally curved at tail region when killed by heat. Testis single, ventrally located, anterior part reflexed to right or left side; spermatogonia arranged in three to five rows in reflexed part, well-developed spermatocytes arranged as three to four rows in anterior two-thirds of main branch, mature amoeboid spermatids arranged in multiple rows in proximal part of gonad. *Vas deferens* not clearly separated from other parts of gonad. Posterior end of *vas deferens* and rectum fused to form a cloacal tube. Spicules paired, separate; spicules smoothly curved in ventral view, adjacent to each other for distal third of their length, each smoothly

tapering to pointed distal end; spicule in lateral view smoothly ventrally arcuate, giving spicule about 100° curvature, oval manubrium at anterior end, lamina/calomus complex smoothly tapering to pointed distal end. Gubernaculum conspicuous, about one-third of spicule length, broad anteriorly such that dorsal wall is slightly recurved with dorsal and ventral walls separate at 50-60° angle at posterior end; dorsal side of gubernaculum possessing single, membranous, anteriorly directed process and lateral pair of more sclerotized, anteriorly and obliquely ventrally directed processes. In lateral view, anterior half of gubernaculum with two successive curves separated by anteriorly and obliquely ventrally directed process, with anterior terminal curvature highly concave and almost closed, with deep posterior curvature being one-third of gubernaculum length; posterior half forming tube-like process enveloping spicules. Cloacal opening (co) slit-like in ventral view; one small, ventral, single genital papilla (vs) on anterior cloacal lip. All nine paired genital papillae papilliform. Tail conoid with short spike. The paired papillae and the phasmid are arranged as <v1, v2d, v3, co, v4, ad, ph, (v5, v6, v7, pd)>, where v1 located about a little more than 1 CBD anterior to co; v2d midway between v1 and co; v3 1/4–1/5 CBD anterior to co; v4 at 1/4–1/5 CBD posterior to co; ad about 1 CBD posterior to co; ph at midway between ad and the root of tail spike; v5–v7 forming triplet, between ph and the root of tail spike; and pd at same level as v6 or v7. v1, v3, v4 and ph subventral, v2d and ad lateral, v5–7 ventral, pd subdorsal in male tail. Bursa or bursal flap absent.

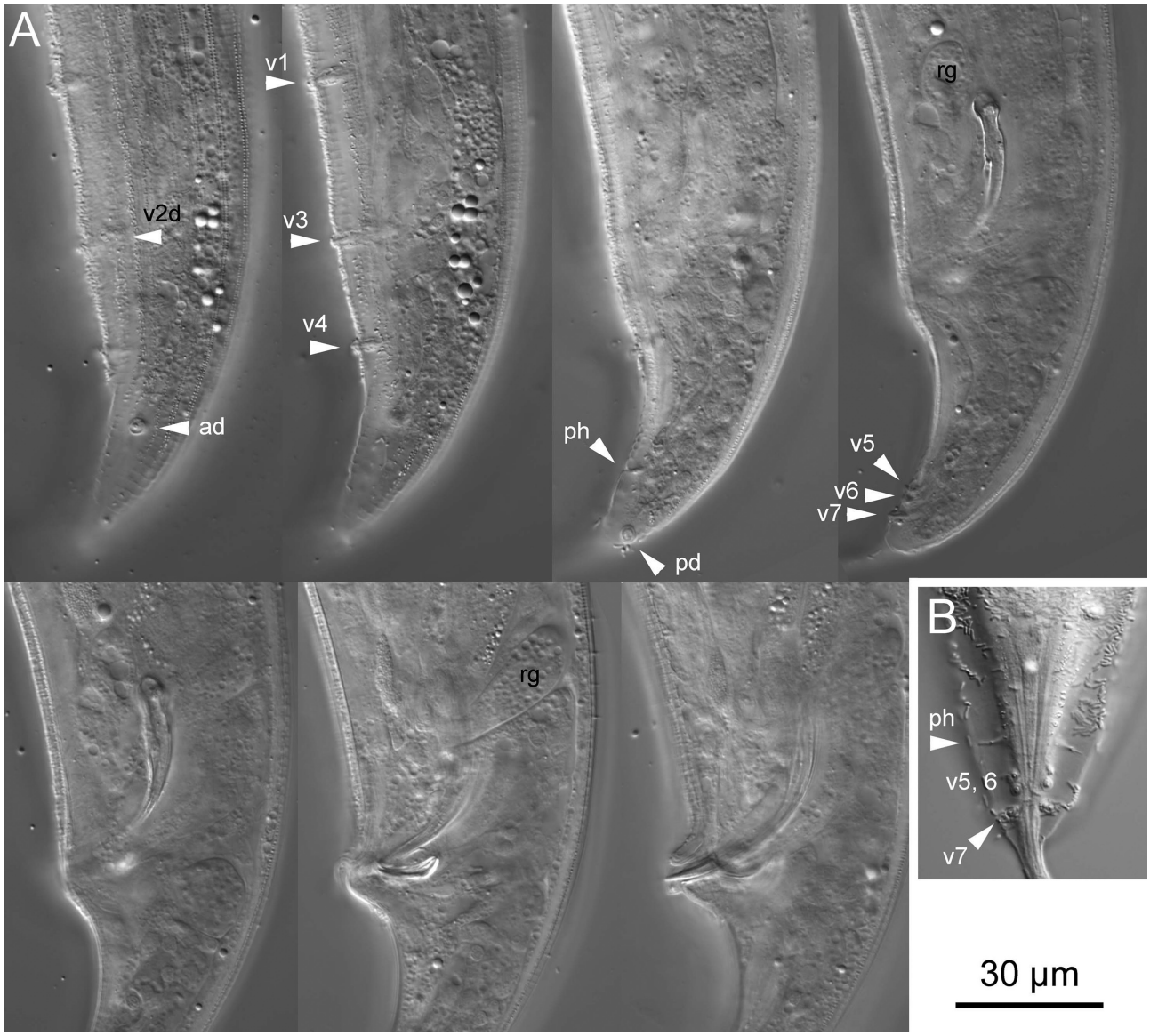


FIGURE 16. Male tail characters of *Pristionchus chinensis* **n. sp.** A: Left lateral view of male tail in seven different focal planes, where genital papillae are labelled according to the terminology by Sudhaus & Fürst von Lieven (2003), and laterally located papillae are indicated with "d", and rectal gland (rg) is labelled. B: Ventral view of the posterior part.

TABLE 2.

	P. chinensis (RS6023)	
character	stenostomatous male	stenostomatous hermaphrodite
n	5	5
L	907 ± 175.8 (703–1123)	1090 ± 122.3 (986–1289)
L’	799 ± 168.8 (604–1008)	929 ± 120.6 (829–1129)
a	14 ± 2.7 (11–17)	13 ± 1.2 (11–14)
b	5.8 ± 0.7 (5.0–6.4)	6.5 ± 0.8 (5.9–7.9)
c	8.4 ± 1.1 (7.1–9.8)	6.8 ± 0.8 (6.2–8.0)
c’	2.4 ± 0.4 (1.9–3.0)	4.0 ± 0.5 (3.4–4.5)
ant. stoma length (cheilo- + gymnostom)	6.7 ± 0.8 (5.8–7.4)	7.9 ± 0.8 (6.9–8.9)
total stoma length	10.9 ± 1.7 (8.3–12.3)	13.5 ± 0.6 (12.6–14.2)
stoma width	7.7 ± 0.8 (6.5–8.5)	8.4 ± 0.4 (8.0–9.0)
ant. pharynx length (pro- + metacorpus)	78 ± 4.6 (75–86)	93 ± 2.4 (91–97)
post. pharynx length (isthmus + basal bulb)	67 ± 9.1 (56–77)	67 ± 3.0 (64–71)
total length pharynx	145 ± 12.1 (133–163)	160 ± 3.6 (155–163)
ant./total pharynx %	54 ± 3.0 (50–58)	58 ± 1.3 (56–60)
median bulb diameter	24 ± 1.6 (23–27)	29 ± 0.8 (29–30)
terminal bulb diameter	25 ± 2.3 (22–27)	28 ± 1.6 (26–30)
neck length (incl. stoma)	156 ± 13.4 (141–175)	168 ± 3.6 (164–171)
excretion pore	137 ± 18.7 (113–156)	157 ± 10.2 (148–173)
nerve ring	105 ± 10.2 (93–119)	120 ± 2.6 (118–124)
testis length	551 ± 141.5 (404–765)	-
anterior gonad length	-	245 ± 32.7 (212–291)
posterior gonad length	-	296 ± 32.1 (251–337)
ant. end to vulva distance	-	533 ± 36.1 (498–586)
vulva to anus distance	-	397 ± 89.5 (327–549)
T or V	60 ± 4.5 (57–68)	49 ± 2.1 (45–51)
max. body diameter	68 ± 9.7 (52–79)	88 ± 18.9 (74–121)
cloacal or anal body diameter	46 ± 5.9 (38–51)	41 ± 4.7 (35–47)
tail length	107 ± 7.1 (99–115)	161 ± 8.8 (155–176)
spicule length (curve)	50 ± 5.2 (43–55)	-
spicule length (chord)	43 ± 4.1 (36–47)	-
gubernaculum length	16 ± 1.9 (14–19)	-

Hermaphrodites. Body relaxed or weakly ventrally arcuate when killed by heat. Gonad didelphic, amphidelphic; each gonadal system arranged from vulva/vagina as uterus, oviduct, and ovary; anterior gonad right of intestine, with uterus and oviduct extending ventrally and anteriorly on right of intestine and with totally reflexed (=antidromous reflexion) ovary extending dorsally on left of intestine; oocytes mostly arranged in three to four or more rows in distal two-thirds of ovary and in double or single row in rest of ovary, distal tips of each ovary reaching oviduct of opposite gonad branch; anterior end of oviduct (=junction tissue between ovary and oviduct) consists of rounded cells; anterior part of oviduct consists of rounded cells, forming a simple tube; middle part of oviduct serving as spermatheca, consists of roundish and relatively large cells. Eggs in single to multiple-cell stage or even further developed at posterior part of oviduct (=uterus), in young females being composed of squared or angular cells, long enough to contain one well-developed oocyte. *Receptaculum seminis* not observed, *i.e.*, the organ is not independent, and a part of oviduct/uterus works as the organ; vaginal glands present but obscure; vagina perpendicular to body surface, surrounded by sclerotized tissue; vulva slightly protuberant in lateral view, pore-like in ventral view; rectum about one anal body diameter (ABD) long, intestine/rectum junction surrounded by well-developed sphincter muscle. Anus in form of dome-shaped slit, posterior anal lip slightly protuberant; phasmid about 1.5 ABDs posterior to anus. Tail conoid with or without elongated posterior half; and tail terminus sharply pointed.

Diagnosis and relationships. *Pristionchus chinensis* **n. sp.** is characterized by having the right subventral stegostomal ridge of stenostomatous form, with three bluntly pointed minute denticles, its relatively short tail in males and hermaphrodites, *i.e.* conoid with short (less than 1 CBD) spike in male and conoid with elongated posterior half in hermaphrodite, arrangement of male genital papillae, <v1, v2d, v3, co, v4, ad, ph, (v5, v6, v7, pd)>, and its characteristic reproductive mode, androdioecy.

Pristionchus chinensis **n. sp.** is typologically and phylogenetically close to *P. fissidentatus*, *i.e.*, these two species share the stomatal characters of stenostomatous form, relatively short male tail and the androdioecy, and further, both species belong to the basal clade of the genus. However, the new species can be distinguished from *P. fissidentatus* by its stomal structure of eurystomatous form, although these two species share the claw-like dorsal tooth and left subventral denticles with many minute cusps, the right subventral sector possessing a large claw-like tooth *vs.* ridge with three tips. Further, *P. chinensis* **n. sp.** can be distinguished from *P. fissidentatus* by its male tail characters. Although the characters of spontaneous males are sometimes inconsistent within species, the difference in the arrangement of genital papillae, <v1, v2d, v3, co, v4, ad, ph, (v5, v6, v7, pd)> *vs.* <v1, (co, v2), v3d, v4, ad, Ph, (v5, v6, v7), pd> and the length of tail spike less than *vs.* more than 1 CBD is consistent for these two species.

Further, the new species is distinguished from all other species by mating experiments and also characterized by a *ca.* 1,600-bp fragment of the SSU rRNA gene (GenBank accession number MW017217), the sequence of which is distinct from that of all other *Pristionchus* species.

Type host and locality. From *Mimela* sp. (Coleoptera: Rutelidae) collected at "Skytree" near Bubeng Fieldstation CAS, Yunnan province, PRC.

Type material and type strain. Type strain RS6023 frozen at the nematode collection of the MPI Tübingen and available as living culture upon request. Voucher specimens sent to the following museums: Holotype male, Paratype hermaphrodite: Museum für Naturkunde Karlsruhe, Germany; Paratype hermaphrodite: Swedish Natural History Museum, Stockholm, Sweden; Paratype hermaphrodite: University of California in Riverside Nematode Collection (UCRNC), Riverside, CA, USA.

Species description of four *Pristionchus triformis*-group species

Common characters. The basal species groups described above are highly variable in the basic typological characters, *e.g.*, the presence/absence of stomal dimorphism, cuticle structure, pharyngeal structure and position of the nerve ring. In comparison, species in the *triformis*-group are uniform, and they share common characters as described below, based on re-observation of four previously described species (*P. triformis* Ragsdale, Kanzaki, Röseler, Herrmann, Sommer, 2013, *P. hoplostomus*, *P. fukusimae* and *P. yamagatae*) and three new species. Several diagnostic characters are summarized in Table 3.

Adult. Body cylindrical, stout, *i.e.*, body length-maximum body diam. ratio (a value) is usually ranging from 10–18, depending on culture condition; cuticle with fine annulation, clear longitudinal striations and indistinct lateral field consisting of two lines, weakly separated from the body striations by presence of deirid, and absence of clear annulations. Head without apparent lips, six mound-like anteriorly directed expansions, one on each sector.

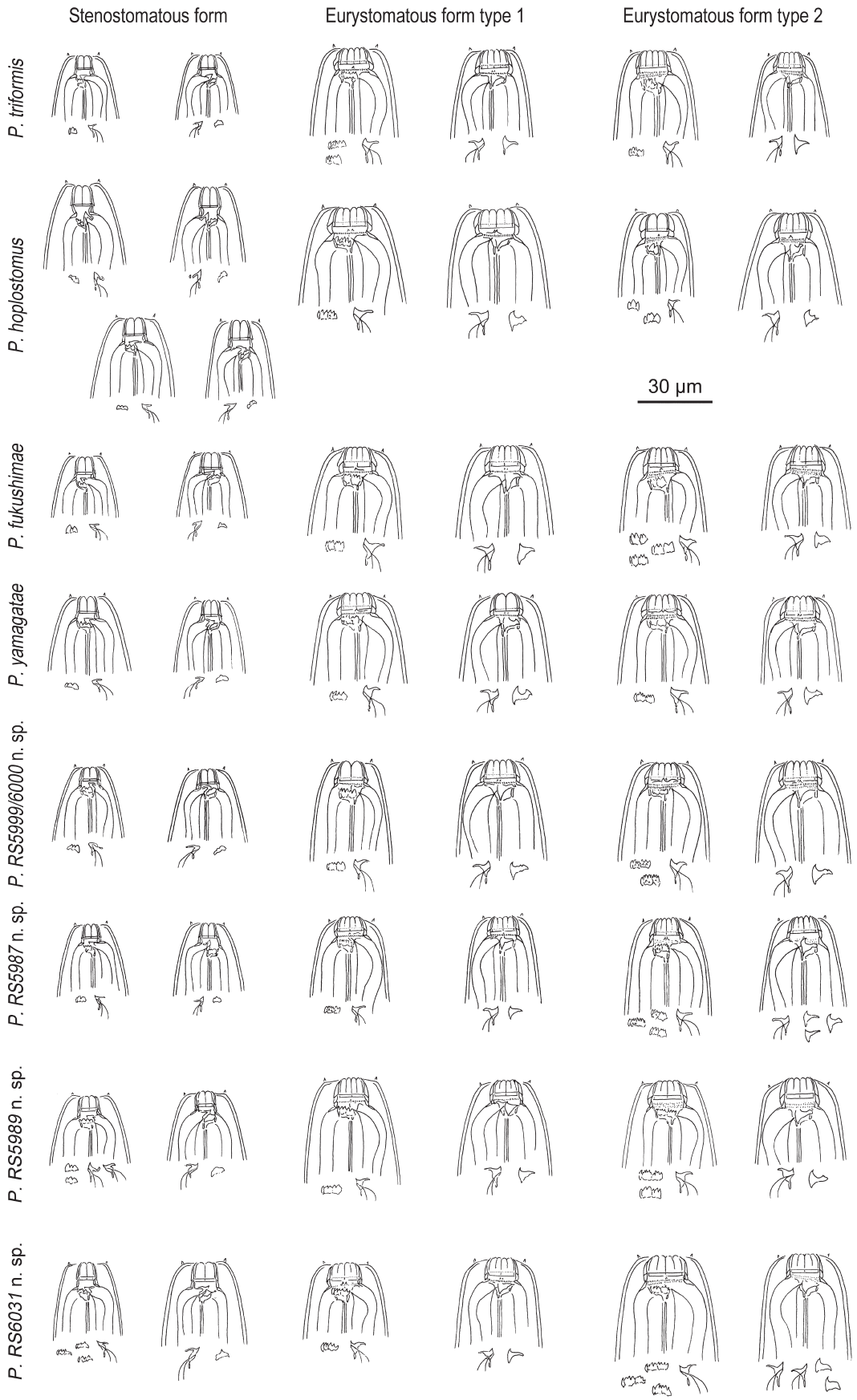


FIGURE 17. Stomal morphology of six *triformis*-group species. The left and right lateral views of three morphotypes are shown. Instead of the six-plated eurystomatous form, the 12-plated (megastomatous) form is drawn for *P. triformis*.

TABLE 3. Summary of stomal and male tail characters of eight *Pristionchus triformis* group species.

Species	Reproductive mode	Stenostomatous form		Eurystomatous form		Genital papillae	Other characteristic feature(s)
		Right subventral plate	Left subventral plate	Right subventral tooth	Left subventral cusps		
triformis	H	Single peak, sometimes very well-developed	Three or four bumps	Claw-like, single peak	Three plates with various number of cusps	v1-v2d ≤ v2d-v4; v2d and v3 clearly separated; v4-ad < ad-pd	12-plated (megastomatous) form present
hoplostomus	G	1-3 peak(s)	Three bumps	Claw-like, 1-2 peak(s)	Three plates with various number of cusps	v1-v2d ≥ v2d-v4; v2d and v3 close to each other; v4-ad = ad-pd	Dorsal tooth of stenostomatous form sometimes dorsally curved and anteriorly directed. Twelve-plated stenostomatous form present, but rare.
fukushimae	G	0-3 peak(s)	Three bumps	Claw-like, single peak	Three plates with various number of cusps	v1-v2d ≥ v2d-v4; v2d and v3 close to each other; v4-ad = ad-pd	
yamagatae	G	Single peak	Three bumps	Claw-like, 1-2 peak(s)	Three plates with various number of cusps	v1-v2d ≥ v2d-v4; v2d and v3 close to each other; v4-ad = ad-pd	Dorsal tooth of stenostomatous form somewhat blunt
magnoliae n. sp.	G	Two peaks	Three bumps	Claw-like, 1-2 peak(s)	Three plates with various number of cusps	v1-v2d > v2d-v4; v2d and v3 close to each other; v4-ad = ad-pd	Dorsal tooth of stenostomatous form somewhat blunt.
musae n. sp.	G	Two peaks	Three bumps	Claw-like, 1-2 peak(s)	Three plates with various number of cusps	v1-v2d > v2d-v4; v2d and v3 close to each other; v4-ad = ad-pd	Dorsal tooth of stenostomatous form sometimes dorsally curved and anteriorly directed.
auriculatae n. sp.	G	One or two blunt peak(s)	Three bumps	Claw-like, 1-2 peak(s)	Three plates with various number of cusps	v1-v2d = v2d-v4; v2d and v3 close to each other; v4-ad = ad-pd	Dorsal tooth of stenostomatous form sometimes dorsally curved and anteriorly directed.
passalidorum n. sp.	G	Single peak	Three or more denticles (pointed)	Claw-like, single peak	Three plates with various number of cusps	v1-v2d > v2d-v4; v2d and v3 close to each other; v4-ad = ad-pd	Dorsal tooth of stenostomatous form sometimes dorsally curved and anteriorly directed.

Six labial sensilla one on each lip sector, papilliform, and four male cephalic papillae papilliform, on each of right and left subventral and right and left dorsal lip sectors, smaller than labial sensilla. Stomal dimorphism (or polymorphism) present, and its details are described below. Anterior pharynx 1.5 times as long as posterior pharynx, muscular composed of tube-shaped procorpus and well-developed oval-shaped metacorpus (median bulb); posterior pharynx glandular with nerve ring at mid-isthmus length. Pharyngo-intestinal junction (cardia) clearly observed, well-developed. Intestine simple tube, not forming pre-rectum, extended posteriorly from cardia to rectum; three (two subventral and one dorsal) rectal gland cells observed at distal end of intestine (margin between intestine and rectum). Secretory-excretory pore not conspicuous, ventrally located at level of isthmus to pharyngo-intestinal junction, excretory duct extending anteriad and reflexed back to position of pore; two large secretory-excretory cells around the excretory duct. Deirid observed laterally on lateral field, located at the level around the posterior end of basal bulb to pharyngo-intestinal junction to a half body diameter posterior to the junction, ca 0.5–1 body diam. posterior to secretory-excretory pore. Hemizonid not observed. Lateral glands (small pores connected to secretory cell) on lateral body surface, with positions inconsistent among individuals, numbering 5 to 8 for males and 9 to 13 for females. Postdeirid at anterior part of *vas deferens* in male and the posterior end of posterior gonad in female, on the same striation with deirid (= lateral field) or on the adjacent striation or the second dorsally neighboring striation to lateral field.

Stenostomatous form. Cheilostom consisting of six per- and interradial plates. Each plate is sometimes separated into two plates, forming a rare 12-plated morph. Incision between plates not always easily distinguished. Anterior end of each plate rounded and elongated to project from stomal opening forming small flap. Gymnostom short, cuticular ring-like anterior end overlapping cheilostom internally. Stegostom separated into three subsections: pro-meso, meta, and telostegostom. Pro-mesostegostom forming a weakly cuticularized ring internally overlapping with gymnostoma to connect gymnostom and metastegostom. Metastegostom bearing conspicuous and movable triangular or flint-shaped dorsal tooth with strongly sclerotized surface giving an appearance of an inverted V-shape in light microscopy in lateral view; pointed left subventral ridge with three minute adventitious denticles on a plate, most ventral denticle often masked by remaining two in lateral view; pointed right subventral ridge, often with distinct distal adventitious denticle(s). Shape of dorsal tooth, right and left subventral ridges are variable within and among species. Telostegostom weakly sclerotized cup-like cavity connecting stoma and pharynx.

Eurystomatous form. Cheilostom divided into six well-distinguished per- and interradial plates. Anterior end of each plate rounded and elongated to protrude from stomal opening and form a small flap. Plates often partially or completely separated to form large plates with split tip or two narrow plates. 6–12 cheilostomatal flaps derived from 6–12 plates. 12-plated eurystomatous form of *P. triformis* is regarded as 'megastomatous form'. Gymnostom with thick cuticle, forming short, ring-like tube being thicker posteriorly; finely serrated anterior end of gymnostom internally overlapping posterior end of cheilostomatal plates. Structural variation within each species described below. Stegostom separated into three subsections: pro-meso, meta, and telostegostom. Pro-mesostegostom variable among individuals in all examined species, variation described below. Metastegostom bearing large claw-like dorsal tooth, and large, claw-like or pointed right subventral tooth. Left subventral sector of metastegostom bearing three triangular ridges; tip of each ridge sometimes split into two or more fine tips, shape varies within and among species. Dorsal tooth and right subventral tooth movable. Movement not observed in left subventral denticles. Telostegostom weakly sclerotized cup-like cavity connecting stoma and pharynx.

Variation in gymnostom and pro-mesostegostom of eurystomatous form.

Pro-mesostegostom of *triformis*-group well-developed compared to other species in genus, two different types of stomal structure in eury- and megastomatous forms. Relatively well sclerotized pro-mesostegostom internally overlap with posterior end of gymnostom. In one type, referred to as 'type 1', gymnostom relatively thick and short, pro-mesostegostom relatively short, bearing coarse and short serrates at anterior end (Fig. 21D). 'Type 2' with thin and long gymnostom, pro-mesostegostom loses anterior serrates, but often bears sparse or 2–3 rows of fine serrates on inner wall of elongated pro-mesostegostom. These two types are not always easily distinguishable by light microscopy. Further, intermediate types, *e.g.*, lacking both anterior and posterior serrates, occur. Both types are treated as variation within the eury/megastomatous forms.

Male. Whole body ventrally arcuate, strongly ventrally curved at tail region when killed by heat. Testis single, ventrally located, anterior part reflexed to right or left side; spermatogonia arranged in three to five rows in reflexed part, well-developed spermatocytes arranged as three to four rows in anterior two-thirds of main branch, mature amoeboid spermatids arranged in multiple rows in proximal part of gonad. *Vas deferens* not clearly separated from

other parts of gonad. Posterior end of *vas deferens* and rectum fused to form a cloacal tube. Spicules paired, separate; spicules smoothly curved in ventral view, adjacent to each other for distal third of their length, each smoothly tapering to pointed distal end; spicule in lateral view smoothly ventrally arcuate, giving spicule about 100° curvature, oval manubrium at anterior end, lamina/calomus complex smoothly tapering to pointed distal end. Gubernaculum conspicuous, about one-third of spicule length, broad anteriorly such that dorsal wall is slightly recurved with dorsal and ventral walls separate at 50–60° angle at posterior end; dorsal side of gubernaculum possessing single, membranous, anteriorly directed process and lateral pair of more sclerotized, anteriorly and obliquely ventrally directed processes. In lateral view, anterior half of gubernaculum with two successive curves separated by anteriorly and obliquely ventrally directed process, with anterior terminal curvature highly concave and almost closed, with deep posterior curvature being one-third of gubernaculum length; posterior half forming tube-like process enveloping spicules. Cloacal opening (co) slit-like in ventral view; one small, ventral, single genital papilla (vs) on anterior cloacal lip. All nine paired genital papillae papilliform. Tail conoid with a long spike. Arrangement of paired papillae and phasmid is variable among species, but size and structure of papillae are consistent with what has been described above. Bursa or bursal flap absent.

Female/hermaphrodite. Body relaxed or weakly ventrally arcuate when killed by heat. Gonad didelphic, amphidelphic; each gonadal system arranged from vulva/vagina as uterus, oviduct, and ovary; anterior gonad right of intestine, with uterus and oviduct extending ventrally and anteriorly on right of intestine and with totally reflexed (=antidromous reflexion) ovary extending dorsally on left of intestine; oocytes mostly arranged in three to four or more rows in distal two-thirds of ovary and in double or single row in rest of ovary, distal tips of each ovary reaching oviduct of opposite gonad branch; anterior end of oviduct (=junction tissue between ovary and oviduct) consists of rounded cells; anterior part of oviduct consists of rounded cells, forming a simple tube; middle part of oviduct serving as spermatheca, consists of roundish and relatively large cells. Eggs in single to multiple-cell stage or even further developed at posterior part of oviduct (=uterus), in young females being composed of squared or angular cells, long enough to contain one well-developed oocyte. *Receptaculum seminis* not observed, *i.e.*, the organ is not independent, and a part of oviduct/uterus works as the organ; vaginal glands present but obscure; vagina perpendicular to body surface, surrounded by sclerotized tissue; vulva slightly protuberant in lateral view, pore-like in ventral view; rectum about one anal body diameter (ABD) long, intestine/rectum junction surrounded by well-developed sphincter muscle. Anus in form of dome-shaped slit, posterior anal lip slightly protuberant. Tail elongate conoid with filiform terminus. The position of phasmid described for each species.

***Pristionchus magnoliae* n. sp.**

urn:lsid:zoobank.org:act:ADA242D3-23FE-4ED0-AE51-6A537385AF79

Etymology. The species name is derived from the associated substrate from which the species was recovered. The strain was isolated from rotting *Magnolia grandiflora* fruits with some soil material.

Measurements. See Table 4

Adult. General characters are as described above for the *triformis*-group.

Stenostomatous form. Cheilostom, gymnostom, pro-mesostegostom and telostegostom as described above. 12-plated form not found. Arrangement of tooth and denticles in metastegostom as described above; dorsal movable tooth somewhat narrow, *i.e.*, the angle of anteriorly detected tip is steeper, triangular and slightly anteriorly directed in male, flint-shaped with somewhat blunt tip in female, both have strongly sclerotized surface giving an appearance of an inverted V-shape in lateral view; left subventral ridge with three minute, blunt adventitious denticles on plate; right subventral ridge with three distal rounded adventitious denticles, plate slightly narrower in male than female, thus males seems to have two pointed denticles.

Eurystomatous form. Cheilostom as described above, anterior half of each cheilostomal plate often split into two tips to form 12-flapped form. Six-flapped form (without split cheilostomal plate) rare. Gymnostom in both types 1 and 2 short and thick, forming cuticular ring. Pro-mesostegostom well-developed, internally overlapping with the posterior end of gymnostom; type 1 form short, somewhat flattened and bearing weak serrates at anterior end; type 2 with 2–3 rows of small spines on inner surface. Arrangement of tooth and denticles in metastegostom as described above; dorsal movable tooth claw-like as typical of the genus; left subventral ridge with three large plates, each often has split tips and extra denticles on middle forming spiny plate; right subventral movable tooth claw-like and often bears an extra peak on the ventral side. Telostegostom as described above.

Male. Paired papillae and phasmid are arranged as <v1, v2d, v3, co, v4, ad, ph, (v5, v6, v7), pd>, where v1 located about a little more than 1 CBD anterior to co; v2d less than 1/4 CBD anterior to co; v3 adcloacal; v4 at 1/4–1/5 CBD posterior to co, *i.e.*, v2d, v3, co and v4 are close to each other; ad about 1 CBD posterior to co; ph at midway between ad and root of tail spike; v5–v7 forming triplet, between ph and the root of tail spike; and pd around level of v7. v1, v3, v4 and ph subventral, v2d and ad lateral, v5–7 ventral, pd subdorsal in male tail. General shape of spicule and gubernaculum as described above.

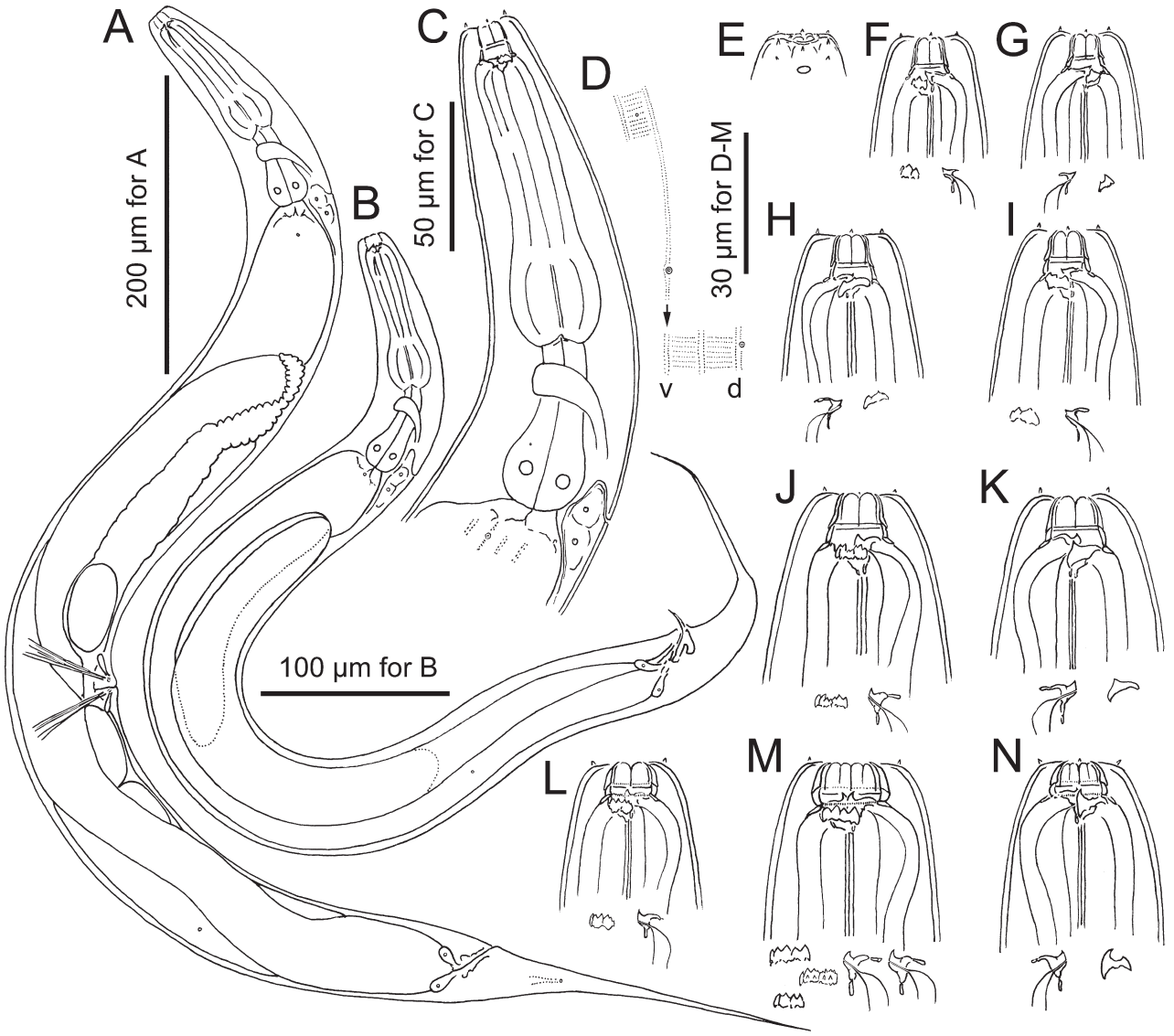


FIGURE 18. *Pristionchus magnoliae* **n. sp.** A: Right lateral view of adult female. B: Right lateral view of adult male. C: Anterior part of adult female in right lateral view. D: Body surface structure showing lateral gland, deirid, postdeirid and relative position of deirid and postdeirid. E: Left lateral view of male head region showing labial sensilla, cephalic papillae and amphid. F, G: Stomal region of stenostomatous male in left (F) and right (G) lateral view. H, I: Stomal region of stenostomatous female in left (H) and right (I) lateral view. J, K: Stomal region of eurystomatous female with pro-mesostegostomatal spines (type 2) in left (J) and right (K) lateral view. L–N: Stomal region of eurystomatous female with pro-mesostegostomatal serrates (type 1) in left (L, M) and right (N) lateral view where six-flapped form (L) is also shown. Morphological variations of teeth and ridges are separately drawn in each subfigure as left subventral ridge (left) and dorsal tooth (right) (F), dorsal (left) and right subventral ridge (right) (G), dorsal tooth (left) and right subventral ridge (right) (H), left subventral ridge (left) and dorsal tooth (rght) (I), left subventral ridge (left) and dorsal tooth (right) (J), dorsal (left) and right subventral (right) teeth (K), left subventral ridge (left) and dorsal tooth (right) (L), left subventral ridge (left) and dorsal tooth (middle and right) (M), and dorsal (left) and right subventral (right) teeth (N).

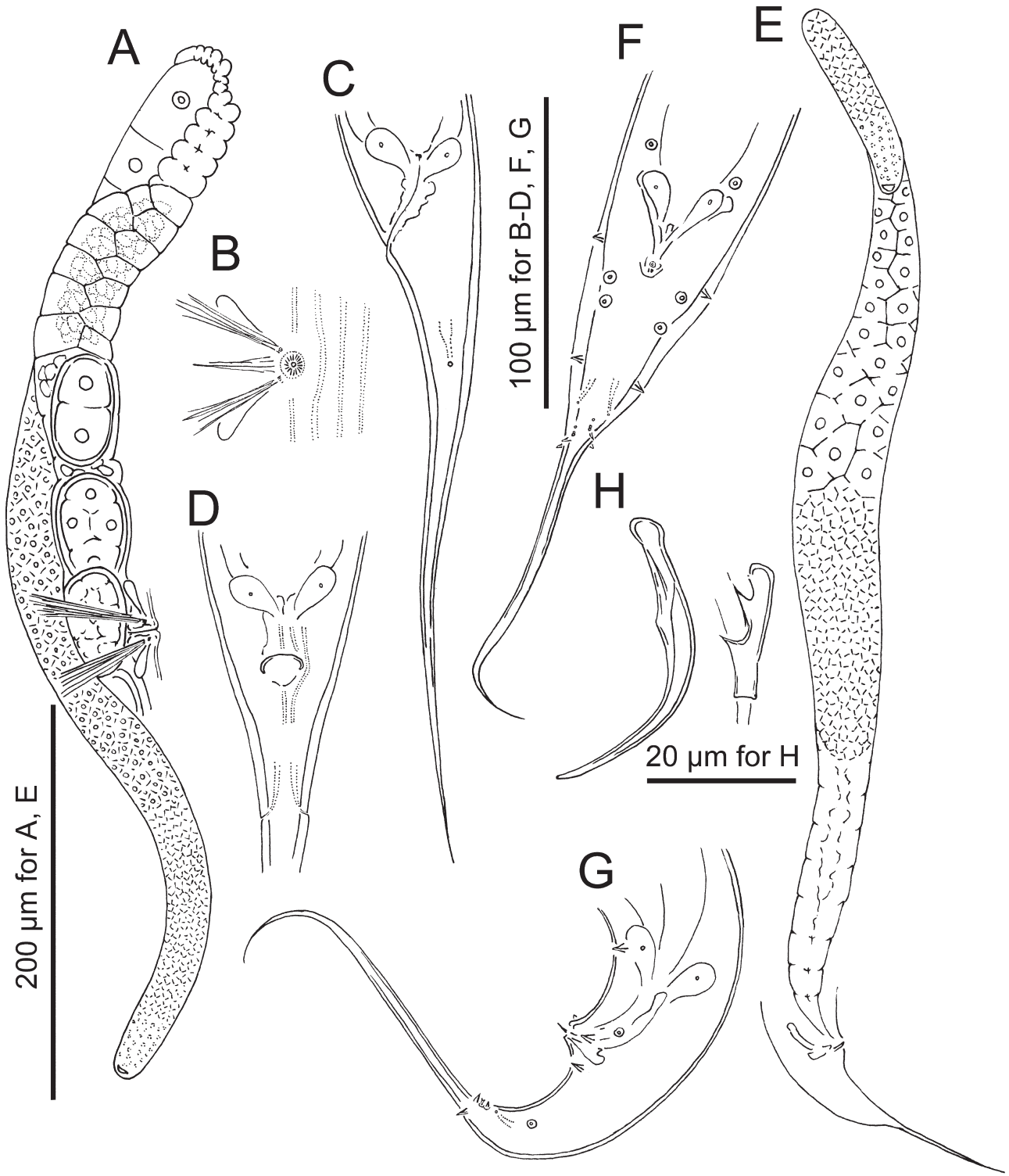


FIGURE 19. *Pristionchus magnoliae* **n. sp.** A: Anterior female gonad of well-developed individual in right lateral view. B: Vulval region in ventral view. C, D: Female tail in left lateral (C) and ventral (D) view. E: Male gonad in right lateral view. F, G: Male tail in ventral (F) and left lateral (G) view. H: Spicule and gubernaculum in left lateral view.

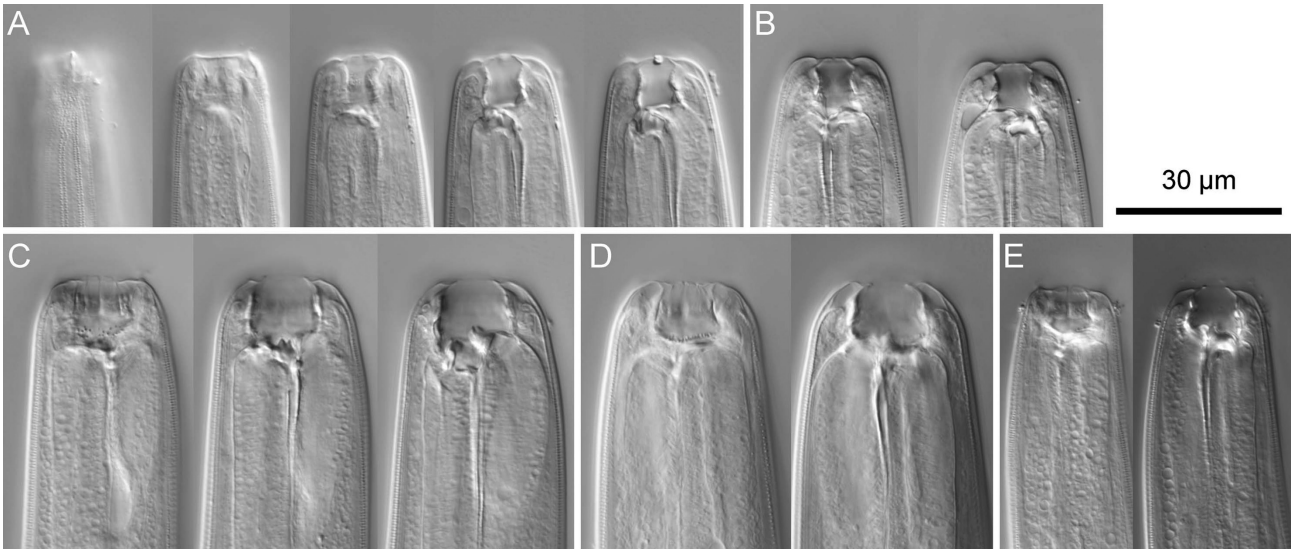


FIGURE 20. Stomal region of *Pristionchus magnoliae* **n. sp.** A: Left lateral view of stenostomatous form in five different focal planes. B: Right lateral view of stenostomatous form in two different focal planes. C: Left lateral view of type 2 eurystomatous form in three different focal planes. D: Right lateral view of type 1 eurystomatous form in two different focal planes. E: Left lateral view of six-plated eurystomatous form in two different focal planes.

Female. Gonadal characters of female as described above. Tail elongate conoid with slightly filiform terminus, *i.e.*, the posterior half of tail more elongated compared with anterior part. Phasmid ventro-laterally located at about 1.0–1.5 ABD posterior to anal opening.

Diagnosis and relationships. *Pristionchus magnoliae* **n. sp.** is characterized by the somewhat blunt and flint-shaped dorsal tooth and the right and left subventral ridges with three blunt denticles of stenostomatous form, right subventral stegostomal tooth which often has blunt peak on ventral side in eurystomatous form, well-serrated promesostegostomal wall in type 2 eurystomatous form, arrangement of male genital papillae, <v1, v2d, v3, co, v4, ad, ph, (v5, v6, v7), pd> where v2d, v3, co and v4 are close to each other, and a long spike occupying more than 2/3 of tail length. *Pristionchus magnoliae* **n. sp.** is typologically close to *P. yamagatae* and *P. hoplostomus*. The new species and the other two species share the arrangement of genital papillae, <v1, v2d, v3, co, v4, ad, ph, (v5, v6, v7), pd> where v2d, v3, co and v4 are close to each other, and right subventral tooth in eurystomatous form, often having an extra peak on its ventral side. In addition, somewhat blunt and flint-shaped dorsal tooth in the stenostomatous form is common in the new species and *P. yamagatae*. However, the new species is distinguished from *P. hoplostomus* by the absence *vs.* presence of 12-plated stenostomatous form and dorsal tooth of stenostomatous form, blunt *vs.* pointed. The typological characters of *P. magnoliae* **n. sp.** are almost identical with those of *P. yamagatae*, distinguished only by the relative position of ph and v5, close but clearly separated *vs.* very close and sometimes overlapping. Further, the new species is distinguished from all other species by mating experiments and also characterized by a *ca.* 1,600-bp fragment of the SSU rRNA gene (GenBank accession number MW017219), the sequence of which is distinct from that of all other *Pristionchus* species.

Type host and locality. Rotting *Magnolia grandiflora* fruits in Changfeng Park, Shanghai, PRC.

Type material and type strain. Type strain RS5999, other strain RS 6000 from the same location, frozen at the nematode collection of the MPI Tübingen and available as living culture upon request. Voucher specimens sent to the following museums: Holotype male, Paratype male and female: Museum für Naturkunde Karlsruhe, Germany; Paratype male and female: Swedish Natural History Museum, Stockholm, Sweden; Paratype male and female: University of California in Riverside Nematode Collection (UCRNC), Riverside, CA, USA.

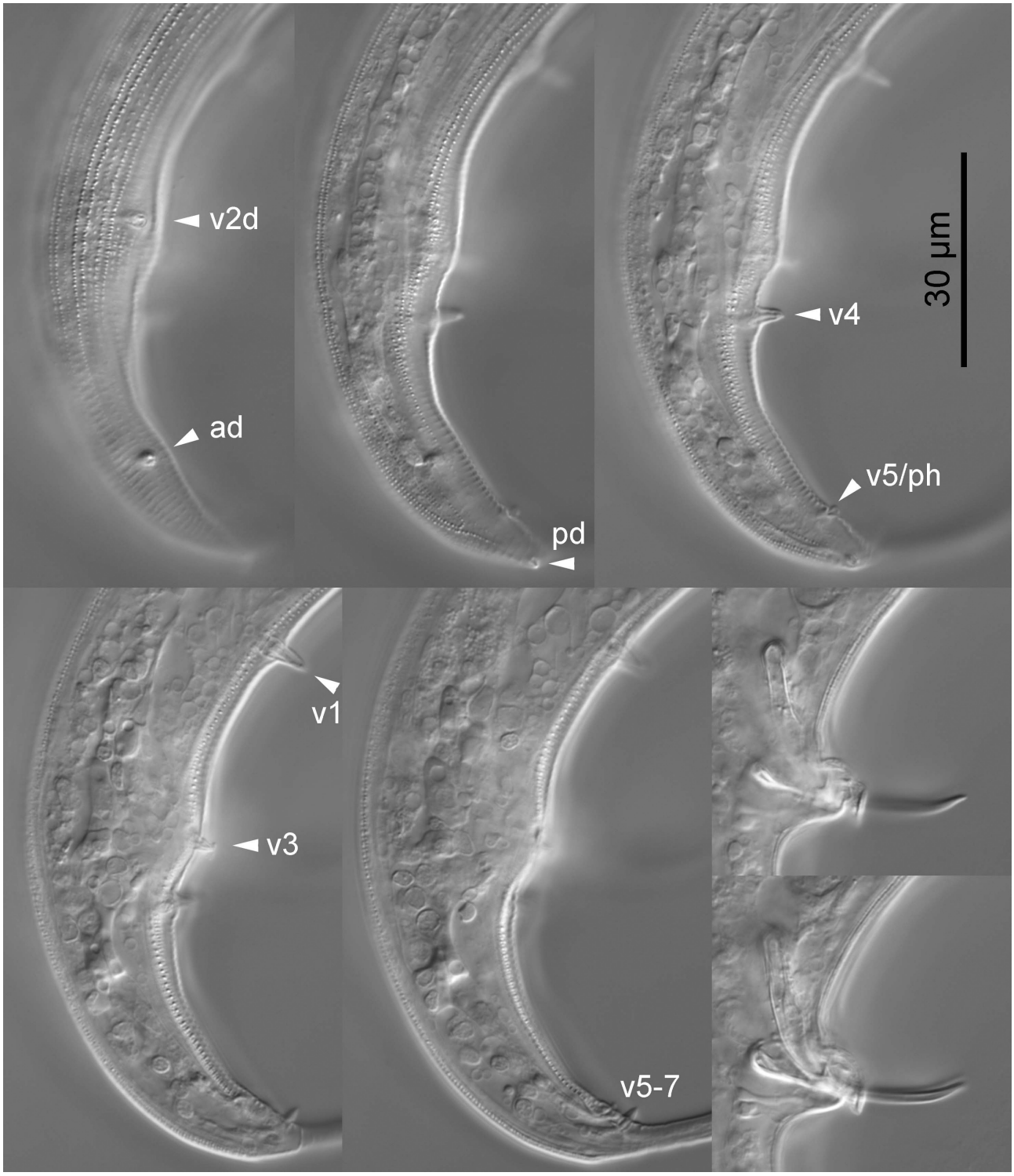


FIGURE 21. Male tail characters of *Pristionchus magnoliae* **n. sp.** Left lateral view of whole tail in five different focal planes and of spicule and gubernaculum in two different focal planes. Genital papillae are labelled using to the terminology of Sudhaus & Fürst von Lieven (2003), and laterally located papillae are indicated with "d".

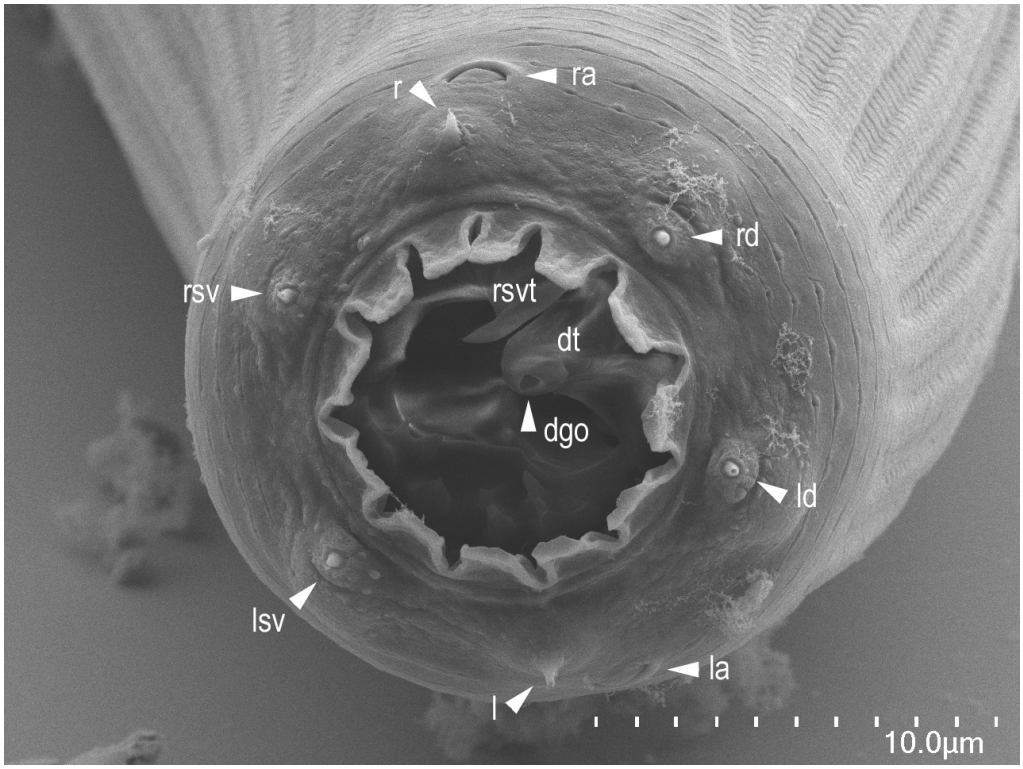


FIGURE 22. Scanning electron photomicrographs of 12-flapped eurystomatous female of *Pristionchus magnoliae* **n. sp.** *en face* view with left and right amphids (la and ra) and left and right subventral (lsv and rsv), lateral (l) and left and right dorsal (ld and rd) labial sensillae, dorsal (dt) and right subventral (rsvt) teeth and dorsal gland orifice (dgo) indicated. Form type is not specified because the deeper part of the stoma is not seen.

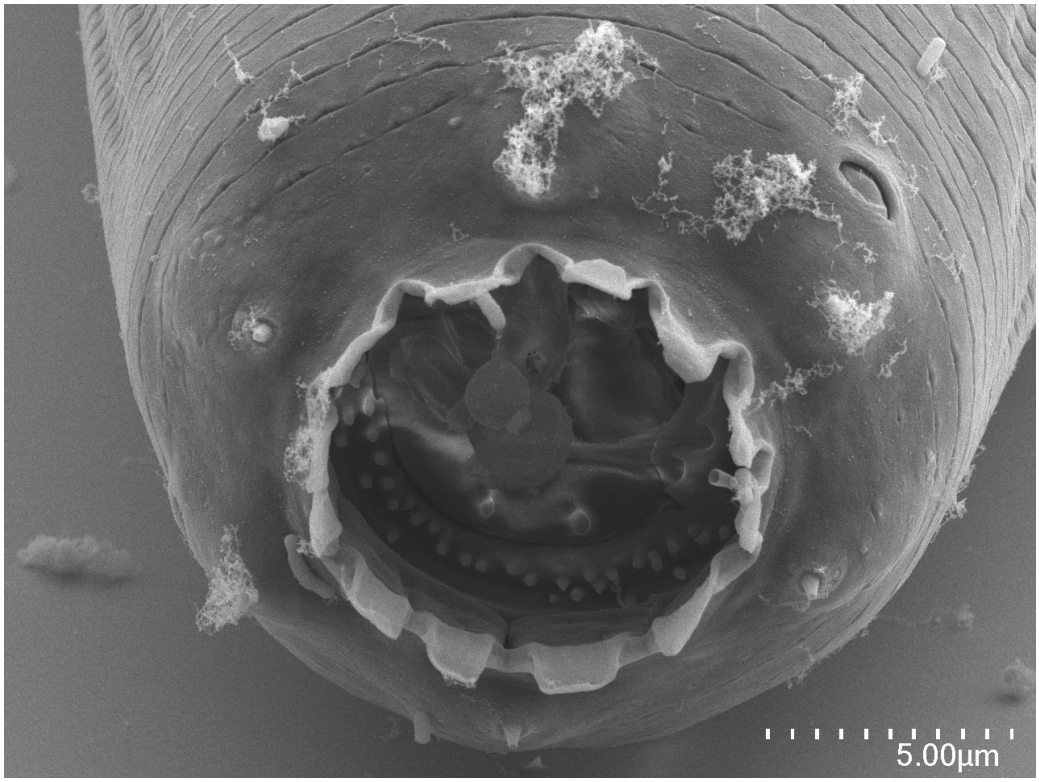


FIGURE 23. Scanning electron photomicrograph of 12-flapped 'type 1' eurystomatous female of *Pristionchus magnoliae* **n. sp.** showing serrated anterior edge of pro-meso stegostom.

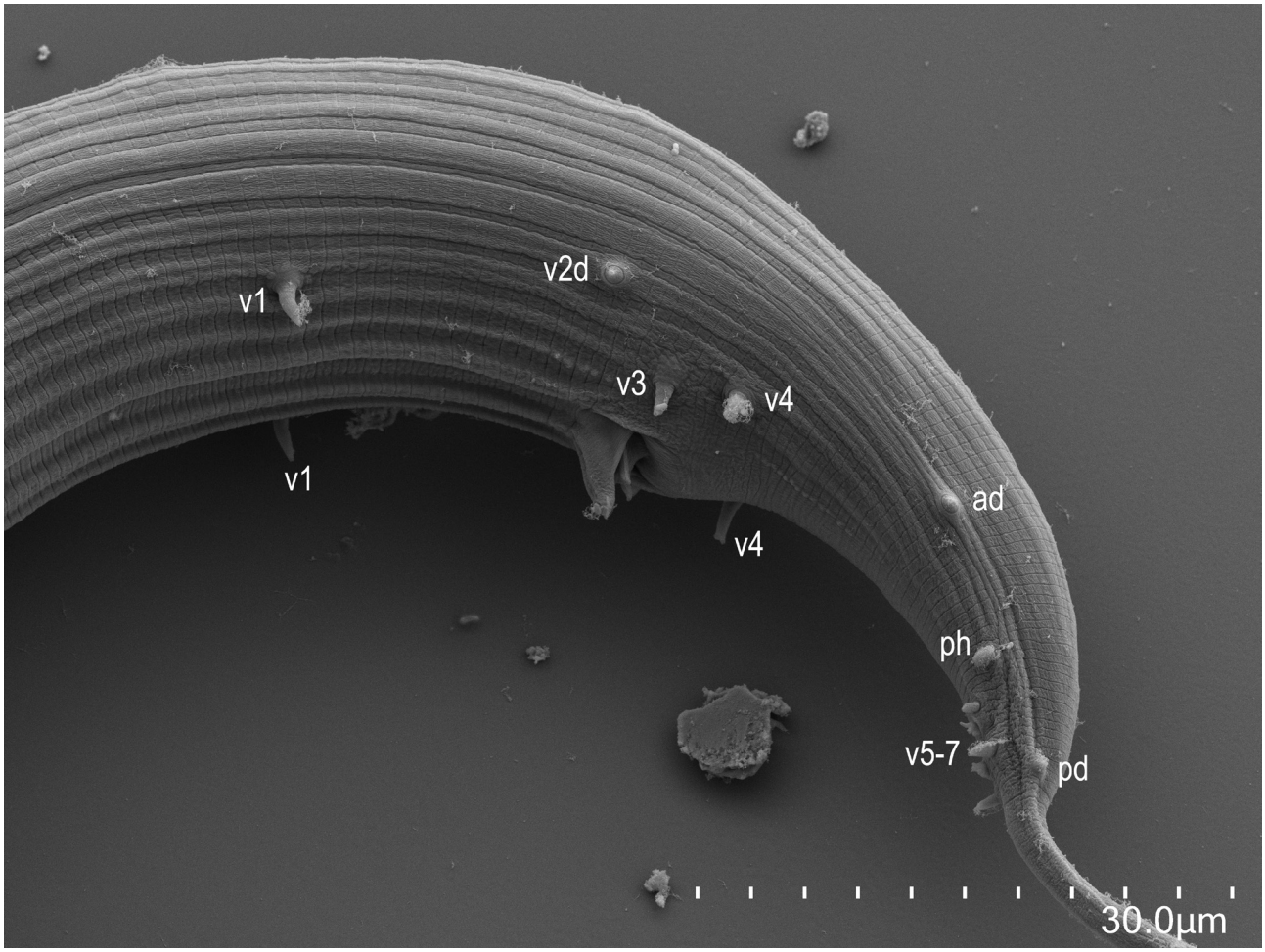


FIGURE 24. Scanning electron photomicrograph of *Pristionchus magnoliae* **n. sp.** male tail with the position of genital papillae according to the terminology of Sudhaus and Fürst von Lieven (2003).

***Pristionchus musae* n. sp.**
urn:lsid:zoobank.org:act:E015F8AB-4F6F-46B1-A96A-85F724C7448F

Etymology. The species name is derived from the associated substrate from which the species was recovered. The strain was isolated from a rotting pseudostem of a plant of the banana family (*Musa* sp.) with soil material.

Measurements. See Table 4

Adult. General characters are as described above for the *triformis*-group.

Stenostomatous form. Cheilostom, gymnostom, pro-mesostegostom and telostegostom as described above. 12-plated form was not found during observation. Arrangement of tooth and denticles in metastegostom as described above; dorsal movable tooth triangular, anterior end slightly curved, tooth often directed anteriorly; left subventral ridge with three minute, blunt or pointed adventitious denticles on a plate; right subventral ridge with three distal rounded or blunt adventitious denticles. Whole stoma including the tooth and denticles is narrow and more pointed in males.

Eurystomatous form. Cheilostom as described above, anterior half of each cheilostomal plate often split into two tips to form 12-flapped form. Six-flapped form (without split cheilostomal plate) rare. Gymnostom in both types 1 and 2 short and thick, forming cuticular ring. Pro-mesostegostom well-developed, internally overlapping with posterior end of gymnostom; type 1 form short, somewhat flattened and bearing weak serrates at anterior end; type 2 with small spines sparsely on inner surface. Arrangement of tooth and denticles in metastegostom as described above; dorsal movable tooth claw-like as typical of the genus; left subventral ridge with three large plates, each often has split tips and extra denticles on middle forming spiny plate; right subventral movable tooth claw-like and often bears extra peak on the ventral side. Telostegostom as described above.

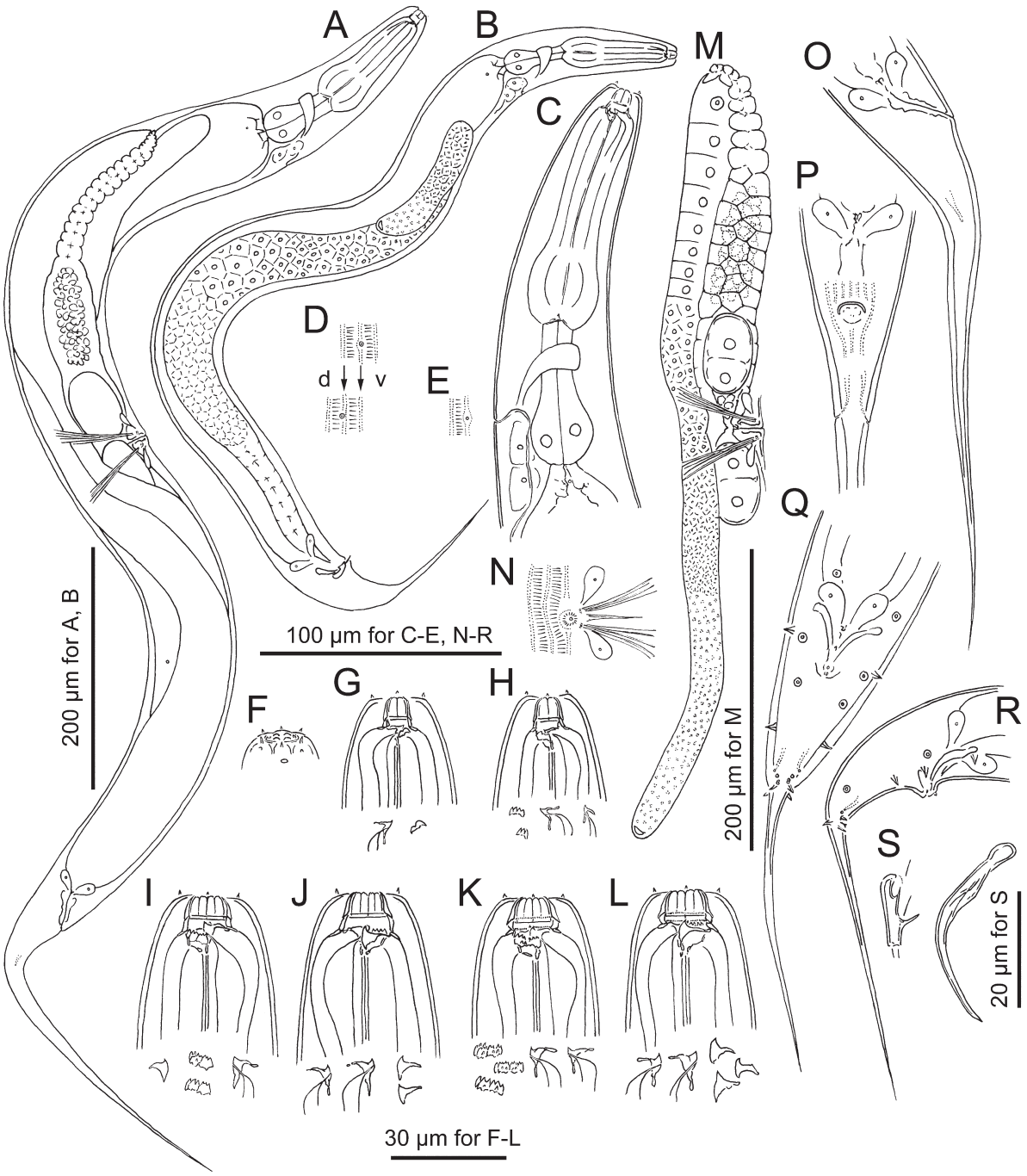


FIGURE 25. *Pristionchus musae* **n. sp.** A: Right lateral view of adult female. B: Right lateral view of adult male. C: Anterior part of adult female in left lateral view. D: Body surface structure of deirid and postdeirid region showing the relative position of deirid and postdeirid where 'd' and 'v' indicates dorsal and ventral sides, respectively. E: Ventral view of secretory-excretory pore opening. F: Left lateral view of male head region showing labial sensilla, cephalic papillae and amphid. G, H: Stomal region of adult female in left lateral (G) and right lateral (H) view. I, J: Stomal region of eurystomatous female with pro-meso-stegostomatal spines (type 2) in left (I) and right (J) lateral view. K, L: Stomal region of eurystomatous female with pro-meso-stegostomatal serrates (type 1) in left (K) and right (L) lateral view. Morphological variations of teeth and ridges are separately drawn in each subfigure as dorsal (left) and right subventral ridge (right) (G), left subventral ridge (left) and dorsal tooth (middle and right) (H), right subventral tooth (left), left subventral ridge (middle) and dorsal tooth (right) (I), dorsal (left and middle) and right subventral (right) teeth (J), left subventral ridge (left) and dorsal tooth (middle and right) (K), and dorsal (left and middle) and right subventral (right) teeth (L). M: Anterior female gonad in right lateral view. N: Vulval region in ventral view. O, P: Female tail in right lateral (O) and ventral (P) view. Q, R: Male tail in ventral (Q) and right lateral (R) view. S: Spicule and gubernaculum in right lateral view.

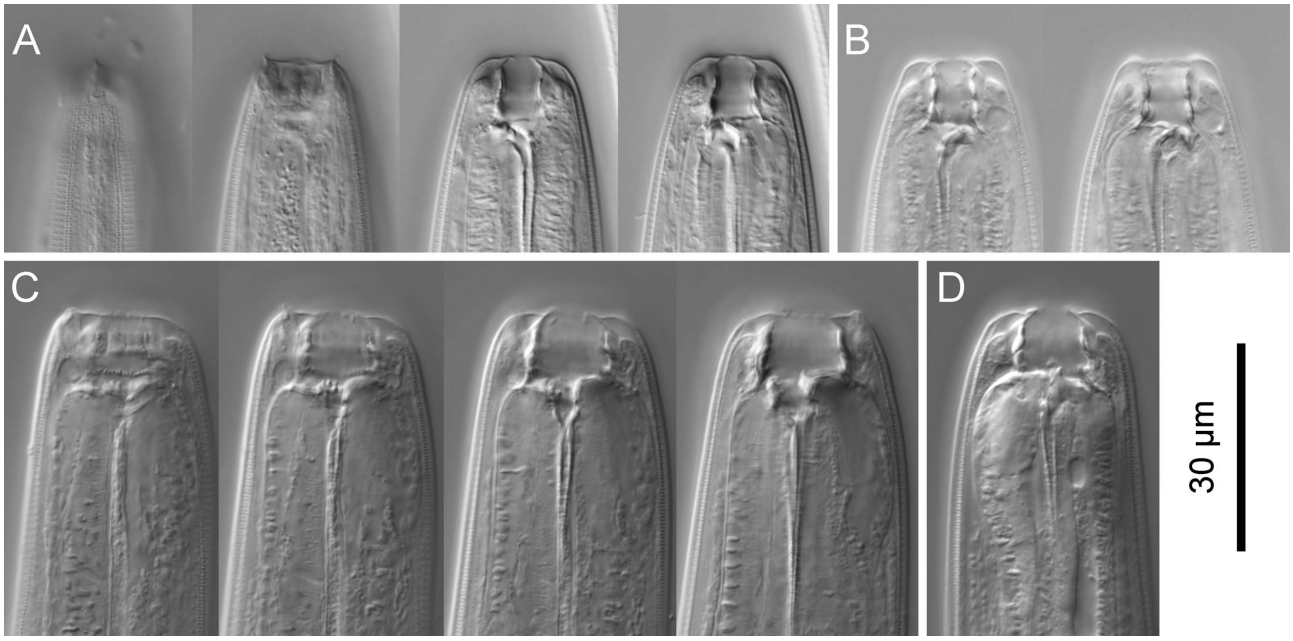


FIGURE 26. Stomal region of *Pristionchus musae* **n. sp.** A: Left lateral view of stenostomatous form in four different focal planes. B: Right lateral view of stenostomatous form in two different focal planes. C: Left lateral view of type 1 eurystomatous form in four different focal planes. D: Right lateral view of type 1 eurystomatous form.

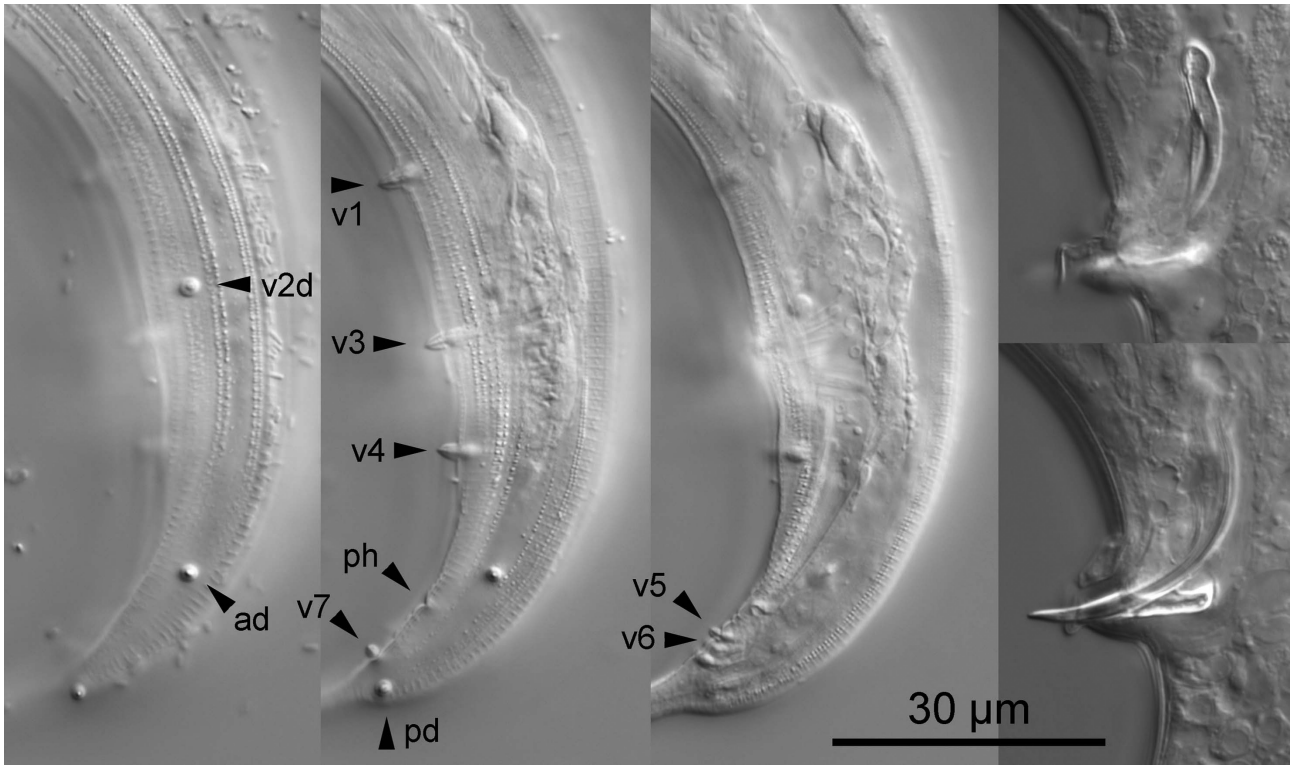


FIGURE 27. Male tail characters of *Pristionchus musae* **n. sp.** Left lateral view of whole tail in three different focal planes and spicule and gubernaculum two different focal planes. Genital papillae are labelled using the terminology of Sudhaus & Fürst von Lieven (2003), and laterally located papillae are indicated with "d".

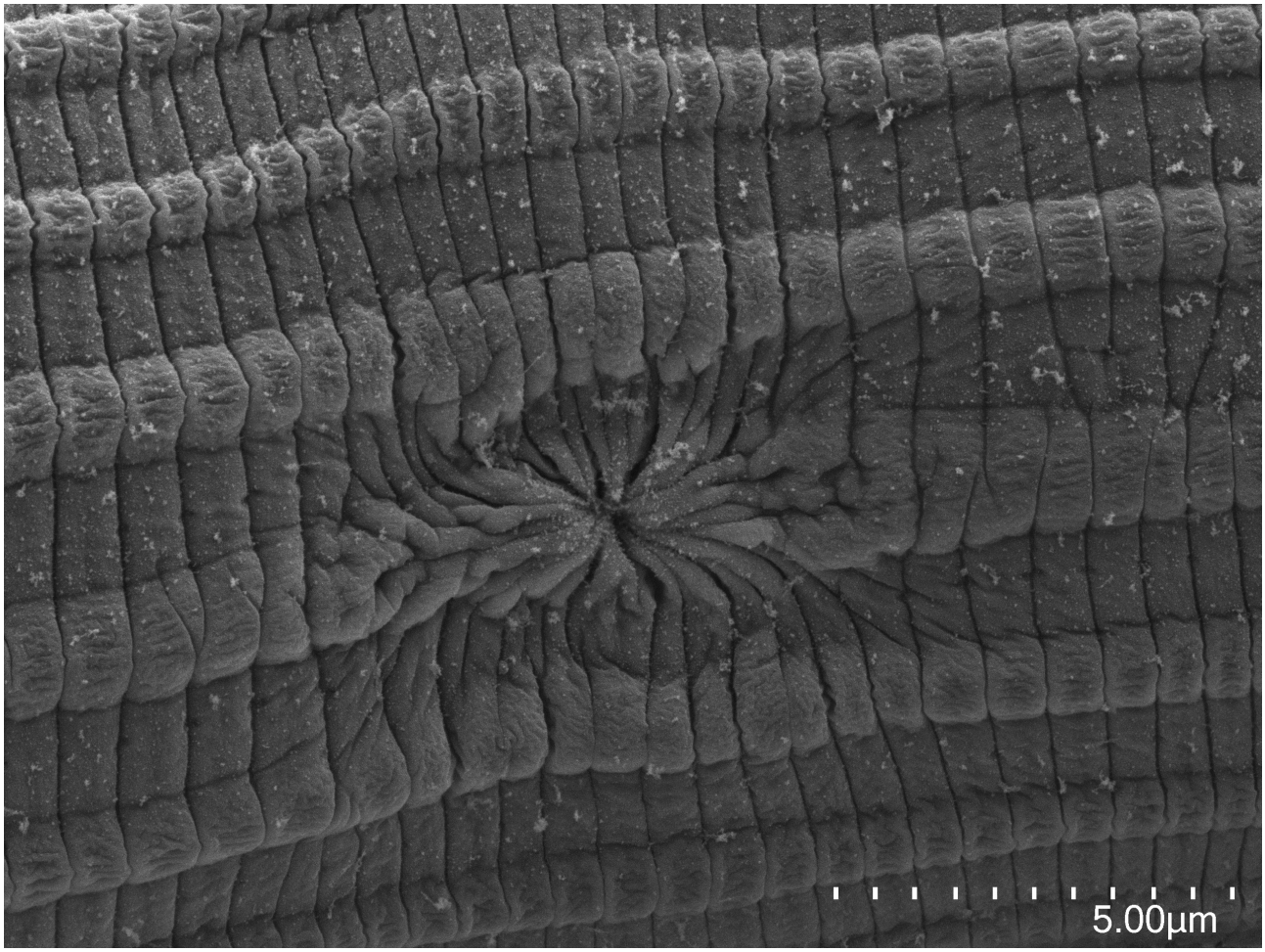


FIGURE 28. Scanning electron photomicrograph of vulval opening region of *Pristionchus musae* **n. sp.**

Male. Paired papillae and phasmid are arranged as <v1, v2d, v3, co, v4, ad, (ph, v5, v6, v7, pd)>, where v1 located about 1 CBD anterior to co; v2d less than 1/5 CBD anterior to co; v3 adcloacal; v4 at 1/3 CBD posterior to co, *i.e.*, v2d, v3, co and v4 are close to each other; ad about 1 CBD posterior to co; ph at 2/3 way from ad and the root of tail spike; v5–v7 forming triplet, just posterior to ph; and pd around level of v7, *i.e.*, ph, triplet papillae and pd are very close to each other. v1, v3, v4 and ph subventral, v2d and ad lateral, v5–7 ventral, pd subdorsal in male tail. General shape of spicule and gubernaculum as described above.

Female. Gonadal characters of female as described above. Tail elongate conoid with slightly filiform terminus, *i.e.*, the posterior half of tail is more elongated compared with anterior part. Phasmid ventro-laterally located at about 1.5 ABD posterior to anal opening.

Diagnosis and relationships. *Pristionchus musae* **n. sp.** is characterized by the slightly anteriorly curved triangular dorsal tooth and the right and left subventral ridges with three denticles of stenostomatous form, right subventral stegostomal tooth which often has blunt peak on the ventral side in eurystomatous form, sparsely serrated pro-mesostegostomal wall in type 2 eurystomatous form, the arrangement of male genital papillae, <v1, v2d, v3, co, v4, ad, (ph, v5, v6, v7, pd)> where v2d, v3, co and v4 and ph, v5, v6, v7, pd are close to each other, respectively, and a long spike of male tail. *Pristionchus musae* **n. sp.** is typologically close to *P. hoplostomus*. The new species and *P. hoplostomus* share a triangular and pointed dorsal tooth in stenostomatous form, right subventral tooth in eurystomatous form, often have an extra peak on its ventral side, and the arrangement of genital papillae, closely located v2d, v3, co and v4. The new species is distinguished from *P. hoplostomus* by the absence *vs.* presence of 12-plated stenostomatous form, and dorsal tooth of stenostomatous form, *i.e.*, the tooth is somewhat larger and more anteriorly directed in *P. hoplostomus* than *P. musae* **n. sp.** Further, the new species is distinguished from all other species by mating experiments and also characterized by a *ca.* 1,600-bp fragment of the SSU rRNA gene (GenBank accession number MW017220), the sequence of which is distinct from that of all other *Pristionchus* species.

TABLE 4.

	***P. magnoliae* (RS5999)**		***P. musae* (RS5987)**	
character	**stenostomatous male**	**stenostomatous female**	**stenostomatous male**	**stenostomatous female**
n	5	5	5	5
L	852 ± 20.4 (822-877)	1011 ± 89.3 (922-1111)	853 ± 77.4 (749-949)	1015 ± 84.7 (940-1120)
L’	689 ± 12.9 (674-708)	778 ± 77.8 (708-879)	688 ± 65.2 (605-762)	787 ± 59.5 (712-849)
a	14 ± 0.9 (14-16)	13 ± 0.8 (13-14)	16 ± 1.4 (14-18)	16 ± 1.1 (15-18)
b	6.4 ± 0.2 (6.2-6.8)	6.8 ± 0.3 (6.4-7.2)	5.9 ± 0.5 (5.3-6.6)	6.4 ± 0.6 (5.6-7.3)
c	5.2 ± 0.4 (4.8-5.7)	4.4 ± 0.4 (4.0-4.9)	5.2 ± 0.3 (4.7-5.6)	4.5 ± 0.4 (4.1-5.0)
c’	4.7 ± 0.7 (3.8-5.6)	7.4 ± 1.2 (6.3-9.2)	4.8 ± 0.5 (4.2-5.3)	7.0 ± 1.0 (5.9-8.1)
ant. stoma length (cheilo- + gymnostom)	7.3 ± 0.4 (6.8-7.7)	8.4 ± 0.9 (7.4-9.3)	7.8 ± 0.8 (6.9-8.4)	8.7 ± 0.5 (8.0-9.4)
total stoma length	11.5 ± 0.3 (11.0-11.8)	14.1 ± 0.4 (13.5-14.6)	12.3 ± 0.7 (11.8-13.4)	13.8 ± 0.7 (12.8-14.5)
stoma width	6.8 ± 0.3 (6.4-7.1)	7.3 ± 0.5 (6.6-7.8)	5.8 ± 0.4 (5.5-6.5)	6.7 ± 0.7 (5.6-7.5)
ant. pharynx length (pro- + metacorpus)	74 ± 2.7 (72-78)	86 ± 6.7 (78-93)	80 ± 7.6 (74-93)	92 ± 4.6 (86-99)
post. pharynx length (isthmus + basal bulb)	47 ± 2.0 (44-49)	54 ± 3.8 (49-59)	53 ± 2.8 (50-57)	58 ± 3.3 (54-61)
total length pharynx	121 ± 4.4 (117-127)	140 ± 8.6 (129-150)	132 ± 10.1 (125-150)	149 ± 7.3 (140-160)
ant./total pharynx %	61 ± 0.7 (60-62)	61 ± 2.2 (58-64)	60 ± 1.3 (59-62)	61 ± 1.0 (60-63)
median bulb diameter	22 ± 0.8 (21-23)	28 ± 0.9 (26-28)	23 ± 1.5 (22-25)	27 ± 2.2 (23-29)
terminal bulb diameter	19 ± 0.9 (18-20)	24 ± 1.5 (23-26)	19 ± 1.3 (18-21)	23 ± 1.4 (21-24)
neck length (incl. stoma)	132 ± 4.5 (128-139)	148 ± 9.3 (136-159)	145 ± 10.6 (137-163)	158 ± 7.3 (149-168)
excretion pore	134 ± 6.3 (129-142)	149 ± 10.9 (131-157)	141 ± 10.2 (127-154)	149 ± 9.7 (135-159)
nerve ring	95 ± 2.9 (91-98)	111 ± 7.0 (104-120)	107 ± 7.8 (99-119)	119 ± 4.2 (114-125)
testis length	475 ± 13.9 (459-495)	-	458 ± 61.8 (399-544)	-
anterior gonad length	-	190 ± 34.2 (159-236)	-	195 ± 27.7 (174-241)
posterior gonad length	-	156 ± 18.7 (137-180)	-	142 ± 34.9 (104-194)
ant. end to vulva distance	-	459 ± 52.6 (410-530)	-	468 ± 33.8 (427-515)
vulva to anus distance	-	322 ± 28.0 (301-359)	-	323 ± 27.3 (289-361)
T or V	56 ± 2.0 (52-58)	45 ± 1.7 (44-48)	54 ± 3.9 (50-60)	46 ± 1.8 (43-48)
max. body diameter	59 ± 2.6 (56-62)	76 ± 2.8 (71-78)	55 ± 6.6 (47-65)	64 ± 7.2 (53-71)
cloacal or anal body diameter	35 ± 3.0 (32-40)	32 ± 2.4 (29-35)	35 ± 2.5 (31-38)	33 ± 2.4 (29-34)
tail length	163 ± 14.3 (149-182)	233 ± 23.7 (201-266)	166 ± 16.3 (145-187)	229 ± 32.5 (190-274)
spicule length (curve)	44 ± 1.6 (42-45)	-	41 ± 2.1 (39-44)	-
spicule length (chord)	33 ± 2.5 (30-36)	-	33 ± 2 (30-35)	-
gubernaculum length	16 ± 0.8 (15-17)	-	15 ± 0.5 (14-15)	-

Type host and locality. Pugao Laozhai village, close to road, Yuanyang, Yunnan province, PRC.

Type material and type strain. Type strain RS5987 frozen at the nematode collection of the MPI Tübingen and available as living culture upon request. Voucher specimens sent to the following museums: Holotype male, Paratype male and female: Museum für Naturkunde Karlsruhe, Germany; Paratype male and female: Swedish Natural History Museum, Stockholm, Sweden; Paratype male and female: University of California in Riverside Nematode Collection (UCRNC), Riverside, CA, USA.

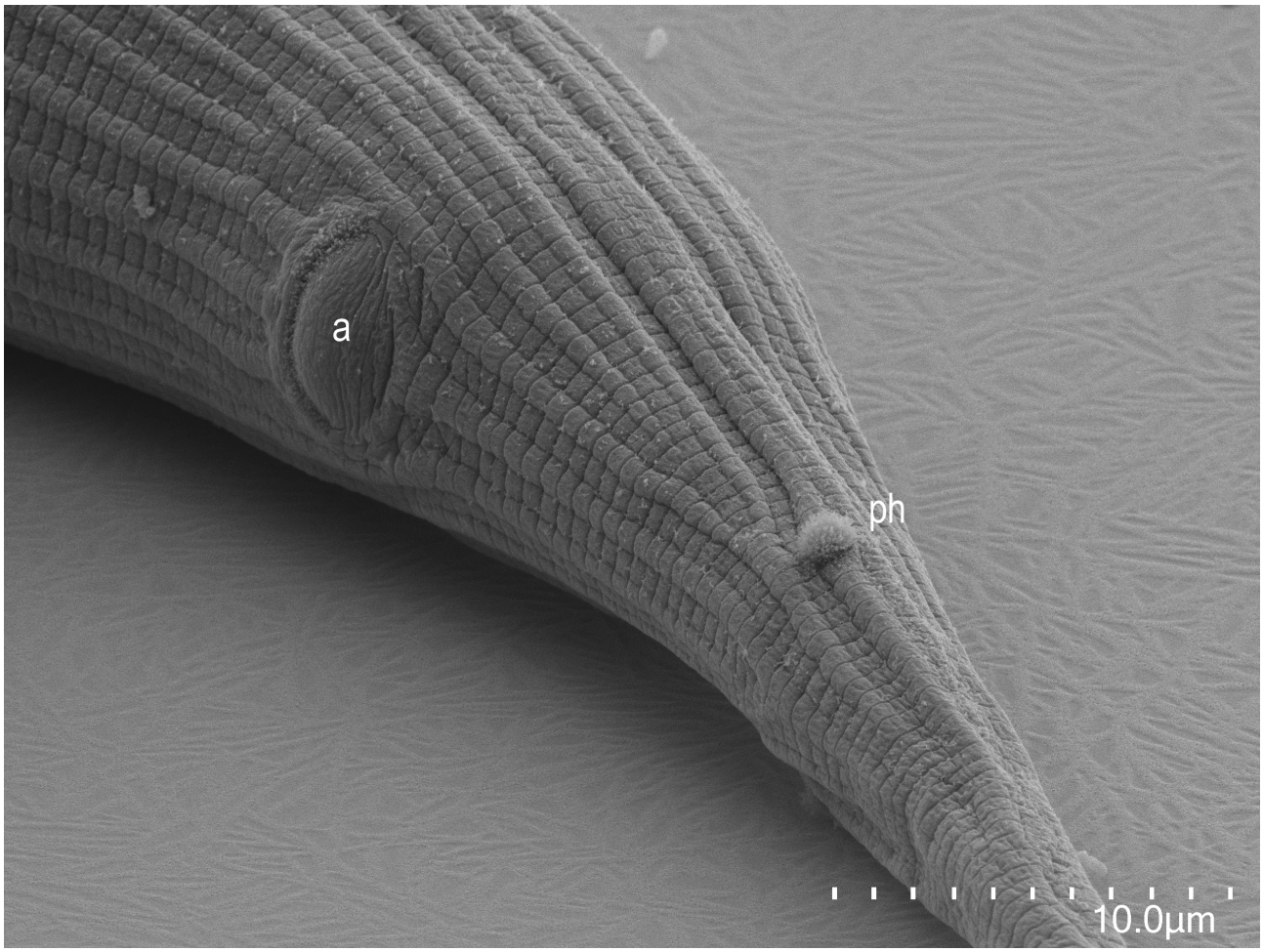


FIGURE 29. Scanning electron photomicrograph of *Pristionchus musae* **n. sp.**, showing anus (a) and phasmid (ph).

***Pristionchus auriculatae* n. sp.**
urn:lsid:zoobank.org:act:57A7DF1B-E33B-4F73-8A6F-6A3ABCBC2600

Etymology. The species name is derived from the substrate from which the species was recovered. The type strain was isolated from a rotting fig of *Ficus auriculata* from soil substrate.

Measurements. See Table 5

Adult. General characters are as described above for the *triformis*-group.

Stenostomatous form. Cheilostom, gymnostom, pro-mesostegostom and telostegostom as described above. 12-plated form not found. Arrangement of tooth and denticles in metastegostom as described above; dorsal movable tooth triangular, anterior end slightly curved, tooth sometimes directed anteriorly; left subventral ridge with three minute, blunt or pointed adventitious denticles on a plate; right subventral ridge with one or two distal rounded or blunt adventitious denticles. Whole stoma including the tooth and denticles is narrow and more pointed in males.

Eurystomatous form. Cheilostom as described above, anterior half of each cheilostomal plate often split into two tips to form 12-flapped form. Six-flapped form (without split cheilostomal plate) rare. Gymnostom in both types 1 and 2 short and thick, forming cuticular ring. Pro-mesostegostom well-developed, internally overlapping with posterior end of gymnostom; type 1 form short, somewhat flattened and bearing weak serrates at anterior end;

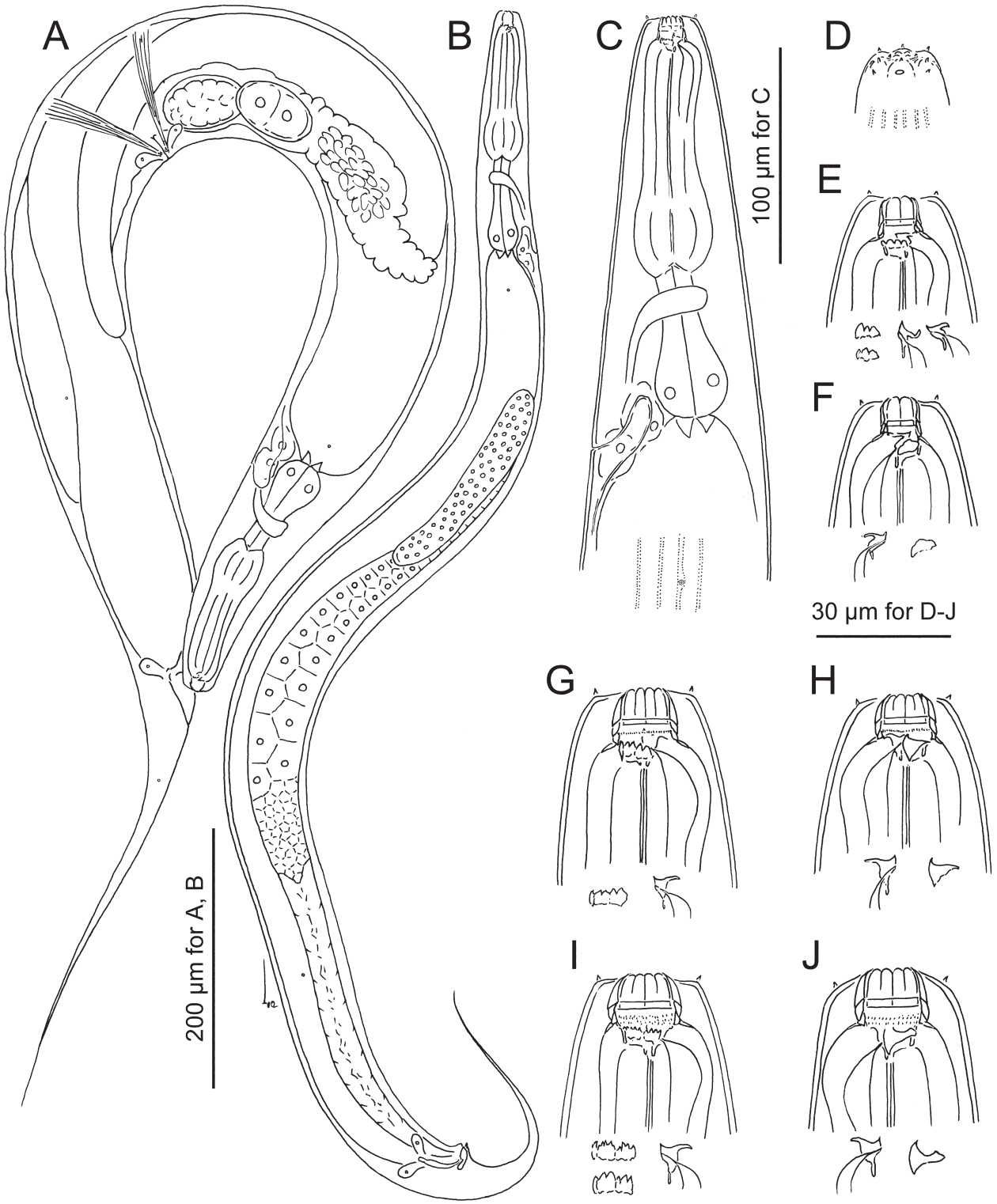


FIGURE 30. *Pristionchus auriculatae* **n. sp.** A: Right lateral view of adult female. B: Right lateral view of adult male. C: Anterior part of adult eurystomatous female in left lateral view. D: Left lateral view of male head region showing labial sensilla, cephalic papillae and amphid. E, F: Stomal region of adult female in left lateral (E) and right lateral (F) view. G, H: Stomal region of eurystomatous female with pro-mesostegostomatal serrates (type 1) in left (G) and right (H) lateral view. I, J: Stomal region of eurystomatous female with pro-mesostegostomatal spines (type 2) in left (I) and right (J) lateral view. Morphological variations of teeth and ridges are separately drawn in each subfigure as left subventral ridge (left) and dorsal tooth (right) (E), dorsal tooth (left) and right subventral widge (right) (F), left subventral ridge (left), and dorsal tooth (right) (G), dorsal (left) and right subventral (right) teeth (H), left subventral ridge (left) and dorsal (right) teeth (I), and dorsal (left) and right subventral (right) teeth (J).

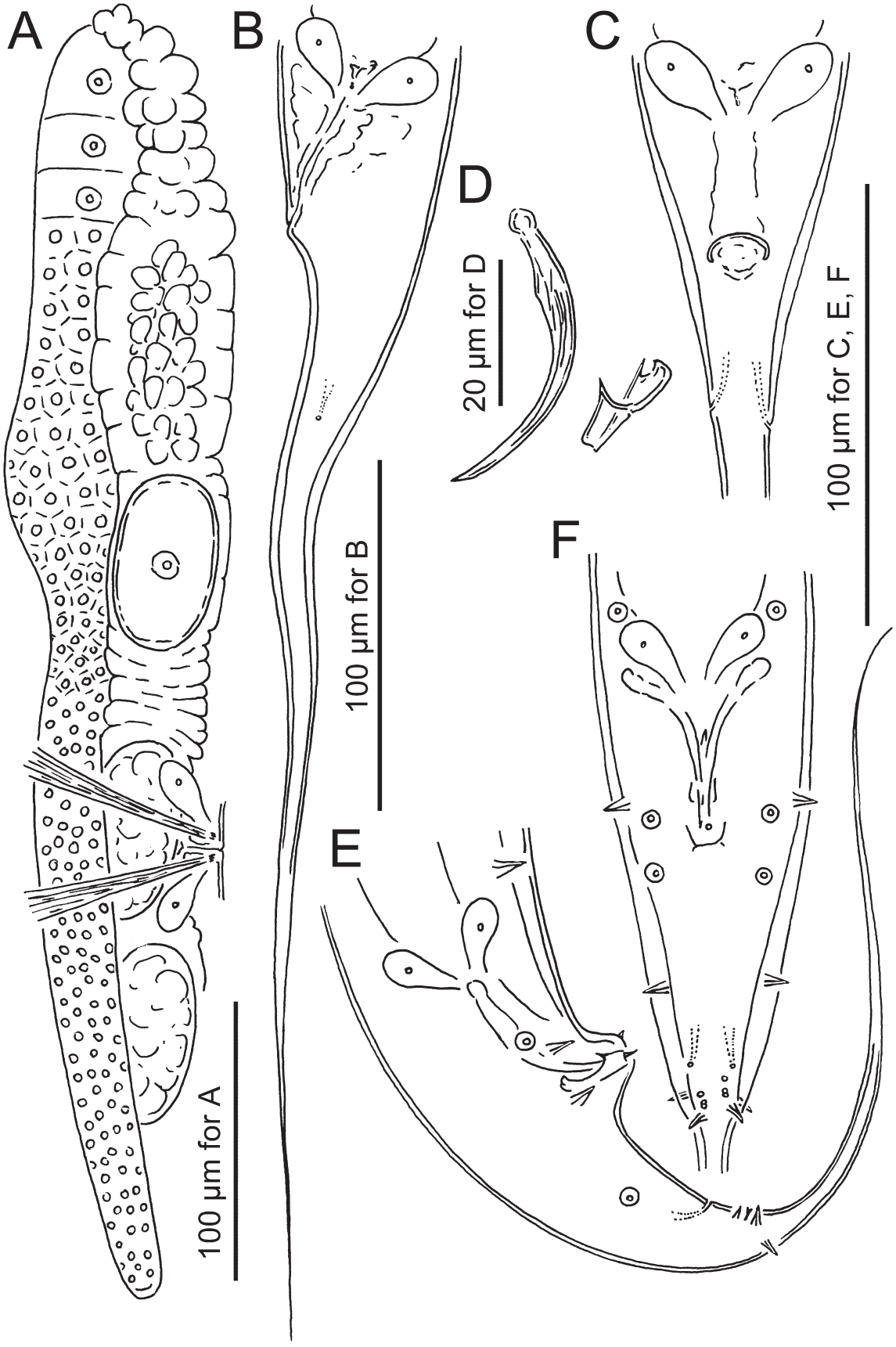


FIGURE 31. *Pristionchus auriculatae* **n. sp.** A: Anterior female gonad in right lateral view. B, C: Female tail in left lateral (B) and ventral (C) view. D: Spicule and gubernaculum in left lateral view: E, F: Male tail in right lateral (E) and ventral (F) view.

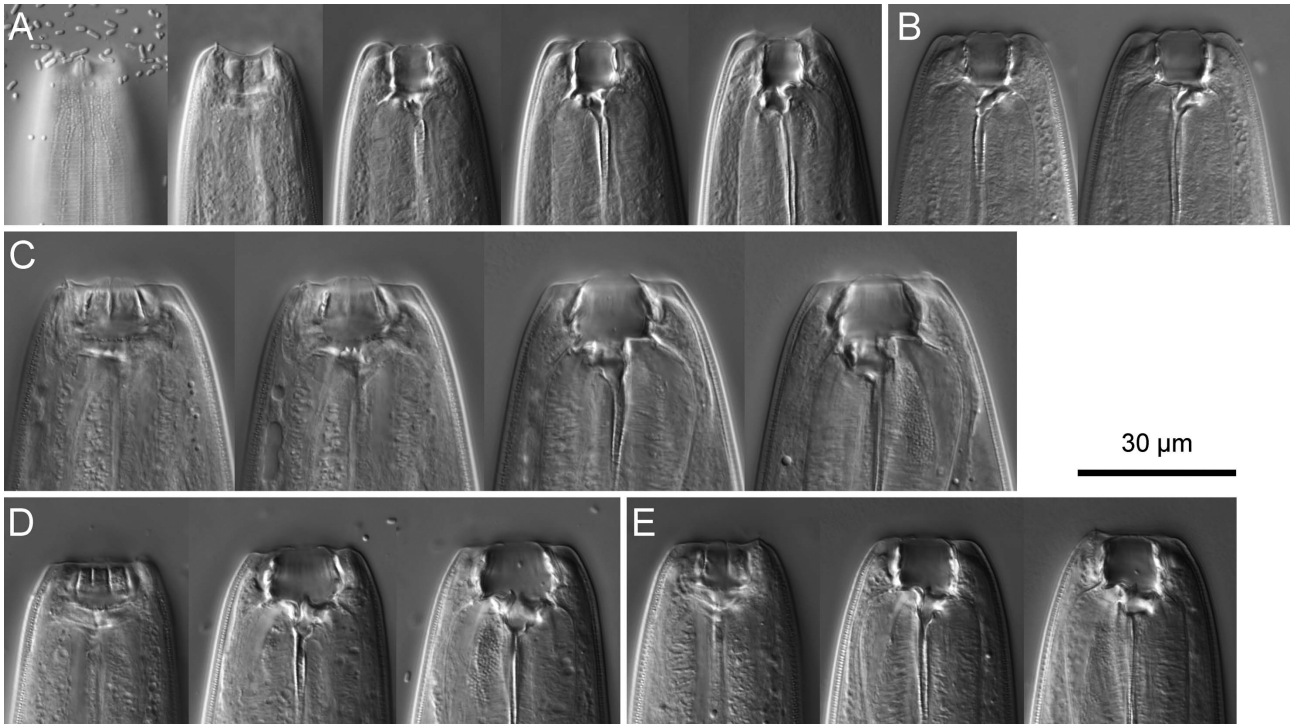


FIGURE 32. Stomal region of *Pristionchus auriculatae* **n. sp.** A: Left lateral view of stenostomatous form in five different focal planes. B: Right lateral view of stenostomatous form in two different focal planes. C: Left lateral view of type 2 eurystomatous form in four different focal planes. D: Right lateral view of type 1 eurystomatous form in three different focal planes. E: Right lateral view of six-plated eurystomatous form in three different focal planes.

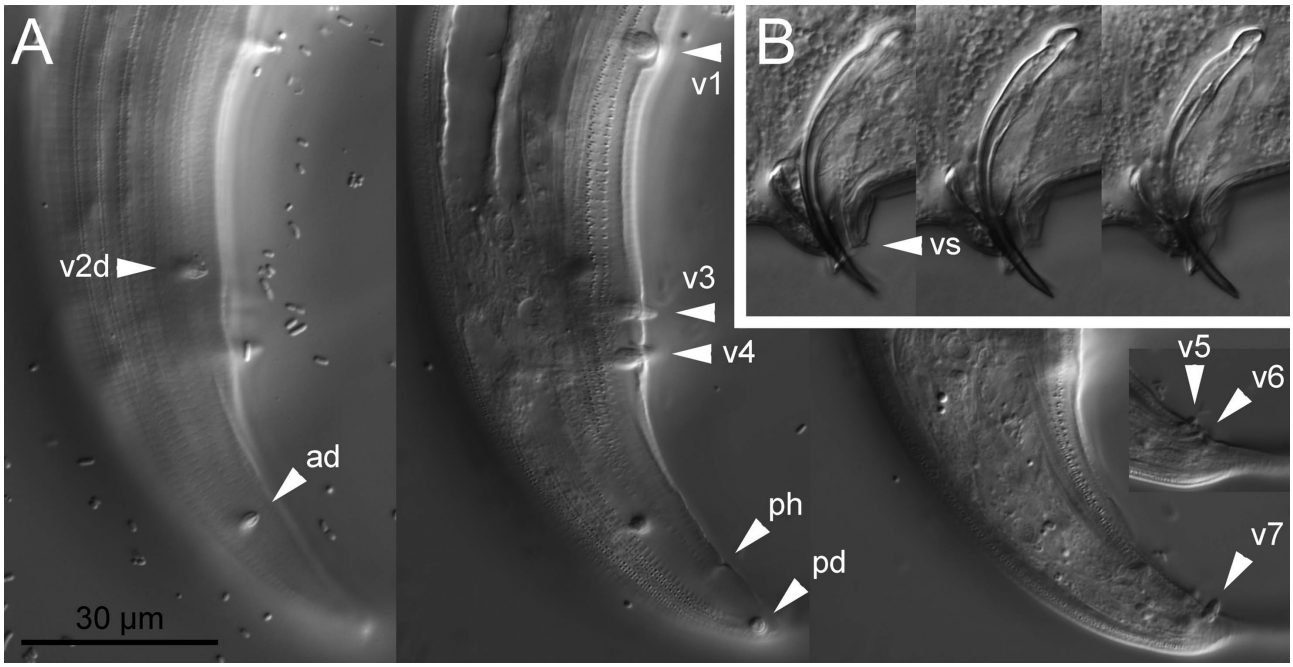


FIGURE 33. Male tail characters of *Pristionchus auriculatae* **n. sp.** A: Right lateral view of tail in four different focal planes. B: Spicule and gubernaculum three different focal planes. Genital papillae are labelled using the terminology of Sudhaus & Fürst von Lieven (2003), and laterally located papillae are indicated with "d".

type 2 with few small spines on inner surface. Arrangement of tooth and denticles in metastegostom as described above; dorsal movable tooth claw-like as typical of the genus; left subventral ridge with three large plates, each often has split tips and extra denticles on middle forming spiny plate; right subventral movable tooth claw-like and sometimes bears an extra peak on the ventral side. Telostegostom as described above.

TABLE 5.

	P. auriculatae (RS5989)		*P. passalidorum* (RS6031)	
character	**stenostomatous male**	**stenostomatous male**	**stenostomatous male**	**stenostomatous female**
n	5	5	5	5
L	1076 ± 92.3 (934–1188)	1076 ± 92.3 (934–1188)	987 ± 87.3 (893–1131)	1404 ± 112.4 (1252–1532)
L’	898 ± 85.8 (777–1000)	898 ± 85.8 (777–1000)	791 ± 88.3 (702–938)	1126 ± 81.5 (1022–1217)
a	12 ± 1.0 (11–13)	12 ± 1.0 (11–13)	12 ± 1.2 (10–13)	13 ± 0.7 (12–14)
b	6.6 ± 0.5 (5.8–7.0)	6.6 ± 0.5 (5.8–7.0)	6.3 ± 0.5 (5.7–7.0)	7.9 ± 0.4 (7.3–8.5)
c	6.1 ± 0.7 (5.3–7.0)	6.1 ± 0.7 (5.3–7.0)	5.1 ± 0.6 (4.4–5.9)	5.1 ± 0.2 (4.8–5.4)
c’	4.2 ± 0.6 (3.7–5.1)	4.2 ± 0.6 (3.7–5.1)	4.7 ± 0.5 (3.9–5.1)	7.4 ± 1.2 (5.9–8.7)
ant. stoma length (cheilo- + gymnostom)	8.7 ± 0.5 (8.0–9.4)	8.7 ± 0.5 (8.0–9.4)	9.4 ± 0.7 (8.3–10.3)	10.3 ± 1.1 (9.1–11.5)
total stoma length	13.7 ± 0.5 (13.1–14.5)	13.7 ± 0.5 (13.1–14.5)	14.3 ± 0.5 (13.7–15.0)	15.3 ± 1.6 (12.7–16.9)
stoma width	6.3 ± 0.6 (5.5–7.0)	6.3 ± 0.6 (5.5–7.0)	6.6 ± 0.7 (5.8–7.6)	8.0 ± 0.8 (7.2–9.0)
ant. pharynx length (pro- + metacorpus)	90 ± 7.4 (84–103)	90 ± 7.4 (84–103)	87 ± 2.7 (84–91)	102 ± 3.4 (99–108)
post. pharynx length (isthmus + basal bulb)	58 ± 2.6 (56–62)	58 ± 2.6 (56–62)	57 ± 4.1 (52–62)	65 ± 4.6 (60–69)
total length pharynx	148 ± 6.3 (142–159)	148 ± 6.3 (142–159)	143 ± 5.2 (139–151)	168 ± 6.6 (161–177)
ant./total pharynx %	61 ± 2.6 (58–65)	61 ± 2.6 (58–65)	61 ± 1.8 (58–63)	61 ± 1.6 (59–63)
median bulb diameter	26 ± 1.7 (24–28)	26 ± 1.7 (24–28)	28 ± 1.5 (27–30)	33 ± 1.2 (32–35)
terminal bulb diameter	24 ± 1.4 (22–26)	24 ± 1.4 (22–26)	24 ± 2.4 (21–27)	30 ± 1.8 (27–32)
neck length (incl. stoma)	162 ± 6.3 (156–172)	162 ± 6.3 (156–172)	158 ± 4.8 (153–164)	178 ± 6.9 (172–189)
excretion pore	156 ± 10.6 (141–168)	156 ± 10.6 (141–168)	152 ± 13.3 (142–175)	172 ± 14.4 (155–191)
nerve ring	112 ± 2.4 (109–115)	112 ± 2.4 (109–115)	116 ± 4.5 (110–122)	125 ± 10.9 (114–140)
testis length	660 ± 66.6 (574–747)	660 ± 66.6 (574–747)	552 ± 68.4 (492–665)	-
anterior gonad length	-	-	-	313 ± 32.3 (286–363)
posterior gonad length	-	-	-	271 ± 54.5 (233–366)
ant. end to vulva distance	-	-	-	631 ± 53.2 (573–691)
vulva to anus distance	-	-	-	492 ± 32.4 (451–536)
T or V	61 ± 1.8 (59–63)	61 ± 1.8 (59–63)	56 ± 2.6 (52–59)	45 ± 1.0 (43–46)
max. body diameter	88 ± 9.1 (79–100)	108 ± 5.7 (99–113)	85 ± 9.2 (76–98)	99 ± 10.0 (83–108)
cloacal or anal body diameter	42 ± 2.5 (40–46)	38 ± 2.4 (34–40)	42 ± 2.3 (40–44)	37 ± 1.7 (35–40)
tail length	179 ± 20.0 (158–203)	279 ± 32.3 (230–315)	196 ± 16.6 (175–221)	301 ± 16.5 (288–327)
spicule length (curve)	49 ± 2.8 (44–51)	-	43 ± 1.2 (41–44)	-
spicule length (chord)	41 ± 2.3 (38–43)	-	36 ± 0.7 (35–37)	-
gubernaculum length	18 ± 0.8 (17–19)	-	17 ± 0.7 (16–17)	-

Male. Paired papillae and the phasmid are arranged as <v1, v2d, v3, co, v4, ad, (ph, v5, v6, v7, pd)>, where v1 located about 1 CBD anterior to co; v2d just less than 1/5 CBD anterior to co; v3 adcloacal; v4 at 1/3 CBD posterior to co, *i.e.*, v2d, v3, co and v4 are close to each other; ad about 1 CBD posterior to co; ph 2/3 distance from ad and root of tail spike; v5–v7 forming triplet, just posterior to ph; and pd at level of triplet, *i.e.*, ph, triplet papillae and pd are close to each other. v1, v3, v4 and ph subventral, v2d and ad lateral, v5–7 ventral and pd subdorsal in male tail. General shape of spicule and gubernaculum as described above.

Female. Gonadal characters of female as described above. Tail elongate conoid with slightly filiform terminus, *i.e.*, posterior half of tail is more elongated compared to anterior part. Phasmid ventro-laterally located at about 1.5 ABD posterior to anal opening.

Diagnosis and relationships. *Pristionchus auriculatae* **n. sp.** is characterized by the slightly anteriorly curved triangular dorsal tooth and the right and left subventral ridges with three denticles of stenostomatous form, right subventral stegostomal tooth may have a blunt peak on the ventral side in eurystomatous form, sparsely serrated pro-mesostegostomal wall in type 2 eurystomatous form, the arrangement of male genital papillae, <v1, v2d, v3, co, v4, ad, (ph, v5, v6, v7, pd)> where v2d, v3, co and v4 and ph, v5, v6, v7, pd are close to each other, respectively, and a long spike occupying more than 2/3 of tail length of male tail. In addition, the arrangement of v1, v2d and ad genital papillae, where the distance between v1 and v2d is almost same as that between v2d and ad is characteristic to this species.

Pristionchus auriculatae **n. sp.** is typologically close to *P. hoplostomus*. The new species and *P. hoplostomus* share triangular and pointed dorsal tooth in stenostomatous form, right subventral tooth in eurystomatous form, sometimes have an extra peak on its ventral side, and the arrangement of genital papillae, closely located v2d, v3, co and v4 and pd overlapping with v5–7 triplet. The new species is distinguished from *P. hoplostomus* by the absence *vs.* presence of 12-plated stenostomatous form, right subventral ridge in stenostomatous form, with blunt *vs.* pointed denticles, and the arrangement of posterior four pairs of genital papillae, *i.e.*, pd is overlapping with triplet *vs.* posterior to v7.

Further, the new species is distinguished from all other species by mating experiments and also characterized by a *ca.* 1,600-bp fragment of the SSU rRNA gene (GenBank accession number MW017216), the sequence of which is distinct from that of all other *Pristionchus* species.

Type host and locality. Isolated from rotting *Ficus auriculata* fruits from soil substrates collected at the Botanical garden, Shanghai - conservatory 1, Shanghai, PRC.

Type material and type strain. Type strain RS5989 frozen at the nematode collection of the MPI Tübingen and available as living culture upon request. Voucher specimens sent to the following museums: Holotype male, Paratype male and female: Museum für Naturkunde Karlsruhe, Germany; Paratype male and female: Swedish Natural History Museum, Stockholm, Sweden; Paratype male and female: University of California in Riverside Nematode Collection (UCRNC), Riverside, CA, USA.

Pristionchus passalidorum n. sp.

urn:lsid:zoobank.org:act:DE5D7130-E1CE-4856-BD75-CAA18814BA00

Etymology. The species name is derived from the family name of the host, bess beetle (Coleoptera: Passalidae).

Measurements. See Table 5

Adult. General characters are as described above for the *triformis*-group.

Stenostomatous form. Cheilostom, gymnostom, pro-mesostegostom and telostegostom as described above. 12-plated form not found. Arrangement of tooth and denticles in metastegostom as described above; dorsal movable tooth triangular, anterior end slightly curved, tooth often directed anteriorly; left subventral ridge with three minute, blunt or pointed adventitious denticles on a plate; right subventral ridge with a distal pointed adventitious denticle. Whole stoma including tooth and denticles narrow and more pointed in males.

Eurystomatous form. Cheilostom as described above, anterior half of each cheilostomal plate often split into two tips to form 12-flapped form. Six-flapped form (without split cheilostomal plate) rare. Gymnostom short and thick in both types 1 and 2, forming cuticular ring. Pro-mesostegostom well-developed, internally overlapping with posterior end of gymnostom; type 1 form short flattened with weak serrates at anterior end; type 2 bearing few small spines on inner surface. Arrangement of tooth and denticles in metastegostom as described above; dorsal movable

tooth claw-like as typical of the genus; left subventral ridge with three large plates, each often has split tips and extra denticles on middle forming spiny plate; right subventral movable tooth claw-like often with extra peak on ventral side. Telostegostom as described above.

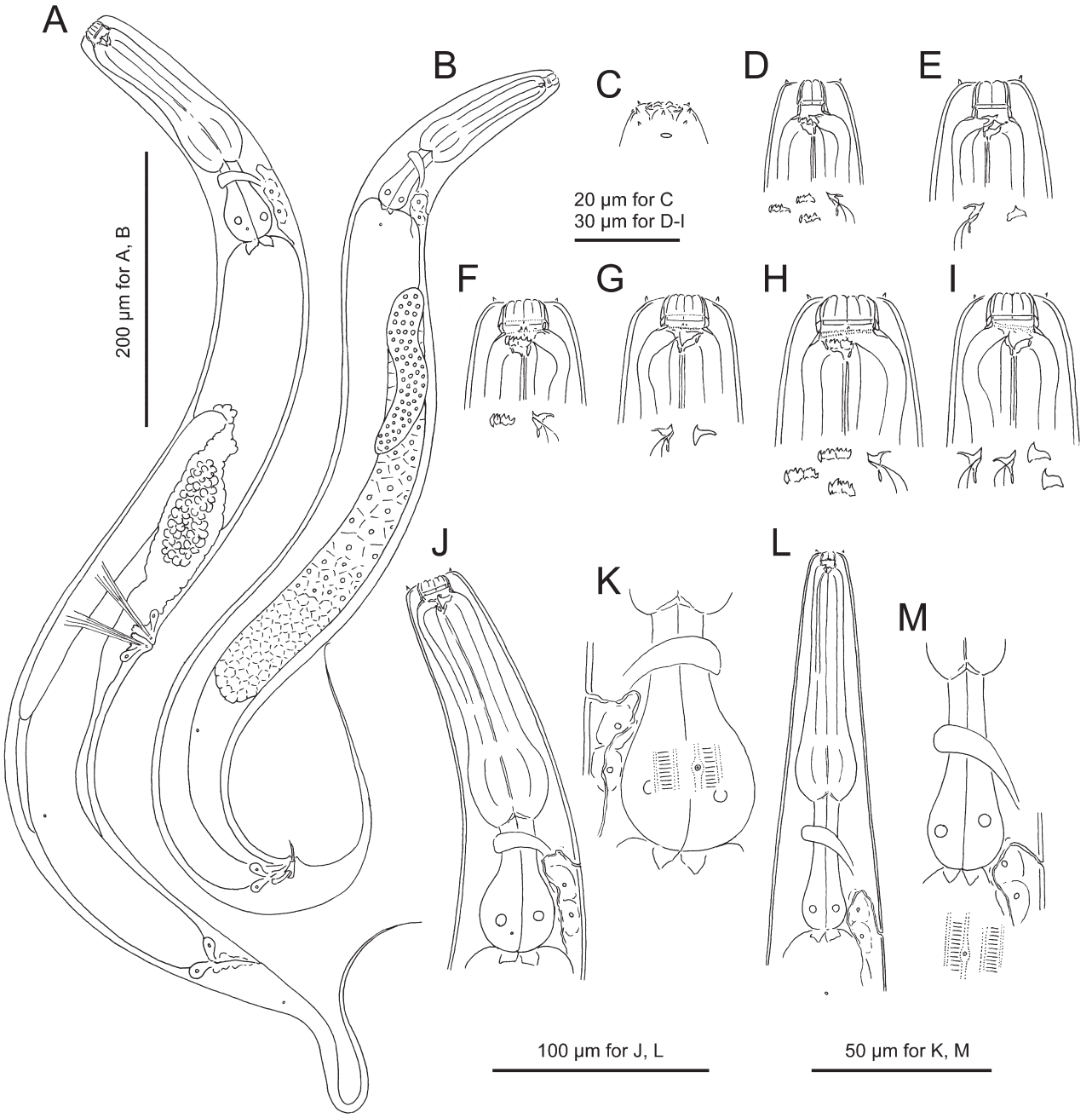


FIGURE 34. *Pristionchus passalidorum* **n. sp.** A: Right lateral view of adult female. B: Right lateral view of adult male. C: Left lateral view of male head region showing labial sensilla, cephalic papillae and amphid. D, E: Stomal region of adult female in left lateral (D) and right lateral (E) view. F, G: Stomal region of eurystomatous female with pro-mesostegostomatal serrates (type 1) in left (F) and right (G) lateral view. H, I: Stomal region of eurystomatous female with pro-mesostegostomatal spines (type 2) in left (H) and right (I) lateral vie. Morphological variations of teeth are separately drawn in each subfigure as left subventral ridge (left and middle) and dorsal tooth (right) (D), dorsal tooth (left) and right subventral ridge (right) (E), left subventral ridge (left), and dorsal tooth (right) (F), dorsal (left) and right subventral (right) teeth (G), left subventral ridge (left and middle) and dorsal tooth (right) (H), and dorsal (left and middle) and right subventral (right) teeth (I). J: Anterior part of adult eurystomatous female in right lateral view. K: Relative position of deirid and secretory-excretory pore of eurystomatous female. L: Anterior part of adult stenostomatous male in right lateral view. M: Relative position of deirid and secretory-excretory pore of stenostomatous male.

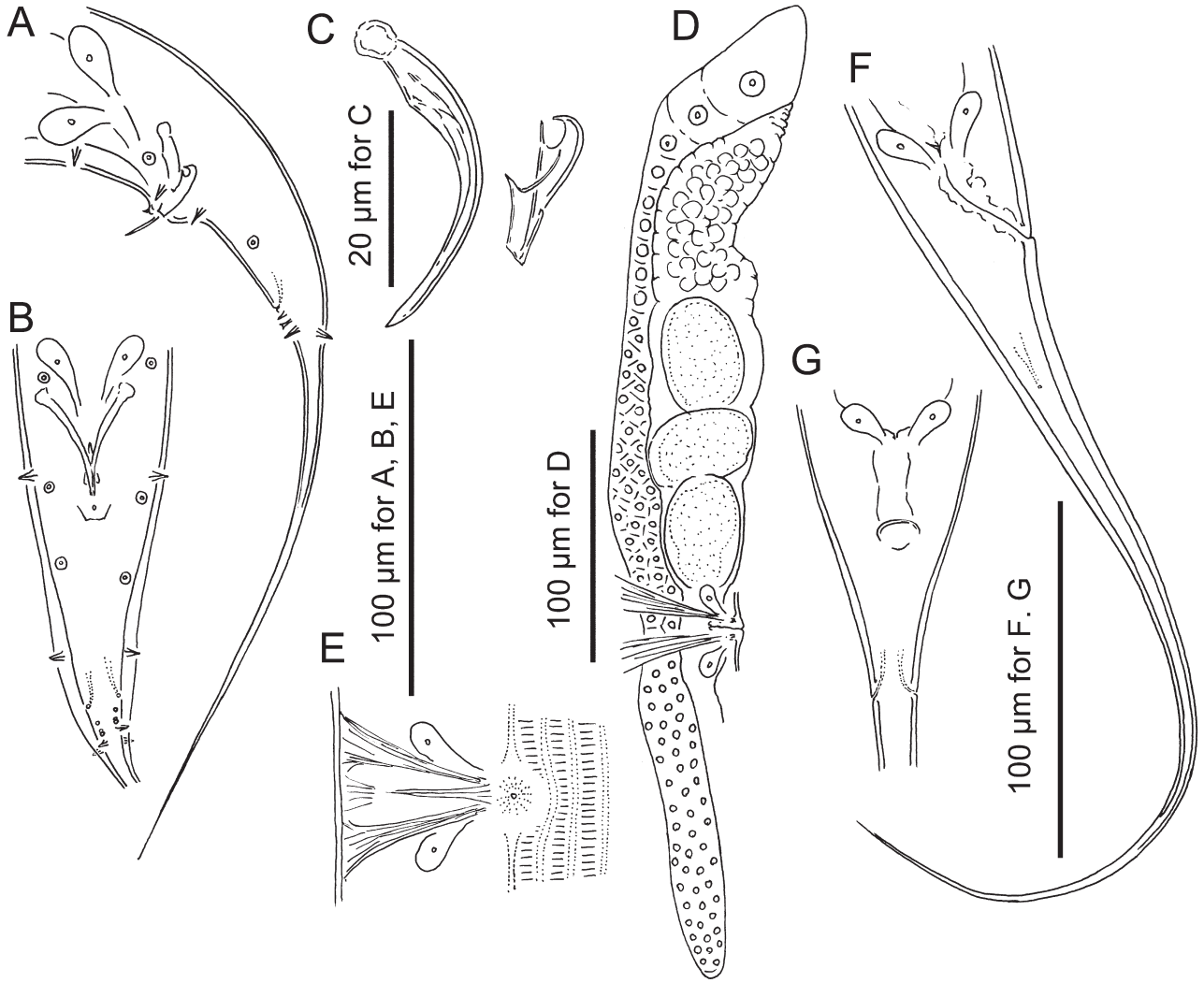


FIGURE 35. *Pristionchus passalidorum* **n. sp.** A: Male tail in left lateral (A) and ventral (B) views. C: Spicule and gubernaculum in left lateral view. D: Anterior female gonad in right lateral view. E: Vulval region in ventral view. F, G: Female tail in right lateral (F) and ventral (G) views.

Male. Paired papillae and phasmid are arranged as <v1, v2d, v3, co, v4, ad, (ph, v5, v6, v7, pd)>, where v1 located about 1 CBD anterior to co; v2d just less than 1/5 CBD anterior to co; v3 adcloacal; v4 at 1/3 CBD posterior to co, *i.e.*, v2d, v3, co and v4 are close to each other; ad about 1 CBD posterior to co; ph 2/3 distance from ad and rest of tail spike; v5–v7 forming triplet, just posterior to ph; and pd around the level of v7, *i.e.*, ph, triplet papillae and pd close to each other. v1, v3, v4 and ph subventral, v2d and ad lateral, v5–7 ventral, pd subdorsal in the male tail. General shape of spicule and gubernaculum as described above.

Female. Gonadal characters of female as described above. Tail elongate conoid with slightly filiform terminus, *i.e.*, the posterior half of tail is more elongated compared to anterior part. Phasmid ventro-laterally located at about 1.5 ABD posterior to anal opening.

Diagnosis and relationships. *Pristionchus passalidorum* **n. sp.** is characterized by the slightly anteriorly curved triangular dorsal tooth, the right and left subventral plates with one and three denticle(s) in stenostomatous form, right subventral stegostomal tooth without extra peak in eurystomatous form, sparsely serrated pro-mesostegostomal wall in type 2 eurystomatous form, and the arrangement of male genital papillae, <v1, v2d, v3, co, v4, ad, (ph, v5, v6, v7, pd)> where v2d, v3, co and v4 and ph, v5, v6, v7, pd are close to each other, respectively, and a long spike occupying more than 2/3 of tail length of male tail. *Pristionchus passalidorum* **n. sp.** is typologically similar to *P. hoplostomus* and *P. musae* **n. sp.** The new species and these two species share a triangular and pointed dorsal

tooth with slightly anteriorly curved tip in stenostomatous form and the arrangement of genital papillae, closely located v2d, v3, co and v4. The new species is distinguished from *P. hoplostomus* by the absence *vs.* presence of 12-plated stenostomatous form, left subventral ridges of stenostomatous form, clearly pointed *vs*. bluntly pointed, right subventral plate of stenostomatous form, single *vs*. multiple peaks, right subventral tooth of eurystomatous form, without *vs*. with extra peak. The new species is distinguished from all other species by mating experiments and also characterized by a *ca.* 1,600-bp fragment of the SSU rRNA gene (GenBank accession number MW017223), the sequence of which is distinct from that of all other *Pristionchus* species.

Type host and locality. Bessbug (Coleoptera: Passalidae) at Ailaoshan Field Station CAS, Yunnan province, PRC.

Type material and type strain. Type strain RS6031, frozen at the nematode collection of the MPI Tübingen and available as living culture upon request. Voucher specimens sent to the following museums: Holotype male, Paratype male and female: Museum für Naturkunde Karlsruhe, Germany; Paratype male and female: Swedish Natural History Museum, Stockholm, Sweden; Paratype male and female: University of California in Riverside Nematode Collection (UCRNC), Riverside, CA, USA.

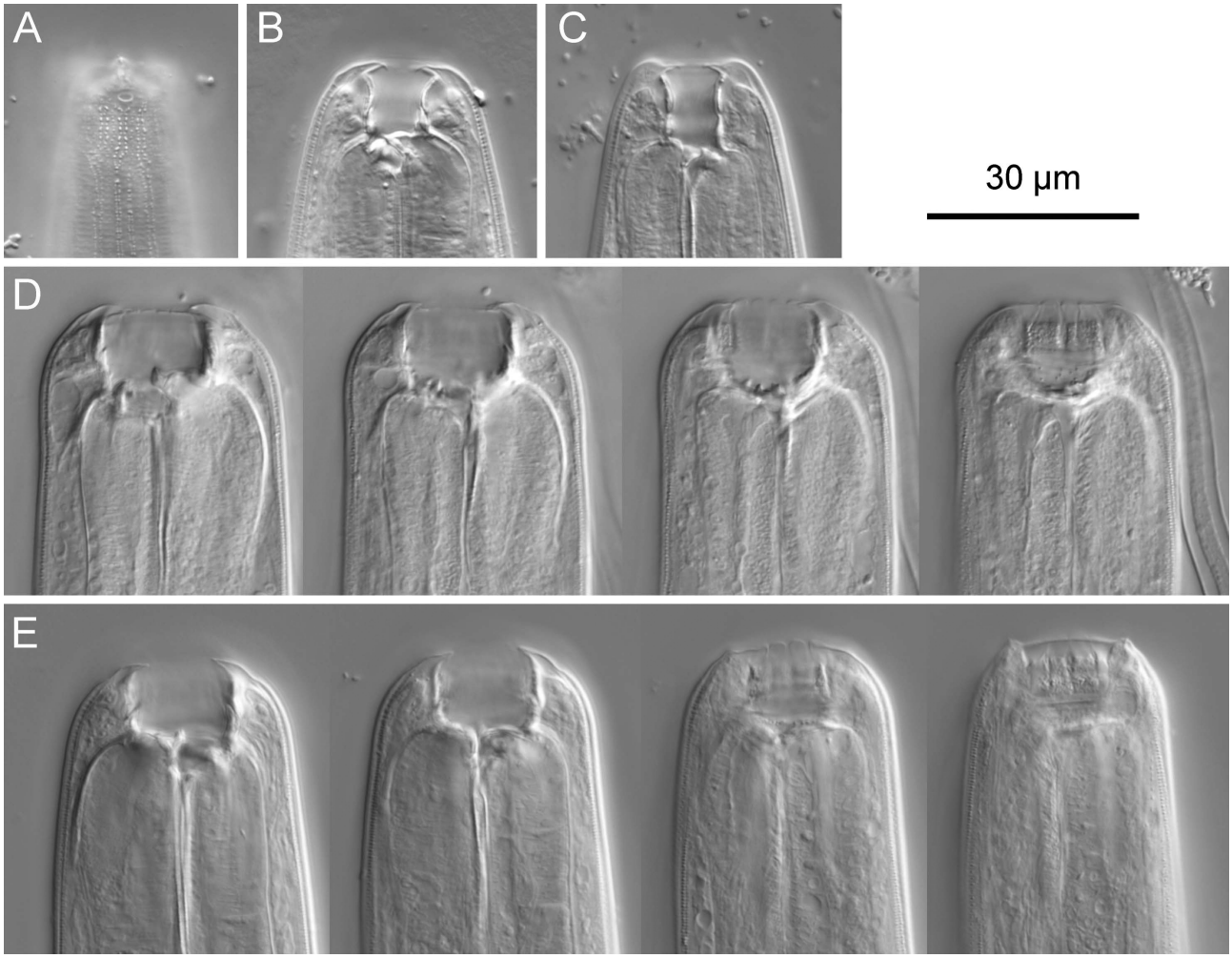


FIGURE 36. Stomal region of *Pristionchus passalidorum* **n. sp.** A: Surface of stenostomatous female in left lateral view. B: Dorsal tooth and left subventral ridge of stenostomatous in left lateral view. C: Dorsal tooth and right subventral ridge of stenostomatous form in right lateral view Right lateral view of stenostomatous form in two different focal planes. D: Left lateral view of type 2 eurystomatous form in four different focal planes. E: Right lateral view of type 1 eurystomatous form in four different focal planes.

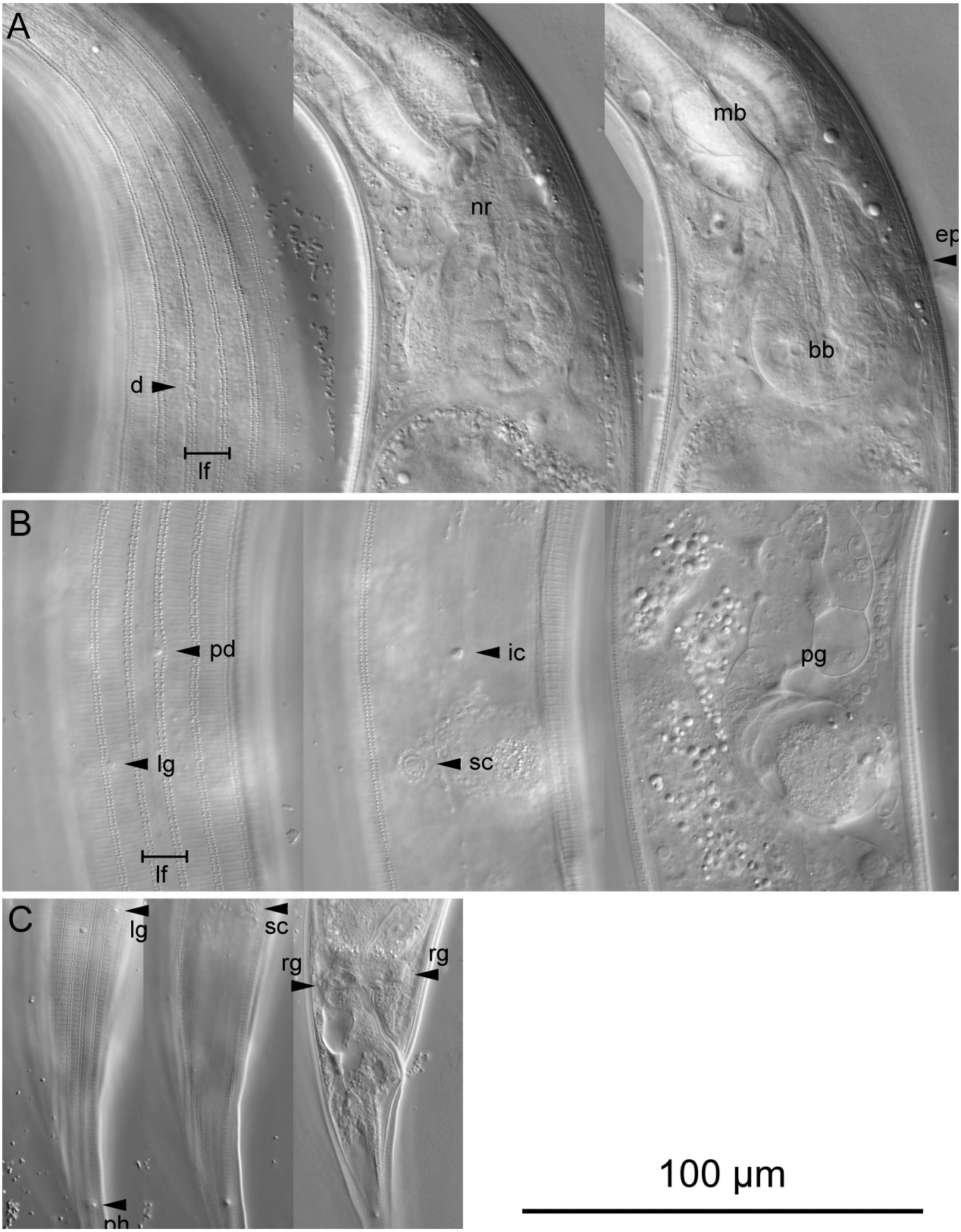


FIGURE 37. *Pristionchus passalidorum* **n. sp.** A: Right lateral view of posterior pharynx region in three different focal planes where deirid (d), lateral field (lf), nerve ring (nr), median bulb (mb) secretory-excretory pore (ep) and basal bulb (bb) are suggested. B: Left lateral view of postdeirid region in three different focal planes where postdeirid (pd) and its internal connection (ic), lateral gland opening (lg) and its associated secretory cell (sc) and posterior gonad (pg) are suggested. C: Female anal region in three different focal planes where lateral gland and its associated secretory cell, phasmid (ph) and rectal glands (rg) are indicated.

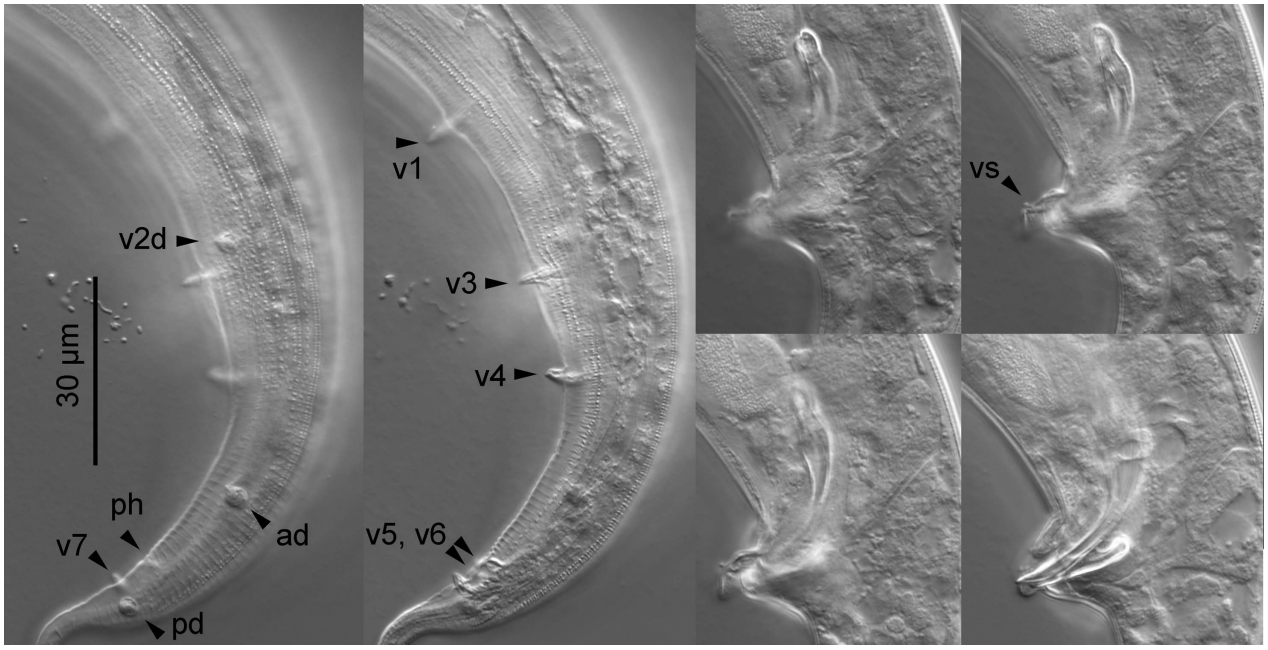


FIGURE 38. Male tail characters of *Pristionchus passalidorum* **n. sp.** Left lateral view of tail in two different focal planes and spicule and gubernaculum four different focal planes. Genital papillae are labelled using the terminology of Sudhaus & Fürst von Lieven (2003), and laterally located papillae are indicated with "d".

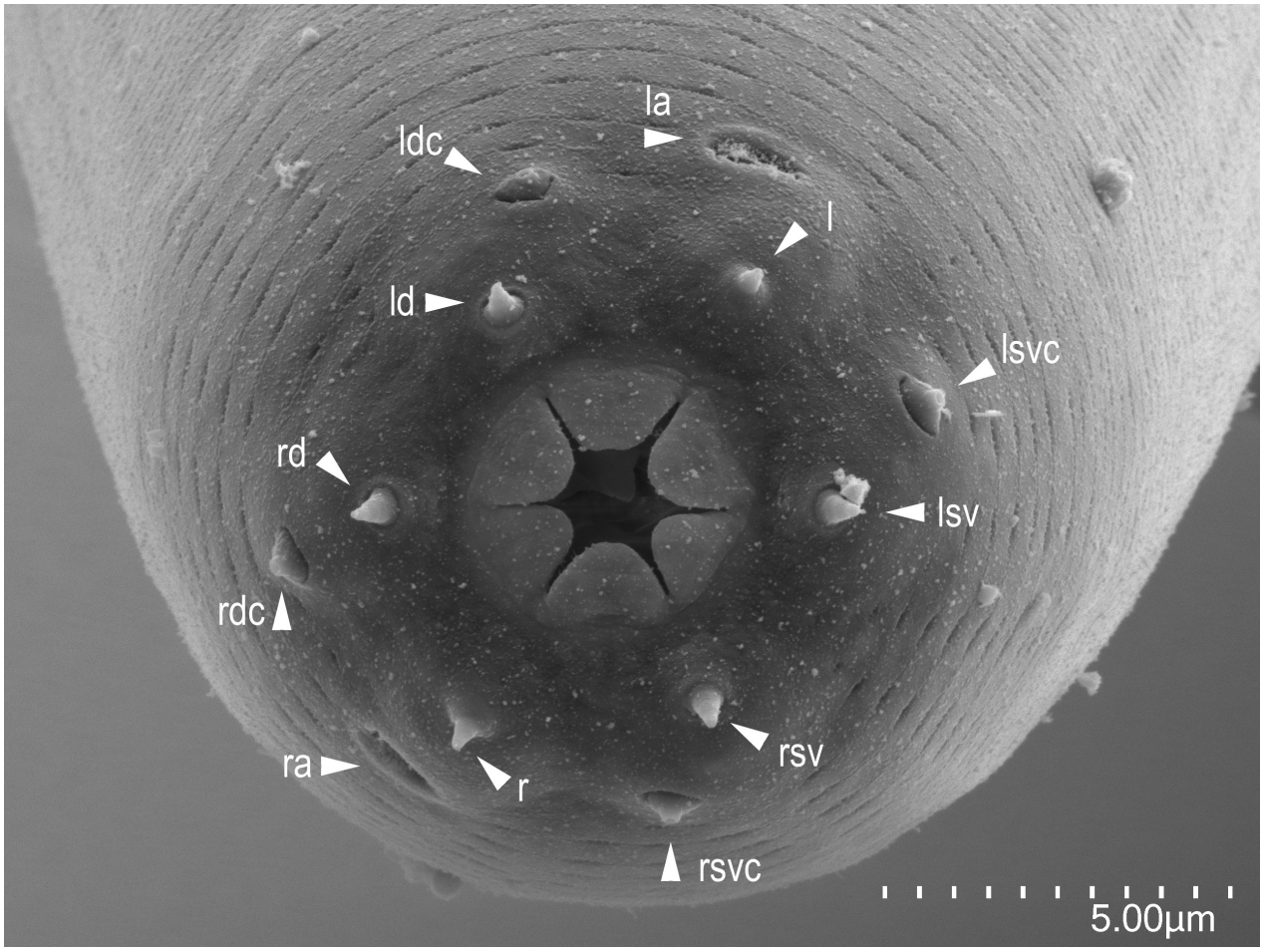


FIGURE 39. Scanning electron photomicrograph of *Pristionchus passalidorum* **n. sp.** stenostomatous male *en face* view where left and right amphids (la and ra), left and right subventral (lsv and rsv), left and right lateral (l and r) and left and right dorsal (ld and rd) labial sensilla, and left and right subventral (lsvc and rsvc) and dorsal (ldc and rdc) cephalic sensilla are indicated.

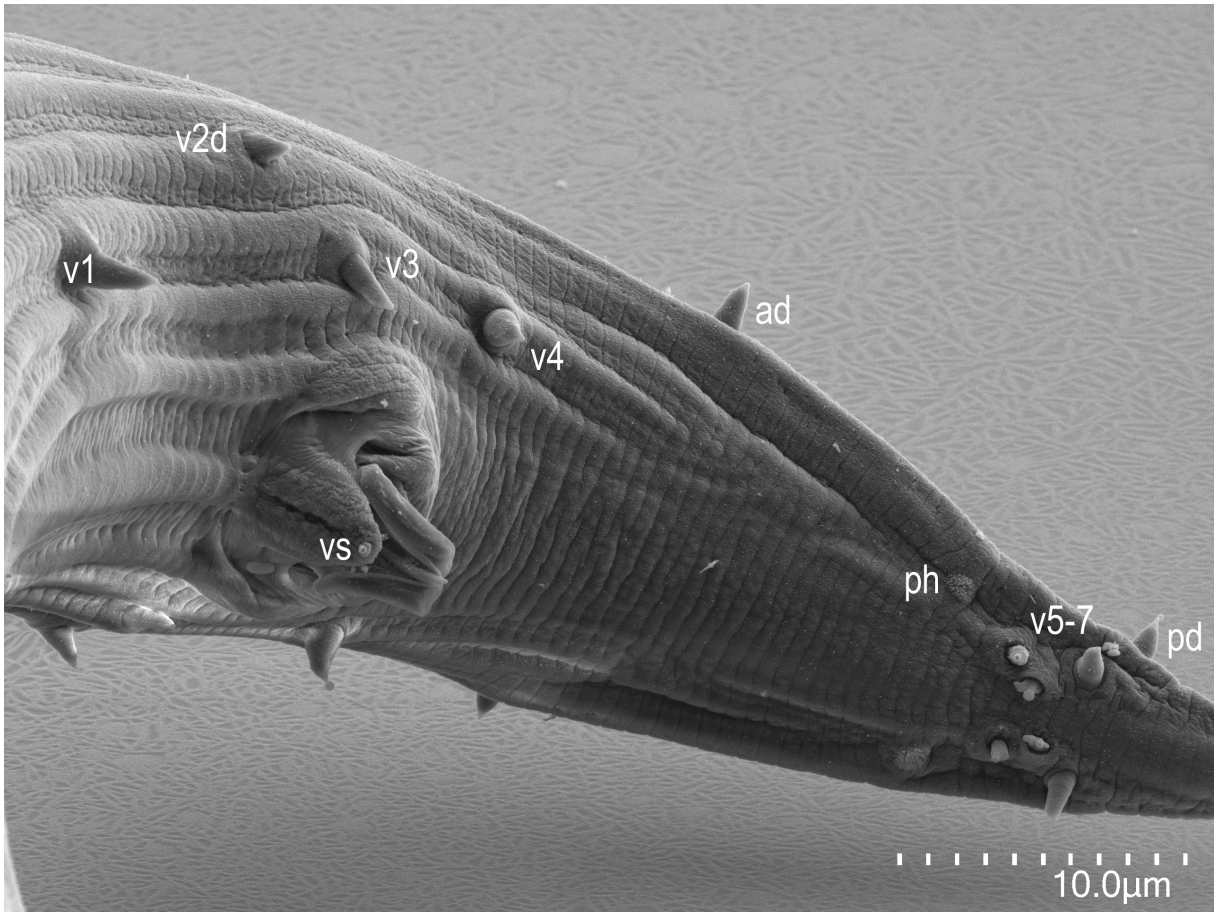


FIGURE 40. Scanning electron photomicrograph of *Pristionchus passalidorum* **n. sp.** male tail with the position of genital papillae shown according to the terminology of Sudhaus and Fürst von Lieven (2003).

Species description of *P. dorci* n. sp. and *P. purgamentorium* n. sp. of the *Pristionchus maupasi*-group

Adult. Body cylindrical, stout, *i.e.*, body length-maximum body diam. ratio (a value) is usually ranging from 10–14, depending on culture condition; cuticle thick with fine annulation and longitudinal striations. Head without apparent lips, six mound-like anteriorly directed expansions, one on each sector. Six labial sensilla, one on each lip sector (on the mound); four small papilliform cephalic papillae in males on right and left subventral and right and left dorsal sectors; amphidial apertures located on lateral sector, slightly dorsally shifted, at level of margin of cheilo- and gymnostom. Stomal dimorphism present, and detailed morphology of the stoma is described below. Dorsal pharyngeal gland clearly observed, penetrating dorsal tooth to gland opening. Anterior pharynx longer than posterior pharynx; procorpus muscular, stout, occupying half to two-thirds of corresponding body diameter; metacorpus muscular, forming median bulb; isthmus narrow, not muscular; basal bulb glandular. Pharyngo-intestinal junction (cardia) clearly observed, well-developed. Intestine simple tube, not forming pre-rectum, extended posteriorly from cardia to rectum; three (two subventral and one dorsal) rectal gland cells observed at distal end of intestine (margin between intestine and rectum). Secretory-excretory pore not conspicuous, ventrally located at level of isthmus to pharyngo-intestinal junction, excretory duct extending anteriad and reflexed back to position of pore; two large secretory-excretory cells around the excretory duct. Deirid observed laterally on lateral field, located at the level around the posterior end of basal bulb to pharyngo-intestinal junction to a half body diameter posterior to the junction, ca. 0.5–1 body diam. posterior to secretory-excretory pore. Hemizonid not observed. Lateral glands (small pores connected to secretory cell) on lateral body surface, with positions inconsistent among individuals, numbering 5 to 8 for males and 9 to 13 for females. Postdeirid at anterior part of *vas deferens* in male and the posterior end of posterior gonad in female, on the same striation with deirid (= lateral field) or on the adjacent striation or the second dorsally neighboring striation to lateral field.

Stenostomatous form. Cheilostom consisting of six per- and interradial plates. Incision between plates not always distinguished. Anterior end of each plate rounded and elongated to project from stomaal opening and form a small flap. Split cheilostomal plate not observed. Gymnostom short, cuticular ring-like anterior end overlapping cheilostom internally; dorsal gymnostomal wall with mound-like expansion probably derived from metastegostom which gives an appearance of thickened dorsal gymnostomal wall compared to ventral side. Stegostom separated into three subsections: pro-meso, meta, and telostegostom. Pro-mesostegostom forming a weakly cuticularized ring internally overlapping with gymnostoma to connect gymnostom and metastegostom. Metastegostom bearing conspicuous and movable triangular or flint-shaped dorsal tooth with strongly sclerotized surface giving an appearance of an inverted V-shape in light microscopy in lateral view; pointed left subventral ridge with three minute adventitious denticles on a plate, most ventral denticle often masked by remaining two in lateral view; pointed right subventral ridge, often with distinct distal adventitious denticle(s). Telostegostom forming weakly sclerotized cup-like cavity connecting stoma and pharynx. Whole stoma including tooth and denticles narrow and more pointed in males.

Eurystomatous form. Cheilostom divided into six well-distinctive per- and interradial plates. Anterior end of each plate rounded and elongated to stick out from stomal opening and form a small flap. Split cheilostomal plate not observed. Gymnostom with thick cuticle, forming short, ring-like tube with more heavily sclerotized wall in the posterior; anterior end of gymnostom internally overlapping posterior end of cheilostomatal plates; lacking the serrates at anterior end of gymnostom. Pro-mesostegostom forming a weakly cuticularized ring connecting gymnostom and metastegostom. Metastegostom bearing large claw-like dorsal tooth; claw-like right subventral tooth; and three left subventral denticles, where the tip of each denticle sometimes splits into two or three small ridges. Telostegostom forming weakly sclerotized cup-like cavity connecting stoma and pharynx.

Male. Whole body ventrally arcuate, strongly ventrally curved at tail region when killed by heat. Testis single, ventrally located, anterior part reflexed to right or left side; spermatogonia arranged in three to five rows in reflexed part, well-developed spermatocytes arranged as three to four rows in anterior two-thirds of main branch, mature amoeboid spermatids arranged in multiple rows in proximal part of gonad. *Vas deferens* not clearly separated from other parts of gonad. Posterior end of *vas deferens* and rectum fused to form a cloacal tube. Spicules paired, separate; spicules smoothly curved in ventral view, adjacent to each other for distal third of their length, each smoothly tapering to pointed distal end; spicule in lateral view smoothly ventrally arcuate, giving spicule about 100° curvature, oval manubrium at anterior end, lamina/calomus complex smoothly tapering to pointed distal end. Gubernaculum conspicuous, about one-third of spicule length, broad anteriorly such that dorsal wall is slightly recurved with dorsal and ventral walls separate at 40–50° angle at posterior end; dorsal side of gubernaculum possessing single, membranous, anteriorly directed process and lateral pair of more sclerotized, anteriorly and obliquely ventrally directed processes. In lateral view, anterior half of gubernaculum with two successive curves separated by anteriorly and obliquely ventrally directed process, with anterior terminal curvature highly concave and almost closed, with deep posterior curvature being one-third of gubernaculum length; posterior half forming tube-like process enveloping spicules. Cloacal opening (co) slit-like in ventral view; one small, ventral, single genital papilla (vs) on anterior cloacal lip. All nine paired genital papillae papilliform. Paired papillae and phasmid arranged as <v1, v2, v3d, co, v4, ad, ph, (v5, v6, v7, pd)>, where v1 located about 1.5 CBD anterior to co; v2 about 0.5 CBD anterior to co; v3d less than 1/10 CBD posterior to v2, *i.e.*, about 1/3 CBD anterior to co; v4 at about 1/3 CBD posterior to co; ad about 1 CBD posterior to co; ph midway between ad and root of tail spike; v5–v7 forming triplet, just posterior to ph; and pd level of ventral triplet (v5–v7). v1, v2, v4 and ph subventral, v3d and ad lateral, v5–7 ventral and pd subdorsal in the male tail. General shape of spicule and gubernaculum as described above, spicule relatively thin, gubernaculum relatively low (flattened). Male tail spike long.

Female. Body relaxed or weakly ventrally arcuate when killed by heat. Gonad didelphic, amphidelphic; each gonadal system arranged from vulva/vagina as uterus, oviduct, and ovary; anterior gonad right of intestine, with uterus and oviduct extending ventrally and anteriorly on right of intestine and with totally reflexed (=antidromous reflexion) ovary extending dorsally on left of intestine; oocytes mostly arranged in three to four or more rows in distal two-thirds of ovary and in double or single row in rest of ovary, distal tips of each ovary reaching oviduct of opposite gonad branch; anterior end of oviduct (=junction tissue between ovary and oviduct) consists of rounded cells; anterior part of oviduct consists of rounded cells, forming a simple tube; middle part of oviduct serving as spermatheca, consists of roundish and relatively large cells. Eggs in single to multiple-cell stage or even further developed at posterior part of oviduct (=uterus), in young females being composed of squared or angular cells, long

enough to contain one well-developed oocyte. *Receptaculum seminis* not observed, *i.e.*, the organ is not independent, and a part of oviduct/uterus works as the organ; vaginal glands present but obscure; vagina perpendicular to body surface, surrounded by sclerotized tissue; vulva slightly protuberant in lateral view, pore-like in ventral view; rectum about one anal body diameter (ABD) long, intestine/rectum junction surrounded by well-developed sphincter muscle. Anus in form of dome-shaped slit, posterior anal lip slightly protuberant. Tail elongate conoid with long and filiform terminus. Phasmid ventro-laterally located at about 1.5–2.0 ABD posterior to anal opening.

***Pristionchus dorci* n. sp.**

urn:lsid:zoobank.org:act:55E69BA4-6D38-48A3-9D3F-A73D7E3F5E5A

Etymology. The species name is derived from the generic name of the host Lucanid beetle (*Dorcus davidis* [Fairmaire, 1887]).

Measurements. See Table 6

Adult. General characters are as described above for *maupasi*-group.

Stenostomatous form. Dorsal movable tooth flint-shaped with anterior end slightly curved; left subventral ridge with three minute, blunt or pointed and adventitious denticles on a plate, most dorsal denticle often masked by middle denticle; right subventral ridge with a distal pointed adventitious denticle.

Eurystomatous form. Dorsal movable tooth claw-like as typical of the genus; left subventral ridge with three large plates, often with two or three split tips; right subventral movable tooth claw-like and extra peak not observed.

Male. As described above for *maupasi*-group.

Female. As described above for *maupasi*-group.

Diagnosis and relationships. *Pristionchus dorci* **n. sp.** is characterized by the slightly anteriorly curved flint-shaped dorsal tooth, the left subventral ridge with three pointed or blunt denticles and right subventral ridge with a small pointed denticle of stenostomatous form, claw-like right subventral stegostomal tooth without an extra peak in eurystomatous form, the arrangement of male genital papillae, <v1, v2, v3d, co, v4, ad, ph, (v5, v6, v7, pd)> where posterior four paired papillae are forming cluster, and a long tail of both sexes. The *maupasi*-group of *Pristionchus* currently contains 15 species. These species are typologically similar with each other, but can be distinguished by the stomatal characters, *e.g.*, number of denticles, and the arrangement of genital papillae (summarized in Kanzaki *et al.* 2018). *P. dorci* **n. sp.** shares its single-peaked right subventral ridge in the stenostomatous form and the laterally directed third paired male genital papillae (v3d) with *P. laevicollis* Kanzaki, Herrmann, Yoshida, Weiler, Rödelsperger & Sommer, 2018, *P. japonicus* Kanzaki, Ragsdale, Herrmann, Mayer & Sommer, 2012 and *P. hongkongensis* Kanzaki, Herrmann, Yoshida, Weiler, Rödelsperger & Sommer, 2018, which have single-peaked right subventral ridge and v3d papillae. The new species is also similar to *P. maxplancki* Kanzaki, Ragsdale, Herrmann, Röseler & Sommer, 2013 and *P. quartusdecimus* Kanzaki, Ragsdale, Herrmann, Röseler & Sommer, 2013 which has 1–2 peaks on the right subventral ridge and v3d papillae. However, *P. dorci* **n. sp.** is distinguished from *P. laevicollis* by the relative position of genital papillae, v2 and v3d two pairs are close to each other *vs.* clearly separated, the v3d is at 1/3 CBD anterior to co *vs.* very close to co and sometimes adcloacal, and the ventral triplet papillae and pd overlap each other *vs.* pd is located at just posterior to the triplet papillae. *P. dorci* **n. sp.** is distinguished from *P. japonicus* by the relative position of genital papillae, v1 is at 1.5 CBD *vs.* 1.0 CBD anterior to co, v1–v2 distance is almost same as *vs.* obviously shorter than v2–v4 distance, and female tail elongated conoid with *vs.* without filiform terminus. The new species is also readily distinguished from *P. hongkongensis* by the relative position of genital papillae, v1 is at 1.5 CBD *vs.* 1 CBD anterior to co, v1–v2 equivalent *vs.* shorter than v2–v4 distance, and the spicule shape, the spicule of *P. dorci* **n. sp.** is slenderer than that of *P. hongkongensis*. In addition, *P. hongkongensis* is characterized by its stomal morphology of eurystomatous form, *i.e.*, large and barrel-shaped stoma bearing left subventral ridges with many cusps, and this character is clearly different from the stomal morphology of new species. *P. dorci* **n. sp.** is distinguished from *P. quartusdecimus* by the right subventral ridge of stenostomatous form, with single *vs.* 1-2 peaks, and relative position of genital papillae, v1 is at 1.5 CBD *vs.* 1.0 CBD anterior to co, v1–v2 distance is almost same as *vs.* obviously shorter than v2–v4 distance, and the ventral triplet papillae and pd overlap to each other *vs.* pd is located at just posterior to the triplet papillae. Typological characters, *e.g.*, the arrangement of genital papillae, of new species is very similar to those of *P. maxplancki*. However, *P. dorci* **n. sp.** is distinguished from *P. maxplancki* by the right subventral ridge of stenostomatous form, with single *vs.* 1–2 peaks, and the relative position of poste-

rior four paired papillae, the ventral triplet papillae and pd overlap to each other *vs.* pd is located at just posterior to the triplet papillae. Further, the new species is distinguished from all other species by mating experiments and also characterized by a *ca.* 1,600-bp fragment of the SSU rRNA gene (GenBank accession number MW017218), the sequence of which is distinct from that of all other *Pristionchus* species.

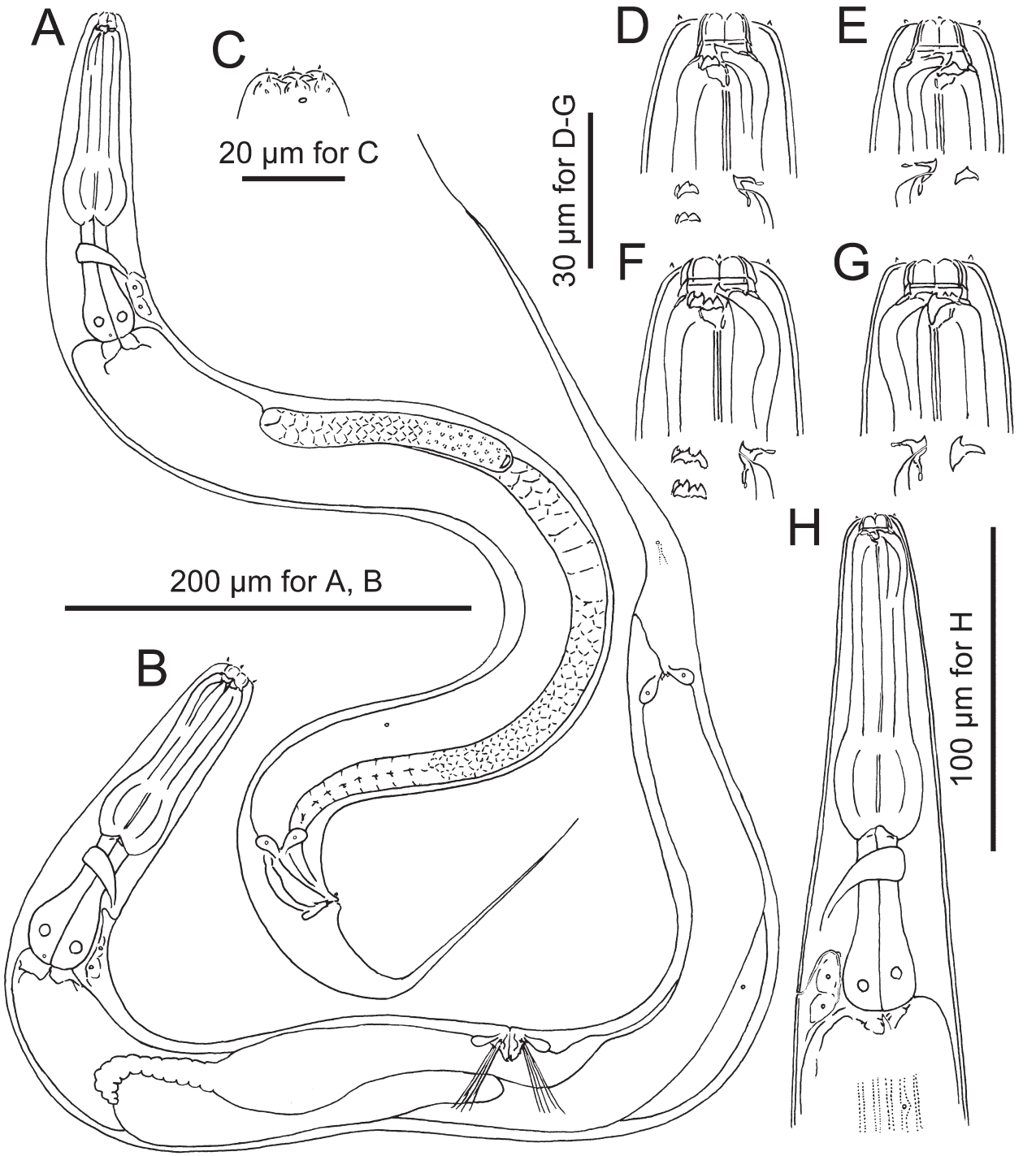


FIGURE 41. *Pristionchus dorci* **n. sp.** A: Right lateral view of adult male. B: Right lateral view of adult female. C: Left lateral view of male head region showing labial sensilla, cephalic papillae and amphid. D, E: Stomal region of adult female in left lateral (D) and right lateral (E) view. F, G: Stomal region of eurystomatous female in left (F) and right (G) lateral view. Morphological variations of teeth and ridges are separately drawn in each subfigure as left subventral ridge (left) and dorsal tooth (right) (D), dorsal tooth (left) and right subventral ridge (right) (E), left subventral ridge (left) and dorsal tooth (right) (F), and dorsal (left) and right subventral (right) teeth (G). H: Anterior part of adult female in left lateral view.

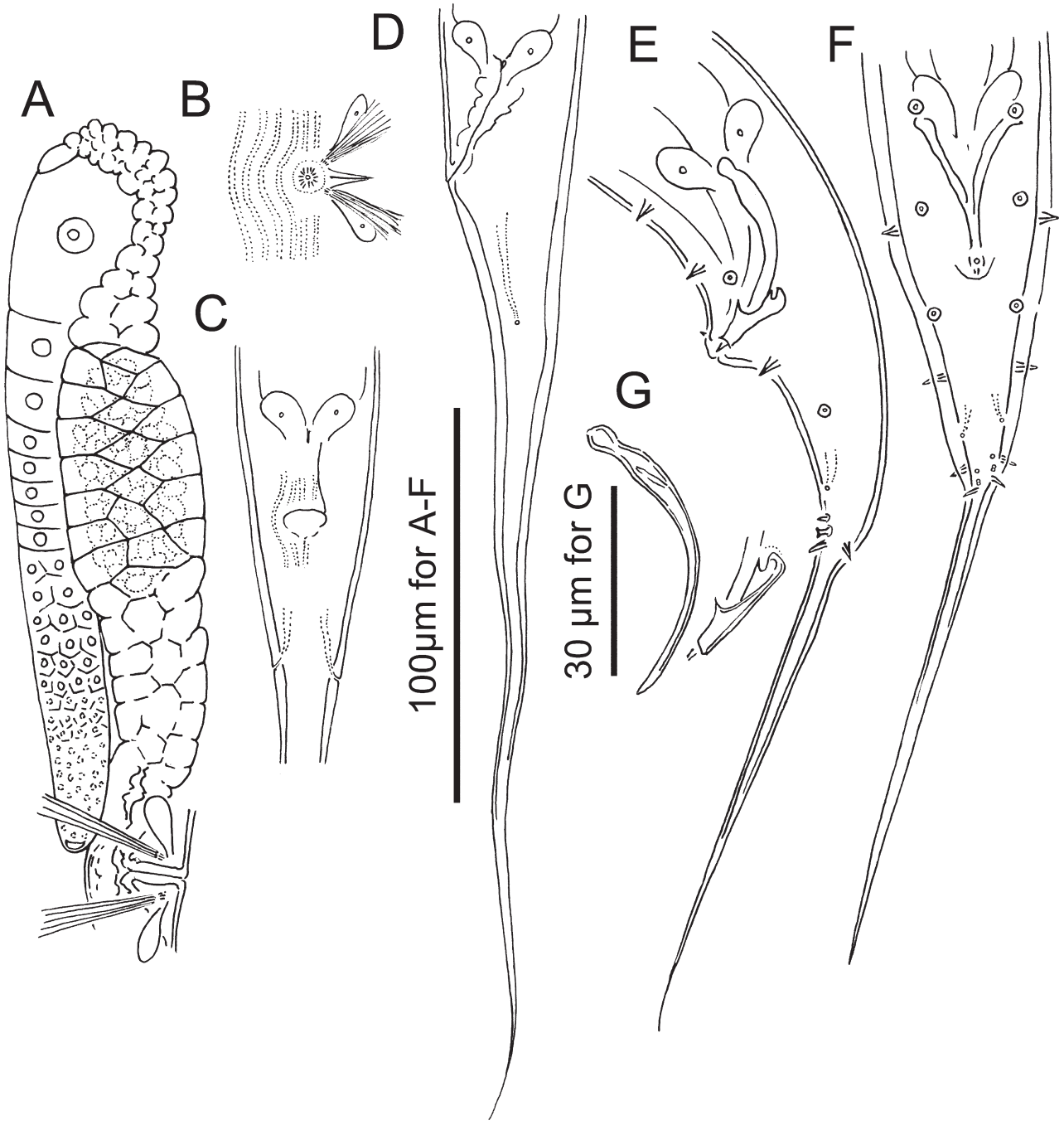


FIGURE 42. *Pristionchus dorci* **n. sp.** A: Anterior gonad of female in right lateral view. B: Vulval region in ventral view. C, D: Female tail in ventral (C) and left lateral (D) view. E, F: Male tail in left lateral (E) and ventral (F) view. G: Spicule and gubernaculum in left lateral view.

Type host and locality. Isolated from a specimen of the lucanid beetle *Dorcus davidis* collected at a dumpster near Ganquan, Shasnxi province, PRC.

Type material and type strain. Type strain RS 6134, other strain RS5992; frozen at the nematode collection of the MPI Tübingen and available as living culture upon request. Voucher specimens sent to the following museums: Holotype male, Paratype male and female: Museum für Naturkunde Karlsruhe, Germany; Paratype male and female: Swedish Natural History Museum, Stockholm, Sweden; Paratype male and female: University of California in Riverside Nematode Collection (UCRNC), Riverside, CA, USA.

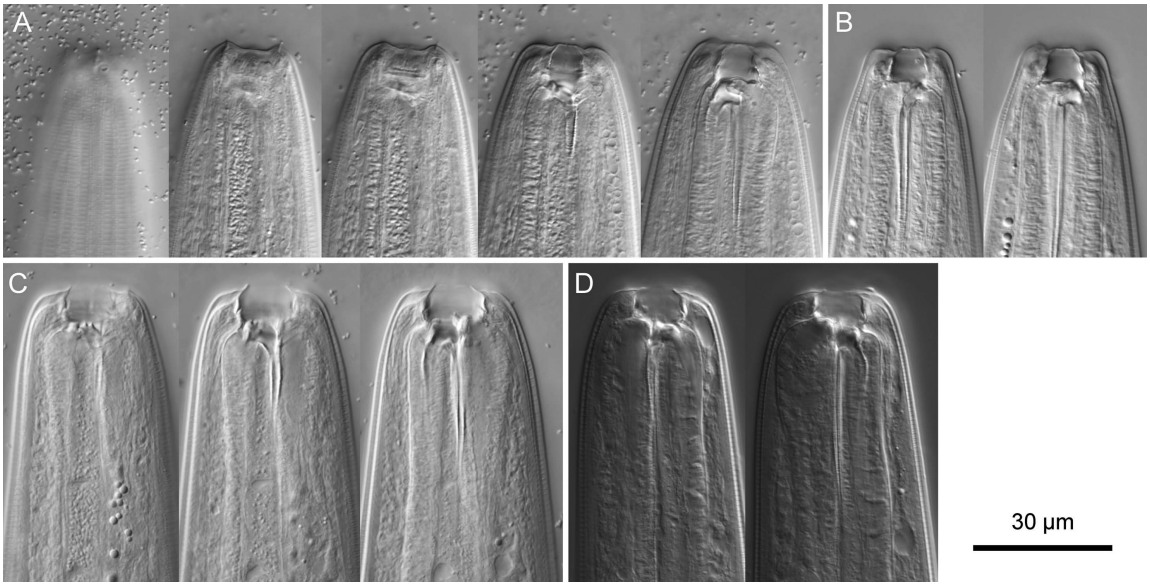


FIGURE 43. Stomal region of *Pristionchus dorci* **n. sp.** A: Left lateral view of stenostomatous form in five different focal planes. B: Right lateral view of stenostomatous form in two different focal planes. C: Left lateral view of eurystomatous form in three different focal planes. D: Right lateral view of eurystomatous form in two different focal planes.

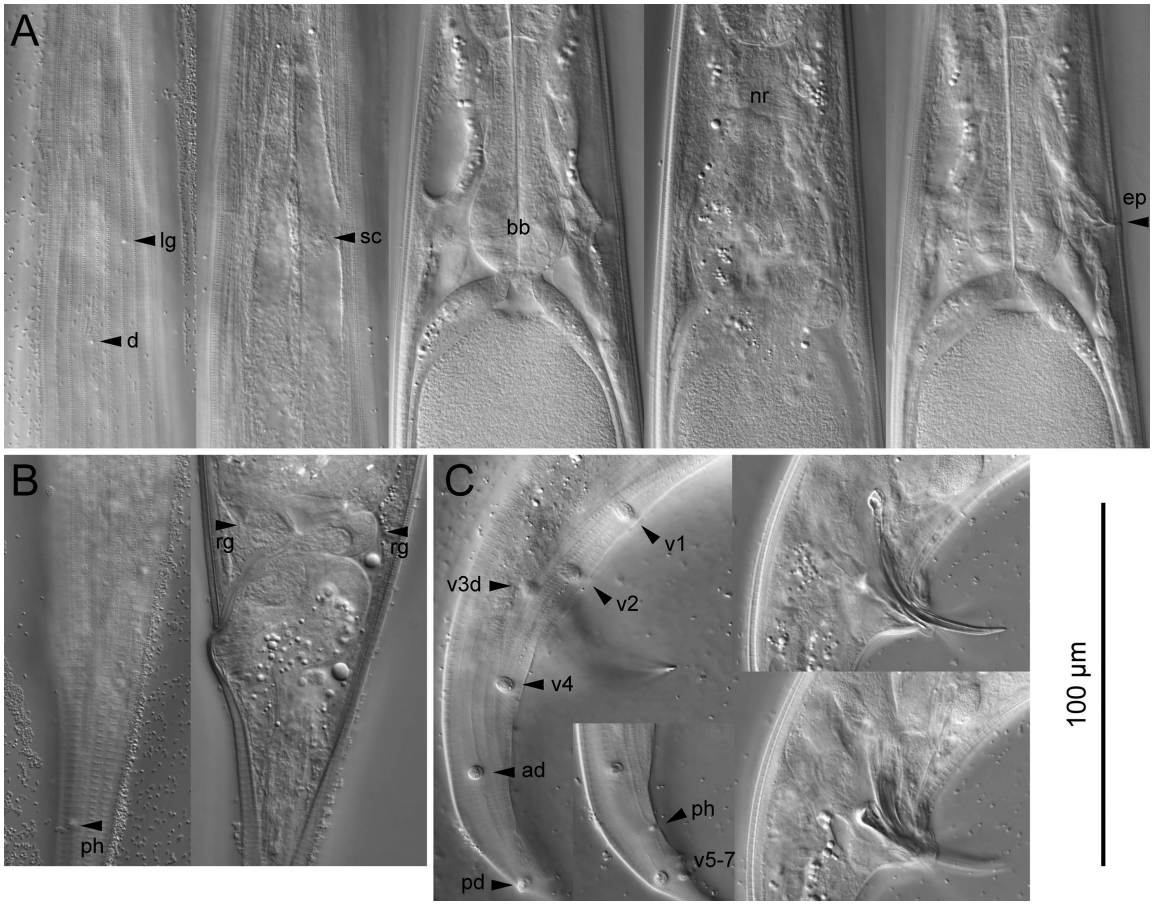


FIGURE 44. *Pristionchus dorci* **n. sp.** A: Right lateral view of posterior pharynx region in five different focal planes where deirid (d), lateral gland opening (lg) and its associated secretory cell (sc), nerve ring (nr), secretory-excretory pore (ep) and basal bulb (bb) are indicated. B: Left lateral view od female anal region in two different focal planes where phasmid (ph) and rectal glands (rg) are indicated. C: Right lateral view of whole tail in two different focal planes and spicule and gubernaculum in two different focal planes. Genital papillae are labelled using the terminology of Sudhaus & Fürst von Lieven (2003), and laterally located papillae are indicated with "d".

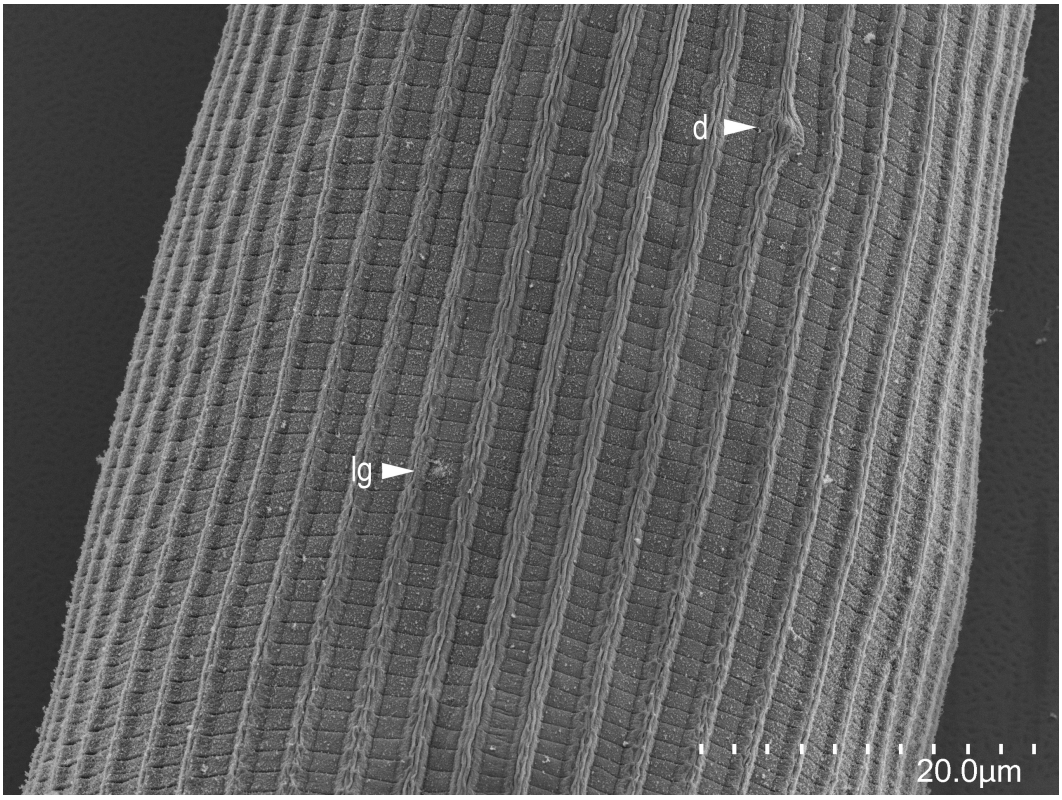


FIGURE 45. Scanning electron photomicrograph of *Pristionchus dorci* **n. sp.** surface structure with deirid (d) and lateral galnd (lg) indicated.

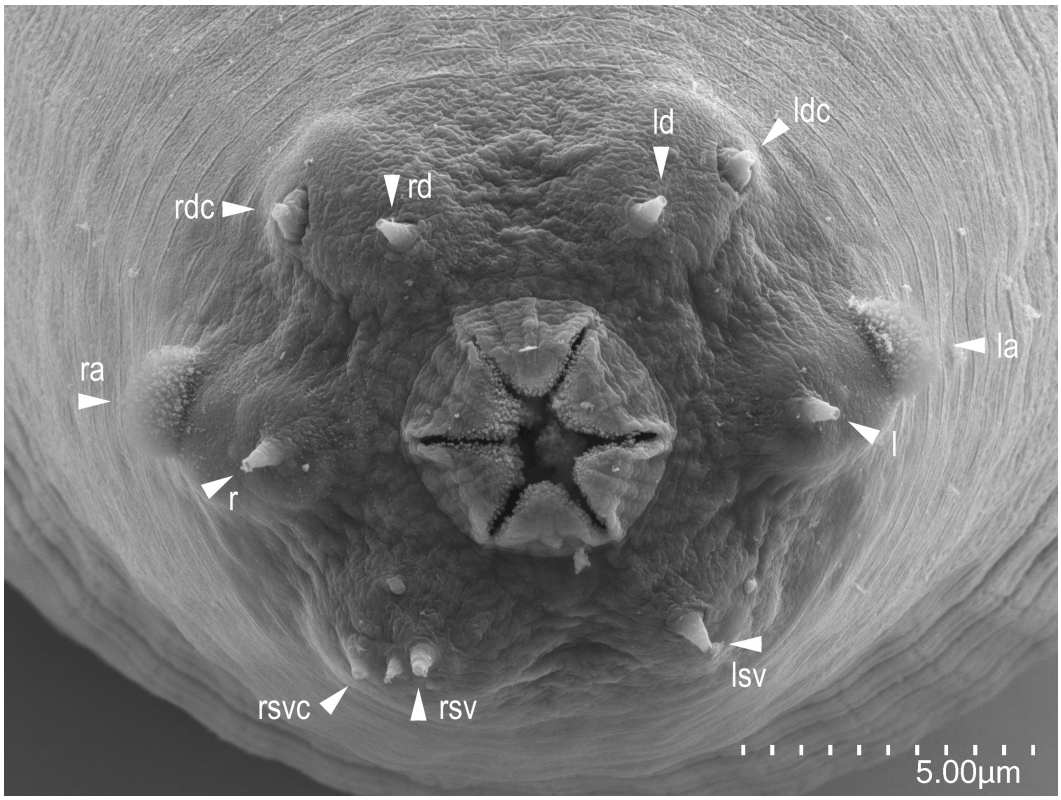


FIGURE 46. Scanning electron photomicrograph of *Pristionchus dorci* **n. sp.** stenostomatous male *enface* view where left and right amphids (la and ra), left and right subventral (lsv and rsv), left and right lateral (l and r) and left and right dorsal (ld and rd) labial sensilla, and right subventral (rsvc) and dorsal (ldc and rdc) cephalic sensilla are indicated. Left subventral cephalic sensillum is not seen.

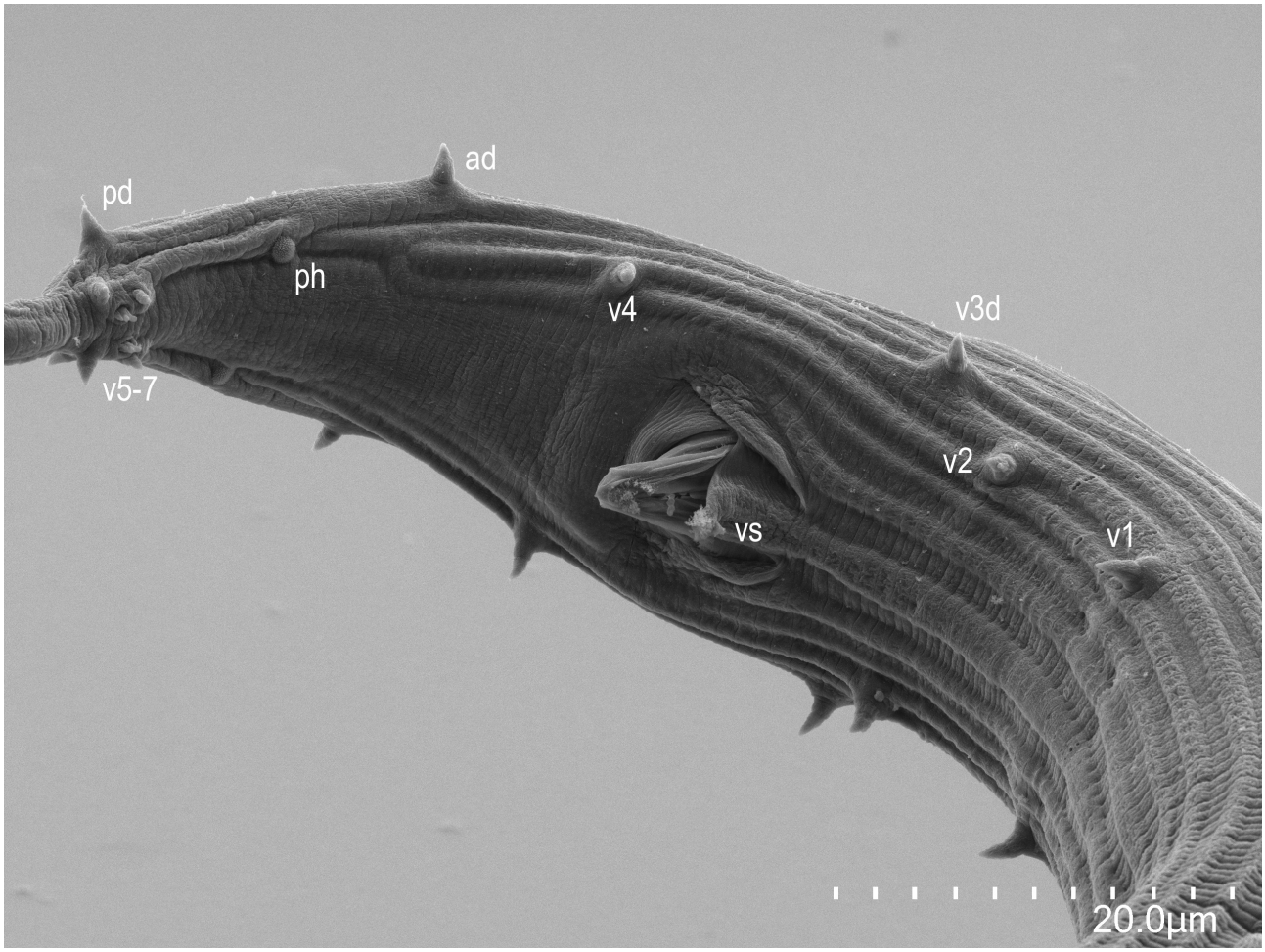


FIGURE 47. Scanning electron photomicrograph of *Pristionchus dorci* **n. sp.** male tail with the position of genital papillae according to the terminology of Sudhaus and Fürst von Lieven (2003).

***Pristionchus purgamentorium* n. sp.**
urn:lsid:zoobank.org:act:EE7F188B-438B-4E11-BF8E-5D389514932A

Etymology. The species name is derived from the finding circumstances. The beetle was collected on a dumpster (lat. purgamentum = trash).

Measurements. See Table 6

Adult. General characters are as described above for *maupasi*-group.

Stenostomatous form. Dorsal movable tooth flint-shape, anterior end slightly curved; left subventral ridge with three or sometimes more minute pointed and adventitious denticles on a plate, most dorsal denticle often masked by middle denticle; right subventral ridge with two or three bluntly pointed adventitious denticles.

Eurystomatous form. Dorsal movable tooth claw-like as typical of the genus; left subventral ridge with three large plates, each often has two or three split tips; right subventral movable tooth claw-like, an indistinctive extra peak sometimes seen.

Male. As described above for *maupasi*-group.

Female. As described above for *maupasi*-group.

Diagnosis and relationships. *Pristionchus purgamentorium* **n. sp.** is characterized by the slightly anteriorly curved flint-shaped dorsal tooth, the left subventral ridge with three or sometimes more pointed denticles and right subventral ridge with two or three small bluntly pointed denticles of stenostomatous form, claw-like right subventral stegostomal tooth sometimes with an indistinctive extra peak in eurystomatous form, the arrangement of male genital papillae, <v1, v2, v3d, co, v4, ad, ph, (v5, v6, v7, pd)> where posterior four paired papillae form a cluster,

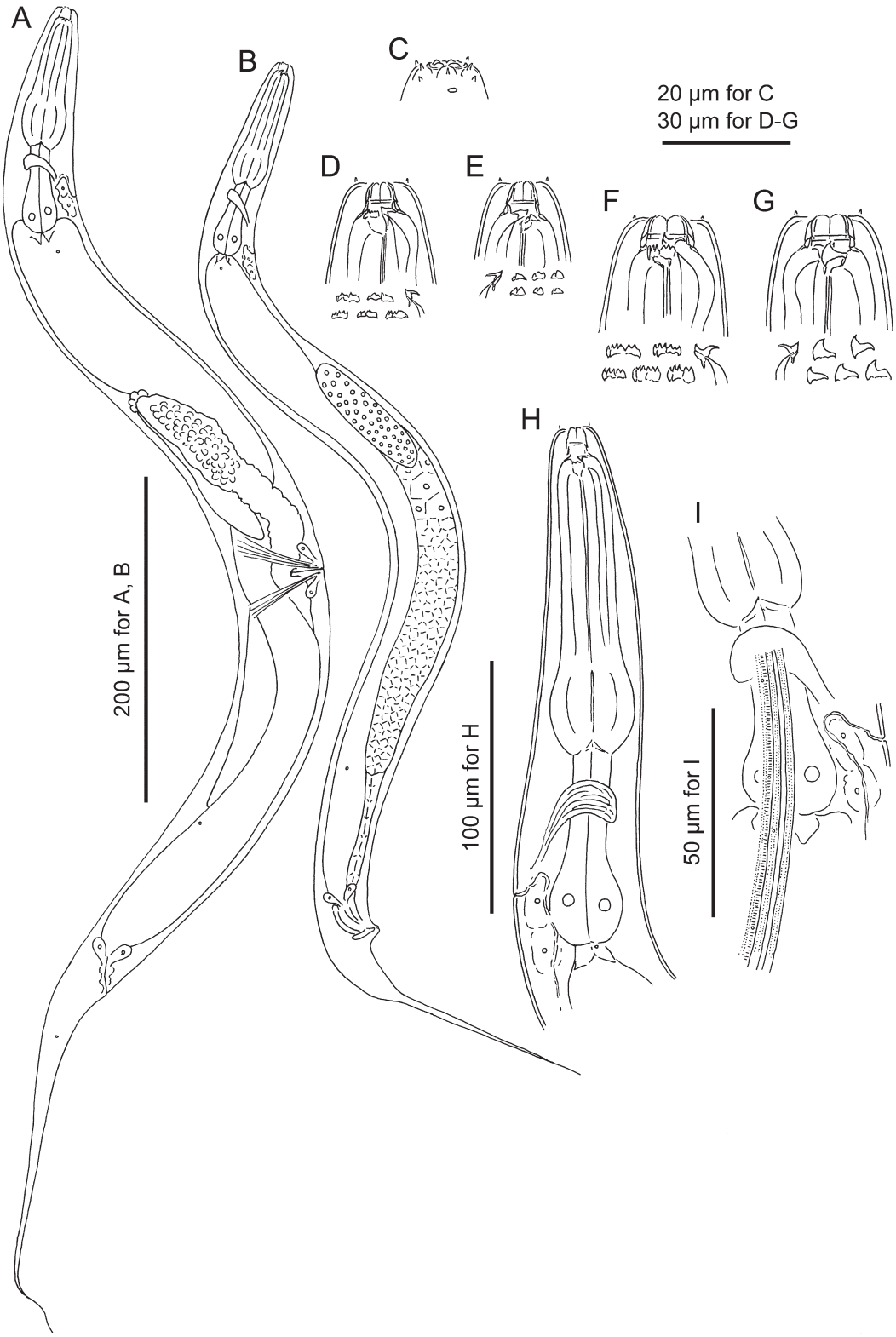


FIGURE 48. *Pristionchus purgamentorium* **n. sp.** A: Right lateral view of adult female. B: Right lateral view of adult male. C: Left lateral view of male head region showing labial sensilla, cephalic papillae and amphid. D, E: Stomal region of adult female in left lateral (D) and right lateral (E) view. F, G: Stomal region of eurystomatous female in left (F) and right (G) lateral view. Morphological variations of teeth and ridges are separately drawn in each subfigure as left subventral ridge (left and middle) and dorsal tooth (right) (D), dorsal tooth (left) and right subventral ridge (middle and right) (E), left subventral ridge (left and middle) and dorsal tooth (right) (F), and dorsal (left) and right subventral (middle and right) teeth (G). H: Anterior part of adult stenostomatous female in left lateral view; I: relative position of secretory-excretory pore, deirid and lateral glands.

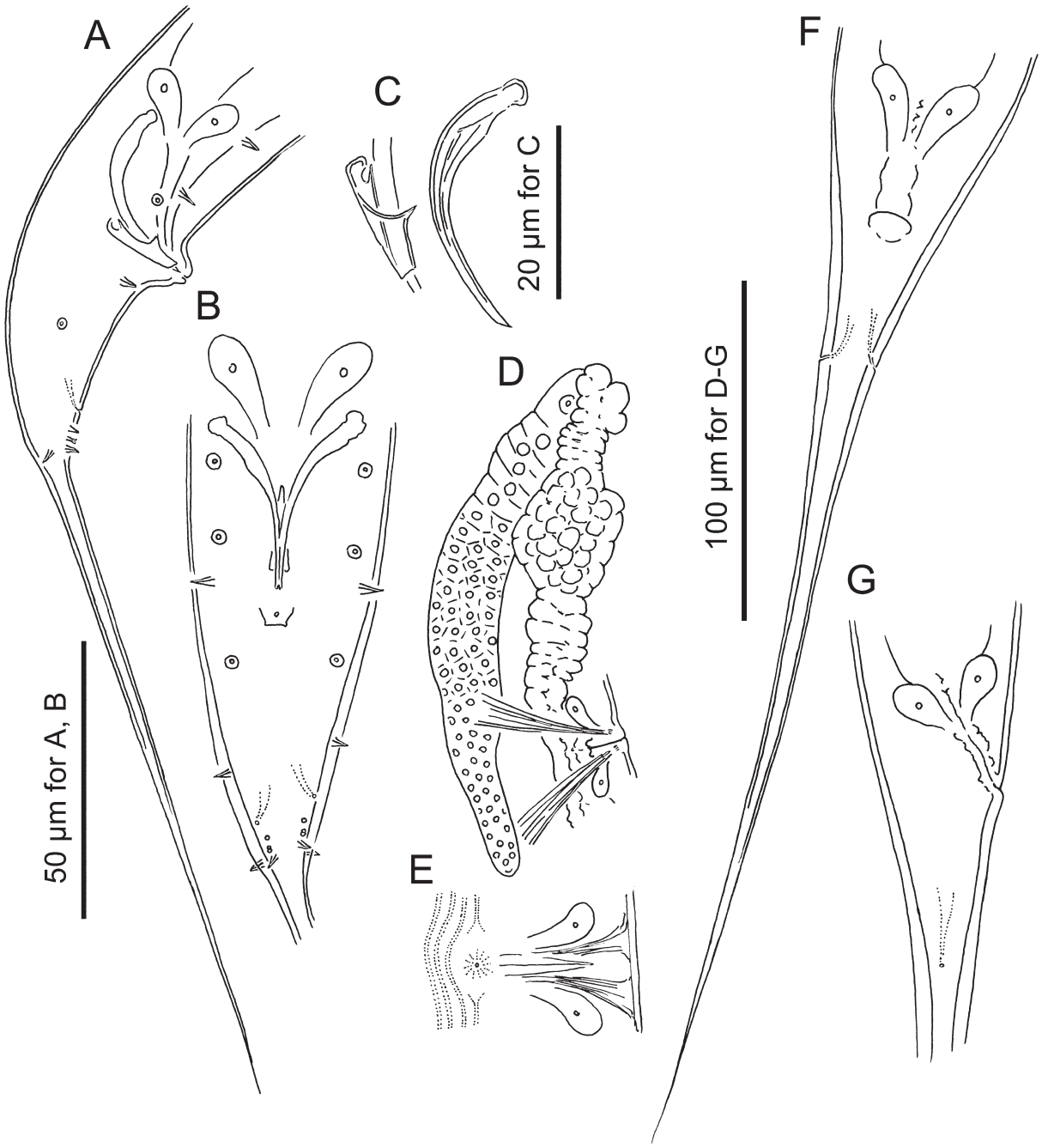


FIGURE 49. *Pristionchus purgamentorium* **n. sp.** A, B: Male tail in right lateral (E) and ventral (F) views. C: Spicule and gubernaculum in right lateral view. D: Anterior gonad of female in right lateral view. E: Vulval region in ventral view. F, G: Female tail in ventral (C) and right lateral (D) views.

and distinctive and long tail of both sexes. The *maupasi*-group currently contains 16 species including *P. dorci* **n. sp.**. These species are typologically similar to each other, but can be distinguished by the stomatal characters, *e.g.*, number of denticles, and the arrangement of genital papillae (summarized in Kanzaki *et al.*, 2018). *Pristionchus purgamentorium* **n. sp.** shares its multiple-peaked right subventral plate in the stenostomatous form and the laterally directed third paired male genital papillae (v3d) with *P. maxplancki*, which have one to two-peaked right subventral ridge and v3d papillae. The new species also shares multiple-peaked right suvbentral plate of stenostomatous form with *P. riukiariae* Kanzaki, Herrmann, Yoshida, Weiler, Rödelsperger & Sommer, 2018. However, *P. purgamentorium* **n. sp.** is distinguished from *P. maxplancki* by its right subventral plate of stenostomatous form, with two to three ridges *vs.* one to two ridges. *P. purgamentorium* **n. sp.** is also distinguished from *P. riukiariae* by right subventral plate of stenostomatous form, with two to three ridges *vs.* two ridges, and the arrangement of genital papillae, third

vs. second pair directed laterally (v3d *vs.* v2d), and second and third pairs are separated *vs.* very close to each other. Further, the new species is distinguished from all other species by mating experiments and also characterized by a *ca.* 1,600-bp fragment of the SSU rRNA gene (GenBank accession number MW017224), the sequence of which is distinct from that of all other *Pristionchus* species.

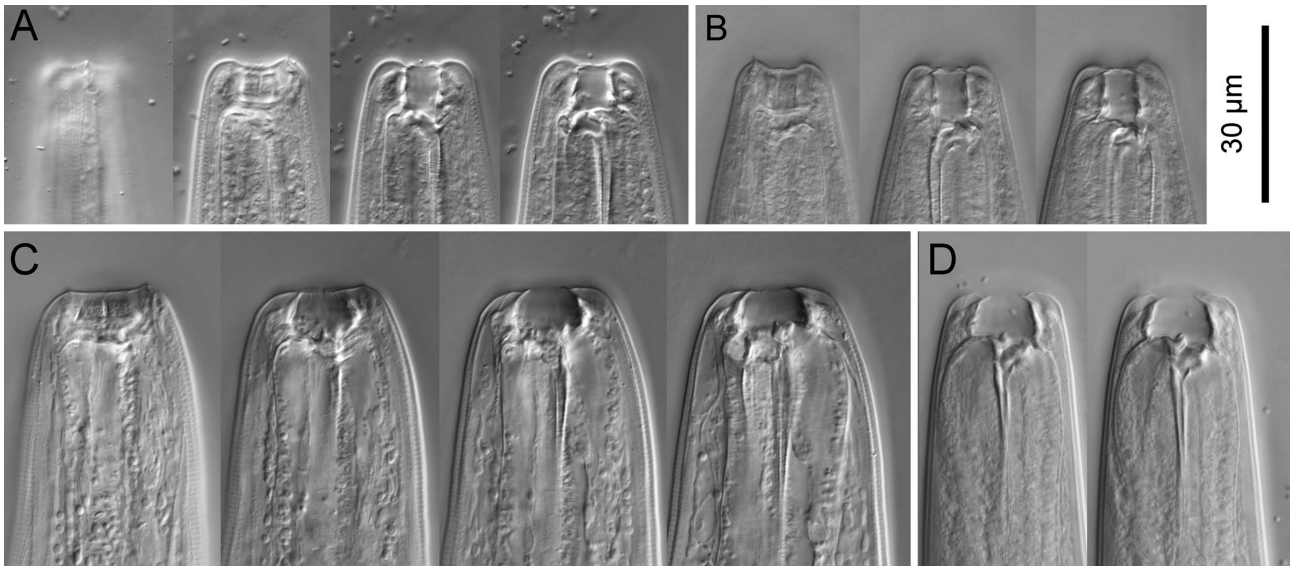


FIGURE 50. Stomal region of *Pristionchus purgamentorium* **n. sp.** A: Left lateral view of stenostomatous form in four different focal planes. B: Right lateral view of stenostomatous form in three different focal planes. C: Left lateral view of eurystomatous form in four different focal planes. D: Right lateral view of eurystomatous form in two different focal planes.

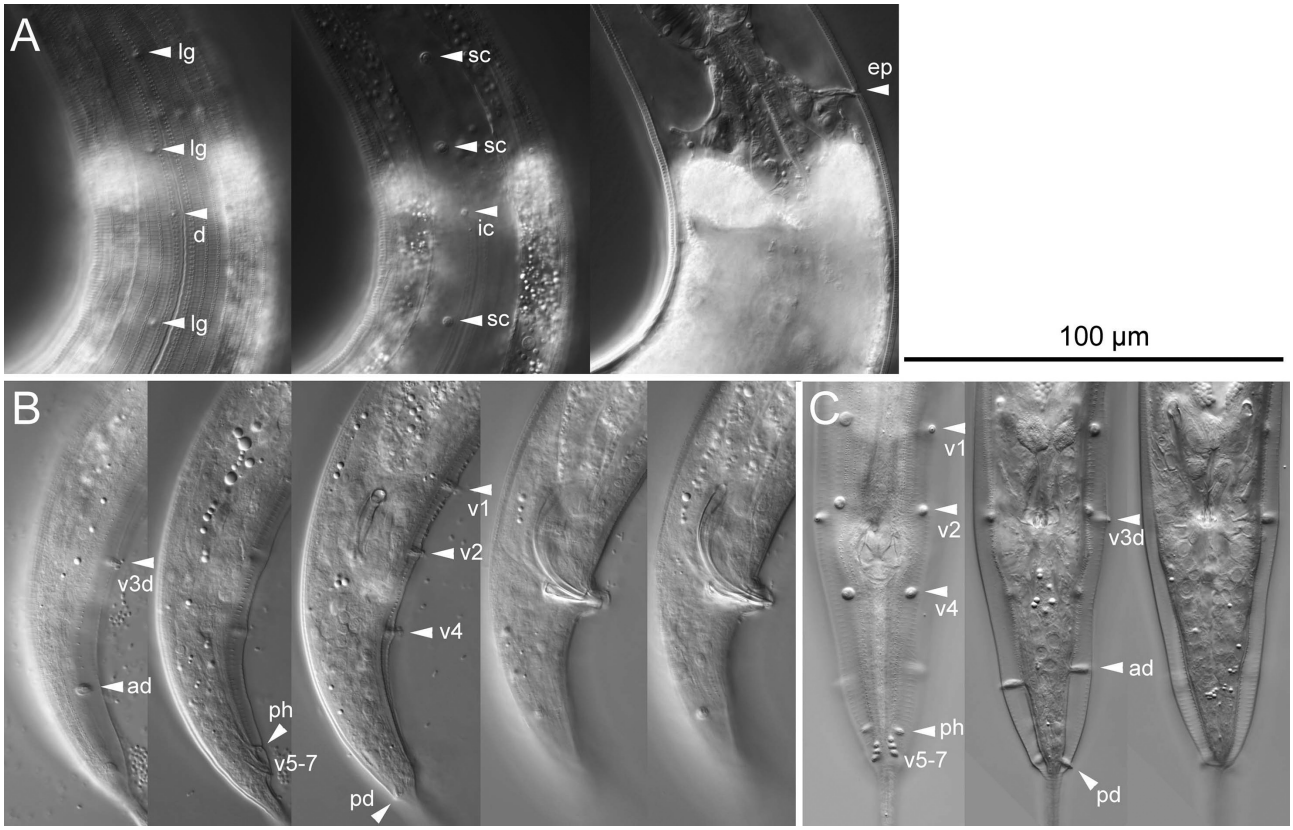


FIGURE 51. *Pristionchus purgamentorium* **n. sp.** A: Right lateral view of posterior pharynx region in three different focal planes where deirid (d), lateral gland opening (lg) and its associated secretory cell (sc) and secretory-excretory pore (ep) are labelled. B: Right lateral view of male tail in five different focal planes. C: Ventral view of male tail in three different focal planes. Genital papillae are labelled according to the terminology of Sudhaus & Fürst von Lieven (2003), and laterally located papillae are indicated with "d".

TABLE 6.

	P. dorci (RS6134)		*P. purgamentorium* (RS6138)	
character	**stenostomatous male**	**stenostomatous female**	**stenostomatous male**	**stenostomatous female**
n	5	5	5	5
L	973 ± 91.4 (874–1105)	1126 ± 109.9 (1023–1313)	688 ± 51.2 (648–768)	905 ± 52.7 (844–985)
L’	799 ± 87.7 (707–912)	870 ± 88.2 (792–1021)	542 ± 47.2 (501–615)	686 ± 57.0 (633–773)
a	12 ± 1.8 (10–14)	14 ± 0.5 (13–14)	13 ± 1.0 (12–14)	12 ± 0.2 (12–12)
b	6.2 ± 0.6 (5.7–7.2)	7.4 ± 0.7 (6.6–8.6)	5.1 ± 0.2 (5.0–5.4)	6.1 ± 0.3 (5.7–6.5)
c	5.6 ± 0.8 (4.7–6.9)	4.4 ± 0.1 (4.2–4.5)	4.7 ± 0.2 (4.4–5.0)	4.2 ± 0.5 (3.4–4.6)
c’	4.1 ± 0.5 (3.3–4.5)	7.1 ± 0.6 (6.6–8.1)	4.2 ± 0.2 (4.0–4.4)	6.7 ± 1.1 (4.9–7.5)
ant. stoma length (cheilo- + gymnostom)	7.5 ± 0.8 (6.9–8.6)	7.3 ± 1.0 (6.0–8.7)	6.4 ± 0.6 (5.7–7.1)	8.0 ± 0.6 (7.4–9.0)
total stoma length	11.2 ± 0.4 (10.7–11.7)	12.5 ± 0.9 (11.6–13.7)	10.5 ± 1.1 (9.6–12.3)	13.0 ± 0.8 (12.1–14.2)
stoma width	7.0 ± 0.5 (6.2–7.5)	7.7 ± 1.0 (6.4–8.9)	5.2 ± 0.3 (4.8–5.6)	6.3 ± 0.9 (5.0–7.3)
ant. pharynx length (pro- + metacorpus)	83 ± 3.2 (78–87)	86 ± 4.1 (83–93)	78 ± 2.0 (76–80)	87 ± 3.4 (84–91)
post. pharynx length (isthmus + basal bulb)	62 ± 4.4 (55–66)	59 ± 3.4 (55–62)	46 ± 4.4 (41–51)	54 ± 3.5 (51–59)
total length pharynx	145 ± 3.3 (142–150)	145 ± 2.9 (141–149)	124 ± 5.1 (119–131)	142 ± 5.1 (135–148)
ant./total pharynx %	57 ± 2.5 (55–61)	59 ± 2.3 (57–63)	63 ± 2.2 (60–66)	62 ± 1.7 (59–64)
median bulb diameter	29 ± 1.2 (28–31)	27 ± 2.9 (25–32)	26 ± 1.3 (25–29)	31 ± 1.3 (29–33)
terminal bulb diameter	26 ± 2.6 (25–31)	23 ± 2.8 (20–27)	22 ± 1.2 (21–23)	28 ± 1.6 (26–30)
neck length (incl. stoma)	156 ± 3.2 (153–161)	153 ± 2.6 (149–155)	134 ± 5.7 (130–143)	150 ± 5.2 (143–156)
excretion pore	147 ± 14.2 (129–167)	148 ± 7.9 (142–161)	112 ± 9.5 (102–123)	122 ± 8.5 (111–135)
nerve ring	106 ± 6.1 (95–110)	111 ± 4.4 (103–114)	96 ± 4.8 (91–103)	108 ± 2.9 (104–111)
testis length	483 ± 72.4 (398–577)	-	344 ± 40.4 (308–410)	-
anterior gonad length	-	188 ± 27.8 (149–224)	-	163 ± 34.1 (135–222)
posterior gonad length	-	214 ± 35.0 (187–274)	-	144 ± 12.4 (127–158)
ant. end to vulva distance	-	503 ± 47.3 (452–581)	-	414 ± 37.4 (376–475)
vulva to anus distance	-	368 ± 36.8 (344–432)	-	271 ± 29.3 (223–293)
T or V	49 ± 2.9 (46–52)	45 ± 0.6 (44–45)	50 ± 2.6 (47–53)	46 ± 1.6 (45–48)
max. body diameter	84 ± 10.0 (72–94)	82 ± 11.0 (73–101)	54 ± 6.0 (46–61)	75 ± 4.7 (70–81)
cloacal or anal body diameter	43 ± 4.2 (37–49)	36 ± 4.9 (31–44)	35 ± 1.6 (33–37)	34 ± 6.0 (28–43)
tail length	174 ± 20.9 (142–193)	256 ± 22.3 (231–292)	147 ± 5.2 (139–153)	220 ± 24.0 (201–261)
spicule length (curve)	59 ± 2.0 (57–62)	-	43 ± 1.8 (41–45)	-
spicule length (chord)	48 ± 2.6 (45–51)	-	35 ± 1.9 (33–37)	-
gubernaculum length	21 ± 1.6 (18–23)	-	17 ± 0.5 (16-17)	-

Type host and locality. Isolated from a beetle of the genus *Mimela* (Coleoptera: Rutelidae) at a dumpster near Ganquan, Shaanxi province, PRC.

Type material and type strain. Type strain RS6138 frozen at the nematode collection of the MPI Tübingen and available as living culture upon request. Voucher specimens sent to the following museums: Holotype male, Paratype male and female: Museum für Naturkunde Karlsruhe, Germany; Paratype male and female: Swedish Natural History Museum, Stockholm, Sweden; Paratype male and female: University of California in Riverside Nematode Collection (UCRNC), Riverside, CA, USA.

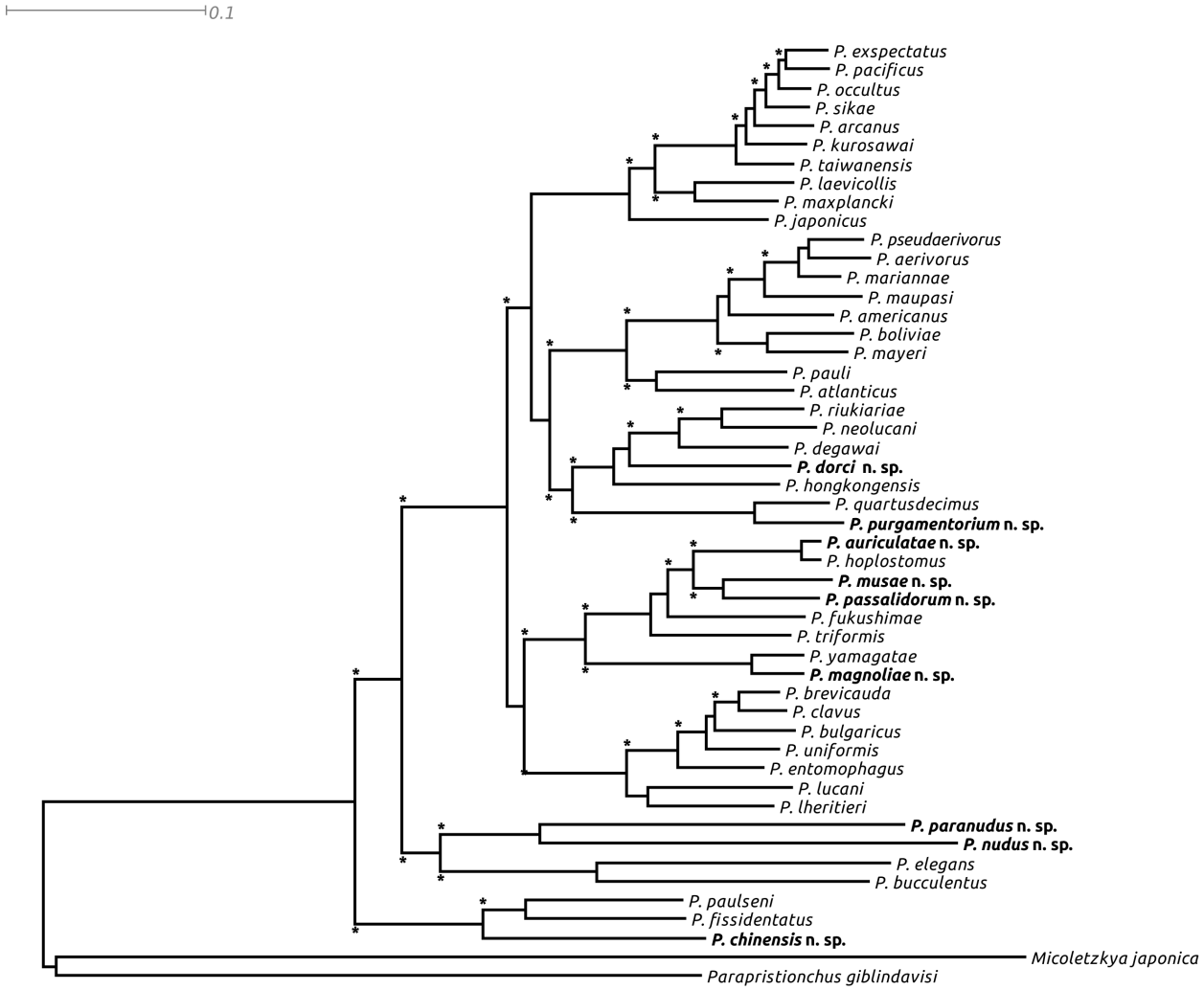


FIGURE 52. The phylogeny represents a maximum likelihood tree that was computed from a concatenated alignment of more than 800 orthologous proteins with the LG substitution model. The stars denote branches with full support after 100 bootstrap pseudoreplicates.

Discussion

This study describes nine novel *Pristionchus* species from Mainland China. As a result, the genus *Pristionchus* currently contains 48 species with whole transcriptome-based phylogenetic support and living cultures for active research. The fact that only two restricted sampling trips to two provinces in China resulted in an increase in the number of nominal species in the genus of roughly 20% strongly suggests that *Pristionchus* species sampling is far from saturation. Thus, additional collecting trips to China will likely add more interesting new species and life history traits to the *Pristionchus* clade. In contrast to previous sampling trips in Taiwan and Hong Kong, the species described in this study cover almost all phylogenetic clades within the genus, with the exception of the *pacificus*-group *senso stricto*.

In the present study, the phylogenetic relationship, *i.e.*, the topology of the phylogenetic tree is different from the previous study (Rödelsperger *et al.*, 2018). One particular subclade, previously part of the *pacificus*-group is transferred to the *maupasi*-group. This example indicates that additional samplings and the identification of new species can have an influence on topology and therefore, no given phylogenetic tree should be considered ultimate. Still, the phylogenetic groups are largely in accordance to their geographic distribution, *i.e.*, *maupasi* and *entomophagus*-groups are distributed mostly in Europe and North-America, respectively, whereas the other groups are mostly from Asia. However, the Asian *pacificus*-group *s.l.* is now split, and a subclade including *P. dorci* **n. sp.** and *P. purgamentorium* **n. sp.** collected from China, are closer to the American clade (*maupasi*-group *s.s.*) than to the *pacificus*-group *s.s*. Thus, further sampling from these regions are necessary to provide additional information about their species diversification in relation to biogeography.

The species described in this study contain two unexpected but exciting findings. First, with *P. nudus* **n. sp.**, this study includes a novel species with unique morphological features. Specifically, the elongation of lateral sensilla, the apparent loss of cephalic papillae in males and the reduction of longitudinal striation resulting in an almost 'naked' appearance of the cuticle, indicate that novel morphological features can be identified in an already heavily sampled genus of nematodes. From a developmental perspective, this finding also indicates that conserved morphological features can be subject to evolutionary change in individual species or lineages. Thus, the phenotypic space of *Pristionchus* and other free-living nematodes is likely far from being fully explored. Second, *P. chinensis* **n. sp.** is the eighth hermaphroditic species in the genus *Pristionchus*. With one sixth (8/48) of *Pristionchus* species being androdioecious, the genus shows a remarkable plasticity in sex determination, unprecedented in any other nematode taxon. Taken together, China represents a potentially lucrative and promising region to search for new *Pristionchus* species and their associated life history traits, probably still harboring a plethora of yet undescribed nematodes that should be studied in more detail.

Acknowledgements

A huge endeavor such as the one reported here could not have been successful without the help of many heads and hands. Foremost, we would like to thank Dr. Ming Bai from the Chinese Academy of Science. Collecting was done in collaboration with him and under his permit(s). Without his help the whole project could never have been realized.

We thank Heike Haussmann who spent hours and hours freezing all the nematode strains for our nematode collection in liquid nitrogen and Marie-Anne Félix for sharing of nematode strains and associated information.

References

Baskaran, P., Rödelsperger, C., Prabh, N., Serobyan, V., Markov, G.V., Hirsekorn, A. & Dieterich, C. (2015) Ancient gene duplications have shaped developmental stage-specific expression in *Pristionchus pacificus*. *BMC Evolutionary Biology*, 15, 185.
https://doi.org/10.1186/s12862-015-0466-2

Dieterich, C., Clifton, S.W., Schuster, L.N., Chinwalla, A., Delehaunty, K., Dinkelacker, I., Fulton, L., Fulton, R., Godfrey, J., Minx, P., Mitreva, M., Roeseler, W., Tian, H., Witte, H., Yang, S.P., Wilson, R.K. & Sommer, R.J. (2008) The *Pristionchus pacificus* genome provides a unique perspective on nematode lifestyle and parasitism. *Nature Genetics,* 40, 1193–1198.
https://doi.org/10.1038/ng.227

Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, 32, 1792–1797.
https://doi.org/10.1093/nar/gkh340

Ekseth, O.K., Kuiper M. & Mironov, V. (2014) orthAgogue: an agile tool for the rapid prediction or orthology relations. *Bioinformatics*, 30, 734–736.
https://doi.org/10.1093/bioinformatics/btt582

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N. & Regev, A. (2011) Full length transcriptomics assembly from RNA-seq data without a reference genome. *Nature Biotechnology*, 29, 644–652.

https://doi.org/10.1038/nbt.1883
Herrmann, M., Mayer, W.E. & Sommer, R.J. (2006a) Nematodes of the genus *Pristionchus* are closely associated with scarab beetles and the Colorado potato beetle in Western Europe. *Zoology*, 109, 96–108.
https://doi.org/10.1016/j.zool.2006.03.001
Herrmann, M., Mayer, W.E. & Sommer, R.J. (2006b) Sex, bugs and Haldane's rule: the nematode genus *Pristionchus* in the United States. *Frontiers in Zoology*, 3, 14.
https://doi.org/10.1186/1742-9994-3-14
Herrmann, M., Mayer, W.E., Hong, R.L., Kienle, S., Minasaki, R. & Sommer, R.J. (2007) The nematode *Pristionchus pacificus* (Nematoda: Diplogastridae) is associated with the Oriental beetle *Exomala orientalis* (Coleoptera: Scarabaeidae) in Japan. *Zoological Science*, 24, 883–889.
https://doi.org/10.2108/zsj.24.883
Herrmann, M., Weiler, C., Rödelsperger, C., Kanzaki, N. & Sommer, R.J. (2016) Two new *Pristionchus* species (Nematoda: Diplogastridae) from Taiwan are part of a species-cluster representing the closest known relatives of the model organism *P. pacificus*. *Zoological Studies*, 55, 48.
https://doi.org/10.6620/ZS.2016.55-48
Herrmann, M., Kanzaki, N., Weiler, C., Yoshida, K., Rödelsperger, C. & Sommer, R.J. (2019) Two new Species of *Pristionchus* (Nematoda: Diplogastridae) include the Gonochoristic Sister Species of *P. fissidentatus*. *Journal of Nematology*, 51, 1–14.
https://doi.org/10.21307/jofnem-2019-024
Hooper, D.J. (1986) Handling, fixing, staining and mounting nematodes. *In*: Southey, J.F. (Ed.), *Laboratory methods for work with plant and soil nematodes*. Her Majesty's Stationally Office, London, pp. 59–80.
Kanzaki, N. (2013) Simple methods for morphological observation of nematodes. *Nematological Research*, 43, 15–17.
https://doi.org/10.3725/jjn.43.15
Kanzaki, N. & Giblin-Davis, R.M. (2015) Diplogastrid systematics and phylogeny. *In*: Sommer, R.J. (Ed.), Pristionchus pacificus—*a nematode model for comparative and evolutionary biology* Brill, Leiden. pp.43–76.
https://doi.org/10.1163/9789004260306_004
Kanzaki, N., Ragsdale, E., Herrmann, M. & Sommer, R.J. (2012a) Two new species of *Pristionchus* (Rhabditida: Diplogastridae) *P. fissidentatus* n. sp. from Nepal and La Réunion island and *P. elegans* n. sp. from Japan. *Journal of Nematology*, 44, 80–90. [PMC free article, PubMed, Google Scholar]
Kanzaki, N., Ragsdale, E. J., Herrmann, M., Mayer, W.E. & Sommer, R.J. (2012b) Description of three *Pristionchus* species (Nematoda: Diplogastridae) from Japan that form a cryptic species complex with the model organism *P. pacificus*. *Zoological Science*, 29, 403–417.
https://doi.org/10.2108/zsj.29.403
Kanzaki, N., Ragsdale, E.J., Herrmann, M., Mayer, W.E., Tanaka, R. & Sommer, R.J. (2012c) *Parapristionchus giblindavisi* n. gen., n. sp. (Rhabditida: Diplogastridae) isolated from stag beetles (Coleoptera: Lucanidae) in Japan. *Nematology*,14: 933–947.
https://doi.org/10.1163/156854112X635878
Kanzaki, N., Ragsdale, E.J., Herrmann, M., Röseler, W. & Sommer, R.J. (2013a) Two new species of *Pristionchus* (Nematoda: Diplogastridae) support the biogeographic importance of Japan for the evolution of the genus *Pristionchus* and the model system *P. pacificus*. *Zoological Science*, 30, 680–692.
https://doi.org/10.2108/zsj.30.680
Kanzaki, N., Ragsdale, E.J., Herrmann, M., Roeseler, W. & Sommer, R.J. (2013b) *Pristionchus bucculentus* n. sp. (Rhabditida: Diplogastridae) isolated from a shining mushroom beetle (Coleoptera: Scaphidiidae) in Hokkaido, Japan. *Journal of Nematology*, 45, 78–86.
Kanzaki, N., Ragsdale, E.J., Herrmann, M., Susoy, V., & Sommer, R.J. (2013c) Two androdioecious and one dioecious new species of *Pristionchus* (Nematoda: Diplogastridae): new reference points for the evolution of reproductive mode. *Journal of Nematology*, 45, 172–194.
Kanzaki, N., Ragsdale, E.J., Herrmann, M., & Sommer, R.J. (2014) Two new and two recharacterized species resulting from a radiation of *Pristionchus* (Nematoda: Diplogastridae) in Europe. *Journal of Nematology*, 46, 60–74.
Kanzaki, N., Herrmann, M., Yoshida ,K., Weiler, C., Rödelsperger, C. & Sommer, R.J. (2018) Samplings of Millipedes in Japan and Scarab Beetles in Hong Kong result in five new Species of *Pristionchus* (Nematoda: Diplogastridae). *Journal of Nematology*, 50, 4.
https://doi.org/10.21307/jofnem-2018-044
Kreis, H.A. (1932) Beiträge zur Kenntnis pflanzenparasitischer Nematoden. *Zeitschrift für Parasitenkunde*, 5, 184–194.
https://doi.org/10.1007/BF02120641
Li, P., Dai, C., Bao, H., Chen, L., Gao, D., Wang, G., Wang, J., Wang, H., Yedid, G. & Zhang, K. (2015) A new species of *Pristionchus* (Rhabditida: Diplogastridae) and its bacterial symbiont from Yixing, China. *Journal of Nematology*, 47, 190–197.
Morgan, K., McGaughran, A., Villate, L., Herrmann, M., Witte, H., Bartelmes, G., Rochat, J. & Sommer, R.J. (2012) Multi locus analysis of *Pristionchus pacificus* on La Réunion Island reveals an evolutionary history shaped by multiple introductions, constrained dispersal events and rare out-crossing. *Molecular Ecology*, 21, 250–266.

https://doi.org/10.1111/j.1365-294X.2011.05382.x
Ragsdale, E., Kanzaki, N., Roeseler, W., Herrmann, M. & Sommer, R.J. (2013) Three new species of *Pristionchus* (Nematoda: Diplogastridae) show morphological divergence through evolutionary intermediates of a novel feeding polymorphism. *Zoological Journal of the Linnean Society*, 168, 671–698.
https://doi.org/10.1111/zoj.12041
Ragsdale, E.J., Kanzaki, N. & Herrmann, M. (2015) Taxonomy and natural history: the genus *Pristionchus*. *In:* Sommer, R.J. (Ed.), Pristionchus pacificus—*a nematode model for comparative and evolutionary biology*. Brill, Leiden, pp. 77–120.
https://doi.org/10.1163/9789004260306_005
Rödelsperger, C., Meyer, J.M., Prabh, N., Lanz, C., Bemm, F. & Sommer, R.J. (2017) Single-molecule sequencing reveals the chromosome-scale genomic architecture of the nematode model organism *Pristionchus pacificus*. *Cell Reports*, 21, 834–844.
https://doi.org/10.1016/j.celrep.2017.09.077
Rödelsperger, C., Röseler, W., Prabh, N., Yoshida, K., Weiler, C., Herrmann, M. & Sommer, R.J. (2018) Phylotransciptomics of *Pristionchus* nematodes reveals parallel gene loss in six hermaphroditic lineages. *Current Biology*, 28: 3123–3127.
https://doi.org/10.1016/j.cub.2018.07.041
Schlager, B., Wang, X., Braach, G. & Sommer, R.J. (2009) Molecular cloning of a dominant roller mutant and establishment of DNA-mediated transformation in the nematode model *Pristionchus pacificus*. *Genesis*, 47, 300–304.
https://doi.org/10.1002/dvg.20499
Sommer, R.J. (2009) The future of evo-devo: model systems and evolutionary theory. *Nature Revue Genetics*, 10, 416–422.
https://doi.org/10.1038/nrg2567
Sommer, R.J. (2015) Integrative evolutionary biology and mechanistic approaches in comparative biology. *In*: Sommer, R.J. (Ed.), Pristionchus pacificus—*a nematode model for comparative and evolutionary biology*. Brill, Leiden, pp. 19–41.
https://doi.org/10.1163/9789004260306_003
Sommer, R.J., Carta, L.K., Kim, S.Y. & Sternberg, P.W. (1996) Morphological, genetic and molecular description of *Pristionchus pacificus* n. sp. (Nematoda: Neodiplogastridae). *Fundamental and Applied Nematology*, 19, 511–521.
Sommer, R.J., Dardiry, M., Lenuzzi, M., Namdeo, S., Renahan, T., Sieriebriennikov, B. & Werner, M. (2017) The genetics of phenotypic plasticity in nematode feeding structures. *Open Biology*, 7, 160332.
https://doi.org/10.1098/rsob.160332
Stamatakis, A. (2014) RaxML version 8: atool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30: 1312–1313.
https://doi.org/10.1093/bioinformatics/btu033
Sudhaus, W. & Fürst von Lieven, A. (2003) A phylogenetic classification and catalogue of the Diplogastridae (Secernentea, Nematoda). *Journal of Nematode Morphology and Systematics*, 6, 43–90.
Susoy, V., Ragsdale, E.J., Kanzaki, N. & Sommer, R.J. (2015) Rapid diversification associated with a macroevolutionary pulse of developmental plasticity. *eLIFE*, 4, e05463.
https://doi.org/10.7554/eLife.05463
Susoy, V., Herrmann, M., Kanzaki, N., Kruger, M., Nguyen, C.N., Rödelsperger, C., Röseler, W., Weiler, C., Giblin-Davis, R.M., Ragsdale, E.J. & Sommer, R.J. (2016) Large-scale diversification without genetic isolation in nematode symbionts of figs. *Science Advance*, 2, e1501031.
https://doi.org/10.1126/sciadv.1501031
Tian, H., Schlager, B., Xiao, H. & Sommer, R.J. (2008) Wnt signaling induces vulva development in the nematode *Pristionchus pacificus*. *Current Biology*, 18, 142–146.
https://doi.org/10.1016/j.cub.2007.12.048
Weadick, C.J. & Sommer, R.J. (2016) Mating system transitions drive lifespan evolution in *Pristionchus* nematodes. *American Naturalist*, 187, 517–529.
https://doi.org/10.1086/685283
Witte, H., Moreno, E., Rödelsperger, C., Kim, J., Kim, J.-S., Streit, A. & Sommer, R. J. (2015) Gene inactivation using the CRISPR/Cas9 system in the nematode *Pristionchus pacificus*. *Development, Genes & Evolution*, 225, 55–62.
https://doi.org/10.1007/s00427-014-0486-8
Yoshida, K., Herrmann, M., Kanzaki, N., Weiler, C., Rödelsperger, C. & Sommer, R.J. (2018) Two new species of *Pristionchus* (Nematoda: Diplogastridae) from Taiwan and the definition of the *pacificus* species-complex *sensu stricto*. *Journal of Nematology*, 50, 355–368.
https://doi.org/10.21307/jofnem-2018-015

OXFORD GENETICS

GENETICS, 2022, 220(3), iyab233
https://doi.org/10.1093/genetics/iyab233
Advance Access Publication Date: 23 December 2021
Investigation
Highlighted Article

The oscillating Mucin-type protein DPY-6 has a conserved role in nematode mouth and cuticle formation

Shuai Sun [1,†], Tobias Theska [1,†], Hanh Witte[1], Erik J. Ragsdale [2], and Ralf J. Sommer [1,*]

[1]Department for Integrative Evolutionary Biology, Max Planck Institute for Biology Tübingen, 72076 Tübingen, Germany,
[2]Department of Biology, Indiana University, Bloomington, IN 47405, USA
[†]These authors contributed equally to the work.

*Corresponding author: Department for Integrative Evolutionary Biology, Max Planck Institute for Biology Tübingen, Max-Planck Ring 9, Tübingen 72076, Germany. Email: ralf.sommer@tuebingen.mpg.de

Abstract

Nematodes show an extraordinary diversity of mouth structures and strikingly different feeding strategies, which has enabled an invasion of all ecosystems. However, nearly nothing is known about the structural and molecular architecture of the nematode mouth (stoma). *Pristionchus pacificus* is an intensively studied nematode that exhibits unique life history traits, including predation, teeth-like denticle formation, and mouth-form plasticity. Here, we used a large-scale genetic screen to identify genes involved in mouth formation. We identified *Ppa-dpy-6* to encode a Mucin-type hydrogel-forming protein that is macroscopically involved in the specification of the cheilostom, the anterior part of the mouth. We used a recently developed protocol for geometric morphometrics of miniature animals to characterize these defects further and found additional defects that affect mouth form, shape, and size resulting in an overall malformation of the mouth. Additionally, *Ppa-dpy-6* is shorter than wild-type with a typical *Dumpy* phenotype, indicating a role in the formation of the external cuticle. This concomitant phenotype of the cheilostom and cuticle provides the first molecular support for the continuity of these structures and for the separation of the cheilostom from the rest of the stoma. In *Caenorhabditis elegans*, *dpy-6* was an early mapping mutant but its molecular identity was only determined during genome-wide RNAi screens and not further investigated. Strikingly, geometric morphometric analysis revealed previously unrecognized cheilostom and gymnostom defects in *Cel-dpy-6* mutants. Thus, the Mucin-type protein DPY-6 represents to the best of our knowledge, the first protein involved in nematode mouth formation with a conserved role in cuticle deposition. This study opens new research avenues to characterize the molecular composition of the nematode mouth, which is associated with extreme ecological diversification.

Keywords: nematode stoma; *Pristionchus pacificus*; *Caenorhabditis elegans*; *dpy-6*; developmental plasticity; evo devo

Introduction

Nematodes are extremely abundant, estimated to account for 80% of all land animals and are found in all ecosystems and on all continents (van den Hoogen et al. 2019). Nonetheless, their cylindrical body is surprisingly uniform with the exception of the head and mouth region (stoma), which has undergone striking adaptations during evolution (Wright 1976; Malakhov 1994). This resulted in a diversity of feeding structures and strategies in free-living and predatory species as well as in parasites of plants, animals, and humans. For example, the human hook worm *Ancylostoma duodenale* has a mouth structure with 2 plates that form 2 strong teeth each, which enable the parasite to attach to the villi in the small intestine causing severe tissue damage (Anderson 2000). These and other examples indicate that the mouth and head structures are the site of major evolutionary innovation in nematodes. However, little is known about the structural architecture of the nematode mouth and associated genetic and molecular processes in any species. This is largely because in *Caenorhabditis elegans*, the prime nematode model organism, the mouth structures have found little attention except for associated neuronal and behavioral traits. Also, the stoma is formed in parts by miniature cells that escape visual observation by differential interference contrast (DIC) microscopy, the usual microscopical technique used in *C. elegans* research (Hall and Altun 2008; Burr and Baldwin 2016).

In general, the cuticle of nematodes is formed by a syncytial layer of cells, inconsistently named epidermal or hypodermal (Wright 1976; Lee 2002). The nematode cuticle is an exoskeleton that in most species consists of 3 layers of fibrillar proteins with many evolutionary adaptations. In *C. elegans*, the major proteins in the cuticle are collagen-like proteins encoded by a large gene family with more than 150 members (Cohen and Sundaram 2020). When mutated, some of the collagen-encoding genes or collagen-modifying enzymes result in a *Dumpy* (shorter than wild-type) (*Dpy*) phenotype. There are more than 30 *Dpy* mutants in *C. elegans*, which have played an important role for genetic mapping (Brenner 1974). Other components of the cuticle are cuticulins, glycoproteins, and lectins, many of which have been

Received: June 15, 2021. **Accepted:** December 13, 2021

investigated in parasitic species as potential drug targets (Page et al. 2014). Thus, a substantial knowledge exists about the composition of the nematode cuticle, which is in strong contrast to the composition of the mouth.

The mouth of nematodes is the major opening of the cuticle but traditionally, the mouth is considered to be fully separated from the external cuticle (Wright 1976). This view is supported by the fact that no defects in mouth structures have been reported in *C. elegans Dpy* mutants. In general, the stoma consists of the "buccal cavity" as the lumen that extends from the mouth opening to the anterior end of the pharynx and the "buccal capsule" as its cuticular lining (Fig. 1) (Fürst von Lieven and Sudhaus 2000). The cellular architecture underlying these structures has been elucidated from transmission electron microscopy after original work in *C. elegans* (Albertson and Thomson 1976; Wright and Thomson 1981). Many structural modifications have been identified from comparative studies, which are in parts controversial as the identity and homology of epidermal and pharyngeal muscle cells that constitute the buccal capsule remains uncertain (De Ley et al. 1995; Baldwin et al. 1997; Fürst von Lieven and Sudhaus 2000; Ragsdale and Baldwin 2010). In the buccal capsule 6 regions can be distinguished (Fig. 1) (De Ley et al. 1995). First, the anterior most "cheilostom" is surrounded by labial cells and is thought to be continuous with the cuticle of the worm. According to this interpretation, the cheilostom is separated from the more posterior parts of the buccal cavity. However, no molecular support for this hypothesis is available and no *C. elegans* cuticle mutants are known to have defects in the cheilostom. Second, the "gymnostom" is the middle part of the buccal cavity and is surrounded by arcade epidermis. Finally, the "stegostom" is the posterior-most part of the buccal cavity, which is further subdivided into "pro"-, "meso"-, "meta"-, and "telo"-stegostom (De Ley et al. 1995). This region is formed by interradial muscle cells and often overlapping because these regions are telescoped. Importantly, the stegostom is a region of major evolutionary diversification as it forms teeth-like denticles in *Pristionchus pacificus* and other Diplogastridae that enable predation (Fig. 1, a and c) (Fürst von Lieven and Sudhaus 2000; Susoy et al. 2015).

Pristionchus pacificus was introduced as a model organism for comparative and evolutionary studies with genetic and

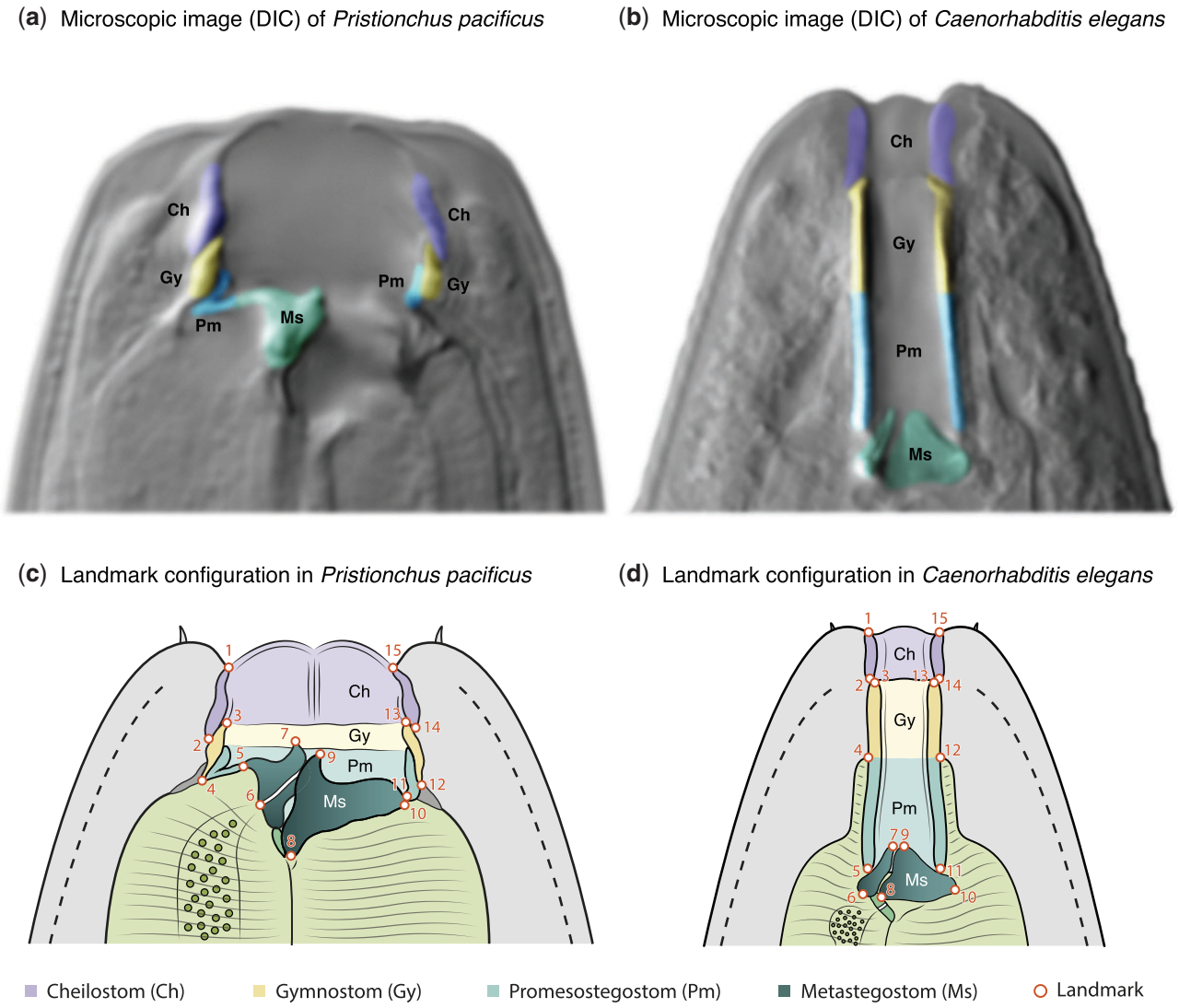


Fig. 1. Structural overview of the nematode stoma. a) DIC image of the *Pristionchus pacificus* mouth in right-lateral perspective (sagittal plane). b) DIC image of the *Caenorhabditis elegans* mouth in right-lateral perspective (sagittal plane). c) Schematic illustration of the *P. pacificus* mouth in right-lateral perspective, including the landmark configuration used for geometric morphometric analysis. d) Schematic illustration of the *C. elegans* mouth in right-lateral perspective, including the landmark configuration used for geometric morphometric analysis.

molecular tools in the mid-1990s (Sommer et al. 1996). This self-fertilizing species has a 4-day generation time when fed *Escherichia coli* under laboratory conditions and has a technological platform related to *C. elegans*. Its genome was originally sequenced in 2008 with single molecule re-sequencing facilitating reverse genetic and genome editing (Dieterich et al. 2008; Rödelsperger et al. 2017; Han et al. 2020; Nakayama et al. 2020). In the last decade, much research on *P. pacificus* focused on the formation of teeth-like denticles in the mouth that result in intraguild predation (Bento et al. 2010; Quach and Chalasani 2020). Importantly, these teeth-like denticles represent an example of phenotypic plasticity as they are expressed in 2 alternative forms. The "stenostomatous" (St) mouth form has a single dorsal tooth with a narrow buccal cavity, whereas the "eurystomatous" (Eu) mouth form has 2 teeth and a broader buccal cavity (Fig. 1, a and c). Specifically, Eu animals have a large right ventrosublateral tooth, curved dorsal tooth, and the anterior tip of their promesostegostom is located posterior to the anterior tip of their gymnostom, whereas the anterior tips of both elements are approximately on the same level in St animals. While Eu animals are omnivorous feeders that can use bacteria, fungi, and nematodes as food source, St animals are strict bacterial feeders. Many studies have identified (1) the environmental influences, (2) the developmental switches, and (3) the gene regulatory network associated with mouth-form plasticity and the associated evolutionary processes (Ragsdale et al. 2013; Werner et al. 2017; Sieriebriennikov et al. 2018, 2020; Bui and Ragsdale 2019; Casasa et al. 2021). Here, we present a large-scale genetic screen to identify genes involved in mouth formation and identified *Ppa-dpy-6* to encode a Mucin-type gel-forming protein that is involved in the specification of the stoma.

Materials and methods

Maintenance of worm cultures and genetic crosses

Stock cultures of all strains were reared at room temperature (20–25°C) on nematode growth medium (NGM) in 6 cm Petri dishes, as outlined before (Sommer et al. 1996). *Escherichia coli* OP50 was used as food source. Bacteria were grown overnight at 37°C in LB medium, and 400 µl of the overnight culture was pipetted on NGM agar plates and left for several days at room temperature to grow bacterial lawns. Nematodes were passed on these lawns and propagated by passing various numbers of mixed developmental stages.

Mouth-form phenotyping

Adult hermaphrodites were immobilized on 5% Noble Agar pads with 0.3% NaN_3 added as an anaesthetic, and examined using DIC microscopy. Animals that had a large right ventrosublateral tooth, curved dorsal tooth, and the anterior tip of the promesostegostom posterior to the anterior tip of the gymnostom were classified as Eu morphs. Animals that did not exhibit these 3 characters simultaneously were classified as St morphs.

Worm size measurement

To study body length of wild-type and mutant animals, we used 7 development stages after bleaching of J1/eggs: 24 h (J2), 36 h (J3), 48 h (J4), 72 h (1-day-old adults), 96 h (2-day-old adults), 120 h (3-day-old adults), 144 h (4-day-old adults), and performed between 10 and 25 replicates. We found a strong increase in body size during the 3 juvenile stages (J2, J3, and J4) but little growth during adulthood (Supplementary Fig. 3a). The maximum body length (1.06 mm) was reached at 4 days adulthood. For the comparison of body length between *Dpy* mutants and wild-type, we singled out the 5-day-old adults to NGM plates without bacteria. Bright field images of the worms were taken using 0.63× objective of ZEISS SteREO Discovery V16 and the AxioCam camera with parameters "magnification 40×, field of view 5.8 mm, resolution 1.1 µm, depth of field 26 µm." Images were analyzed using the Wormsizer plug-in for Image J/Fiji (Schindelin et al. 2012). The raw data of worm size measurement in this study can be found in Supplementary Data 1.

Quantification of morphological change in mutant animals

We quantified morphological differences in the mouths of wild-type and mutant worms, using our recently published protocol (Theska et al. 2020), which combines landmark-based geometric morphometrics with model-based clustering (a detailed account of our analysis is provided in the Supplementary Material). In short, we obtained image stacks of the nematode mouths in lateral position and recorded the X and Y coordinates of 15 fixed landmarks (Fig. 1, c and d and Supplementary Table 4) which capture the overall form of the mouth using FIJI (ver. 2.1.0) (Schindelin et al. 2012). We generated landmark data for 193 animals and created 2 separate data sets: (1) a *P. pacificus* (exclusively Eu animals) data set ($n = 118$) comprising the wild-type strain (PS312), a frameshift mutant of *Ppa-dpy-6* (*tu1645*), and a "rescue" strain RS4003 (*tu1696*); and (2) a *C. elegans* data set ($n = 75$) comprising the wild-type N2 and the *Cel-dpy-6* mutant CB14. All steps of the geometric morphometric analysis of mouth form, shape, and size were performed in R (ver. 4.1.1) (R Core Team 2021), using a combination of the GEOMORPH (ver. 4.0.1) (Adams and Otárola-Castillo 2013; Adams et al. 2021) and MORPHO (ver. 2.9) (Schlager et al. 2017, 2021) packages. Shape and form differences were visualized using PCA; differences in mouth size were approximated by the centroid size and visualized in boxplots. To test whether mouth form, shape, and/or size changed due to a mutation in *dpy-6*, we performed distance-based permutational MANOVA (PERMANOVA) (Anderson 2001, 2014) with a randomized residual permutation procedure (RRPP) (Collyer et al. 2015; Collyer and Adams 2018, 2021), using linear model fits produced with the *procD.lm* function of GEOMORPH. Pairwise comparisons among strains were performed using the *pairwise* function of the RRPP package (ver. 1.1.0; Collyer et al. 2015, 2018, 2021). To identify groups of specimens with similar morphology, we performed model-based clustering with the *Mclust* function of the MCLUST R package (ver. 5.4.7) (Li and Durbin 2009; Fraley et al. 2021). We only used "meaningful" principal components of shape or form variation (identified with MORPHO's *getMeaningfulPCs* function in Schlager 2017; Schlager et al. 2021) as input variables for clustering, in order to avoid overparameterization of the multivariate normal models. To assess static allometry in the adult nematode mouths, we generated common and group-specific allometry model fits with *procD.lm* and compared them in a homogeneity of slopes (HOS) test using the *anova* function of the R STATS package (R Core Team 2021). The compatibility of our data with a respective allometry hypothesis was tested using the aforementioned PERMANOVA-RRPP approach (Collyer et al. 2015). Allometric trajectories were visualized by plotting the first principal component (PC1) of a matrix containing predicted shapes from a multivariate regression of shape on size against the log-transformed centroid size of each specimen, using the *plotAllometry* function of GEOMORPH (Adams and Otárola-Castillo 2013; Adams et al. 2021).

EMS mutagenesis

To obtain Dumpy mutants with a short body size, we performed a large-scale genetic screen using ethyl methanesulfonate (EMS) mutagenesis in *P. pacificus* PS312. Specifically, we incubated a mixture of J4 larvae and young adults in M9 buffer (3 g/l KH_2PO_4, 6 g/l Na_2HPO_4, 5 g/l NaCl, 1 mM $MgSO_4$) with 47 mM EMS for 4 h (Sommer et al. 1996). A total of 1,500 healthy young adults were singled out on new plate in several rounds of mutagenesis. After animals had laid approximately 20 eggs, parents were removed and F1 progeny were allowed to reach maturity. A total of 5–10 F1 animals (which contain heterozygous mutants) from each P0 plate were then singled out. In total, 9,000 F1 animals were isolated and F2 progeny (which contained a mixture of genotypes, including homozygous mutants) were allowed to develop to adulthood. We isolated *Dpy* mutants that are shorter than wild-type using a discovery V20 stereomicroscope (Zeiss). Around 200 *Dpy* mutants were isolated and re-screened for mouth-form defects in the F3 or F4 generation using DIC microscopy. In total, we isolated 6 EMS mutants displaying a simultaneous *Dpy* and mouth-form phenotype (Fig. 2a and Supplementary Fig. 1a).

Mapping of *tu743*

For mapping, the EMS mutant *tu743* was backcrossed 3 times to the PS312 wild-type strain. The backcrossed line was sequenced by generating a NGS library using Low Input Library Prep kit (Clontech) and sequenced on the Illumina HiSeq3000 platform. Raw Illumina reads of the *tu743* mutant and of a mapping panel were aligned to the El Paco assembly of the *P. pacificus* genome (Dieterich et al. 2008; Rödelsperger et al. 2017) by using the BWA software package (ver. 0.6) with parameter "bwa mem -M" (Li and Durbin 2009). Mutations were called by using the samtools (ver. 1.10) software with the parameter "bcftools mpileup—max-depth 1000 | bcftools call -cv" (Li et al. 2009). In total, 98 nonsynonymous/nonsense mutations (Supplementary Table 2) and 7 mutations in introns near splice sites of annotated genes (El Paco gene annotations v3) were identified in the candidate interval by a previously described custom variant classification software (Rae et al. 2012). The single-worm transcriptome of entire postembryonic stage in *P. pacificus* was downloaded from the European Nucleotide Archive with the study accession number PRJEB42613 (https://www.ebi.ac.uk/ena/browser/view/PRJEB42613).

CRISPR/Cas9 mutagenesis

We followed the previously published protocol for *P. pacificus* with subsequently introduced modifications (Han et al. 2020; Nakayama et al. 2020). All target-specific CRISPR RNAs (crRNAs) were designed to target 20 bp upstream of the protospacer adjacent motifs (PAMs). A total of 10 µl of the 100 µM stock of crRNA (CRISPR/Cas9 RNA; IDT) was combined with 10 µl of the 100 µM stock of tracrRNA (catalog # 1072534; IDT), denatured at 95°C for 5 min, and allowed to cool down to room temperature and anneal. The hybridization product was combined with Cas9 protein (catalog# 1081058; IDT) and incubated at room temperature for 5 min. The mix was diluted with Tris-EDTA buffer to a final concentration of 18.1 µM for the RNA hybrid and 2.5 µM for Cas9. For the induction of specific site-directed mutations via CRISPR/Cas9, a ssDNA oligo template was included in

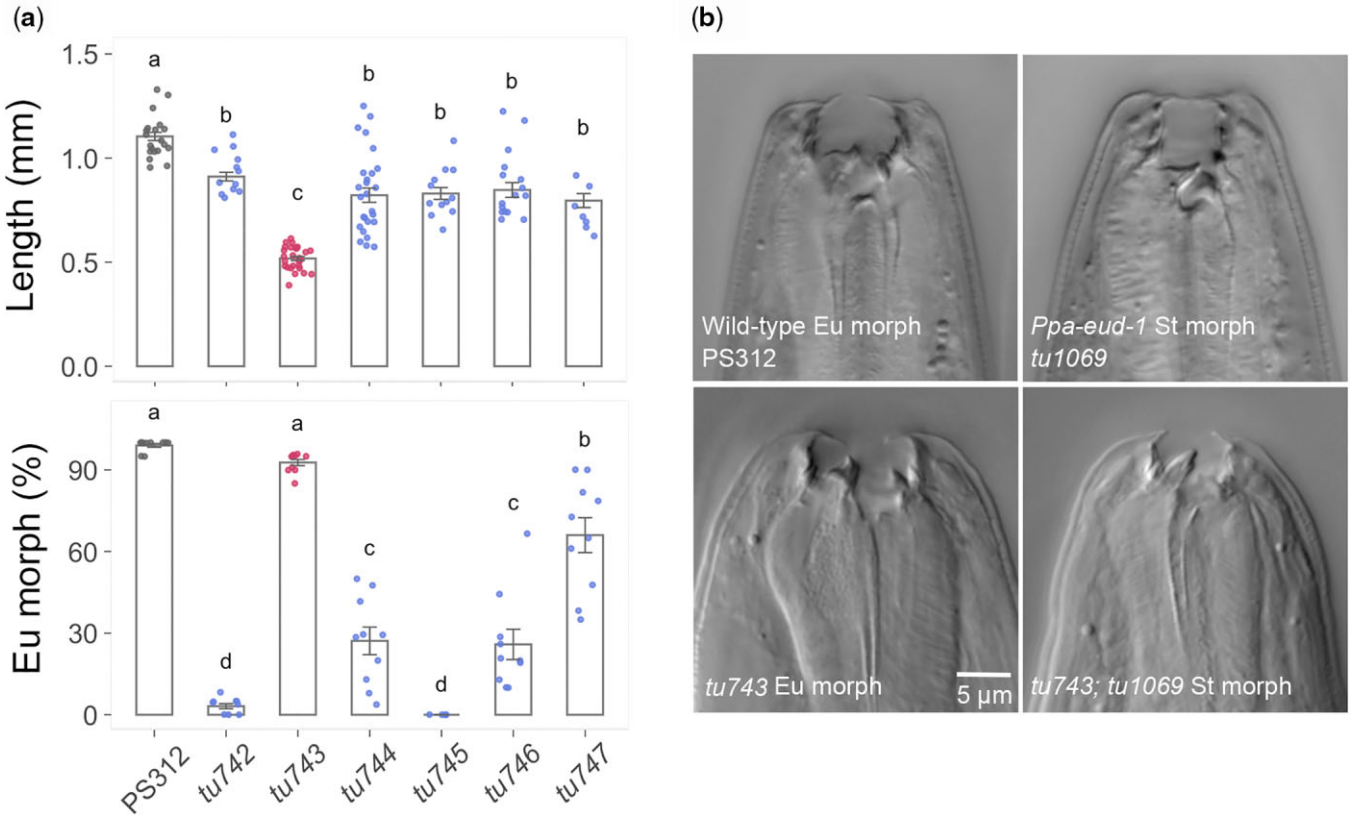


Fig. 2. Dumpy and mouth-form phenotypes of 6 novel *P. pacificus* mutants. a) The upper bar plot shows differences in the length of the worm body among wild-type (PS312) and the 6 mutant strains. The lower bar plot displays changes in mouth-form ratios between strains. Turkeys multiple comparison were tested by the Turkeys HSD test, with letters upon each volume indicating significant differences between the means, *P*-adjusted value < 0.05, $n > 20$. b) Mouth form of wild-type Eu, St morph of the *Ppa-eud-1* mutant (*tu1069*), Eu morph with reduced cheilostom of the *tu743* mutant strain, and the St morph with a reduced cheilostom of *tu743* in a *Ppa-eud-1* double mutant. Additional images in 2 focal planes are shown in Supplementary Fig. 1b.

the mix at a concentration of 4 μM. The repair template contained the desired modifications flanked by 50 bp homology arms either side of the edited sequences. The plasmid carrying the *Ppa-eft-3* promoter and modified TurboRFP sequences (Han et al. 2020) was used as co-injection marker. sgRNAs, associated primers and repair templates for generating mutants and specific targeted knock ins utilized in this study can be found in Supplementary Table 5.

Injections were performed on a Zeiss Axiovert microscope (Zeiss, Germany) coupled to an Eppendorf TransferMan micromanipualtor and Eppendorf FemtoJet injector (Eppendorf AG., Hamburg, Germany). The microinjection mixture was injected in the gonad rachis of approximately 1-day-old adult hermaphrodites. Eggs laid by injected animals within a 12–16-h period post injection were recovered. After 2 days, the P0 plates containing the F1 animals with fluorescent signal of co-injection marker were isolated, and 8–10 F1 progenies from the isolated P0 plates were singled out on individual plates. After F1 animals have laid eggs, they were placed in 10 μl of single worm lysis buffer (10 mM Tris–HCl at pH 8.3, 50 mM KCl, 2.5 mM $MgCl_2$, 0.45% NP-40, 0.45% Tween 20, 120 μg/ml Proteinase K), and incubated in a thermocycler at 65°C for 1 h, followed by heat deactivation of the proteinase at 95°C for 10 min. The resulting lysate was used as a template in subsequent PCR steps. The genotype of the F1 animals were subsequently analyzed via Sanger sequencing and mutations identified before re-isolation in homozygosis.

Genetic transformation

Transcriptional reporter lines of *Ppa-dpy-6* were generated by PCR amplification of a 2.2-kb upstream region of the predicted start codon (Supplementary Table 6) using Gibson assembly (New England Biolabs). The resulting product was cloned into a pUC19 vector containing GFP (Han et al. 2020) and the *Ppa-rpl-23* 3′UTR. For injection, we used 10 ng/ml of the NotI-digested *Ppa-dpy-6*::GFP plasmid with 60 ng/ml of genomic carrier DNA, also digested with NotI. We found the same expression pattern in 2 independent lines. For the reporter construct of *Cel-dpy-6*, we cloned a 2.2-kb upstream region of the predicted start codon in N2 (Supplementary Table 6) into a pUC19 vector containing gfp and *Cel-rpl-23* 3′UTR via Gibson assembly (New England Biolabs). The injection mix for transformation contained 10 ng/μl of the reporter construct and 60 ng/μl genomic DNA. The same expression pattern was confirmed in 3 independent lines.

Analysis of oscillatory gene expression of *Ppa-dpy-6* and *Cel-dpy-6*

To profile the temporal gene expression of *Ppa-dpy-6* and *Cel-dpy-6* during postembryogenesis, we extracted the FPKM of *Ppa-dpy-6* at 38 time points with 3 biological replicates (Sun et al. 2021) and the RPKM of *Cel-dpy-6* at 48 time points (Meeuse et al. 2020). Mean FPKM at each time point was calculated for profiling the gene expression of *Ppa-dpy-6*, and RPKM at 48 time points was used to display the gene expression of *Cel-dpy-6*. The meta2d_Base and meta2d_AMP of oscillating gene set, which was generated from the results of the METACYCLE package, were used to both, profile the distribution of the mean gene expression and amplitude of all oscillating genes and the comparison of the mean expression and amplitude between *dpy-6* and the average of oscillating gene sets in both species.

Comparison of gene expression of *Ppa-dpy-6* between wild-type and *eud-1* mutant animals

We collected the transcriptome data at 20 h posthatching for PS312 wild-type and *eud-1(tu1069)* mutant animals from European Nucleotide Archive under the study accession number PRJEB42633 to compare the gene expression of *Ppa-dpy-6* between *Ppa-eud-1* mutant and wild-type (Sun et al. 2021). The RNA-seq data for the comparison of *Ppa-dpy-6* expression between the *Ppa-nhr-40* mutant and wild-type, and the *Ppa-nhr-1* mutant and wild-type, were downloaded from the European Nucleotide Archive with the study accession number PRJEB34615. We reperformed differential expression analysis using DESEQ2 (ver. 1.18.1) (Love et al. 2014) and applied an adjusted *P*-value cut-off of 0.05 and a fold change cut-off of 2 to identify differentially expressed genes. To compare the expression of *Ppa-dpy-6* between mouth-form mutants and wild-type, we extracted the FPKM values of *Ppa-dpy-6* (Supplementary Table 7) from each data set and applied a comparison by using the Tukey HSD test function in the AGRICOLAE package (version 1.3.3) with default parameters.

Results

A genetic screen for Dumpy mutants with accompanying mouth-form defects

To isolate genetic mutants defective in *P. pacificus* stomatal structures, we performed a combinatorial screen by first obtaining *Dpy* mutants that are shorter than wild-type, followed by re-screening these mutants for mouth-form defects. This rationale was based on preliminary findings that some available *Dpy* and *Uncoordinated Movement* (*Unc*) mutants showed concomitant defects of stomatal structures (Müller 2010). From a screen of 1,500 mutagenized animals, a total of 9,000 F1 progeny were singled out to individual plates. This resulted in the isolation of around 200 *Dpy* mutants with a phenotypic range similar to previous studies (Kenning et al. 2004). During re-screening, we found 6 mutants to have a concomitant mouth-form phenotype (Fig. 2a, Supplementary Fig. 1a, and Supplementary Tables 1 and 2). We focus on the mutant RS3887*(tu743)*, which is severely *Dpy* and has a strong morphological phenotype in the stoma. First inspection by light microscopy of RS3887*(tu743)* mutant animals indicated a strong reduction in the size of the cheilostom in these otherwise Eu animals (Fig. 2b). To investigate the role of RS3887*(tu743)* in St animals, we crossed this line into the *Ppa-eud-1* mutant background that results in an all-St phenotype (Ragsdale et al. 2013). In *eud-1(tu1069)*; *tu743* double mutants, we found similar defects in the cheilostom (Fig. 2b and Supplementary Figs. 1b and 3b). Thus, RS3887*(tu743)* affects the cheilostom in Eu and St animals and represents the first mutant with a specific defect in the nematode stoma.

RS3887 carries a mutation in the Mucin-type gene *Ppa-dpy-6*

To identify the molecular lesion responsible for the cheilostom defect of RS3887, we backcrossed this mutant line to wild-type animals and performed whole-genome sequencing (see Materials and Methods for detailed methodology). We found a total of 109 single-nucleotide polymorphisms (SNPs) resulting in nonsynonymous or nonsense mutations in annotated genes or intronic changes near splice sites in these backcrossed animals (Supplementary Fig. 2a and Supplementary Table 3). This included 1 SNP in the *P. pacificus* ortholog of *Cel-dpy-6*. As RS3887 shows a severe *Dpy* phenotype, *dpy-6* was a strong candidate for the causal gene behind the combined *Dpy* and cheilostom-defective phenotype. Note that the C>T SNP in *Ppa-dpy-6* is located at the 3′-splice site of intron 10 (Supplementary Fig. 2c). Therefore, we used CRISPR/Cas9-mediated gene knockouts and engineering to confirm the phenotypic consequences of this

Table 1. *Ppa-dpy-6* associated phenotypes.

	Genotype	Mutagen	Mutation		Body shape	Cheilostom	% Eu
a	PS312	n.a.	n.a.	n.a.	WT	WT	>95
b	RS3887 (*tu743*)	EMS	C>T intronic	11th exon	Dumpy	Short	>95
c	*Ppa-dpy-6* (*tu1643*)	CRISPR	Frameshift	3rd exon	Dumpy	Short	>95
	Ppa-dpy-6 (*tu1644*)	CRISPR	Frameshift	3rd exon	Dumpy	Short	>95
	Ppa-dpy-6 (*tu1645*)	CRISPR	Frameshift	3rd exon	Dumpy	Short	>95
d	*Ppa-dpy-6* (*tu1646*)	CRISPR	C>T intronic	11th exon	Dumpy	Short	>95
e	*Ppa-dpy-6* (*tu1658*)	CRISPR	Frameshift	11th exon	Dumpy	Short	>95
	Ppa-dpy-6 (*tu1659*)	CRISPR	Frameshift	11th exon	Dumpy	Short	>95
	Ppa-dpy-6 (*tu1660*)	CRISPR	Frameshift	11th exon	Dumpy	Short	>95
f	RS3887 reversion	CRISPR	T>C intronic	11th exon	WT	WT	>95
g	*Ppa-dpy-6* & *Ppa-eud-1*	EMS, CRISPR and cross	C>T intronic	11th exon	Dumpy	Short	<5

$N > 50$ for all strains. % Eu, percent eurystomatous animals; n.a., not applicable; WT, wild-type.

mutation. First, we created 3 mutant lines in *Ppa-dpy-6* using a sgRNA in exon 3, all of which resulted in frameshift mutations. All 3 alleles *Ppa-dpy-6* (*tu1643*, *tu1644*, and *tu1645*) are strongly *Dpy*, have a reduced cheilostom, but do not affect the mouth-form ratio (Table 1a–c, Supplementary Fig. 3c, and Supplementary Table 1). Thus, all 3 *Ppa-dpy-6* alleles have a phenotype that is indistinguishable of the original RS3887*(tu743)* mutant. Second, we reconstituted the original C>T mutation in intron 10 by CRISPR/Cas9 engineering and this mutant strain has a similar *Dpy* and cheilostom phenotype (Table 1d and Supplementary Fig. 3c). Three mutant lines with frameshift mutations in exon 11 were obtained from the same injection. All 3 alleles *Ppa-dpy-6* (*tu1658*, *tu1659*, and *tu1660*) have the same phenotype as RS3887 (Table 1e and Supplementary Fig. 3c). Finally, we reverted the C>T mutation back to wild-type in the RS3887 mutant background. The resulting strain RS4003 has a normal body shape and a normal cheilostom (Table 1f and Supplementary Fig. 3c). We conclude that mutations in *Ppa-dpy-6* result in strongly reduced body size and concomitant defects in the cheilostom. Thus, *Ppa-dpy-6* represents the first gene involved in the specification of the morphology of the stoma.

In *C. elegans*, *dpy-6* was cloned as part of genome-wide RNA interference screens, but was not further characterized (Simmer et al. 2003). DPY-6 shows sequence similarities to Mucins, a family of glycosylated hydrogel-forming proteins conserved throughout the animal kingdom. In humans, Mucins play important roles as mucus barriers in wet epithelial surfaces of the body, such as eyes, lungs, and the stomach with a proposed function as biological hydrogel that coats surface structures (Witten and Ribbeck 2017; Wagner et al. 2018), which would be compatible with the phenotype of *Ppa-dpy-6* in the cuticle and cheilostom.

Geometric morphometrics reveal mutant phenotype in mouth form, shape, and size

To analyze the mouth-form defects of *Ppa-dpy-6* with greater accuracy, we used our recently established protocol for geometric morphometric analyses in microscopic animals (Theska et al. 2020) and found a consistent phenotype in mouth shape (see Supplementary Material for detailed procedure). First, a PCA of mouth shape revealed that *Ppa-dpy-6(tu1645)* mutants occupy a separate aspect of the shape space when compared to wild-type (PS312) and rescued animals (RS4003, *tu1696*) (Fig. 3a). Note that wild-type and rescued animals occupy strongly overlapping areas of the shape space, indicating that their mouth shapes are very similar. Looking at the variation described by the individual principal components revealed that PC1 described 77.1% of the total variation, rendering this shape data almost 1-dimensional. However, PC1 is not the only PC that contains biologically relevant information. In fact, we found 3 PCs (PC1 = 77.1%, PC2 = 4.0%, and PC3 = 3.0%) to be "meaningful" according to the *getMeaningfulPCs* function of Morpho (threshold value = 1.29, based on $n = 118$). Our PCA indicates (1) a clearly different mouth shape in *Ppa-dpy-6(1645)* mutants compared to wild-type animals and (2) a practically complete reversal of the mutant shape when the function of *Ppa-dpy-6(1645)* is restored with CRISPR. This is also supported by multivariate statistical analysis, which shows that the independent variable "strain" strongly influences *P. pacificus* mouth shape as a response variable (PERMANOVA-RRPP: shape ~ strain, $Z = 5.9$, and $P = 0.0001$). Pairwise comparisons of distances between the group means of individual strains revealed that our shape data are incompatible with the null hypotheses that no mouth shape differences exist between *Ppa-dpy-6(tu1645)* mutants and wild-type or rescued animals (wild-type strain vs. *Ppa-dpy-6(1645)*: $Z = 3.5$ and $P = 0.0001$, rescue strain vs. *Ppa-dpy-6(1645)*: $Z = 3.6$ and $P = 0.0001$), but compatible with the null hypothesis that no clear differences exist in mouth shape between wild-type and rescued animals (wild-type strain vs. rescue strain: $Z = 1.2$ and $P = 0.1322$). Finally, we used model-based clustering to identify how many separate groups can be identified in the shape data and found only 2 separate clusters, one that contains all *Ppa-dpy-6* mutants and another containing all wild-type and rescued animals (Fig. 3b). Thus, PCA, PERMANOVA, and model-based clustering analyses reveal clear mouth shape alterations in *Ppa-dpy-6* mutants.

Second, we estimated the positive and negative extreme shapes for PC1 using wireframe representations of the *P. pacificus* mouth. We found that the most prominent mouth shape difference is an extremely shortened and slightly posteriorly shifted cheilostom in *Ppa-dpy-6(1645)* mutants. This is also reflected in very large relative contributions of landmarks 1 and 15 (anterior tips of the cheilostom) to PC1 (LM1 = 20.0% and LM15 = 20.7% of the described variation). Another obvious change is the anteriorly shifted (relative to the rest of the buccal cavity), but similarly sized, metastegostomal tooth (LM6 = 7.6% and LM7 = 12.3%, LM 9 = 7.9% of the variation described by PC1) (Fig. 3c).

Third, we investigated whether mouth size influences mouth shape, that is, whether there is allometry. A HOS test revealed that a common allometry model (shape ~ size + strain) is most appropriate to fit the shape data. Thus, we found that the slopes of the allometric trajectories among all 3 strains are similar, with mouth size moderately influencing mouth shape and "strain" being the major factor of influence (PERMANOVA-RRPP: size, $Z = 3.3$, and $P = 0.0001$; strain, $Z = 5.3$, and $P = 0.0001$). We conclude that our data are incompatible with a null hypothesis of isometry (see Supplementary Material for more details), but compatible with the alternative hypothesis of allometry. Additionally,

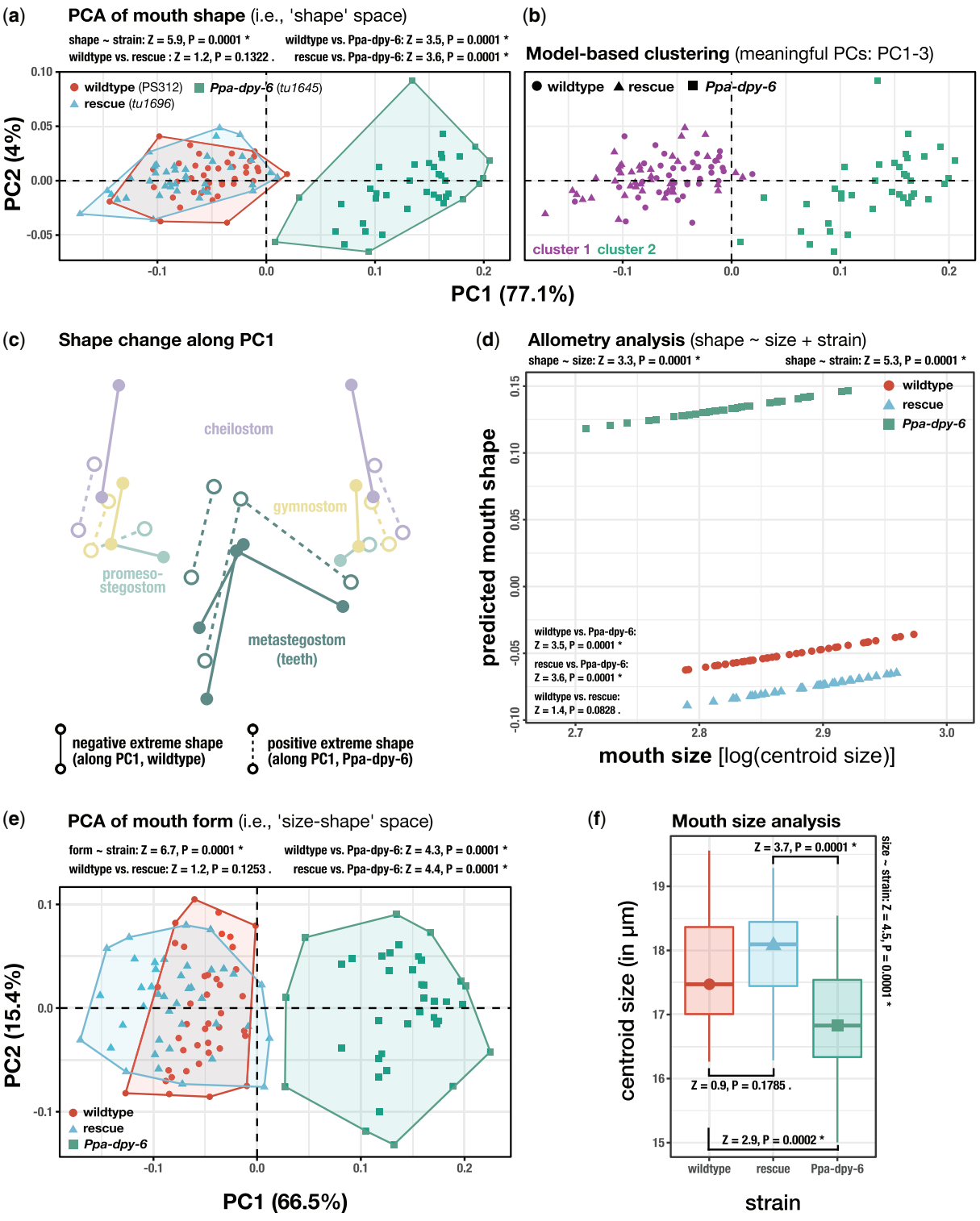


Fig. 3. Geometric morphometric analysis of mouth shape, size, and form in *P. pacificus* (exclusively Eu animals of each strain). a) PCA of the mouth shape for the wild-type (PS312), the *Ppa-dpy-6(tu1645)* mutant, and the rescue strain (RS4003, *tu1696*). The plot shows the distribution of all specimens in a shape space. b) Clusters identified in the *P. pacificus* shape data set. Clusters were found by model-based clustering, using all "meaningful" principal components as input variables. c) Superimposed wireframe representations of extreme shapes along PC1. d) Allometric trajectories for each of the *P. pacificus* strains. The allometry model that was used to fit the data is indicated above the plot. The y-axis corresponds the PC1 of a matrix of predicted values, which were obtained by a multivariate regression of shape on size. e) PCA of the mouth forms of each *P. pacificus* strain. The plot shows the distribution of all specimens in a "size-shape" (i.e. form) space. f) Boxplot showing the centroid sizes for all specimens of each *P. pacificus* strain. The box indicates the interquartile range (IQR, i.e. the range from the first to the third quartile) and the solid line indicates the median centroid size. Lower whisker extends from the lower quartile to the smallest value which is at most 1.5 IQR away from the lower quartile. Upper whisker extends from the upper quartile to the largest value which is at most 1.5 IQR away from the upper quartile. Specimens which fall beyond the whiskers (in either direction) are plotted as individual dots and are regarded to be outliers. Relative effect sizes (*Z*-scores) and *P*-values obtained via PERMANOVA-RRPP are indicated where appropriate. Asterisks "*" indicate incompatibility with a null hypothesis, while periods "." indicate compatibility (see Supplementary Material for details).

pairwise comparisons of LS means among strains revealed that our allometry data are compatible with the null hypothesis of no blatant differences between wild-type and rescued worms (wild-type strain vs. rescue strain: $Z = 1.4$ and $P = 0.0828$), while we could find clear differences in the LS means of *Ppa-dpy-6* mutants on the one hand, and wild-type and rescued animals on the other (wild-type strain vs. *Ppa-dpy-6*: $Z = 3.5$ and $P = 0.0001$, rescue strain vs. *Ppa-dpy-6*: $Z = 3.6$ and $P = 0.0001$). Most importantly, the allometric trajectories of *Ppa-dpy-6* and wild-type animals have very different intercepts along the y-axis (Fig. 3d). Thus, while all 3 strains have similar allometric slopes, they do not share a general underlying trajectory and the mutant mouth shape cannot be obtained by scaling up or down the size of the wild-type mouth shape along such a general allometric trajectory.

Fourth, we analyzed morphometric differences in mouth form. Some clarification is appropriate here: previously, we colloquially used the term "mouth form" to refer to the 2 different mouth morphologies (Eu and St) that can be found in *P. pacificus*. In this section however, we strictly use the morphometric denotation of *form*. Briefly, in traditional morphometrics size and shape are treated as logically separate aspects of a biological structure and it is common practice to analyze variation in size separate from variation in shape (Mitteroecker et al. 2013). When taken together, the shape and the size of a structure define its *form*. Thus, analyses which focus on differences in *form* address the entirety of morphological variation in a biological structure, rather than variation in a specific structural aspect like shape or size. In modern morphometrics, multiple tools are available to investigate biological form and here we performed a "size-shape" PCA *sensu* Mitteroecker et al. (Mitteroecker et al. 2004, see Supplementary Material for additional information). We plotted the first 2 principal components of mouth-form variation against each other to visualize a form space (Fig. 3e). PC1 of the "size-shape" PCA describes less variation (66.5%) than PC1 of the shape PCA (77.1%), while PC2 of the "size-shape" PCA describes much more variation (15.4%) than PC2 of the shape PCA (4%). We were able to recover a clear separation of *Ppa-dpy-6* mutant mouth forms from wild-type and rescued animals, similar to what we found in the shape PCA (Fig. 3, a and e). This is also reflected in the statistical analysis of the mouth-form data set, which indicated that *Ppa-dpy-6* mutants differ from wild-type animals more prominently in mouth form ($Z = 6.7$, Fig. 3e) than in mouth shape ($Z = 5.9$, Fig. 3a). Pairwise comparisons of distances between group means revealed that our form data are incompatible with the null hypothesis that there are no mouth form differences between *Ppa-dpy-6* mutants, and wild-type or rescued animals (wild-type strain vs. *Ppa-dpy-6(1645)*: $Z = 4.3$, $P = 0.0001$, rescue strain vs. *Ppa-dpy-6(1645)*: $Z = 4.4$ and $P = 0.0001$). Similar to the results from our shape analysis, our form data are compatible with the null hypothesis that there are no discernable mouth form differences between wild-type and rescued animals (wild-type vs. rescue strain: $Z = 1.2$, $P = 0.1253$), indicating a reversal of the mutant mouth form to the wild-type situation after CRISPR-repair of the mutant *Ppa-dpy-6(1645)* allele. This is further corroborated by model-based clustering: we were able to identify 2 separate clusters in our form space which almost perfectly correspond to *Ppa-dpy-6(1645)* mutant animals on the one hand, and wild-type and rescued animals on the other (Supplementary Fig. 3).

Lastly, we investigated mouth size using the centroid size as proxy. We found that, on average, wild-type and rescued worms tend to have larger mouths than both *Ppa-dpy-6* mutants (Fig. 3f). The median centroid sizes of wild-type and rescued animals are 17.47 and 18.10 µm, while it is 16.83 µm in *Ppa-dpy-6(1645)* mutants. Thus, the median centroid size of wild-type animals and the first (lower) quartile of rescued animals correspond to the third (upper) quartile of *Ppa-dpy-6(1645)* mutants (17.54 µm) and the median centroid size of *Ppa-dpy-6(1645)* mutants is smaller than the first (lower) quartile of both, wild-type and rescued animals (Fig. 3e). Statistical analysis indicates that our size data are incompatible with the null hypothesis of no mouth size differences among the 3 strains (PERMANOVA-RRPP: mouth size ~ strain: $Z = 4.5$ and $P = 0.0001$). Interestingly, pairwise comparisons of distances between group means revealed that these data are compatible with the null hypothesis that no blatant mouth size differences exist between wild-type and rescued animals ($Z = 0.9$ and $P = 0.1785$), while it is incompatible with the null hypotheses of no mouth size differences between the *Ppa-dpy-6* mutant strain and the wild-type or rescue strain (*Ppa-dpy-6(1645)* vs. wild-type strain: $Z = 2.9$ and $P = 0.0002$, *Ppa-dpy-6(1645)* vs. rescue strain: $Z = 3.7$, $P = 0.0001$). Thus, we were able to fully reverse the observed mutant phenotypes in mouth form, mouth shape, and mouth size by repairing the mutant allele of *Ppa-dpy-6(1645)* with CRISPR. The high resolution of the phenotypic analyses we applied here highlight the power and necessity for modern geometric morphometrics in evo devo studies on miniature animals.

Ppa-dpy-6 is an oscillatory gene, requires *eud-1*-dependent environmental signaling, and is expressed in the hypodermis

Given the prominent stoma phenotype, we next wanted to study gene regulation of *Ppa-dpy-6*. Single-worm transcriptomics (SWT) in *P. pacificus* has recently provided a high-resolution map of the postembryonic developmental transcriptome (Sun et al. 2021). Of the more than 28,000 predicted genes of *P. pacificus*, nearly 3,000 show oscillatory expression with discrete phases, amplitudes, and expression levels (Sun et al. 2021). This includes many of the body-wall collagens that are often transcribed before molting and are involved in cuticle formation. Therefore, we analyzed the expression profile of *Ppa-dpy-6* and found strong oscillation in expression with the highest peak 4 h before the second, third, and fourth molt (Fig. 4a). This expression profile is consistent with a role of *Ppa-dpy-6* in the synthesis of the new cuticle. When we profiled the expression of *Cel-dpy-6* during postembryogenesis using the data of Meeuse et al. (2020), we found that the expression of *Cel-dpy-6* also displayed an oscillating pattern coupled to the molting cycle (Fig. 4b). Interestingly, profiling the distribution of the mean expression and relative amplitude of oscillating genes in both species, we found that *Ppa-dpy-6* and *Cel-dpy-6* exhibited higher gene expression levels and greater oscillating amplitudes than the mean value of the oscillating gene sets (Fig. 4c). Together, these temporally conserved gene expression patterns of *Ppa-dpy-6* and *Cel-dpy-6* indicate a highly conserved role in cuticle formation.

Next, we wanted to know if *Ppa-dyp-6* is controlled by the mouth-form plasticity gene regulatory network that has recently been identified (Sieriebriennikov et al. 2020). The sulfatase-encoding gene *eud-1* is a major developmental switch involved in environmental sensing and developmental reprogramming (Ragsdale et al. 2013; Sommer 2020). SWT of *eud-1* mutant animals have recently been performed at 20 and 28 h after hatching (Sun et al. 2021). When we compared *Ppa-dpy-6* expression in wild-type and *Ppa-eud-1* mutant animals, we found a strong upregulation of *Ppa-dpy-6* in the mutant condition (Fig. 4d). Thus, *Ppa-dpy-6* expression is sensitive to *Ppa-eud-1*-dependent environmental signaling. In contrast, we found no difference in *Ppa-dpy-6* expression when

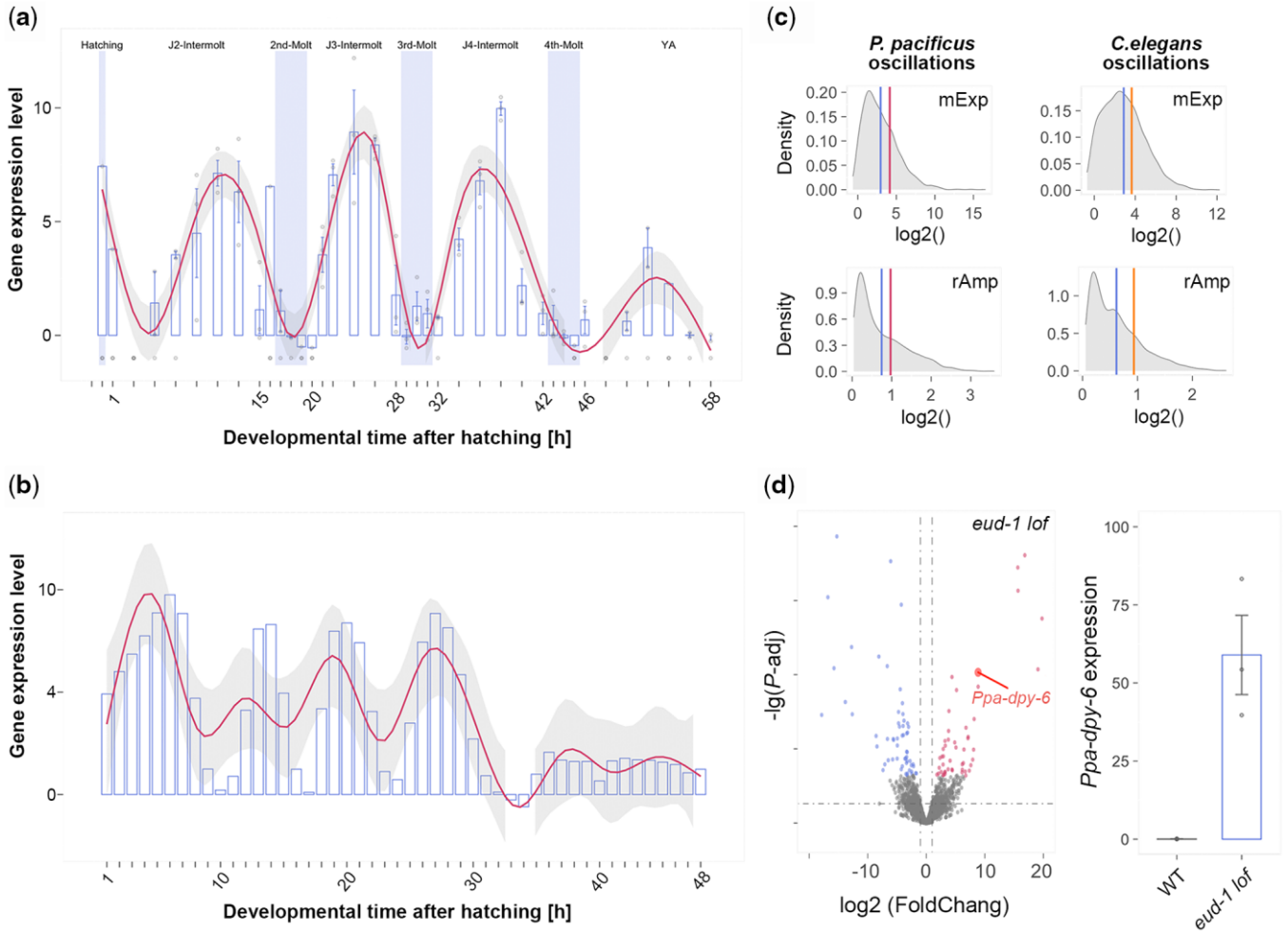


Fig. 4. Temporal expression patterns of *Ppa-dpy-6* and expression of *Ppa-dpy-6* in the *Ppa-eud-1* mutant condition. a) Dynamic expression of *Ppa-dpy-6* during postembryonic development. X-axis represents the 38 time points after hatching, the y-axis is the log2() of FPKM, 3 replicates per time point. The data were extracted from the single-worm transcriptome data set of Sun et al. (2021). b) Bar plot profiles of the dynamic expression of *Cel-dpy-6* during postembryonesis using the data set of Meeuse et al. (2020). c) Line plots profiling the comparison of the mean expression and the relative amplitude between *Ppa-dpy-6* (red) and the mean of the *P. pacificus* oscillatory gene set (blue), and between *Cel-dpy-6* (yellow) and the mean of *C. elegans* oscillating gene set (blue). d) Vocano plot revealing that *Ppa-dpy-6* is one among the strongest differentially expressed genes (DEG) in the comparison between *Ppa-eud-1* mutant and wild-type conditions. Downregulated genes in blue, upregulated genes in red, and nonsignificantly expressed genes in gray. Bar plot displaying differences in gene expression of *Ppa-dpy-6* in *Ppa-eud-1* and wild-type condition. All samples used are at 20 h after hatching. Turkeys multiple comparison were tested by the Turkeys HSD test, with the letters upon each volume indicating significant differences between the means, *P*-adjusted value <0.05.

comparing the transcriptomes of wild-type animals and those of the 2 nuclear hormone receptor (NHR) mutants *Ppa-nhr-1* and *Ppa-nhr-40* (Sieriebriennikov et al. 2020) (Supplementary Fig. 6, a and b). This finding indicates that during evolution toward the *Pristionchus* lineage *dpy-6* has come under the control of the *eud-1* switch locus as this gene does not exist in *C. elegans*. Furthermore, these data suggest that not all *eud-1*-dependent environmental sensing goes through *Ppa-nhr-1* and *Ppa-nhr-40* and that other downstream transcription factors likely exist.

Finally, to study the expression pattern of *Ppa-dpy-6*, we generated a 2-kb promoter fusion construct with GFP. *Ppa-dpy-6* is continuously expressed in postembryogenesis and adult stages (Fig. 5). We observed expression in the hypodermal syncytia hyp1, hyp5, hyp6/7, and the anterior and posterior arcade cells (Fig. 5 and Supplementary Video 1), the characterization of which is based on the recent reconstruction of the head hypodermis in *P. pacificus* (Harry et al. 2021). Interestingly, when we generated a *Cel-dpy-6* reporter construct, we found a similar expression pattern except for *C. elegans* also showing expression in hyp4 (Table 2). We conclude that the oscillatory gene expression and the large overlap in gene expression of *Ppa-dpy-6* and *Cel-dpy-6* are consistent with a conserved role in cuticle formation. Therefore, we finally used geometric morphometrics of *Cel-dpy-6* mutants to search for a stoma phenotype, as no defects in the mouth were previously reported.

Cel-dpy-6 mutants have defects in the cheilostom and gymnostom

We compared the mouth shapes of wild-type (N2) and *Cel-dpy-6(e14)* mutants (CB14) and found a strong mutant phenotype that is reminiscent to the one observed for *Ppa-dpy-6*. Specifically, PCA of mouth shapes revealed that *Cel-dpy-6(e14)* mutants occupy a different part of the shape space than wild-type animals with both strains being almost completely separated along PC1 (Fig. 6a). Similar to our findings in the *Ppa-dpy-6* analysis, the first 3 principal components of shape variation in *C. elegans* (PC1 = 38.5%, PC2 = 23.7%, and PC3 = 7.7%) can be regarded as "meaningful" according to the *getMeaningfulPCs* function of

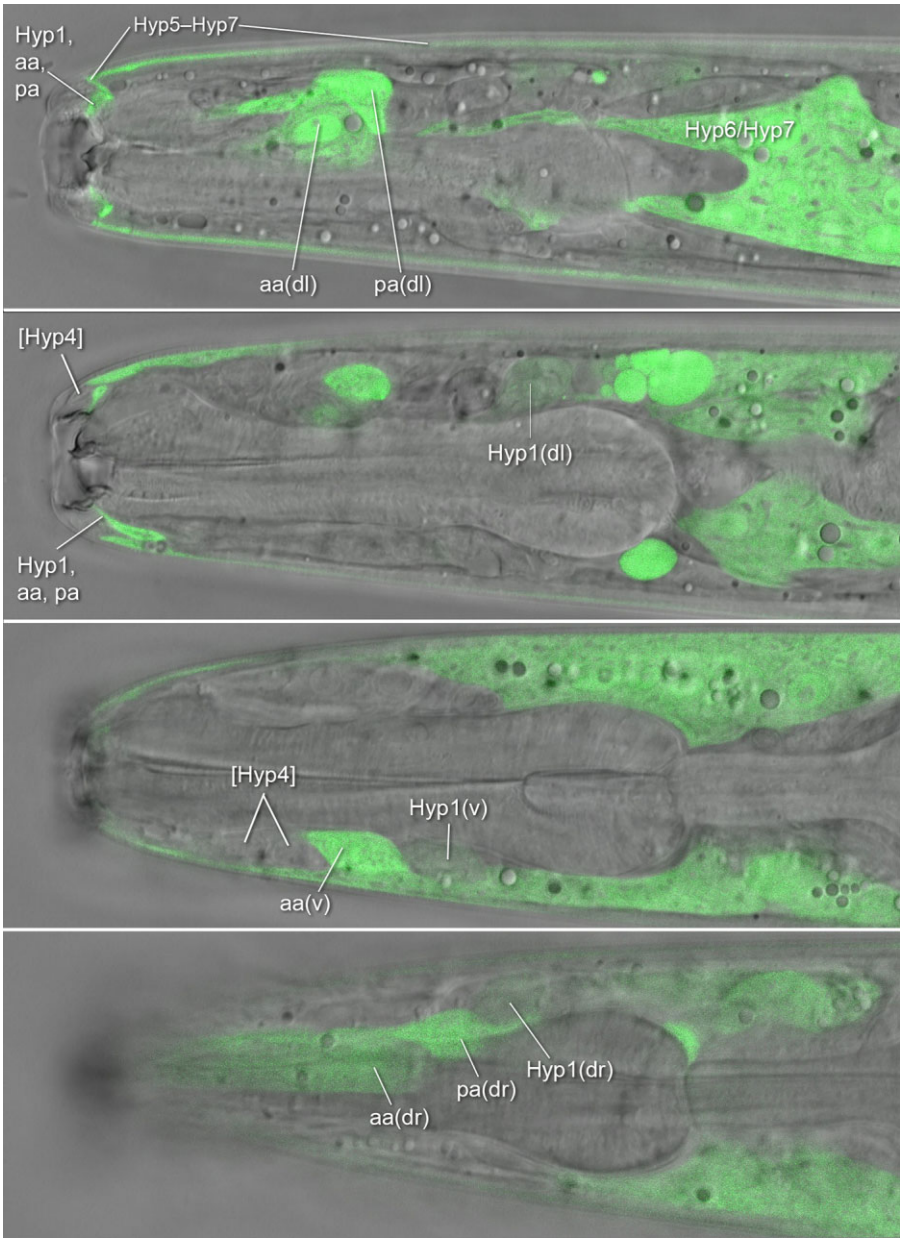


Fig. 5. Spatial expression patterns of *Ppa-dpy-6*. Spatial expression of *Ppa-dpy-6* from a 2.2-kb *Ppa-dpy-6* promoter element driving GFP expression in different focal planes. aa, anterior arcade cell; pa, posterior arcade cell; Hyp, hypodermal syncitia; dr, dorsal right; dl, dorsal left; v, ventral. Images were analyzed using ImageJ software and Arivis Vision 4D software (Arivis AG, Berlin, Germany).

Table 2. *dpy-6* expression in anterior epithelial syncytia in *P. pacificus* and *C. elegans*

	P. pacificus	*C. elegans*
Posterior arcade	+	+
Anterior arcade	+	+
Hyp1	+	+
Hyp2	−	−
Hyp3	−	−
Hyp4	−	+
Hyp5	+	+
Hyp6/7	+	+

Morpho (threshold value = 1.38, based on $n = 75$). The presence of a *Cel-dpy-6(e14)* mutant phenotype in mouth shape is further supported by multivariate statistical analysis (PERMANOVA-RRPP: shape ~ strain, $Z = 5.4$ and $P = 0.0001$). Pairwise comparisons of distances between group means revealed that shape data are incompatible with the null hypothesis of no mouth shape differences between the *Cel-dpy-6(e14)* and wild-type ($Z = 5.1$, $P = 0.0001$). This conclusion is further strengthened by model-based clustering, which identified separate wild-type and mutant clusters in the shape data (Fig. 5b). We conclude that mutating the one-to-one ortholog of *Ppa-dpy-6* in *C. elegans* also results in a clearly identifiable mutant mouth shape defect, just as observed for *Ppa-dpy-6* mutants.

Given this conclusion, we wondered whether the shape differences between wild-type and *Cel-dpy-6* mutants actually stem from morphological changes in the same structures for which we observed shape differences in *P. pacificus*. Comparing the estimated extreme shapes along PC1 revealed that 2 components of the mouth are strongly affected by the genetic perturbation.

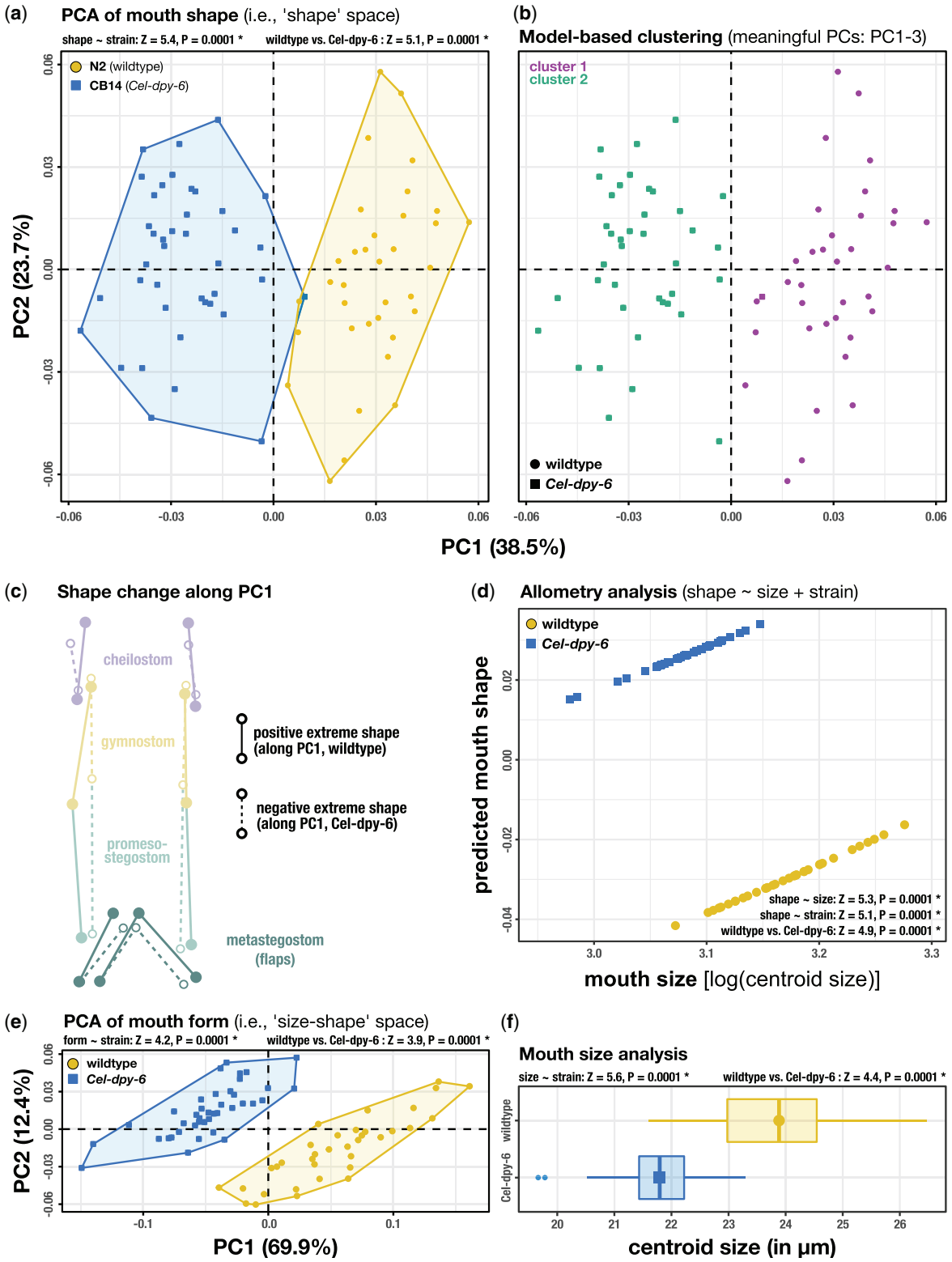


Fig. 6. Geometric morphometric analysis of mouth shape, size and form in *C. elegans*. a) PCA of the mouth shape for the wild-type (N2) and the *Cel-dpy-6* mutant strain (CB14). The plot shows the distribution of all specimens in a shape space. b) Clusters identified in the *C. elegans* shape data set. Clusters were found by model-based clustering, using all "meaningful" principal components as input variables. c) Superimposed wireframe representations of extreme shapes along PC1. d) Allometric trajectories for the *C. elegans* strains. The allometry model that was used to fit the data is indicated above the plot. The y-axis corresponds the PC1 of a matrix of predicted values which were obtained by a multivariate regression of shape on size. e) PCA of the mouth forms of *C. elegans*. The plot shows the distribution of all specimens in a "size-shape" (i.e. form) space. f) Boxplot showing the centroid sizes for all specimens of both *C. elegans* strains. The box indicates the interquartile range (IQR, i.e. the range from the first to the third quartile) and the solid line indicates the median centroid size. Lower whisker extends from the lower quartile to the smallest value which is at most 1.5 IQR away from the lower quartile. Upper whisker extends from the upper quartile to the largest value which is at most 1.5 IQR away from the upper quartile. Specimens which fall beyond the whiskers (in either direction) are plotted as individual dots and are regarded to be outliers. Relative effect sizes (Z-scores) and *P*-values obtained via PERMANOVA-RRPP are indicated where appropriate. Asterisks "*" indicate incompatibility with a null hypothesis, while periods "." indicate compatibility (see Supplementary Material for details).

First, the cheilostom of *Cel-dpy-6(e14)* mutants is shortened along the anterior–posterior axis, and the mouth opening is wider than in wild-type worms. Second, the gymnostom of *Cel-dpy-6(e14)* mutants is shortened along the anterior-posterior axis and the main cuticular tube of the mouth (gymnostom+promesostegostom) is narrower than in wild-type worms (Fig. 6, a–c). These observations are also reflected in very large relative contributions of landmarks 1 and 15 (anterior tips of the cheilostom) to PC1 (LM1 = 17.1% and LM15 = 12.4% of the described variation), and of landmarks 4 and 12 (posterior tips of the gymnostom) also to PC1 (LM4 = 26.4% and LM15 = 8.5% of the described variation). Thus, we were able to identify a similar phenotype in cheilostom shape in *dpy-6* mutants of both nematode species, but only *C. elegans* shows a phenotype in gymnostom shape.

Next, we performed an allometry analysis similar to the one we used for *P. pacificus*. The HOS test revealed a common allometry model as most appropriate to fit *C. elegans* shape data, with both mouth size and strain being highly influential factors on mouth shape as a phenotypic response variable (PERMANOVA-RRPP: shape ~ size + strain; size, $Z = 5.3$ and $P = 0.0001$; strain, $Z = 5.1$ and $P = 0.0001$). We conclude that the *C. elegans* shape data set are incompatible with a null hypothesis of isometry, and instead, is compatible with the alternative hypothesis of allometry (Fig. 5d), a finding in line with our observations for *P. pacificus*. Interestingly, the *dpy-6(e14)* mutation apparently does not change the allometric trend in adult stomatal growth of both nematode species (i.e. the slope of the trajectory; see Supplementary Material for more details), but it changes adult mouth shape and size. The mouth shape phenotype in *Cel-dpy-6(e14)* mutants cannot be explained by an allometric scaling effect only, as this would require similar y-intercepts of mutant and wild-type trajectories, a finding that is congruent with the observation in *P. pacificus* (Figs. 3d and 6d).

Finally, we investigated whether there are phenotypes in mouth form and size in *Cel-dpy-6* mutants. The "size-shape" PCA revealed that wild-type animals and mutants occupy nonoverlapping aspects in the form space created by PC1 and PC2 (Fig. 6e). We found that the first 3 principal components of form variation can be considered "meaningful" (threshold value: 1.38, $n = 75$). Given that PC1 explains such a large amount of form variation (69.9%) and that there is a considerable overlap in shape variation along PC1, we consider the overall mouth form difference between wild-type and mutant to be less obvious than the difference in shape (Fig. 6, a and e). Therefore, the *C. elegans* strains show the opposite pattern to *P. pacificus* in that they differ more prominently in mouth shape ($Z = 5.4$) than in mouth form ($Z = 4.2$, compare Figs. 3, a and e and 6, a and e). Furthermore, pairwise comparison of the *C. elegans* strains revealed that our form data are incompatible with the null hypothesis that there are no mouth form differences between them (wild-type strain vs. *Cel-dpy-6*: $Z = 3.9$, $P = 0.0001$). When analyzing mouth size, we found that (on average) wild-type worms have larger mouths than *Cel-dpy-6(e14)* as the median centroid size of wild-type animals is 23.88 μm, while it is 21.79 μm for the *Cel-dpy-6(e14)* mutants (Fig. 6f). Additionally, the first (lower) quartile of wild-type mouth sizes (22.98 μm) is larger than the third (upper) quartile of mutant mouth sizes (22.22 μm), which we interpret as indicative of a clear mouth size phenotype in mutant animals. This is further supported by the statistical analysis, which found that our data are incompatible with the null hypothesis that there are no differences in mouth size (PERMANOVA-RRPP: mouth size ~ strain: $Z = 5.6$ and $P = 0.0001$). These findings perfectly correspond to the observations for *P. pacificus* (compare Figs. 3f and 6f).

In conclusion, *Ppa-dpy-6* and *Cel-dpy-6* show overlapping defects in all structural aspects of stoma morphology with some minor differences between species as revealed by our geometric morphometric analysis. This overall conclusion supports *dpy-6* as an important and evolutionary conserved gene controlling nematode mouth formation—the first one to be characterized in greater detail.

Discussion

We have identified the Mucin-type hydrogel-forming protein DPY-6 to play a major role in the formation of the cheilostom in both *P. pacificus* and *C. elegans*. The associated morphological defects indicate that the cheilostom and body wall cuticle share common structural components, a finding that supports the traditional notion that the cheilostom is continuous with the external cuticle. In both species, we observed *dpy-6* expression in the hypodermal syncytia hyp1, hyp5, and hyp6/7 and the anterior and posterior arcade cells indicating their involvment in the formation of the cuticle and the cheilostom. Strikingly, however, except for *dpy-6* no other *Dpy* mutants in *C. elegans* show defects of the cheilostom. Thus, the cheilostom is likely different from the cuticle in its exact chemical composition. While the conserved role of DPY-6 in cuticle deposition can easily be deduced from the strong *Dpy* phenotype, the full understanding of its role in stoma formation required sophisticated geometric morphometric analyses. Note that the cheilostom defect of *Ppa-dpy-6* is visible in standard light microscopy, whereas the *Cel-dpy-6* phenotype was more subtle and could be accessed properly only through geometric morphometrics. This highlights the power of modern quantification methods for morphological differences in miniature animals and indicates that these techniques would aid similar research programs in other organisms, such as rotifers or tardigrades. Together, the identification of the first structural component of the nematode stoma will initiate 3 new research directions.

First, our genetic analysis of *dpy-6* is consistent with its gene product serving as an initiation point for the identification of additional proteins of the *Pristionchus* stoma. This will hopefully help to address some of the ultimate questions associated with mouth-form plasticity in *P. pacificus*. Are there qualitative differences—that is novel structural components—in the teeth of *P. pacificus* or other members of the Diplogastridae that are unknown from other free-living nematodes? Also, are the differences between the Eu and the St mouth form of quantitative or qualitative nature? We speculate that qualitative differences between the stoma of *P. pacificus* and other nematodes such as *C. elegans* do exist and future studies will hopefully identify their molecular nature.

Second, the molecular nature of DPY-6 will result in nanotechnological approaches. DPY-6 shows high sequence similarity to human mucins, a group of glycoproteins that shape the barrier function of the mucus (Phani et al. 2018; Werlang et al. 2019). As typical for mucins, DPY-6 is a high-molecular weight protein with a large number of tandemly repeated serine and threonine residues in its backbone that serve as O-glycosylation sites. When hydrated, mucins form hydrogels with crosslinking domains including disulfide bonds supplied by cysteine residues (Wagner et al. 2018). Additionally, the glycosylation will further increase the complexity of protein aggregates. Importantly, multiple structural variations exist in human gel-forming mucins that are not fully understood. Therefore, future work on *Ppa-dpy-6* using the simplicity of the nematode body plan and phenotypic readout

might help to characterize the different domains of mucin proteins. In addition, in vitro studies of *Ppa*-DPY-6 can reveal its nanotechnological properties similar to investigations of the spider's web as another example of a biological hydrogel (Witten and Ribbeck 2017).

Finally, the identification of *Ppa-dpy-6* as a structural component of the stoma extends the knowledge of the gene regulatory network (GRN) involved in *P. pacificus* teeth formation and the evolutionary age structure of these genes. Many genes involved in environmental sensing and subsequent developmental switching have resulted from lineage-specific gene duplications, that is, the *eud-1*/sulfatase of *P. pacificus* is one of 2 duplicates of the single *sul-2* gene in *C. elegans* (Bui and Ragsdale 2019; Casasa et al. 2021). Interestingly and in contrast to these findings, the 2 known transcription factors *Ppa*-NHR-1 and *Ppa*-NHR-40 acting downstream in the GRN are highly conserved in nematode evolution. NHR-1 and NHR-40 of *P. pacificus* have 1:1 orthologs in *C. elegans*, while their known regulatory targets are again rapidly evolving (Sieriebriennikov et al. 2020). Ultimately, the identification of the structural components of teeth formation will represent a major milestone for a full understanding of the GRN of mouth-form plasticity. In conclusion, the identification of the first structural component of the nematode stoma will lead to novel activities in various research areas and provides an important framework for future investigations.

Data availability

Raw and processed data sets from this study have been submitted to the European Nucleotide Archive (ENA; https://www.ebi.ac.uk/ena) under the accession number PRJEB45271. The landmark data used for geometric morphometric analyses are provided in Supplementary Data 1 and 2.

Supplemental material is available at *GENETICS* online.

Acknowledgments

We are grateful to Gabi Eberhardt and Tobias Loschko for their assistance with the mutagenesis screens and Jürgen Berger for taking the SEM image of the *P. pacificus* mouth. We thank Dr Michael Werner, Dr Christian Rödelsperger, Dr Neel Prabh, Dr Adrian Streit, Tess Renahan, and Metta Riebesell for discussion. SS performed the EMS screen, did the mapping experiments, obtained and analyzed the RNA sequencing data, measured worm body sizes, and created Figs. 2 and 4. TT obtained all landmark data, performed the complete geometric morphometric and clustering analyses, and created Figs. 1, 3, and 5. HW generated and injected the plasmid constructs for transgenesis and CRISPR-RNP-complexes for mutagenesis. EJR identified the cells in the transcriptional reporter lines. RJS conceptualized and supervised the project. SS, TT, and RJS wrote the manuscript.

Funding

This study was funded by the Max Planck Society. SS was supported by the China Scholarship Council. TT was supported by the IMPRS "From Molecules to Organisms."

Conflicts of interest statement

The authors declare no conflict of interest.

Literature cited

Adams D, Collyer M, Kaliontzopoulou A, Baken E. Geomorph: Software for Geometric Morphometric Analyses. R Package Version 4.0.1; 2021. https://cran.r-project.org/package=geomorph.

Adams DC, Otárola-Castillo E., geomorph: an R package for the collection and analysis of geometric morphometric shape data. Methods Ecol Evol. 2013;4(4):393–399.

Albertson DG, Thomson JN. The pharynx of *Caenorhabditis elegans*. Philos Trans R Soc Lond B Biol Sci. 1976;275(938):299–325.

Anderson MJ. A new method for non-parametric multivariate analysis of variance. Austral Ecol. 2001;26(1):32–46.

Anderson MJ. Permutational multivariate analysis of variance (PERMANOVA). In: Balakrishnan N, editor. New Jersey: Wiley & Sons. Wiley Statsref: Statistics Reference Online; 2014. p. 1–15. https://onlinelibrary.wiley.com/doi/full/10.1002/9781118445112.stat07841.

Anderson RC. Nematode Parasites of Vertebrates. Oxford, UK: CABI Publishing; 2000.

Baldwin JG, Giblin-Davis RM, Eddleman CD, Williams DS, Vida JT, Thomas WK. The buccal capsule of *Aduncospiculum halicti* (Nematoda: Diplogastridae): an ultrastructural and molecular phylogenetic study. Can J Zool. 1997;75(3):407–423.

Bento G, Ogawa A, Sommer RJ. Co-option of the endocrine signaling module dafachronic acid-DAF-12 in nematode evolution. Nature 2010;466(7305):494–497.

Brenner S. The genetics of *Caenorhabditis elegans*. Genetics 1974;77(1):71–94.

Bui LT, Ragsdale EJ. Multiple plasticity regulators reveal targets specifying an induced predatory form in nematodes. Mol. Biol. Evol. 2019;36(11):2387–2399.

Burr AHJ, Baldwin JG. The nematode stoma: homology of cell architecture with improved understanding by confocal microscopy of labeled boundaries. J Morphol. 2016;277(9):1168–1186.

Casasa S, Biddle JF, Koutsovoulos GD, Ragsdale EJ. Polyphenism of a novel trait integrated rapidly evolving genes into ancestrally plastic networks. Mol Biol Evol. 2021;38(2):331–343.

Cohen JD, Sundaram MV. *C. elegans* apical extracellular matrix shape epithelia. J Dev Biol. 2020;8(4):23.

Collyer ML, Sekora DJ, Adams DC. A method for analysis of phenotypic change for phenotypes described by high-dimensional data. Heredity (Edinb) 2015;115(4):357–365.

Collyer ML, Adams DC. RRPP: an R package for fitting linear models to high-dimensional data using residual randomization. Methods Ecol Evol. 2018;9(7):1772–1779.

Collyer ML, Adams DC. RRPP: Linear Model Evaluation with Randomized Residuals in a Permutation Procedure. R Package Version 1.1.0; 2021. https://cran.r-project.org/package=RRPP.

De Ley P, Van de Velde MC, Mounport D, Baujard P, Coomans A. Ultrastructure of the stoma in Cephalobidae, Panagrolaimidae and Rhabditidae, with a proposal for a revised stoma terminology in Rhabditida (Nematoda). Nematologica 1995;41(1–4):153–182.

Dieterich C, Clifton SW, Schuster LN, Chinwalla A, Delehaunty K, Dinkelacker I, Fulton L, Fulton R, Godfrey J, Minx P, *et al.* The *Pristionchus pacificus* genome provides a unique perspective on nematode lifestyle and parasitism. Nat Genet. 2008;40(10):1193–1198.

Fraley C, Raftery A, Scrucca L, Murphy TB, Fop M. mclust: Normal Mixture Modeling for Model-Based Clustering, Classification, and Density Estimation. R Package Version 5.4.7.; 2021. https://cran.r-project.org/web/packages/mclust.

Fürst von Lieven A, Sudhaus W. Comparative and functional morphology of the buccal cavity of Diplogastridae (Nematoda) and a

first outline of the phylogeny of this taxon. J Zool. Syst. 2000; 38(1):37–63.
Hall DH, Altun ZF. C. elegans Atlas. Cold Spring Harbor: Cold Spring Harbor Laboratory Press; 2008.
Han Z, Lo W-S, Lightfoot JW, Witte H, Sun S, Sommer RJ. Improving trangenesis efficiency and CRISPR-associated tools through codon optimization and native intron addition in nematodes. Genetics 2020;216(4):947–956.
Harry CJ, Messar SM, Ragsdale EJ. Comparative reconstruction of the predatory feeding structure of the polyphonic nematode Pristionchus pacificus. bioRxiv 2021. https://doi.org/10.1101/2021.10.15.464383.
Kenning C, Kipping I, Sommer RJ. Mutations with altered gross-morphology in the nematode *Pristionchus pacificus*. Genesis 2004; 40(3):176–183.
Lee DL. The Biology of Nematodes. London: Taylor & Francis; 2002.
Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 2009;25(14): 1754–1760.
Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R; 1000 Genome Project Data Processing Subgroup. The sequence alignment/map format and SAMtools. Bioinformatics 2009;25(16):2078–2079.
Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550.
Malakhov VV. Nematodes. Washington: Smithsonian Institution Press; 1994.
Meeuse MW, Hauser YP, Morales Moya LJ, Hendriks G-J, Eglinger J, Bogaarts G, Tsiairis C, Großhans H. Developmental function and state transitions of a gene expression oscillator in *Caenorhabditis elegans*. Mol. Syst. Biol. 2020;16:1–21.
Mitteroecker P, Gunz P, Bernhard M, Schaefer K, Bookstein FL. Comparison of cranial ontogenetic trajectories among great apes and humans. J Hum Evol. 2004;46(6):679–698.
Mitteroecker P, Gunz P, Windhager S, Schaefer K. A brief review of shape, form, and allometry in geometric morphometrics, with applications to human facial morphology. HYSTRIX. 2013;24: 59–66.
Müller MR. Genetic aspects of the development of the mouth form dimorphism in *Pristionchus pacificus*. Diploma thesis, University of Tübingen, Tübingen; 2010.
Nakayama K-I, Ishita Y, Chihara T, Okumura M. Screening for CRISPR/Cas9-induced mutations using a co-injection marker in the nematode *Pristionchus pacificus*. Dev Genes Evol. 2020;230(3): 257–264.
Page AP, Stepek G, Winter AD, Pertab D. Enzymology of the nematode cuticle: a potential drug target? Int J Parasitol Drugs Drug Resist. 2014;4(2):133–141.
Phani V, Shivakumara TN, Davies KG, Rao U. Knockdown of a mucin-like gene in *Meloidogyne incognita* (Nematoda) decreases attachment of endospores of *Pasteuria penetrans* to the infective juveniles and reduces nematode fecundity. Mol Plant Pathol. 2018;19(11):2370–2383.
Quach KT, Chalasani SH. Intraguild predation between *Pristionchus pacificus* and *Caenorhabditis elegans*: a complex interaction with the potential for aggressive behavior. J Neurogenetics. 2020. https://doi.org/10.1080/01677063.2020.1833004.
Rae R, Witte H, Rödelsperger C, Sommer RJ. The importance of being regular: *Caenorhabditis elegans* and *Pristionchus pacificus* defecation mutants are hypersusceptible to bacterial pathogens. Int J Parasitol. 2012;42(8):747–753.
Ragsdale EJ, Baldwin JG. Resolving phylogenetic incongruence to articulate homology and phenotypic evolution: a case study from Nematoda. Proc Biol Sci. 2010;277(1686):1299–1307.
Ragsdale EJ, Müller MR, Rödelsperger C, Sommer RJ. A developmental switch coupled to the evolution of plasticity acts through a sulfatase. Cell 2013;155(4):922–933.
R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing; European Environment Agency, 2021.
Rödelsperger C, Meyer JM, Prabh N, Lanz C, Bemm F, Sommer RJ. Single-molecule sequencing reveals the chromosome-scale genomic architecture of the nematode model organism *Pristionchus pacificus*. Cell Rep. 2017;21(3):834–844.
Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, *et al.* Fiji: an open-source platform for biological-image analysis. Nat Methods. 2012;9(7):676–682.
Schlager S. Morpho and Rvcg–shape analysis in R: R-packages for geometric morphometrics, shape analysis and surface manipulations. In: G Zheng, S Li G Szekely, editors. Statistical Shape and Deformation Analysis. Academic Press; 2017. p. 217–256.
Schlager S, Jefferis G, Ian D. Morpho: Calculations and Visualisations Related to Geometric Morphometrics. R Package Version 2.9; 2021. https://cran.r-project.org/web/packages/Morpho.
Sieriebriennikov B, Prabh N, Dardiry M, Witte H, Röseler W, Kieninger MR, Rödelsperger C, Sommer RJ. A developmental switch regulating phenotypic plasticity is a part of a conserved multi-gene locus. Cell Rep. 2018;23(10):2835–2843.
Sieriebriennikov B, Sun S, Lightfoot JW, Witte H, Moreno E, Rödelsperger C, Sommer RJ. Conserved hormone-receptors controlling a novel plastic trait target fast-evolving genes expressed in a single cell. PLoS Genet. 2020;16(4):e1008687.
Simmer F, Moorman C, van der Linden AM, Kuijk E, van den Berghe PVE, Kamath RS, Fraser AG, Ahringer J, Plasterk RHA. Genome-wide RNAi of *C. elegans* using the hypersensitive rrf-3 strain reveals novel gene functions. PLoS Biol. 2003;1(1):e12.
Sommer RJ. Phenotypic plasticity: from theory and genetics to current and future challenges. Genetics 2020;215(1):1–13.
Sommer RJ, Carta LK, Kim S-Y, Sternberg PW. Morphological, genetic and molecular description of *Pristionchus pacificus* sp. n. (Nematoda, Diplogastridae). Fundam. Appl. Nematol. 1996;19: 511–521.
Sun S, Rödelsperger C, Sommer RJ. Single worm transcriptomics identifies a developmental core network of oscillating genes with deep conservation across nematodes. Genome Res. 2021;31(9): 1590–1601.
Susoy V, Ragsdale EJ, Kanzaki N, Sommer RJ. Rapid diversification associated with a macroevolutionary pulse of developmental plasticity. eLIFE 2015;4:e05463.
Theska T, Sieriebriennikov B, Wighard SS, Werner M, Sommer RJ. Geometric morphometrics of microscopic animals as exemplified by model nematodes. Nat Protoc. 2020;15(8):2611–2644.
van den Hoogen J, Geisen S, Routh D, Ferris H, Traunspurger W, Wardle DA, de Goede RGM, Adams BJ, Ahmad W, Andriuzzi WS, *et al.* A global database of sol nematodes abundance and functional group composition. Nature 2019;572(7768):194–198.
Wagner CE, Wheeler KM, Ribbeck K. Mucins and their role in shaping the function of mucus barriers. Annu Rev Cell Dev Biol. 2018;34: 189–215.
Werlang C, Carcarmo-Oyarce G, Ribbeck K. Engineering mucus to study and influence the microbiome. Nat Rev Mater. 2019;4(2): 134–145.

Werner MS, Sieriebriennikov B, Loschko T, Namdeo S, Lenuzzi M, Dardiry M, Renahan T, Sharma DR, Sommer RJ. Environmental influence on *Pristionchus pacificus* mouth-form through different culture methods. Sci Rep. 2017;7(1):7207.

Witten J, Ribbeck K. The particle in the spider's web: transport through biological hydrogels. Nanoscale 2017;9(24):8080–8095.

Wright KA. Functional organization of the nematodés head. In:NA Croll, editor. The Organization of Nematodes. London (UK): Academic Press; 1976. p. 71–106.

Wright KA, Thomson JN. The buccal capsule of *Caenorhabditis elegans* (Nematoda: Rhabditidae): an ultrastructural study. Can J Zool. 1981;59(10):1952–1961.

Communicating editor: B. Grant

Supplementary Material

The oscillating Mucin-type protein DPY-6 has a conserved role in nematode mouth and cuticle formation

Shuai Sun, Tobias Theska, Hanh Witte, Erik J. Ragsdale & Ralf J. Sommer

Content

Detailed methods for morphological analyses:

In order to quantify phenotypic differences in the mouths of wild-type and mutant worms, we used our recently published protocol for shape analysis in microscopic animals (Theska et al. 2020). We generally follow the exact descriptions provided in this protocol, with only a few minor modifications. In the following paragraphs we provide a detailed account on how we handled our data, which settings we used to analyze morphological differences in this data, and which aspects we added to the previously published pipeline.

Replicability of our landmark data

First, we assessed the replicability of the landmark configurations we defined for both species (see Table S4) by randomly sampling 10 adult hermaphrodites of both *P. pacificus* [Eu] and *C. elegans*, encrypting their names in the image stacks, and obtaining landmark data from each of these specimens in three technical replicates. Our landmark annotator took 15 min breaks in between labeling individual worms and each worm was labeled only once per day. The three technical replicates of each landmark configuration were acquired on three subsequent days (i.e., after 24h breaks). We then calculated the measurement error $(\%ME)^2$ for each of the two species of interest, following the procedure described by Claude (Claude 2008). Additionally, we perform a GPA on these 'replicate' data sets and visualize the amount of shape variation among replicates in a PCA plot (Figure S5). We found that our landmark configurations are highly similar (i.e., the replicates of each specimen are very close to each other in the shape space, while the main shape variation results from natural shape differences among worms) and we were able to confirm a small %ME for both *P. pacificus* (16.3%), and *C. elegans* (14.4%). This is further supported by multivariate statistical analysis: for both species the PERMANOVA-RRPP (see below) approach revealed that the effect that the factor 'specimen' has on the phenotypic response variable 'mouth shape' is much stronger than the effect that the factor 'replicate' has (*P. pacificus*: *shape ~ specimen*: $Z = 9.2$ and $P = 0.000999$, *shape ~ replicate*: $Z = -2.4$ and $P = 0.993$; *C. elegans*: *shape ~ specimen*: $Z = 8.8$ and $P = 0.000999$, *shape ~ replicate*: $Z = -3.0$ and $P = 1.0$; see Tables S8-11). Thus, we conclude that our landmark data is highly replicable.

Additional information on the geometric morphometric analysis of shapes

Initially, we generated landmark data for 200 animals (40 specimens per strain) and created two separate data sets: (I) a *P. pacificus* data set comprising the wild-type strain (PS312), a frameshift mutant of *Ppa-dpy-6 (tu1645)*, and the 'rescue' strain RS4003 (*tu1696*); and (II) a *C. elegans* data set comprising the wild-type N2 and the *Cel-dpy-6* mutant CB14. Analyses were performed in R (ver. 4.1.1) (R Core Team 2021) on a MacBook Pro (13-inch, 2020, macOS Catalina ver. 10.15.6) equipped with a 2 GHz Quad-Core Intel Core i5 processor, 16GB RAM (3733 MHz LPDDR4X), and an Intel Iris Plus graphics card (1536 MB). General Procrustes analysis (GPA) was performed with the *gpagen* function of the GEOMORPH package (ver. 4.0.1.) (Adams and Otarola-Castill 2013; Adams *et al.* 2021). Procrustes coordinates were projected into a linear tangent space to obtain Euclidean distances for statistical analysis. We performed GPA on each of the two landmark data sets and checked for outlier specimens using the *plotOutliers* function of GEOMORPH. In the *C. elegans* data set, we identified and subsequently removed five outliers among the wild-type animals, and in the *P. pacificus* data set we identified and removed two outliers among the *Ppa-dpy-6* mutants. We re-ran the GPA on these 'outlier-corrected' landmark data sets (*C. elegans*: n=75; *P. pacificus*: n=118, Supplementary Data 2,3) and used the resulting shape data for all downstream analyses. Differences in mouth shape were visualized by plotting the first two principal components (PCs) of variation in a shape space. Additionally, wireframe representations of the nematode mouth were generated for the positive and negative extreme shapes along PC1 to visualize the differences which separate wild-type and mutant mouths in the shape space (Theska *et al.* 2020).

Multivariate statistical analysis of phenotypic data using PERMANOVA with RRPP

Phenotypic differences (i.e., mouth shape, form, and size) were assessed utilizing a distance-based permutational MANOVA (PERMANOVA) (Anderson *et al.* 2014; 2021) with a randomized residual permutation procedure (RRPP) (Collyer *et al.* 2015; 2018; 2021). This test assesses the terms of a statistical model using the Euclidean distances among specimens as a source of information, rather than explained covariance matrices among variables (Anderson *et al.* 2014;2021). RRPP is an approach that uses a resampling experiment to randomize the row vectors of a matrix containing residuals from a reduced model to calculate pseudorandom values for estimation of effects from a full model (Collyer *et al.* 2015; 2018;

2021). For simplicity, we only reported relative effect sizes (*Z*-scores) and *P*-values in the main text, but additional statistics are provided in the Supplementary Tables below. The *Z*-scores for all PERMANOVA effects were calculated as the number of standard deviations of observed *F*-values from empirically generated sample distributions of random values found via RRPP (i.e., a relative effect size of $Z = 2.0$ has a signal that is two times larger than the standard deviation). The obtained *P*-values represent the probability of finding a random value larger than the observed value.

Before statistical analysis, we considered any PERMANOVA effect as 'incompatible with the null hypothesis' if (I) the relative effect size was at least twice as large as the standard deviation (i.e., *Z*-score ≥ 2.0) and (II) the *P*-value was below a type I error rate of $\alpha = 0.05$ (i.e., 5%). We assessed the compatibility of the data with a null hypothesis through RRPP in 10,000 iterations using the *procD.lm* function of GEOMORPH under an arbitrarily set seed of '12345'. Generally, an association (i.e., $Z \geq 2.0$ and $P < 0.05$) between a phenotypic response variable 'Y' (here: mouth shape, mouth (centroid) size, or mouth form) and the independent variable 'X' (here: strain or, for allometry, centroid size – see below) suggests that our data is incompatible with the null hypothesis of 'no difference in Y between the wild-type and the mutant strains' and indicates compatibility with the alternative hypothesis that 'there is a mutant phenotype in Y'. The absence of an association between the response variable of interest and the independent variable 'strain' (i.e., $Z < 2.0$ and $P \geq 0.05$) indicates compatibility of our data with the null hypothesis or a lack of statistical power to detect phenotypic differences in 'Y' between wild-type and mutant animals. In our figures, incompatibility of our phenotypic data with a corresponding null hypothesis is indicated with an asterisk '*' and compatibility with a period '.'. Pairwise comparisons of LS means among groups were performed with the *pairwise* function of the RRPP package (ver. 1.1.0). The full statistics (including degrees of freedom, Sum of Squares, Mean Squares, coefficients of determination, F statistics, relative effect sizes, and *P*-values) for all PERMANOVA-RRPP assessments of mouth form, shape, and size are reported in the table supplements 12-17. Full summary statistics of the corresponding pairwise comparisons are reported in the table supplements 18-23.

Additional information on geometric morphometric analysis of mouth form and size

In addition to the analyses of mouth shape and mouth allometry, we also accessed differences in mouth form among specimens through a 'size-shape' PCA (*sensu* Mitteroecker *et al.*). We simply appended a size variable (here, the log-transformed centroid size) to our shape data matrices, in order to obtain form data sets (form is defined as shape plus size, see Klingenberg 2016). We subsequently performed a PCA on the form data sets to visualize morphological differences in a 'size-shape space'. Additionally, we assessed differences in mouth size using the centroid size of the original shape data set as source of information. The centroid size is the square root of the sum of squared distances of each landmark to the corresponding centroid in the landmark configuration, and thus provides information on the overall size of the nematode mouth (Theska et al. 2020).

Additional information on group structure analysis by clustering

We used model-based clustering to reveal the number and structure of separate groups with similar morphology in our shape and form data sets. This approach is opposite to the PERMANOVA-RRPP in the sense that it does not define groups (e.g., strains) *a prior* to than subsequently test for differences among these predefined groups. The clustering analysis was performed in R (ver. 4.1.1.) using the MCLUST (ver. 5.4.7) package (Scrucca *et al.* 2016; Fraley *et al.* 2021). All procedures were carried out exactly following the previous protocol[1]. To avoid overparameterization of our models, we only used 'meaningful' PCs of each shape data set as input for the clustering approaches. 'Meaningful' PCs were determined with the *getMeaningfulPCs* function of the MORPHO package (ver. 2.9) (Schlager *et al.* 2017; 2021). This function calculates a threshold value that is based on a log-likelihood ratio and dependent on the sample size (e.g., an estimated threshold of 2.0 means that a PC can be considered 'meaningful' if it describes at least two times as much variation as the next PC) (Schlager *et al.* 2021; Bookstein 2014). We calculated the meaningful PCs separately for the shape PCA and size-shape PCAs we performed on both data sets.

Analyses of static allometric trajectories

Allometry, the relationship of biological shape with size, is a major factor driving morphological diversification (Klingenberg 2016; Esquerre *et al.* 2017). Thus, it is not surprising that allometry is frequently found in shape data sets, often underlying the variation

described by PC1 of the PCA performed on the shape data (Mitteroecker *et al.* 2004; Adams and Nistri 2010). Since *dumpy*-mutants are known to exhibit clear differences in adult size relative to wildtype worms, we wondered whether potential alterations of mouth morphology in these mutants can be explained by size as well. Here, we assessed the relationship of the adult mouth shape and its log-transformed centroid size (i.e., static allometry in adult worms) for each of the strains under investigation in form of a homogeneity of slopes (HOS) test based on the same distance-based PERMANOVA approach we used for shape analysis. We followed the exact descriptions to perform this test which are given in the "GEOMORPH version 3.1.0 assistance" vignette provided by M. Collyer (version from 01. April 2021, see 'How to perform allometry analyses' paragraph in https://cran.r-project.org/web/packages/geomorph/vignettes/geomorph.assistance.html). In short, we created linear model fits (using *procD.lm* of GEOMORPH) for a simple null model of allometry (shape ~ size) and two complex allometry model fits: (I) a common allometry model (shape ~ size + strain) and (II) a group-specific allometry model (shape ~ size * strain). Then, we compared the simple null model of allometry to the common allometry model using the *anova* function of the R STATS package. A large *Z*-score and a small *P*-value indicate a compatibility of our shape data with the common allometry model, suggesting that its underlying linear model are more appropriate to fit the data. Thus, in such a case, we favored the common allometry model over the simple null model and we subsequently compared the common allometry model (as the new null model) to the group-specific allometry model, using the same ANOVA approach. Again, we interpreted a large *Z*-score together with a small *P*-value to indicate a higher compatibility of our shape data with the group-specific allometry model (which assumes allometric slopes to be different among the strains), than with the common allometry model (which suggests that different strains have similar allometric growth patterns). Thus, low compatibility of our shape data with the group-specific allometry model (i.e., small *Z*-scores along with large *P*-values) indicates that allometric trajectories are similar across strains or that there is a lack of statistical power to detect group-specific allometric trends.

Our allometry model comparisons (i.e., the HOS test) revealed that for both data sets a common allometry model is the most appropriate (table supplements 24,25 and 28,29) to fit the shape data. Thus, within both species, the slopes of the static allometric trajectories of each strain are similar. Therefore, we chose the common allometry model for each of our two

shape data sets and assessed the compatibility of our data with the respective allometry hypothesis through the aforementioned PERMANOVA-RRPP approach with 10,000 iterations using the *procD.lm* function of GEOMORPH under an arbitrarily set seed of '12345'. An association of shape and size ($Z \geq 2.0$ and $P < 0.05$) was interpreted as incompatible with the null hypothesis of 'isometric growth' and indicative of allometry (i.e., that the mouth shape changes with size), while the absence of an association ($Z < 2.0$ and $P \geq 0.05$) was regarded to be indicative of isometry or a lack of statistical power to detect allometry. The full statistical results of the allometry analysis for both species are provided in table supplements 26 and 30. Given that the HOS test revealed similar allometric slopes, we performed pairwise comparisons of the allometric data for our nematode strains by focusing on the LS means of these groups, rather than on slope features (e.g., slopes length, angle, or gradient). The summary statistics of these pairwise comparisons are reported in the table supplements 27 and 31.

References

Theska, T., Sieriebriennikov, B., Wighard, S. S., Werner, M.S., and R. J. Sommer, 2020 Geometric morphometrics of microscopic animals as exemplified by model nematodes. *Nat. Protoc.* **15**:2611–2644.

Yezerinac, S., Lougheed, S. C, and P. Handford P.1992.Measurement error and morphometric studies: statistical power and observer experience. *Syst. Biol.* **41**:471–482.

Claude, J., 2008 *Morphometrics with R*. (Springer).

R Core Team, 2021 *R: A language and environment for statistical computing* (R Foundation for Statistical Computing). Ver 4.1.1.

Adams, D. C., and E. Otárola-Castillo, 2013 geomorph: an R package for the collection and analysis of geometric morphometric shape data. *Methods Ecol. Evol.* **4**:393–399.

Adams, D., Collyer, M., Kaliontzopoulou,. A, and E. Baken, 2021 Geomorph: Software for geometric morphometric analyses. *R package version* 4.0.1. https://cran.r-project.org/package=geomorph

Anderson, M. J., 2021 A new method for non-parametric multivariate analysis of variance. *Austral Ecol.* **26**:32–46.

Anderson, M. J., 2014 Permutational multivariate analysis of variance (PERMANOVA). in *Wiley Statsref: Statistics Reference Online* (eds. Balakrishnan, N. *et al.*) 1–15. https://onlinelibrary.wiley.com/doi/full/10.1002/9781118445112.stat07841.

Collyer, M. L., Sekora, D. J., and D. C. Adams, 2015 A method for analysis of phenotypic change for phenotypes described by high-dimensional data. *Heredity* 115:357–365.

Collyer, M. L., and D. C. Adams, 2018 RRPP: An R package for fitting linear models to high-dimensional data using residual randomization. *Methods Ecol. Evol.* **9**:1772–1779.

Collyer, M., and D. C. Adams, 2021.RRPP: linear model evaluation with randomized residuals in a permutation procedure. *R package version* 1.1.0. https://cran.r-project.org/package=RRPP.

Mitteroecker, P., Gunz, P., Bernhard, M., Schaefer, K., and F. L. Bookstein, 2004 Comparison of cranial ontogenetic trajectories among great apes and humans. *J. Hum. Evol.* **46**:679–698.

Klingenberg, C. ,2016 P. Size, shape, and form: concepts of allometry in geometric morphometrics. *Dev. Genes Evol.* **226**:113–137.

Scrucca, L., Fop, M., Murphy, T. B., and A. E. Raftery, 2016 mclust 5: clustering, classification and density estimation using Gaussian finite mixture models. *R J.* **8**:289.

Fraley, C., Raftery, A., Scrucca, L., Murphy, T. B., and M. Fop, 2021 mclust: Normal mixture modeling for model-based clustering, classification, and density estimation. *R package version* 5.4.7. https://cran.r-project.org/web/packages/mclust.

Schlager, S. ,2017 Morpho and Rvcg–Shape Analysis in R: R-Packages for geometric morphometrics, shape analysis and surface manipulations. in *Statistical Shape and Deformation Analysis* (eds. Zheng, G., Li, S. & Szekely, G.) 217–256 (Academic Press).

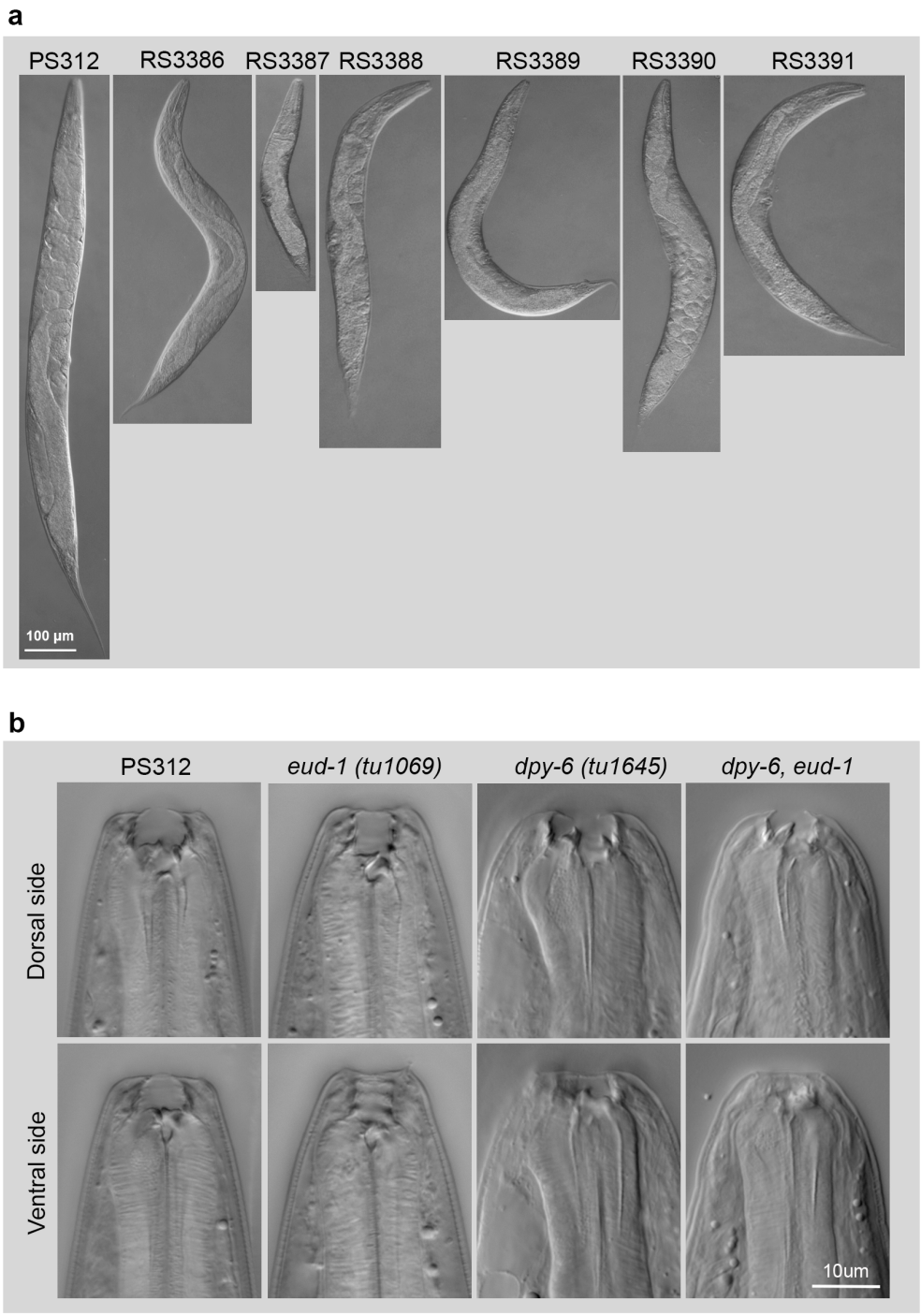
Schlager, S., Jefferis, G., and D. Ian, 2021 Morpho: Calculations and Visualisations Related to Geometric Morphometrics. *R package version* 2.9. https://cran.r-project.org/web/packages/Morpho.

Bookstein, F. L., 2014 *Measuring and reasoning: Numerical inference in the sciences*. (Cambridge University Press).

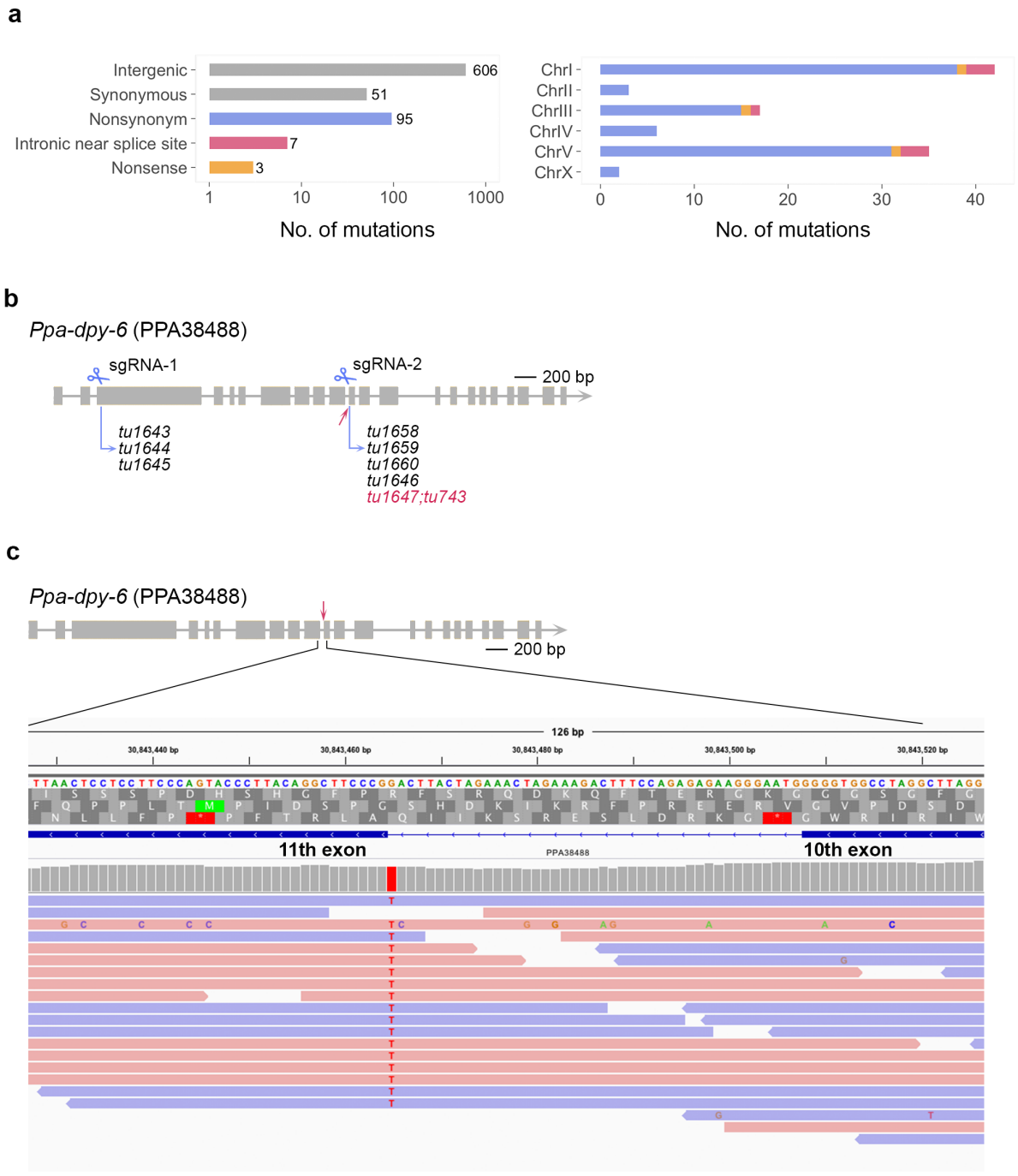
Esquerré, D., Sherratt, E., and J. S. Keogh, 2017 Evolution of extreme ontogenetic allometric diversity and heterochrony in pythons, a clade of giant and dwarf snakes. *Evolution* **71:**2829–2844.

Adams, D. C, and A. Nistri, 2010 Ontogenetic convergence and evolution of foot morphology in European cave salamanders (Family: Plethodontidae). *BMC Evol. Biol.* **10**:1–10.

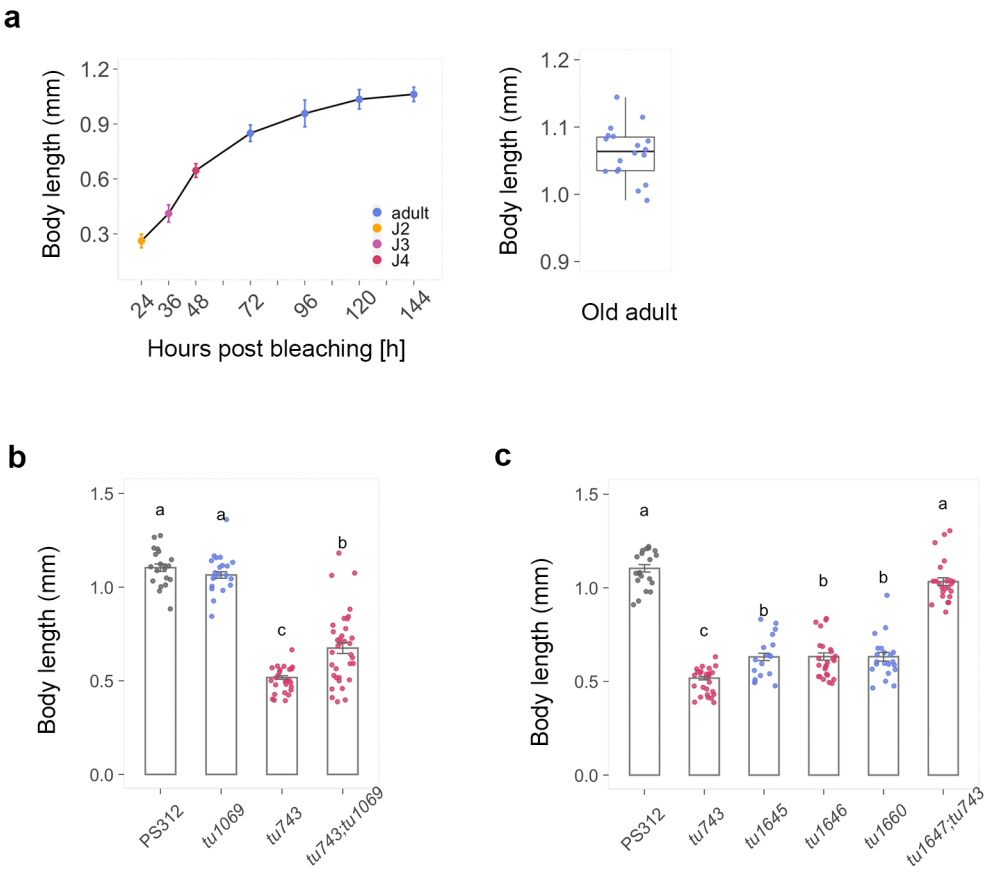
Supplementary Figures



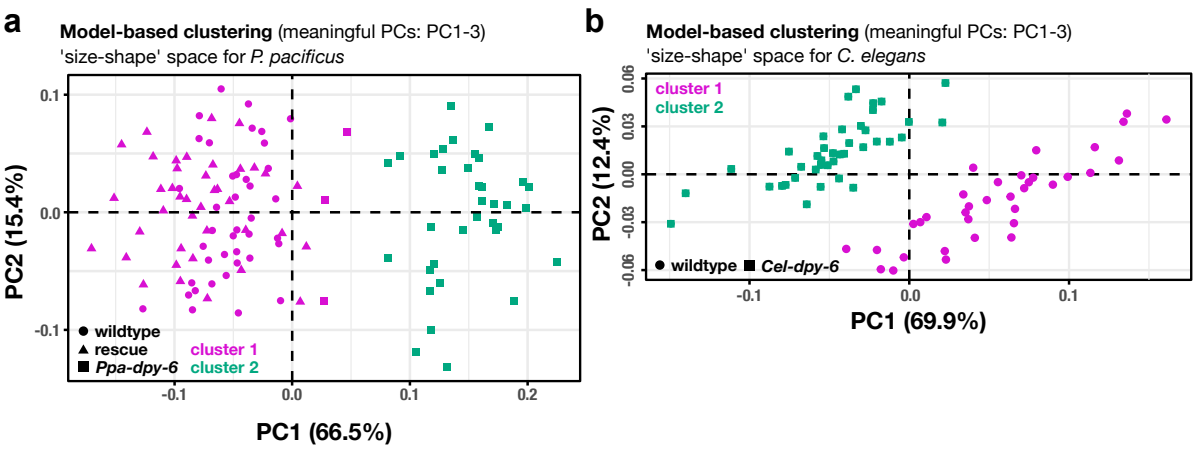
Supplementary Figure 1. Additional images. **a,** Body size of wild-type (PS312) and the six mutant strains (designated as RS3886-RS3891). **b,** Mouth structure of wild-type eurystomatous (Eu) morph, stenostomatous (St) morph of the *Ppa-eud-1* (*tu1069*) mutant, Eu morph with reduced cheilostom of the *Ppa-dpy-6(tu1645)* mutant (strain RS3887), and the St morph of the *Ppa-dpy-6; Ppa-eud-1* double mutant with reduced cheilostom in two focal planes.



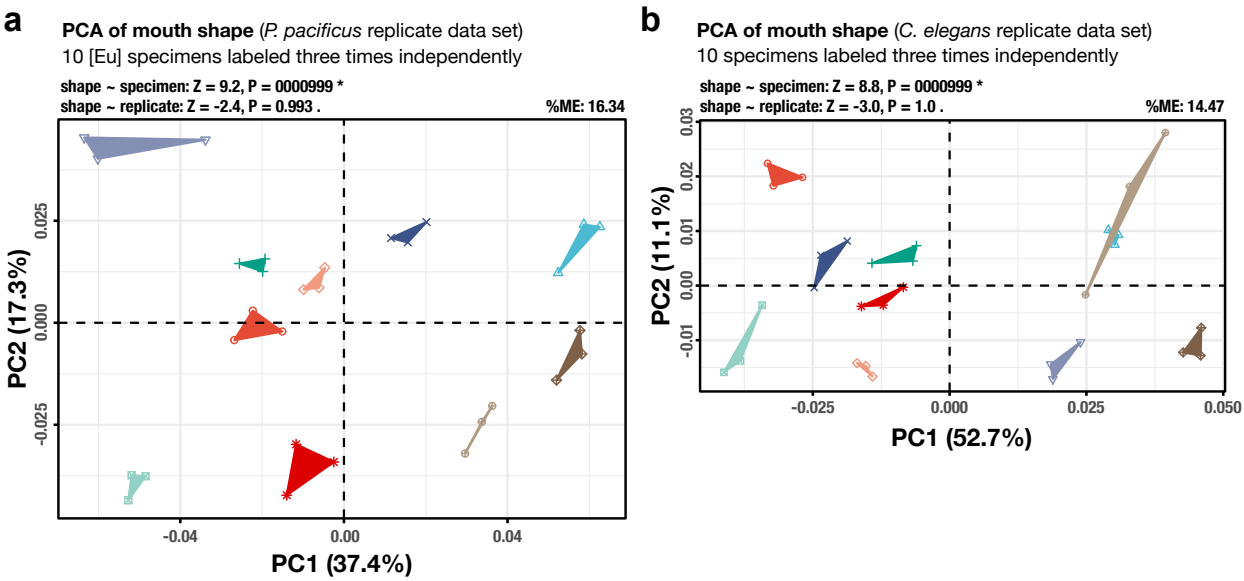
Supplementary Figure 2. Mapping of *tu743* and CRISPR/Cas9 mutagenesis on *Ppa-dpy-6*. **a,** The left bar plot shows the distribution of single nucleotide polymorphisms in the *tu743* mutant relative to PS312 wild-type with five classes, including intergenic, synonymous, nonsynonym, intronic near splice site, and nonsense mutations. The right plot displays the distribution of the three classes, which result in changes in the amino acid sequences among six chromosomes. **b,** Schematic gene model of *Ppa-dpy-6* and the design of two different sgRNAs and CRISPR mutants obtained from these sgRNA. **c,** Screenshot of the IGV profiles of the reads of the RS3887 t*u743* mutant strain at the *Ppa-dpy-6* locus indicating the T > C substitution in intron 10 at the splice site of *Ppa-dpy-6*.



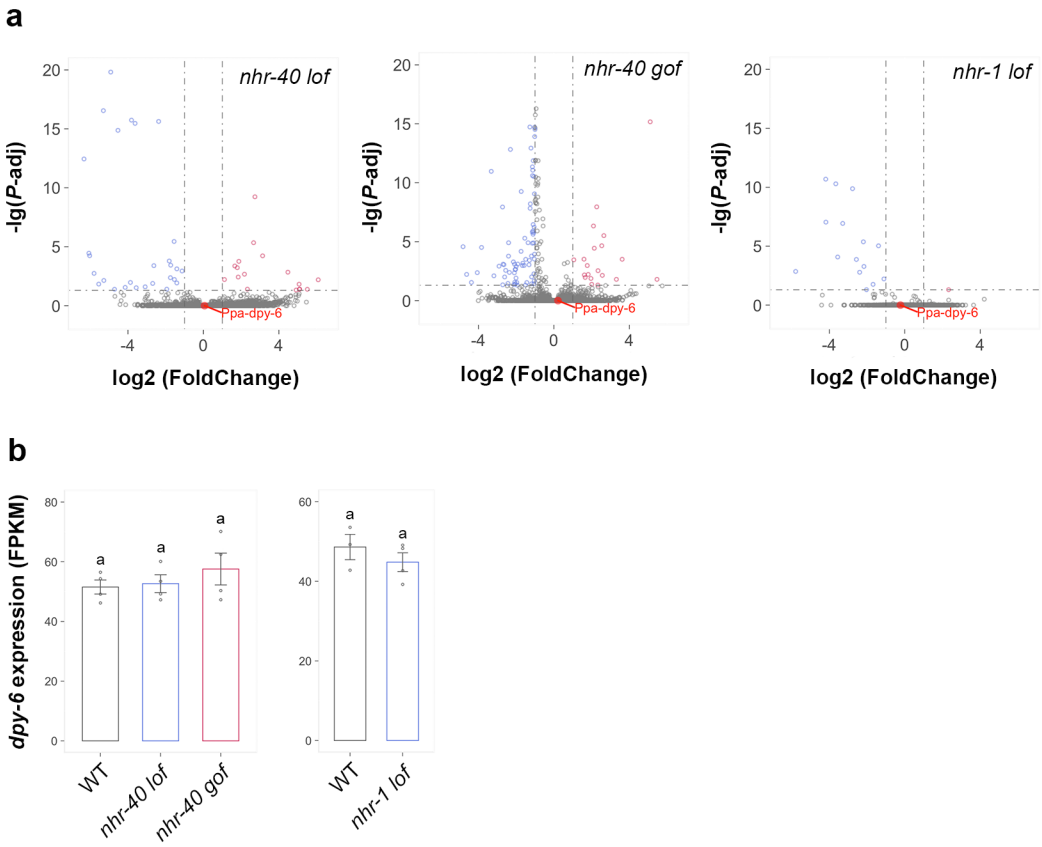
Supplementary Figure 3. Body length of *Ppa-dpy-6* associated mutants. **a,** The left line plot presents the differences in the length of the whole body of wild-type strain (PS312) at seven distinct developmental time points, containing post-bleaching 24 h (J2), 36 h (J3), 48 h (J4), 72 h (one-day old adult), 96 h (two-day old adult), 120 h (three-day old adult), and 144 h (four-day old adult). Worms were cultured on OP50 at 20 °C, more than 20 replicates for each time point. The right box plot displays the distribution of the body length of old adults (> five-day old adult), n > 20. **b,** Bar plot shows the length of worm body between wild-type (PS312), *Ppa-eud-1* mutant (*tu1069*), *Ppa-dpy-6* (*tu743*), and the double mutant (*tu743; tu1069*). **c,** Bar plot displays length of worm body of wild-type (PS312), and the five *Ppa-dpy-6* mutant strains. Turkeys multiple comparison were tested by the Turkeys HSD test, letters upon each volume indicate significant difference between the means, p-adjusted value < 0.05, n > 20.



Supplementary Figure 4. Model-based clustering performed on mouth **form** data sets. **a,** clusters identified in the *P. pacificus* form data set. **b,** clusters identified in the *C. elegans* form data set. Only 'meaningful' principal components of form variation were used as input variables.



Supplementary Figure 5. Replicability of landmark data for the nematode mouths. **a,** PCA of shape variation in the *P. pacificus* replicate data set. **b,** PCA of shape variation in the *C. elegans* replicate data set. Different colors and symbols represent different specimens. Replicates of the same specimens are connected by convex hulls. %ME, percent measurement error according to Yezerinac *et al.*[2] (i.e., the amount of shape variation in the replicate data set which is due to annotation imprecision). *Z*-scores (relative effect sizes) and *P*-values were obtained with PERMANOVA-RRPP (see details above, Supplementary Tables S8-S11).



Supplementary Figure 6. Gene expression of *Ppa-dpy-6* in nuclear-hormone-receptor (NHR) mutant backgrounds. **a,** Vocano plots reveals that *Ppa-dpy-6* is not among the differentially expressed genes (DEG) in *nhr-40* and *nhr-1 loss-of-function* of *nhr-40 gain-of-function* mutants. Downregulated genes in blue, upregulated genes in red, and non-differentially expressed genes in grey. Fold change cutoff is 2, adjusted P-value cut-off is 0.05. **b,** Comparison of gene expression of *Ppa-dpy-6* in *nhr-40 loss-of-function*, *nhr-40 gain-of-function* mutant, and wild-type (left), and in *nhr-1 loss of function* mutant and wild-type (right). All samples used in **a** and **b** are at mouth-form determination period (Sieriebriennikov et al. 2020), *lof = loss-of- function*, *gof = gain-of-function*. Turkeys multiple comparison were tested by the Turkeys HSD test, and the letters upon each volume means significant difference between the means, p-adjusted value < 0.05.

Supplementary Tables

Supplementary Table 1. Detailed information on *Ppa-dpy-6* associated mutants

Strain ID (RS)	Strain ID (tu)	Mutagen	Backgrou d strain	Gene locus (El_Paco_V3	Modifications	Source
RS3886	*tu742*	EMS				This paper
RS3887	*tu743*	EMS				This paper
RS3888	*tu744*	EMS				This paper
RS3889	*tu745*	EMS				This paper
RS3890	*tu746*	EMS				This paper
RS3891	*tu747*	EMS				This paper
RS3899	*tu1643*	CRISPR	PS312	PPA38488	8 bp deletion in the third exon of the *Ppa-dpy-6*	This paper
RS3900	*tu1644*	CRISPR	PS312	PPA38488	28 bp insertion in the third exon of the *Ppa-dpy-6*	This paper
RS3905	*tu1645*	CRISPR	PS312	PPA38488	2 bp deletion in the third exon of the *Ppa-dpy-6*	This paper
RS3910	*tu1646*	CRISPR	PS312	PPA38488	create the C>T substitution of the *Ppa-dpy-6* in PS312	This paper
RS3901	*tu1647;tu743*	CRISPR	*tu743*	PPA38488	rescue *Ppa-dpy-6* in *tu743* strain	This paper
RS3914	*tu1658*	CRISPR	PS312	PPA38488	1 bp insertion in the 11th exon of the *Ppa-dpy-6*	This paper
RS3915	*tu1659*	CRISPR	PS312	PPA38488	37 bp insertion in the 11th exon of the *Ppa-dpy-6*	This paper
RS3916	*tu1660*	CRISPR	PS312	PPA38488	4 bp deletion in the 11th exon of the *Ppa-dpy-6*	This paper
RS3111	*tu1069*	CRISPR	PS312	PPA43535	10 bp deletion in the third exon of the *Ppa-eud-1*	Sieriebrienniko v et al. 2018
RS3924	*RS3887;tu106 9*	EMS, CRISPR		PPA38488; PPA43535	Introduce the CRISPR mutation of *Ppa-eud-1* into the RS3887	This paper

Supplementary Table 2. Mouth-form ratio of the EMS strains *tu742-tu747*

Strain ID	Individual ID	No. of Eu	No. of St	% Eu
PS312	PS312_ind01	20	0	100.0
PS312	PS312_ind02	19	1	95.0
PS312	PS312_ind03	20	0	100.0
PS312	PS312_ind04	20	0	100.0
PS312	PS312_ind05	20	0	100.0
PS312	PS312_ind06	20	0	100.0
PS312	PS312_ind07	19	1	95.0
PS312	PS312_ind08	20	0	100.0
PS312	PS312_ind09	20	0	100.0
PS312	PS312_ind10	20	0	100.0
tu742	*tu742*_ind01	0	20	0.0
tu742	*tu742*_ind02	1	25	3.8
tu742	*tu742*_ind03	2	22	8.3
tu742	*tu742*_ind04	1	19	5.0
tu742	*tu742*_ind05	0	20	0.0
tu742	*tu742*_ind06	1	19	5.0
tu742	*tu742*_ind07	0	20	0.0
tu742	*tu742*_ind08	1	20	4.8
tu742	*tu742*_ind09	1	21	4.5
tu742	*tu742*_ind10	0	20	0.0
tu743	*tu743*_ind01	18	2	90.0
tu743	*tu743*_ind02	19	1	95.0
tu743	*tu743*_ind03	19	1	95.0
tu743	*tu743*_ind04	17	3	85.0
tu743	*tu743*_ind05	23	1	95.8
tu743	*tu743*_ind06	19	1	95.0
tu743	*tu743*_ind07	20	2	90.9
tu743	*tu743*_ind08	21	1	95.5

Strain ID	Individual ID	No. of Eu	No. of St	% Eu
tu745	*tu745*_ind01	0	20	0.0
tu745	*tu745*_ind02	0	21	0.0
tu745	*tu745*_ind03	0	16	0.0
tu745	*tu745*_ind04	0	20	0.0
tu745	*tu745*_ind05	0	20	0.0
tu745	*tu745*_ind06	0	20	0.0
tu745	*tu745*_ind07	0	20	0.0
tu745	*tu745*_ind08	0	20	0.0
tu745	*tu745*_ind09	0	21	0.0
tu745	*tu745*_ind10	0	20	0.0
tu746	*tu746*_ind01	4	27	12.9
tu746	*tu746*_ind02	6	17	26.1
tu746	*tu746*_ind03	5	21	19.2
tu746	*tu746*_ind04	5	19	20.8
tu746	*tu746*_ind05	6	15	28.6
tu746	*tu746*_ind06	2	18	10.0
tu746	*tu746*_ind07	14	7	66.7
tu746	*tu746*_ind08	4	16	20.0
tu746	*tu746*_ind09	8	10	44.4
tu746	*tu746*_ind10	2	18	10.0
tu747	*tu747*_ind01	16	6	72.7
tu747	*tu747*_ind02	22	6	78.6
tu747	*tu747*_ind03	13	21	38.2
tu747	*tu747*_ind04	7	13	35.0
tu747	*tu747*_ind05	13	7	65.0
tu747	*tu747*_ind06	18	4	81.8
tu747	*tu747*_ind07	18	2	90.0
tu747	*tu747*_ind08	11	7	61.1

tu743	*tu743*_ind09	18	2	90.0
tu743	*tu743*_ind10	19	1	95.0
tu744	*tu744*_ind01	9	9	50.0
tu744	*tu744*_ind02	5	12	29.4
tu744	*tu744*_ind03	8	19	29.6
tu744	*tu744*_ind04	1	25	3.8
tu744	*tu744*_ind05	10	11	47.6
tu744	*tu744*_ind06	6	15	28.6
tu744	*tu744*_ind07	2	23	8.0
tu744	*tu744*_ind08	3	20	13.0
tu744	*tu744*_ind09	10	14	41.7
tu744	*tu744*_ind10	4	16	20.0

tu747	*tu747*_ind09	11	12	47.8
tu747	*tu747*_ind10	18	2	90.0

Supplementary Table 3. List of non-synonymous substitutions, nonsense substitutions and the substitutions in intronic near splice site in *tu743* mutant

Chromosome	Position	Ref	Mut	Gene ID (El Paco V3)	Classification	Putative *C. elegans* ortholog
ChrI	1626620	G	A	ppa_stranded_DN18867_c0_g1_i1	intronic_near_splice_sit	spd-1
ChrI	3084346	C	T	PPA38488	intronic_near_splice_sit	dpy-6
ChrI	3824021	C	T	PPA24475	intronic_near_splice_sit	
ChrIII	129246	G	A	PPA17835	intronic_near_splice_sit	let-805
ChrV	5869190	C	T	PPA03558	intronic_near_splice_sit	F56F10.1
ChrV	7549180	G	A	ppa_stranded_DN18023_c0_g1_i1	intronic_near_splice_sit	ser-5
ChrV	2243076	C	T	PPA23050	intronic_near_splice_sit	rad-8
ChrI	3852296	T	A	PPA29201	nonsense	dhc-3
ChrIII	1287249	A	T	PPA02442	nonsense	glr-2
ChrV	1843872	C	T	PPA10610	nonsense	
ChrI	4742880	C	A	PPA42121	nonsynonym	asd-2
ChrI	6745893	C	T	PPA02066	nonsynonym	
ChrI	7070181	G	A	PPA28207	nonsynonym	
ChrI	8637570	A	G	PPA45065	nonsynonym	
ChrI	9040509	T	C	ppa_stranded_DN30480_c0_g1_i3	nonsynonym	sipa-1
ChrI	9226073	G	A	PPA01881	nonsynonym	C29G2.6
ChrI	1024332	G	A	PPA25009	nonsynonym	CELE_F41G3.10
ChrI	1070290	G	A	PPA10389	nonsynonym	CELE_F40F9.10
ChrI	1138090	G	A	PPA37714	nonsynonym	CELE_T10G3.3
ChrI	1227163	A	T	ppa_stranded_DN12886_c0_g1_i1	nonsynonym	CELE_Y73F8A.35
ChrI	1835444	C	T	PPA18108	nonsynonym	
ChrI	1999454	G	A	Contig0-snapTAU.622	nonsynonym	CELE_T07D3.4
ChrI	2154374	G	A	PPA31671	nonsynonym	lbp-1
ChrI	2296623	G	A	PPA07314	nonsynonym	
ChrI	2464863	A	T	ppa_stranded_DN29041_c0_g1_i1	nonsynonym	CELE_Y4C6B.4
ChrI	2577954	G	T	PPA07095	nonsynonym	pgrn-1
ChrI	2617752	G	A	ppa_stranded_DN27869_c0_g1_i1	nonsynonym	CELE_Y53F4B.9
ChrI	2866459	G	C	ppa_stranded_DN25223_c0_g2_i1	nonsynonym	ivns-1
ChrI	2869864	G	A	ppa_stranded_DN27915_c1_g2_i2	nonsynonym	F46H5.3
ChrI	3092251	T	A	PPA21692	nonsynonym	air-2
ChrI	3115168	C	T	PPA38482	nonsynonym	npp-11
ChrI	3121515	C	T	PPA38476	nonsynonym	CELE_F53B6.4
ChrI	3183154	C	T	ppa_stranded_DN24132_c1_g1_i1	nonsynonym	nas-10
ChrI	3188118	C	T	ppa_stranded_DN11105_c0_g1_i1	nonsynonym	C15C6.2
ChrI	3188119	T	A	ppa_stranded_DN11105_c0_g1_i1	nonsynonym	C15C6.2
ChrI	3192976	C	T	ppa_stranded_DN27249_c0_g1_i1	nonsynonym	tsp-11
ChrI	3228929	C	T	PPA23730	nonsynonym	pgp-10
ChrI	3331629	T	A	ppa_stranded_DN21200_c0_g1_i1	nonsynonym	CELE_W02D9.2
ChrI	3360265	G	A	PPA00025	nonsynonym	CELE_Y57G11C.9
ChrI	3636447	C	T	PPA00476	nonsynonym	nhr-28
ChrI	3685136	C	A	PPA38259	nonsynonym	glh-3
ChrI	3696854	T	A	PPA33913	nonsynonym	
ChrI	3763901	C	A	PPA38349	nonsynonym	zipt-15
ChrI	3831847	G	A	ppa_stranded_DN12571_c0_g1_i1	nonsynonym	
ChrI	3836733	C	T	PPA24496	nonsynonym	
ChrI	3851548	C	T	PPA29201	nonsynonym	dhc-3
ChrI	3852305	C	T	PPA29201	nonsynonym	dhc-3
ChrI	3928562	C	T	ppa_stranded_DN29329_c0_g1_i3	nonsynonym	CELE_F22G12.5
ChrII	1239789	T	C	ppa_stranded_DN30137_c0_g2_i3	nonsynonym	CELE_K07C5.3
ChrII	1798426	C	T	PPA14720	nonsynonym	
ChrII	2116389	T	A	PPA36530	nonsynonym	elo-1
ChrIII	316553	G	C	ppa_stranded_DN28612_c0_g1_i1	nonsynonym	CELE_Y75B8A.7
ChrIII	1510368	C	T	PPA04029	nonsynonym	srsx-35
ChrIII	4294991	C	A	PPA02947	nonsynonym	
ChrIII	5073718	C	T	PPA05259	nonsynonym	mrps-10
ChrIII	7864948	C	T	PPA15345	nonsynonym	
ChrIII	1016633	C	A	ppa_stranded_DN27642_c0_g2_i1	nonsynonym	let-716
ChrIII	1016633	C	A	ppa_stranded_DN27642_c0_g2_i2	nonsynonym	let-716
ChrIII	1032331	A	C	PPA03169	nonsynonym	cdh-4

ChrIII	1055834	C	T	PPA40084	nonsynonym	CELE_T28D6.5
ChrIII	1166376	C	T	PPA25490	nonsynonym	C13B9.2
ChrIII	1236173	C	T	ppa_stranded_DN28219_c0_g1_i1	nonsynonym	doxa-1
ChrIII	1292407	C	T	ppa_stranded_DN16501_c0_g1_i4	nonsynonym	sftb-1
ChrIII	1315230	G	A	PPA36678	nonsynonym	
ChrIII	2026517	G	A	ppa_stranded_DN29773_c2_g1_i3	nonsynonym	R05D3.12
ChrIII	2030171	T	C	ppa_stranded_DN25495_c1_g2_i1	nonsynonym	ztf-18
ChrIV	1825046	C	G	PPA14698	nonsynonym	plc-4
ChrIV	1851637	G	T	ppa_stranded_DN30915_c1_g1_i6	nonsynonym	wdfy-3
ChrIV	2366706	A	G	PPA42887	nonsynonym	
ChrIV	2632444	C	A	PPA42712	nonsynonym	
ChrIV	2720828	T	A	PPA13673	nonsynonym	znf-207
ChrIV	3056791	C	T	Iso_D.12473.1	nonsynonym	
ChrV	1859556	C	G	PPA33781	nonsynonym	pept-1
ChrV	3697367	T	C	PPA30982	nonsynonym	
ChrV	5149268	A	C	PPA33807	nonsynonym	
ChrV	6089722	G	A	PPA01242	nonsynonym	ced-7
ChrV	6356139	T	C	PPA27650	nonsynonym	CELE_Y92H12BR.7
ChrV	6437626	G	A	PPA43437	nonsynonym	
ChrV	7865942	C	G	PPA03142	nonsynonym	CELE_W02A2.5
ChrV	8238020	C	G	ppa_stranded_DN18776_c0_g1_i1	nonsynonym	
ChrV	9333361	C	T	PPA26749	nonsynonym	ubc-25
ChrV	9693645	T	G	PPA31261	nonsynonym	
ChrV	1078286	G	A	ppa_stranded_DN31053_c0_g1_i2	nonsynonym	ccep-290
ChrV	1227677	A	T	ppa_stranded_DN18733_c0_g1_i1	nonsynonym	lin-41
ChrV	1228818	G	A	PPA09266	nonsynonym	ppfr-1
ChrV	1252591	G	A	PPA39971	nonsynonym	slc-25a42
ChrV	1276287	G	A	PPA39682	nonsynonym	abtm-1
ChrV	1324471	C	T	PPA08122	nonsynonym	ant-1.1
ChrV	1410204	C	T	PPA35687	nonsynonym	
ChrV	1559072	G	C	PPA12037	nonsynonym	kin-32
ChrV	1560884	G	T	PPA12041	nonsynonym	cutl-11
ChrV	1582504	C	T	PPA11165	nonsynonym	bath-36
ChrV	1631799	C	T	ppa_stranded_DN31518_c0_g1_i2	nonsynonym	atm-1
ChrV	1649669	C	T	PPA11055	nonsynonym	noah-1
ChrV	1727803	C	T	PPA35811	nonsynonym	
ChrV	1831972	C	T	PPA35854	nonsynonym	
ChrV	1886102	T	A	ppa_stranded_DN24465_c0_g3_i2	nonsynonym	lim-7
ChrV	1982665	C	T	ppa_stranded_DN31403_c0_g1_i5	nonsynonym	
ChrV	1994990	G	T	PPA35952	nonsynonym	
ChrV	2054492	T	A	ppa_stranded_DN29783_c1_g1_i3	nonsynonym	let-607
ChrV	2054492	T	A	ppa_stranded_DN29783_c1_g1_i4	nonsynonym	let-607
ChrV	2209488	T	A	ppa_stranded_DN24948_c0_g1_i1	nonsynonym	rad-54.l
ChrV	2327107	C	A	PPA34021	nonsynonym	
ChrX	18682	A	G	PPA09538	nonsynonym	CELE_K06A9.1
ChrX	20645	A	G	PPA09538	nonsynonym	CELE_K06A9.1

Supplementary Table 4. Description on the 15 fixed (i.e., non-sliding) landmarks used to quantify morphological differences in *Pristionchus pacificus* and *Caenorhabditis elegans* strains.

landmark	*Pristionchus pacificus*	*Caenorhabditis elegans*
1	**anterior tip of the dorsal cheilostom**	**anterior tip of the dorsal cheilostom**
2	**posterior tip of the dorsal cheilostom**	**posterior tip of the dorsal cheilostom**
3	**anterior tip of the dorsal gymnostom**	**anterior tip of the dorsal gymnostom**
4	**posterior tip of the dorsal gymnostom**	**posterior tip of the dorsal gymnostom**
5	**posterior tip of the dorsal promesostegostom**	**posterior tip of the dorsal promesostegostom**
6	**posterior base point of the dorsal tooth (metastegostom)**	**posterior base point of the dorsal flap (metastegostom)**
7	**anterior tip of the dorsal tooth (metastegostom)**	**anterior tip of the dorsal flap (metastegostom)**
8	**lateral base point of the right ventrosublateral tooth (metastegostom)**	**lateral base point of the right ventrosublateral flap (metastegostom)**
9	**anterior tip of the right ventrosublateral tooth (metastegostom)**	**anterior tip of the right ventrosublateral flap (metastegostom)**
10	**ventral base point of the right ventrosublateral tooth (metastegostom)**	**ventral base point of the right ventrosublateral flap (metastegostom)**
11	**posterior tip of the ventral promesostegostom**	**posterior tip of the ventral promesostegostom**
12	**posterior tip of the ventral gymnostom**	**posterior tip of the ventral gymnostom**
13	**anterior tip of the ventral gymnostom**	**anterior tip of the ventral gymnostom**
14	**posterior tip of the ventral cheilostom**	**posterior tip of the ventral cheilostom**
15	**anterior tip of the ventral cheilostom**	**anterior tip of the ventral cheilostom**

Supplementary Table 5. CRISPR/Cas9 sgRNA target sequences and primer information used to generate all *Ppa-dpy-6* (PPA38488) mutants including specific targeted modifications

Strain	Modification	Type of primer	Sequences
PS312	*Ppa-dpy-6* knock out from the third exon	CRISPR sgRNA	ATCGTTGTCGTCCCTTCAGA
		Forward Primer	CATTCCAGATTTCACCATCGTCAAGAGGATAAAG
		Reverse Primer	GGCATCTCTCTGTGCCTTCAACTCATCCAAT
		Sequencing Primer	CATTCCAGATTTCACCATCGTCAAGAGGATAAAG
PS312	*Ppa-dpy-6* knock out from the 10th exon	CRISPR sgRNA	GGGTCATGGGAATGTCCGAA
		Forward Primer	TCAGACTGAGTGGCCATCGGCA
		Reverse Primer	GATGAATAAGTCACTGACGGGTCC
		Sequencing Primer	TCAGACTGAGTGGCCATCGGCA
PS312	Create the substitution of C > T in the intronic splice site of *Ppa-dpy-6* gene	CRISPR sgRNA	GGGTCATGGGAATGTCCGAA
		Forward Primer	TCAGACTGAGTGGCCATCGGCA
		Reverse Primer	GATGAATAAGTCACTGACGGGTCC
		Sequencing Primer	TCAGACTGAGTGGCCATCGGCA
		Repair Templates	CTTTCTTCATCCTCTTCAGCTTGCTTGTCCATGTCGAATTGAGGAGGAAGGGT CATGGGAATGTCCGAAGGGCATGAATGATCTTTGATCTTTCTGAAAGGTCTCT CTTCCCTTACCCCCAC
tu743	Repaire the substitution of T > C in the intronic splice site of Ppa-dpy-6 gene	CRISPR sgRNA	GGGTCATGGGAATGTCCGAA
		Forward Primer	TCAGACTGAGTGGCCATCGGCA
		Reverse Primer	GATGAATAAGTCACTGACGGGTCC
		Sequencing Primer	TCAGACTGAGTGGCCATCGGCA
		Repair Templates	CTTTCTTCATCCTCTTCAGCTTGCTTGTCCATGTCGAATTGAGGAGGAAGGGT CATGGGAATGTCCGAAGGGCCTGAATGATCTTTGATCTTTCTGAAAGGTCTCT CTTCCCTTACCCCCAC

Supplementary Table 6. Description of transgenic constructs

Strain ID (RS)	Strain ID (tu)	Gene ID	Backgroud strain	Promoter	CDS	3'UTR
RS3892	*tuEx340*	*Ppa-dpy-6, PPA38488*	PS312	2,189 bp upstream sequences of the first ATG of *PPA38488*	GFP sequence (Han et al. 2020)	*rpl-23*
RS3893	*tuEx341*	*Ppa-dpy-6, PPA38488*	PS312	2,189 bp upstream sequences of the first ATG of *PPA38488*	GFP sequence (Han et al. 2020)	*rpl-23*
RS3894	*tuEx342*	*Cel-dpy-6, F16F9.2*	N2	2,275 bp upstream sequences of the first ATG of *F16F9.2*	GFP sequence	*rpl-23*
RS3895	*tuEx343*	*Cel-dpy-6, F16F9.2*	N2	2,275 bp upstream sequences of the first ATG of *F16F9.2*	GFP sequence	*rpl-23*
RS3896	*tuEx344*	*Cel-dpy-6, F16F9.2*	N2	2,275 bp upstream sequences of the first ATG of *F16F9.2*	GFP sequence	*rpl-23*

Supplementary Table 7. FPKM values used in the analysis of differential expression gene

Comparison the gene expression of *Ppa-dpy-6* in between *Ppa-eud-1* mutant and WT				
Strain name	Strain description	Replicate_ID	Developmental stage	FPKM
tu1069	*Ppa-eud-1* loss of function strain	Replicate_1	20 hours after hatching	39.7
tu1069	*Ppa-eud-1* loss of function strain	Replicate_2	21 hours after hatching	54.3
tu1069	*Ppa-eud-1* loss of function strain	Replicate_3	22 hours after hatching	82.9
PS312	Wild-type strain	Replicate_1	23 hours after hatching	0.0
PS312	Wild-type strain	Replicate_2	24 hours after hatching	0.0
PS312	Wild-type strain	Replicate_3	25 hours after hatching	0.2
Comparison the gene expression of *Ppa-dpy-6* in between *Ppa-nhr-40* mutants and WT				
Strain name	Strain description	Replicate_ID	Developmental stage	FPKM
tu505	*Ppa-nhr-40* gain of function strain	Replicate_1	Mouth form determination	70.2
tu505	*Ppa-nhr-40* gain of function strain	Replicate_2	Mouth form determination	50.3
iub6	*Ppa-nhr-40* gain of function strain	Replicate_3	Mouth form determination	47.3
iub6	*Ppa-nhr-40* gain of function strain	Replicate_4	Mouth form determination	62.5
tu1418	*Ppa-nhr-40* loss of function strain	Replicate_1	Mouth form determination	49.2
tu1418	*Ppa-nhr-40* loss of function strain	Replicate_2	Mouth form determination	47.2
tu1423	*Ppa-nhr-40* loss of function strain	Replicate_3	Mouth form determination	53.5
tu1423	*Ppa-nhr-40* loss of function strain	Replicate_4	Mouth form determination	60.7
PS312	Wild-type strain	Replicate_1	Mouth form determination	54.3
PS312	Wild-type strain	Replicate_2	Mouth form determination	49.1
PS312	Wild-type strain	Replicate_3	Mouth form determination	46.2
PS312	Wild-type strain	Replicate_4	Mouth form determination	56.5
Comparison the gene expression of *Ppa-dpy-6* in between *Ppa-nhr-1* mutant and WT				
Strain name	Strain description	Replicate_ID	Developmental stage	FPKM
tu1163	*Ppa-nhr-1* loss of function strain	Replicate_1	Mouth form determination	42.7
tu1163	*Ppa-nhr-1* loss of function strain	Replicate_2	Mouth form determination	39.2
tu1164	*Ppa-nhr-1* loss of function strain	Replicate_3	Mouth form determination	49.1
tu1164	*Ppa-nhr-1* loss of function strain	Replicate_4	Mouth form determination	48.3
PS312	Wild-type strain	Replicate_1	Mouth form determination	53.8
PS312	Wild-type strain	Replicate_2	Mouth form determination	49.2
PS312	Wild-type strain	Replicate_3	Mouth form determination	42.8

Mouth form determination: Combined stage of 24 h and 48 h post-bleaching

Supplementary Table 8. PERMANOVA–RRPP summary statistics for **replicate factor of *P. pacificus* mouth shape replicability** (n=10, replicates=3). Linear model used: *shape ~ replicate*. Iterations used: 1,001. Seed used: 12345. Effect type: F. Z, relative effect sizes; P, probability of finding a random value larger than the observed F value (P-value).

source	df	SS	MS	R^2	F	Z	P (>F)
replicate	2	0.002887	0.0014435	0.02475	0.3427	-2.4493	0.993
residuals	27	0.113739	0.0042126	0.97525			
total	29	0.116626					

Supplementary Table 9. PERMANOVA–RRPP summary statistics for **specimen factor of *P. pacificus* mouth shape replicability** (n=10, replicates=3). Linear model used: *shape ~ specimen*. Iterations used: 1,001. Seed used: 12345. Effect type: F. Z, relative effect sizes; P, probability of finding a random value larger than the observed F value (P-value).

source	df	SS	MS	R^2	F	Z	P (>F)
specimen	9	0.102674	0.0114083	0.88037	16.354	9.2699	0.000999
residuals	20	0.013952	0.0006976	0.11963			
total	29	0.116626					

Supplementary Table 10. PERMANOVA–RRPP summary statistics for **replicate factor of *C. elegans* mouth shape replicability** (n=10, replicates=3). Linear model used: *shape ~ replicate*. Iterations used: 1,001. Seed used: 12345. Effect type: F. Z, relative effect sizes; P, probability of finding a random value larger than the observed F value (P-value).

source	df	SS	MS	R^2	F	Z	P (>F)
replicate	2	0.000646	0.00032314	0.01452	0.1989	-3.0027	1.0
residuals	27	0.043859	0.00162443	0.98548			
total	29	0.044506					

Supplementary Table 11. PERMANOVA–RRPP summary statistics for **specimen factor of *C. elegans* mouth shape replicability** (n=10, replicates=3). Linear model used: *shape ~ specimen*. Iterations used: 1,001. Seed used: 12345. Effect type: F. Z, relative effect sizes; P, probability of finding a random value larger than the observed F value (P-value).

source	df	SS	MS	R^2	F	Z	P (>F)
specimen	9	0.039784	0.0044205	0.89392	18.726	8.8816	0.000999
residuals	20	0.004721	0.0002361	0.10608			
total	29	0.044506					

Supplementary Table 12. PERMANOVA–RRPP summary statistics for ***P. pacificus* mouth shapes** (n=118). Linear model used: *shape ~ strain*. Iterations used: 10,000. Seed used: 12345. Effect type: F. Z, relative effect sizes; P, probability of finding a random value larger than the observed F value (P-value).

source	df	SS	MS	R^2	F	Z	P (>F)
strain	2	1.01955	0.50977	0.65646	109.87	5.9632	0.0001
residuals	115	0.53355	0.00464	0.34354			
total	117	1.55310					

Supplementary Table 13. PERMANOVA–RRPP summary statistics for ***C. elegans*** **mouth shapes** (n=75). Linear model used: *shape ~ strain*. Iterations used: 10,000. Seed used: 12345. Effect type: F.

source	df	SS	MS	R^2	F	Z	P (>F)
strain	1	0.059076	0.059076	0.32253	34.754	5.4348	0.0001
residuals	73	0.124089	0.001700	0.67747			
total	74	0.183165					

Supplementary Table 14. PERMANOVA–RRPP summary statistics for ***P. pacificus*** **mouth forms** (n=118). Linear model used: *form ~ strain*. Iterations used: 10,000. Seed used: 12345. Effect type: F.

source	df	SS	MS	R^2	F	Z	P (>F)
strain	2	1.10374	0.55187	0.58068	79.626	6.7103	0.0001
residuals	115	0.79704	0.00693	0.41932			
total	117	1.90078					

Supplementary Table 15. PERMANOVA–RRPP summary statistics for ***C. elegans*** **mouth forms** (n=75). Linear model used: *form ~ strain*. Iterations used: 10,000. Seed used: 12345. Effect type: F.

source	df	SS	MS	R^2	F	Z	P (>F)
strain	1	0.21504	0.215038	0.46015	62.222	4.2637	0.0001
residuals	73	0.25229	0.003456	0.53985			
total	74	0.46732					

Supplementary Table 16. PERMANOVA–RRPP summary statistics for ***P. pacificus*** **mouth sizes** (n=118). Linear model used: *centroid size ~ strain*. Iterations used: 10,000. Seed used: 12345. Effect type: F.

source	df	SS	MS	R^2	F	Z	P (>F)
strain	2	25.122	12.5609	0.23849	18.008	4.5533	0.0001
residuals	115	80.214	0.6975	0.76151			
total	117	105.335					

Supplementary Table 17. PERMANOVA–RRPP summary statistics for ***C. elegans*** **mouth sizes** (n=75). Linear model used: *centroid size ~ strain*. Iterations used: 10,000. Seed used: 12345. Effect type: F.

source	df	SS	MS	R^2	F	Z	P (>F)
strain	1	82.069	82.069	0.54601	87.798	5.6049	0.0001
residuals	73	68.237	0.935	0.45399			
total	74	150.305					

Supplementary Table 18. Summary statistics of pairwise comparisons among ***P. pacificus*** **mouth shapes** (n=118). Pairwise comparisons of LS means were performed with the *pairwise* function of RRPP. Abbreviations: d, distance between the means; UCL, upper confidence limit; Z, relative effect size (Z-score); P, probability of finding a random value larger than the observed distance value (P-value).

pairwise comparisons	d	UCL (95%)	Z	P > d
wildtype vs. rescue	0.03609841	0.04582869	1.217962	0.1322
wildtype vs. *Ppa-dpy-6*	0.18483771	0.04642146	3.593817	0.0001
***Ppa-dpy-6* vs. rescue**	0.20883560	0.04610816	3.642244	0.0001

Supplementary Table 19. Summary statistics of pairwise comparisons between ***C. elegans*** **mouth shapes** (n=75). Abbreviations: d, distance between the means; UCL, upper confidence limit; Z, relative effect size (Z-score); P, probability of finding a random value larger than the observed distance value (P-value).

pairwise comparisons	d	UCL (95%)	Z	P > d
wildtype vs. *Cel-dpy-6*	0.05625639	0.01726215	5.129054	0.0001

Supplementary Table 20. Summary statistics of pairwise comparisons among ***P. pacificus*** **mouth forms** (n=118). Pairwise comparisons of LS means were performed with the *pairwise* function of RRPP. Abbreviations: d, distance between the means; UCL, upper confidence limit; Z, relative effect size (Z-score); P, probability of finding a random value larger than the observed distance value (P-value).

pairwise comparisons	d	UCL (95%)	Z	P > d
wildtype vs. rescue	0.0395137	0.04839635	1.217225	0.1322
wildtype vs. *Ppa-dpy-6*	0.1908064	0.04897315	4.329126	0.0001
***Ppa-dpy-6* vs. rescue**	0.2182532	0.04899521	4.466567	0.0001

Supplementary Table 21. Summary statistics of pairwise comparisons between ***C. elegans*** **mouth forms** (n=75). Pairwise comparisons of LS means were performed with the *pairwise* function of RRPP. Abbreviations: d, distance between the means; UCL, upper confidence limit; Z, relative effect size (Z-score); P, probability of finding a random value larger than the observed distance value (P-value).

pairwise comparisons	d	UCL (95%)	Z	P > d
wildtype vs. *Cel-dpy-6*	0.1073307	0.03117189	3.926253	0.0001

Supplementary Table 22. Summary statistics of pairwise comparisons among ***P. pacificus*** **mouth sizes** (n=118). Pairwise comparisons of LS means were performed with the *pairwise* function of RRPP. Abbreviations: d, distance between the means; UCL, upper confidence limit; Z, relative effect size (Z-score); P, probability of finding a random value larger than the observed distance value (P-value).

pairwise comparisons	d	UCL (95%)	Z	P > d
wildtype vs. rescue	0.2785290	0.4114308	0.9472539	0.1785
wildtype vs. *Ppa-dpy-6*	0.8172373	0.4218938	2.9842151	0.0002
***Ppa-dpy-6* vs. rescue**	1.0957663	0.4193543	3.7446847	0.0001

Supplementary Table 23. Summary statistics of pairwise comparisons between ***C. elegans*** **mouth sizes** (n=75). Pairwise comparisons of LS means were performed with the *pairwise* function of RRPP. Abbreviations: d, distance between the means; UCL, upper confidence limit; Z, relative effect size (Z-score); P, probability of finding a random value larger than the observed distance value (P-value).

pairwise comparisons	d	UCL (95%)	Z	P > d
wildtype vs. *Cel-dpy-6*	2.096794	0.6345145	4.489541	0.0001

Supplementary Table 24. ANOVA summary statistics for a comparison of a simple allometry model (*shape ~ log(size)*) and a common allometry model (*shape ~ log(size) + strain*) for the ***P. pacificus*** data set (n=118). Iterations used: 1,000. Effect type: F. RRPP = T. Size refers to the centroid sizes contained in the shape data. *procD.lm* was used to fit the data, *anova* was used for the model comparison.

model	res.df	df	rSS	SS	MS	R^2	F	Z	P
shape ~ log(size) (Null)	116	1	1.36576			0.00000			
shape ~ log(size) + strain	114	2	0.51853	0.84723	0.42362	0.54551	93.134	5.1362	0.001
total	117		1.55310						

Supplementary Table 25. ANOVA summary statistics for a comparison of a common allometry model (*shape ~ log(size) + strain*) and a group-specific allometry model (*shape ~ log(size) * strain*) for the ***P. pacificus*** data set (n=118). Iterations used: 1,000. Effect type: F. RRPP = T. Size refers to the centroid sizes contained in the shape data. *procD.lm* was used to fit the data, *anova* was used for the model comparison.

model	res. df	df	rSS	SS	MS	R^2	F	Z	P
shape ~ log(size) + strain (Null)	114	1	0.51853			0.000000			
shape ~ log(size) * strain	112	2	0.50132	0.017212	0.0086059	0.011082	1.9227	1.7334	0.044
total	117		1.55310						

Supplementary Table 26. PERMANOVA–RRPP summary statistics for ***P. pacificus*** **mouth allometry** (n=118). Linear model used: *shape ~ log(centroid size) + strain*. Iterations used: 10,000. Seed used: 12345. Effect type: F. c.size = centroid size. *procD.lm* was used to fit the data.

source	df	SS	MS	R^2	F	Z	P (>F)
log(c.size)	1	0.18476	0.18476	0.11896	40.592	3.310	0.0001
strain	2	0.84944	0.42472	0.54693	93.311	5.317	0.0001
residuals	114	0.51889	0.00455	0.33410			
total	117	1.55310					

Supplementary Table 27. Summary statistics of pairwise comparisons among ***P. pacificus*** **mouth allometries** (n=118). Pairwise comparisons of LS means were performed with the *pairwise* function of RRPP. Abbreviations: d, distance between the LS means; UCL, upper confidence limit; Z, relative effect size (Z-score); P, probability of finding a random value larger than the observed distance value (P-value).

pairwise comparisons	d	UCL (95%)	Z	P > d
wildtype vs. rescue	0.03860415	0.04338278	1.450365	0.0828
wildtype vs. *Ppa-dpy-6*	0.19129182	0.04631384	3.556634	0.0001
***Ppa-dpy-6* vs. rescue**	0.21818263	0.04884414	3.625766	0.0001

Supplementary Table 28. ANOVA summary statistics for a comparison of a simple allometry model (shape ~ log(size)) and a common allometry model (shape ~ log(size) + strain) for the ***C. elegans*** data set (n=75). Iterations used: 1,000. Effect type: F. RRPP = T. Size refers to the centroid sizes contained in the shape data. procD.lm was used to fit the data, anova was used for the model comparison.

model	res.df	df	rSS	SS	MS	R^2	F	Z	P
shape ~ log(size) (Null)	73	1	0.143393			0.00000			
shape ~ log(size) + strain	72	1	0.095601	0.04779	0.04779	0.26092	35.993	5.9397	0.001
total	74		0.183165						

Supplementary Table 29. ANOVA summary statistics for a comparison of a common allometry model (*shape ~ log(size) + strain*) and a group-specific allometry model (*shape ~ log(size) * strain*) for the ***C. elegans*** data set (n=75). Iterations used: 1,000. Effect type: F. RRPP = T. Size refers to the centroid sizes contained in the shape data. *procD.lm* was used to fit the data, *anova* was used for the model comparison.

model	res.df	df	rSS	SS	MS	R^2	F	Z	P
shape ~ log(size) + strain (Null)	72	1	0.0956			0.00000			
shape ~ log(size) * strain	71	1	0.0948	0.000765	0.000765	0.00417	0.573	-1.164	0.87
total	74		0.1831						

Supplementary Table 30. PERMANOVA–RRPP summary statistics for ***C. elegans*** **mouth allometry** (n=75). Linear model used: *shape ~ log(centroid size) + strain*. Iterations used: 10,000. Seed used: 12345. Effect type: F. c.size = centroid size. *procD.lm* was used to fit the data.

source	df	SS	MS	R^2	F	Z	P (>F)
log(c.size)	1	0.039762	0.039762	0.21709	30.024	5.3888	0.0001
strain	1	0.048048	0.048048	0.26232	36.280	5.1584	0.0001
residuals	72	0.095355	0.001324	0.52059			
total	74	0.183165					

Supplementary Table 31. Summary statistics of pairwise comparisons among ***C. elegans*** **mouth allometries** (n=75). Pairwise comparisons of LS means were performed with the *pairwise* function of RRPP. Abbreviations: d, distance between the LS means; UCL, upper confidence limit; Z, relative effect size (Z-score); P, probability of finding a random value larger than the observed distance value (P-value).

pairwise comparisons	d	UCL (95%)	Z	P > d
wildtype vs. *Cel-dpy-6*	0.07529769	0.02232566	4.92008	0.0001

nature communications

Article

https://doi.org/10.1038/s41467-023-37734-z

Histone 4 lysine 5/12 acetylation enables developmental plasticity of *Pristionchus* mouth form

Received: 4 May 2022

Accepted: 28 March 2023

Published online: 13 April 2023

Check for updates

Michael S. Werner[1,2], Tobias Loschko[1], Thomas King[2], Shelley Reich[2], Tobias Theska[1], Mirita Franz-Wachtel[3], Boris Macek[3] & Ralf J. Sommer[1]

Development can be altered to match phenotypes with the environment, and the genetic mechanisms that direct such alternative phenotypes are beginning to be elucidated. Yet, the rules that govern environmental sensitivity vs. invariant development, and potential epigenetic memory, remain unknown. Here, we show that plasticity of nematode mouth forms is determined by histone 4 lysine 5 and 12 acetylation (H4K5/12ac). Acetylation in early larval stages provides a permissive chromatin state, which is susceptible to induction during the critical window of environmental sensitivity. As development proceeds deacetylation shuts off switch gene expression to end the critical period. Inhibiting deacetylase enzymes leads to fixation of prior developmental trajectories, demonstrating that histone modifications in juveniles can carry environmental information to adults. Finally, we provide evidence that this regulation was derived from an ancient mechanism of licensing developmental speed. Altogether, our results show that H4K5/12ac enables epigenetic regulation of developmental plasticity that can be stored and erased by acetylation and deacetylation, respectively.

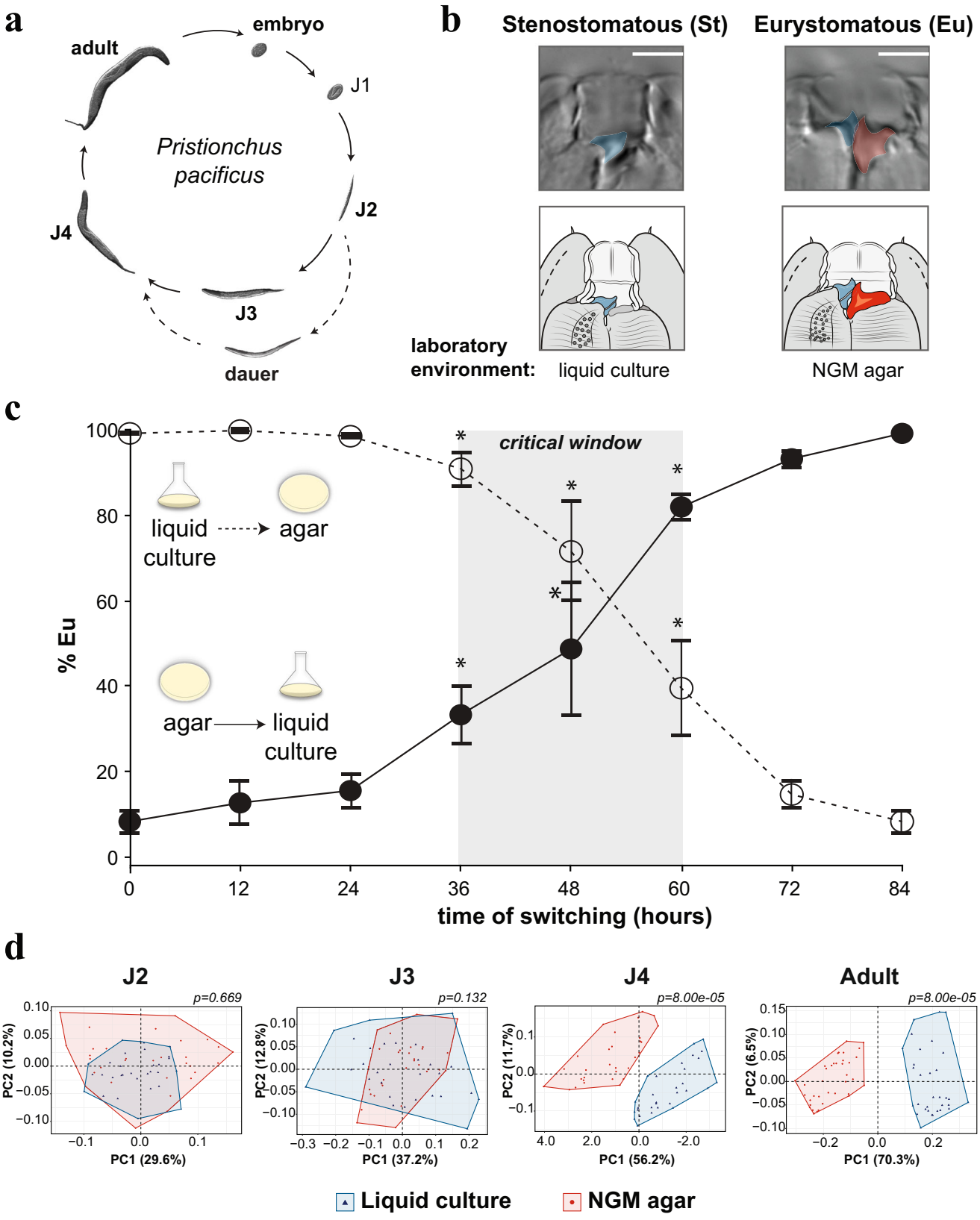
Different environments can elicit distinct phenotypes from a single genotype, referred to as phenotypic plasticity[1,2]. Ecological and theoretical approaches over the last 50 years have formalized the evolutionary implications and significance of plasticity[3–6]. More recently, molecular approaches are homing in on the mechanisms that direct environmental influence[7,8]. In contrast, the mechanisms that provide environmental sensitivity remain unknown. To do so requires an experimentally-tractable system of plasticity that is capable of linking environmental sensitivity, development and gene regulation.

Pristionchus pacificus has a short life cycle and reproduces primarily as self-fertilizing hermaphrodites similar to *Caenorhabditis elegans* (Fig. 1a), allowing the use of forward and reverse genetic tools[9]. Unlike *C. elegans* however, *P. pacificus* exhibits mouth-form plasticity. Specifically, adult worms express either a narrow Stenostomatous (St) morph with a single dorsal 'tooth', or a wide Eurystomatous (Eu) morph with two teeth[10] (Fig. 1b). Morphological plasticity is coupled to behavioral plasticity, as St animals are strict bacterial feeders, while Eu animals can use their opposing teeth to kill other nematodes for food or competitive advantage[11–16]. This binary readout in laboratory model organism has enabled the discovery of a gene regulatory network for each morph[17–20]. In contrast, if and how epigenetic processes are involved in the regulation of plasticity remain poorly understood.

Here, we provide evidence that acetylation of histone 4 lysines 5/12 (H4K5/12ac) enables plasticity to different culture environments. Moreover, pharmacologically preventing deacetylation enforced epigenetic memory of a previous environmental experience; effectively canalizing a developmental trajectory despite changing environments. While several post-translational modifications to histone 3 have been implicated in dynamic gene regulation[21,22], outside of H4K16ac, the function of specific modifications on H4 has been less clear. The role of

[1]Department for Integrative Evolutionary Biology, Max Planck Institute for Biology Tübingen, Tübingen 72076, Germany. [2]School of Biological Sciences, The University of Utah, Salt Lake City, UT, USA. [3]Proteome Center Tübingen, University of Tübingen, Tübingen 72076, Germany.
e-mail: Ralf.Sommer@tuebingen.mpg.de

https://doi.org/10.1038/s41467-023-37734-z



H4K5/12ac identified here also extends beyond gene regulation to an ecologically relevant developmental decision. Finally, we found that acetylation regulates developmental rate across multiple Ecdysozoan species, which may be linked to the establishment of the Eu morph in *P. pacificus*. Thus, our results reveal the molecular determinants of plasticity and their potential evolutionary origin.

Results

Reciprocal transplantation reveals a critical window of environmental influence

In the laboratory, different culture conditions can be harnessed to bias morph development: the wild type *P. pacificus* strain PS312 is predominantly St when grown in liquid S-Medium and Eu when grown on

Fig. 1 | Mouth-form phenotypic plasticity is determined during a critical window in development. a Life cycle of *Pristionchus pacificus*. **b** Stenostomatous (St, left) and Eurystomatous (Eu, right). Scale bar = 5 μm. **c** Switching experiments delineate developmental phases of plasticity. Adult phenotypes are plotted as a function of transferring between environments after synchronization. Statistical significance was assessed by binomial logistic regression on Eu and St counts between a given time-point, and the $t' = 0$- and 84-h phenotype; p-values were adjusted by Bonferroni correction, $^*p < 0.025$ relative to both. Error bars represent S.E.M. for $n = 5$ independent worm populations. Note, intermediate values are intermediate ratios, not intermediate morphs. **d** GMM analysis of mouth-shape differences throughout development. PCA plots show the morphospace that contains shape variation per stage. Effect size (Z) for J2 = 1.403, J3 = 2.172652, J4 = 3.847113, Adult = 5.0949. p-values represent FDR-adjusted pairwise Procrustes Anova. Source data are provided as a Source Data file.

NGM agar plates[23]. We investigated the boundaries of environmental sensitivity by performing reciprocal transplant experiments between liquid and NGM-agar. Transferring between environments at 12 and 24 h after egg synchronization (t'0) led to a complete adoption of the second environment's phenotype, regardless of the direction of the environmental shift (Fig. 1c; $n = 5$, $p < 0.05$). Thus, this period represents a naïve and fully plastic phase of development. In contrast, when animals were switched at ≥36 h they began to retain memory of their previous environment. Specifically, an increasing percentage of animals executed the morph of the first environment (Fig. 1c). Finally, after 60 h, transplanting had little to no effect on mouth-form ratios. These results reveal a critical time window of environmental sensitivity between 36 and 60 h. Prior to that, juveniles are completely plastic while afterwards the decision is irreversible.

Molecular and morphological plasticity underly the critical window

Next, we looked into potential mechanisms that provide plasticity during the critical window and end plasticity after it. First, we investigated the role of 'switch genes', which determine alternative phenotypes depending on their expression above or below a given threshold[24,25]. *eud-1* is a steroid sulfatase that yields 100% Eu animals when constitutively overexpressed and 100% St animals when knocked out[19]. We predicted that transcriptional plasticity of switch genes (i.e., *eud-1*) underlies phenotypic plasticity during the critical window, while invariant expression demarcates deterministic development. To test this prediction, we measured mRNA levels of *eud-1* during normal development and reciprocal transplant experiments. At 12 h, we found strong induction of *eud-1* in NGM-agar (Eu condition) compared to liquid culture (St condition)(Supplementary Fig. 1a, b, Supplementary Data 1). Surprisingly though, modest expression was still observed in liquid culture, indicating some amount of environment-insensitive transcription during the naïve phase. This modest expression is rapidly induced when worms are transferred to NGM-agar at 24 h, or decreased when transferred from NGM-agar to liquid (Supplementary Fig. 1a,b; $p < 0.05$), demonstrating environmental sensitivity. Intriguingly, *eud-1* begins to be repressed after 36 h regardless of the environmental condition. At 60 h *eud-1* mRNA levels are normalized between environments—coinciding with the end of the critical window.

Next, we wondered if plasticity during the critical window is confined to a specific juvenile stage/molt. First, we measured juvenile stages after hypochlorite-synchronization and compared these data to our reciprocal transplant data, which indicated that the critical window is centered in the J3 stage (Supplementary Fig. 1c–g). Second, fitting our reciprocal transplant data to a logistic model revealed inflection points between 48–54 h, which coincides with the J3-J4 molt (Supplementary Fig. 1h, i). Third, quantitative geometric morphometrics (GMM) throughout development revealed significant differences between conditions beginning in the J4 stage (Fig. 1d, $p < 0.05$ and $Z \geq 2.0$; Supplementary Fig. 2), even though mouth dimorphism was previously thought to occur only in adults[26]. Molecular and phenotypic divergence may differ depending on genotype and environmental conditions, and it's unclear if Eu J4s will necessarily become Eu adults after the final molt. Nevertheless, under these standardized conditions with the laboratory strain PS312, a pattern emerged: plasticity correlates with switch-gene transcriptional flexibility, while exit from the critical window correlates with switch-gene repression and morphological differentiation. We note that this pattern provides a plausible mechanism for the establishment of environmentally sensitive critical periods of development.

Morph development is regulated by H4K5/12 acetylation

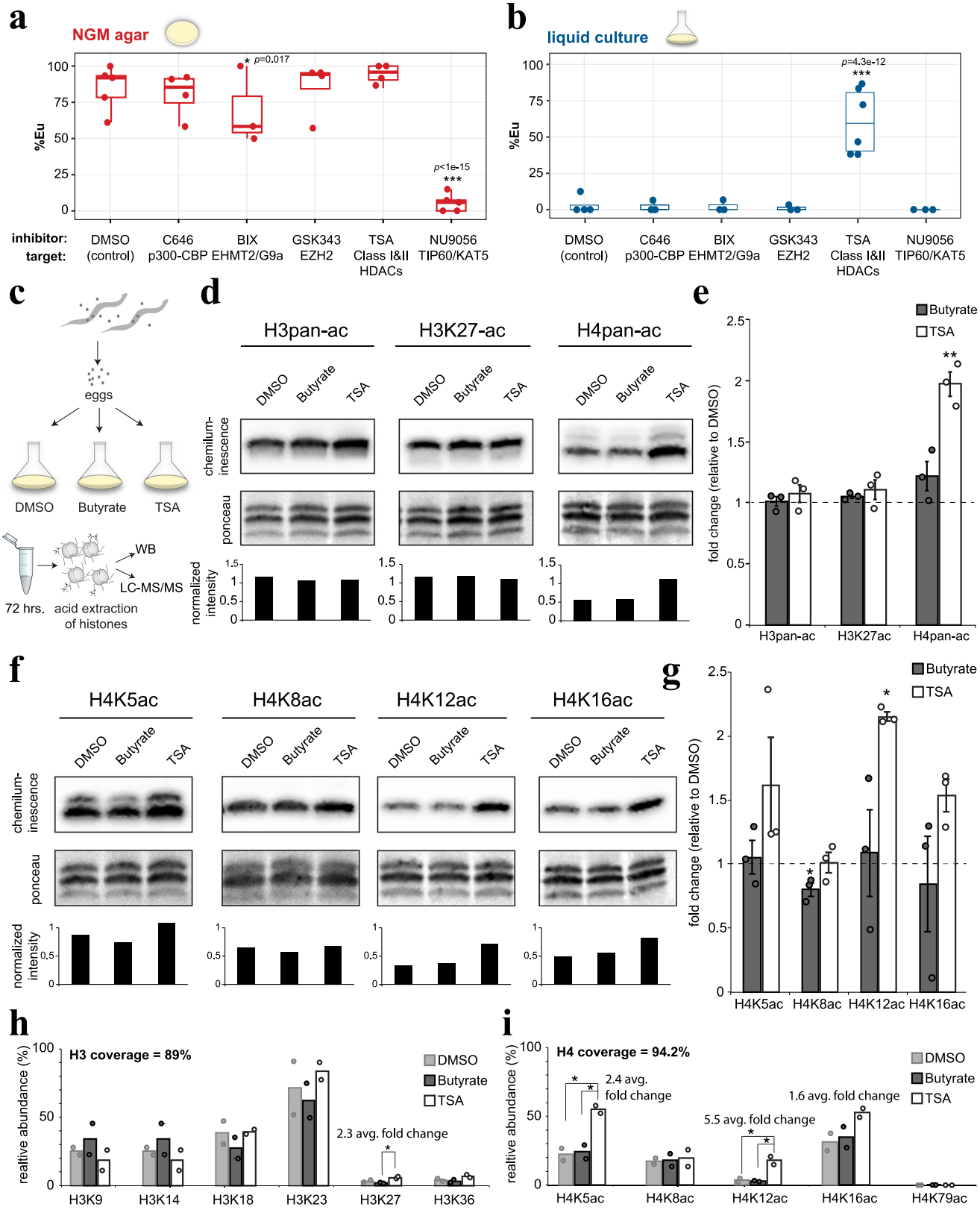
We hypothesized that open and closed chromatin structure—mediated by histone modifications[27–29]—underlies transcriptional plasticity *vs.* repression of the switch gene *eud-1*. First, we queried a panel of inhibitors that target histone-modifying enzymes for a potential effect on mouth form (Fig. 2a, b; see Methods for detailed description of statistics, $n \geq 3$). Interestingly, the Tip-60/KAT5 histone acetyltransferase inhibitor NU9056 converted normally Eu animals on NGM-agar to the St morph ($p = 1e\text{-}15$), while the histone deacetylase (HDAC) inhibitor Trichostatin A (TSA) converted normally St animals in liquid culture to the Eu morph ($p = 4.3e\text{-}12$). Hence, functionally opposite inhibitors yielded correspondingly opposite phenotypes, indicating that histone acetylation/deacetylation has an important role in mouth-form development.

NU9056-treated animals had pleiotropic effects (e.g., reduced fertility, egg laying) consistent with *Ppa-mbd-2* and *Ppa-lys-12* mutants that are egg laying or vulva defective and are St-biased[30]. However, animals that were exposed to TSA appeared wild type except for the effect on mouth form. HDAC inhibitors have also been shown to affect developmental decisions in other organisms, and several are currently in use as chemotherapeutic drugs[28,29,31,32]. Yet, how HDAC inhibition leads to these effects remains poorly understood. Therefore, in the following we focus on TSA and the potential role of histone acetylation/deacetylation in plasticity.

First, we confirmed that TSA also induced the Eu morph in axenic liquid culture, arguing for a direct inhibition of nematode enzymes rather than an indirect effect on their bacterial diet ($p = 1.2e\text{-}6$)(Supplementary Fig. 3a). TSA also increased the proportion of Eu animals in three different *Pristionchus* species (Supplementary Fig. 3b), indicating that acetylation has an evolutionary conserved role in mouth form across the genus. Surprisingly however, although TSA is a pan-deacetylase inhibitor[33], no obvious effect on mouth form was seen after treatment with several other HDAC inhibitors (Supplementary Fig. 3c). Presumably this is due to an unusual degree of specificity between TSA and its target enzyme in *Pristionchus*, and we wondered if that would allow us to investigate the role of discrete acetylated residues in plasticity.

To probe the potential specificity of TSA in *P. pacificus* we performed Western Blots (WB) on acid-extracted histones (Fig. 2c, Supplementary Fig. 4). In addition to DMSO (solvent) we used butyrate as a second negative control because it is an HDAC inhibitor that did not affect mouth form (Supplementary Fig. 3c). H3K27ac is a reproducible marker of active enhancers and promoters[27], and has recently been implicated in behavioral differences between ant castes[28,34]. However, we did not observe an increase in H3 acetylation using a pan-acetyl antibody or with a specific antibody toward H3K27ac ($n = 3$, Fig. 2d, e). In contrast, we observed a 2-fold induction of H4 acetylation (±0.1, $p = 6.3e\text{-}4$ *vs.* DMSO and 8.5e-3 *vs.* butyrate). To verify TSA's effect on H4 and to determine which H4 lysine(s) are hyperacetylated, we repeated WBs with specific H4-acetyl antibodies (Fig. 2f, g). Consistent with the apparent specificity of TSA in *Pristionchus*, we observed only

https://doi.org/10.1038/s41467-023-37734-z



one H4 lysine with significant hyperacetylation relative to both controls: H4K12 (p=5.7e-6, 3.6e-2, respectively).

To further validate this result, we performed Liquid Chromatography Tandem-Mass Spectrometry (LC-MS/MS) ($n = 2$, Fig. 2c, and h,i, Supplementary Data 2). In agreement with our WB data, H4K12ac exhibited the greatest increase in TSA for all queried H3 and H4 lysine residues (avg. fold change = 5.5). The greater sensitivity of LC-MS/MS revealed a smaller yet significant increase in H4K5ac against both DMSO and butyrate (avg. fold change = 2.4). No other modification was statistically significant against both controls (e.g., H3K27ac exhibited an increase relative to butyrate, but not to DMSO). Collectively, both immunostaining and mass spectrometry implicate H4 acetylation, and in particular H4K12ac, as the main effector by which TSA influences mouth-form development.

Fig. 2 | Mouth form is linked to H4K5/12ac. a, **b** Phenotype (%Eu) after exposure to histone-modifying enzyme inhibitors in each environment. Target enzymes indicated below inhibitors. Box-plot minima and maxima represent the 25 and 75% quantile, respectively, middle bar represents 50% quantile (median). Whiskers denote 1.5x the interquartile range. Statistical significance determined by binomial logistic regression on Eu and St counts, $n \geq 3$ biologically independent worm populations. p-values for agar-BIX = 0.017, agar-NU9056 = 1e-15, liquid-TSA = 4.3e-12. **c** Strategy to identify the molecular mechanism of TSA's effect: Bleach-synchronized eggs were seeded into liquid culture with DMSO (negative control for solvent), Butyrate (negative control for HDAC inhibitor that had no effect on phenotype), and TSA. Worm pellets were used for acid-extraction of histones and subject to WB and LC-MS/MS. **d** Representative WBs of *P. pacificus* histones with specified antibodies, total histone staining by ponceau, and resulting relative intensities by quantitative densitometry. **e** Mean fold change vs. DMSO, $n = 3$ biologically independent worm populations. Error bars represent S.E.M. Statistical significance was assessed by a two-tailed student's t-tests between TSA and Butyrate, p-value for H4pan-ac = 0.0085. **f**, **g** Same as **d**, **e** but with H4-specific antibodies, $n = 3$ biologically independent worm populations. p-value for H4K12ac = 0.036. **h** Relative H3 and **i** H4 acetylation of indicated residues compared to total H3 and H4 peptide intensities. Statistical significance was assessed by a 2-sided student's t-test, $n = 2$ biologically independent worm populations. P-values for H3K27ac: TSA vs. DMSO = 0.13, TSA vs. butyrate = 0.033. P-values for H4K5ac: TSA vs. DMSO = 0.025, TSA vs. butyrate = 0.033. P-values for H4K12ac: TSA vs. DMSO = 0.047, TSA vs. butyrate = 0.036.*$p < 0.05$, **$p < 0.01$. Source data are provided as a Source Data file.

Genome-wide patterns of H4 acetylation and a role in switch-gene expression

The role of H4K12ac in development is not well understood, but there are links to aging[35], hormone-dependent cancer[36], and learning;[37] processes that are all connected to dynamic gene regulation. Previous Chromatin Immunoprecipitation and sequencing (ChIP-seq) experiments in human cells showed that H4K12ac is distributed more throughout gene bodies than other acetylated lysines, which primarily demarcate enhancers or promoters[27,38], hinting at a role for H4K12ac in transcription elongation rather than initiation. We investigated whether, and how, TSA effects switch-gene expression in *P. pacificus*. Using RT-qPCR, we found that TSA does increase *eud-1* expression compared to DMSO (Fig. 3a, Supplementary Fig. 3d, $p < 0.05$). However, instead of inducing *eud-1*, TSA appeared to prevent the repression that normally occurs after 36 h. These data suggest that H4K5/12ac does not have a role in environmental induction per se. Instead, hyperacetylation caused by TSA maintained the baseline of transcription beyond the critical window (36–48 h). These findings appear to differentiate H4K5/12ac from other histone acetylation marks that are thought to directly promote transcriptional initiation[39].

To explore this possibility further, and to examine comprehensive patterns of H4 acetylation, we performed ChIP-seq with antibodies targeting H4K5ac, H4K8ac, H4K12ac, and H4K16ac at 48 h in +/- TSA conditions. Consistent with our original biochemical results (Fig. 2), H4K12ac exhibited the greatest induction in TSA (Fig. 3b-e). H4K5ac exhibited the second largest induction, and correlated with H4K12ac abundance (Pearson correlations 0.82 and 0.84 per replicate; Supplementary Fig. 5a). H4K12ac peaks were significantly enriched in regulatory regions compared to genome-wide chromatin state distributions[40] ($p < 0.01$, Fisher's Exact Test; Supplementary Fig. 5b), and Gene Set Enrichment Analysis[41] of genes adjacent to hyperacetylated regions. (>1.5-fold change) revealed a significant association with 'positive regulation of programmed cell death' and an RNAi phenotype of 'cell proliferation increased' (Supplementary Fig. 5c–e; e.g. the putative vacuolar ATPase subunit *vha-10*). Programmed cell death is one of several competing models by which TSA and other HDAC inhibitors have been proposed to underly antitumor effects in humans[42–44]. Our results in *P. pacificus* suggest a deeply conserved role for acetylation in controlling pro-apoptotic gene expression and indicate that nematodes may be a useful model for investigating the chemotherapeutic mechanisms of HDAC inhibition.

An H4 acetylation/deacetylation timer determines the critical window

Interestingly, although H4K12ac peaks were enriched in promoters and enhancers (Supplementary Fig. 5b), TSA led to H4K12 hyperacetylation primarily over gene bodies (Fig. 3d and Supplementary Fig. 5e). This pattern also appeared to be true at the locus encompassing the *eud-1* switch gene, which is part of a four-member 'supergene' on the X chromosome where each gene contributes to mouth form regulation[45]. Our ChIP-seq data revealed that the super-gene locus is bordered by large peaks of H4 acetylation (Fig. 3f, Supplementary Data 3), and that TSA led to an increase in H4K12ac between these peaks, including across *eud-1*, rather than at the peaks themselves (Fig. 3f, g).

Next, we performed ChIP-qPCR at 24 and 48 h to see if we could track deacetylation across *eud-1* during the critical window. Indeed, we found a significant decrease in H4K12ac at 3/3 qPCR primer locations spanning the *eud-1* gene locus (Fig. 3h–j, $p < 0.05$). However, adding TSA partially maintained the early juvenile levels of H4K12ac at 48 h (Fig. 3h–j, Supplementary Fig. 6c); arguing that TSA prevents *eud-1* repression by preventing H4 deacetylation. In contrast, we did not detect an increase in either H3 acetylation or H4K16ac at *eud-1* in the presence of TSA, confirming the specificity of our previous results (Fig. 3k). Interestingly, we also observed a broad distribution of the facultative heterochromatin mark H3K27me3 over the supergene locus flanked by the H4K12ac peaks (Supplementary Fig. 6a). However, in contrast to H4K12ac, we did not observe significant changes in H3K27me3 across development at *eud-1* (Supplementary Fig. 6b). Collectively, these data lead toward a model of early H4K5/12ac deposition in a facultative heterochromatin domain that permits switch-gene induction by the environment. During the J3-J4 transition, an unknown developmental mechanism leads to deacetylation, turning off switch-gene expression and effectively closing the critical window (Supplementary Fig. 7a).

To test this model, we first asked whether TSA requires *eud-1* to induce the Eu morph. Indeed, *eud-1* mutant animals remained 100% St even in the presence of TSA ($n = 3$, $p < 0.05$). Next, we wondered if the sulfatase inhibitor STX-64[46] can be used to inactivate EUD-1, which would allow temporal control of EUD-1 activity. Application of 1 μg/ml of STX-64 caused a 100% St phenotype on NGM-agar, mirroring the *eud-1* phenotype (Supplementary Fig. 7b). With this tool in hand, we grew a *eud-1* constitutive expression line (Ex[*eud-1*])[19] on STX-64 plates until 48 h, and then switched to DMSO plates. In theory, transferring off STX-64 plates should remove inhibition, and thereby test whether EUD-1 activity after 48 h is sufficient to induce the Eu morph. Unswitched control Ex[*eud-1*] worms exhibited a 49% Eu phenotype, yet switched Ex[*eud-1*] worms exhibited a 98% Eu phenotype (Supplementary Fig. 7c; $n = 3$, $p < 0.05$). Thus, EUD-1 activity after the critical window is sufficient to induce the Eu morph—consistent with our model that prolonged *eud-1* expression by hyperacetylation induced the Eu morph in our TSA experiments.

Epigenetic memory can be enforced by HDAC inhibition

The role of histone modifications in plasticity provides the potential for epigenetic gene regulation; a long-held but poorly supported hypothesis to connect plasticity to evolution[5]. To determine whether histone acetylation can provide long-term memory, we assessed if preventing deacetylation during the critical window would 'fix' or 'freeze' an initial developmental trajectory despite shifting to a different environment. In principle, this would provide compelling evidence that histone modifications can carry long-term environmental

https://doi.org/10.1038/s41467-023-37734-z

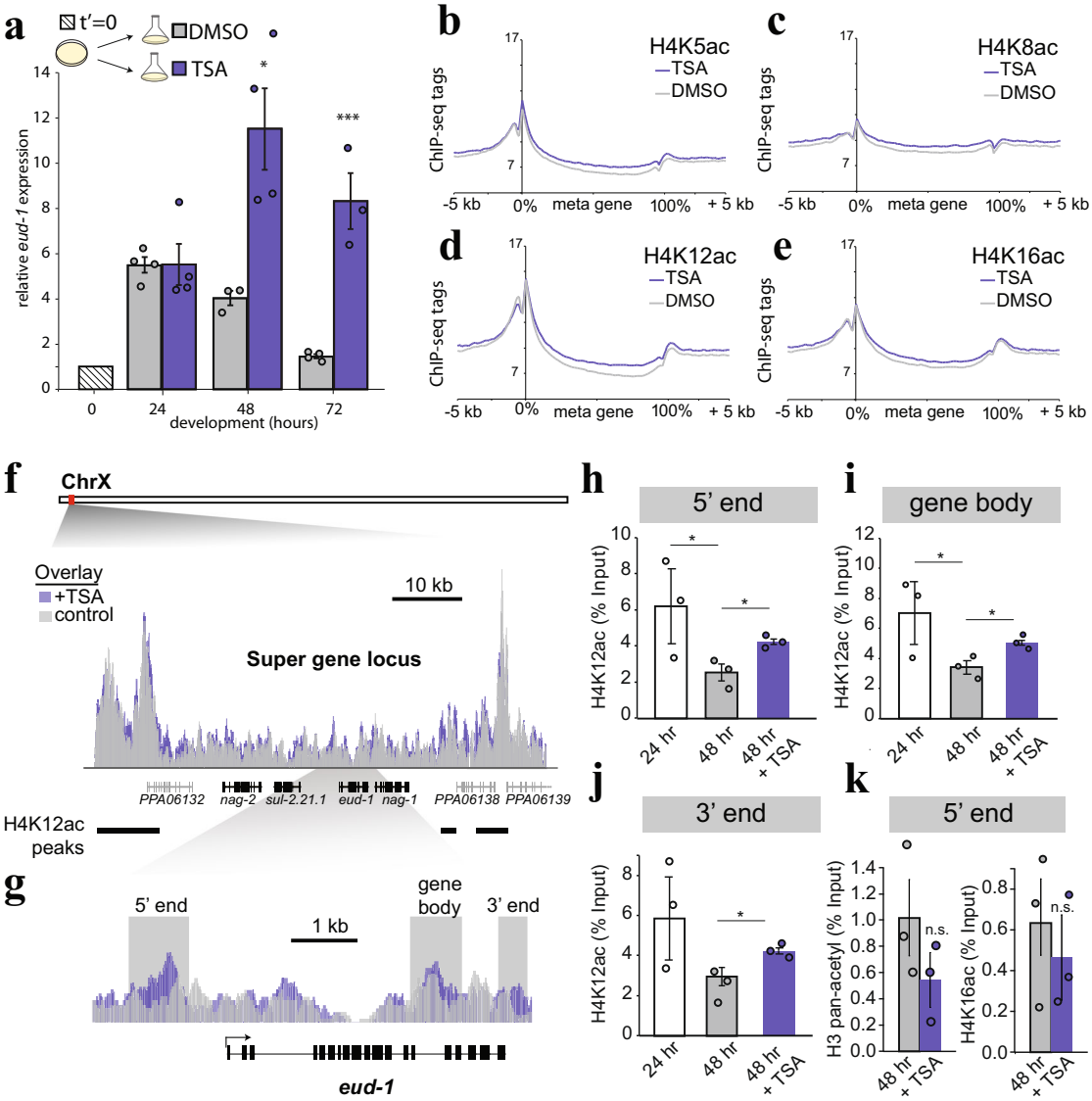


Fig. 3 | H4K12ac acetylation/deacetylation at switch genes determines plasticity and canalization, respectively. a, *eud-1* expression at indicated time points by RT-qPCR with either DMSO or TSA added directly after bleach synchronization. Error bars represent S.E.M. of 4 biologically independent replicates except for DMSO 48 h and TSA 72 h where $n = 3$. Statistical significance was determined by student's 1-tailed *t*-test. *P*-values for 48-h comparisons = 0.0089, and 0.00062 for 72-h comparisons; $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. **b**–**e** ChIP-seq meta-gene profiles of H4-acetylated residues at 48 h in +/- TSA. y-axis represents average input-normalized ChIP-seq tag counts per bp per gene, and the *x*-axis represents 5' to 3' gene coordinates divided into 200 bins +/− 5 kb, $n = 2$. **f** ChIP-seq tag density of H4K12ac over the multi-gene locus (black genes) that control mouth-form plasticity, which are flanked by large peaks. y-axis = 0-72 depth-normalized density for both tracks. Black bands indicate H4K12ac peaks in liquid culture determined by MACS2. **g** Zoom in on *eud-1* and regions targeted by qPCR in **h**–**k**. **h**–**j**, ChIP-qPCR of H4K12ac across *eud-1* at 24 h, 48 h and 48 h when grown in TSA, $n = 3$ biologically independent worm populations. **k** ChIP-qPCR of H3-pan-acetyl and H4K16ac at the 5' end of *eud-1*, $n = 3$ biologically independent worm populations. Statistical significance in '**h**–**k**' was assessed by a 1-tailed student's *t*-test. *P*-values for '**h**': 24-h vs. 48-h = 0.043, 48-h vs. 48-h+TSA = 0.013. *P*-values for '**i**': 24-h vs. 48-h = 0.042, 48-h vs. 48-hr+TSA = 0.016. *P*-values for '**j**': 48 h vs. 48 h + TSA = 0.017. $^{*}p < 0.05$ Source data are provided as a Source Data file.

information to affect future developmental decisions. To test this premise, we combined transplant experiments with TSA treatment. Specifically, we transferred worms between NGM-agar and liquid at 24 h, which normally leads to the St phenotype (Fig. 1c), but this time added TSA at the time of switching. Remarkably, these worms were phenotypically similar to having experienced the agar environment for their entire development (90.8% Eu ±1.2)(Fig. 4). Not only were environmentally shifted TSA-treated worms significantly different from DMSO controls, but they were also significantly different from worms treated with TSA at $t' = 0$ without switching (60.9% Eu ±8.0). These results demonstrate a combined effect of the environment with HDAC inhibition that is consistent with fixing the initial Eu developmental trajectory. We then performed a series of experiments to attempt to refute this interpretation (Fig. 4). First, to rule out that simply adding TSA at a later time point has a greater effect on mouth form than at $t' = 0$, perhaps due to degradation, we added TSA at 24 or 48 h without environmental transfer. These experiments resulted in decreasingly intermediate ratios of Eu animals similar to, and less than adding TSA at $t' = 0$, respectively ($p < 0.05$). We also repeated transplant experiment at 24 h, but waited until 48 h to add TSA. Again, this experiment failed to induce a typical NGM-agar phenotypic ratio (28.5% Eu ±3.8). These results show that there is an additive effect of the environment with TSA, and that this effect depends on adding TSA at the precise time of environmental transfer. We conclude that preventing histone deacetylation after a temporary juvenile experience in NGM-agar fixed the Eu developmental trajectory; 48 h and two molts prior to adult

https://doi.org/10.1038/s41467-023-37734-z

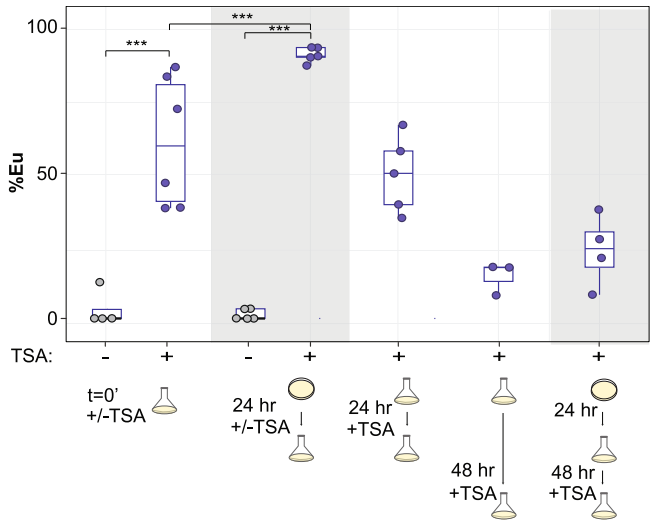


Fig. 4 | H4 acetylation can provide long term memory. Transplant experiments between NGM-agar and liquid culture +/- TSA at different time points. When TSA is not added (-), the solvent DMSO was added. The first two experimental conditions (DMSO and TSA) are from Fig. 2a. Individual data points are shown for all conditions; $n = 4$ independent worm populations x DMSO, 6 x TSA, 5 × 24 h switch +/-TSA, 5 × 24 h TSA, 3 × 48 h TSA, 4 × 24 h switch + 48 h TSA. Box-plot minima and maxima represent the 25 and 75% quantile, respectively, middle bar represents 50% quantile (median). Whiskers denote 1.5x the interquartile range. Statistical significance determined by binomial logistic regression on Eu and St counts. *P*-value for TSA vs. DMSO = 8.69e-13, <2e-16 for TSA vs. DMSO 24 h switch, and 7.67e-09 for TSA vs. TSA 24 h switch, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. Source data are provided as a Source Data file.

differentiation. Thus, memory of environmental exposure can be stored and erased by acetylation and deacetylation, respectively.

TSA treatment delays development in worms and flies

The extended *eud-1* expression in TSA was reminiscent of repeated or delayed developmental programs seen in some heterochronic gene mutants[47]. It was also previously shown that the Eu morph develops ~6 h slower than the St morph[48]. Therefore, we speculated that acetylation/deacetylation might also affect plasticity by affecting developmental speed. We coupled hypochlorite-treatment with temporary starvation to obtain highly synchronized cultures, and compared developmental rates in the presence or absence of HDAC inhibition. With this higher-resolution staging we found that TSA prolonged development by several hours (Fig. 5a, $p < 0.05$, two-way ANOVA). Most stressful or toxic treatments that delay development (i.e., excess DMSO, ethanol, low temperatures[49]) lead to St phenotypes, arguing for a more specific relationship between acetylation, development, and the Eu morph. To explore the generality of this phenomenon, we repeated these experiments in *Caenorhabditis elegans*, and again found a similar developmental delay (Fig. 5b, $p < 0.05$, two-way ANOVA). Importantly, *C. elegans* lacks mouth-form plasticity, implying that the effect on development is ancestral to the effect on mouth form. Recent data suggest that the divergence time between *C. elegans* and *P. pacificus* is 100-200 million years ago[50], hinting at a deeply conserved mechanism to control development timing. Indeed, when searching the literature, we found a similar effect of TSA on development had been reported in *Drosophila melanogaster* in 2001[51], which we independently confirmed (Fig. 5c, $p < 0.05$, two-way ANOVA). Thus, HDAC inhibition delays development in three highly diverged species of Ecdysozoa, a superphylum of molting animals spanning hundreds of millions of years[50].

Histone acetylation may regulate plasticity by licensing developmental speed

Finally, we examined whether H4 hyperacetylation is the cause of delayed development. We found a 2.2-fold induction of H4K12ac in the presence of TSA in *C. elegans*, and a 2.6-fold induction in *D. melanogaster* (Fig. 5d–g, $p < 0.05$), in good agreement with our results from *P. pacificus*. However, we also saw an increase in H3 acetylation in *C. elegans* and *D. melanogaster*, consistent with previous reports that TSA has pan-deacetylase activity[33]. Given that both mouth form and developmental timing in *P. pacificus* are regulated by deacetylation, but control of timing appears to be ancestral, we speculate that developmental licensing was co-opted during the evolution of mouth-form plasticity. In this model, slowing down development would enable a longer period of switch-gene expression prior to the critical window.

Discussion

Developmental plasticity allows organisms to adjust development to match their environment[5]. Environmental sensitivity is often limited to a specific stage of development, referred to as the critical 'window' or 'period'. Here, we show that entry and exit of the critical window determining *Pristionchus* mouth form is defined by H4K5/12 acetylation and de-acetylation. Moreover, we demonstrate that acetylation in juveniles can determine adult phenotypes.

HDAC inhibition also affects ant caste behavior[28,52], beetle horn size[29] and ocular dominance in mammals[53], suggesting that acetylation may be a common mechanism of regulating developmental plasticity. Yet, identifying which specific modification(s) control plasticity has been complicated by the number of acetyllysine sites on both histone 3 and 4. The role of H4K5/12ac in mouth form was uncovered in part due to our unbiased biochemical approach and in part to the unusual specificity of TSA in *P. pacificus*. H4 N-terminal tail acetylation is broadly correlated with gene activation[54], however, the functions of most individual acetyllysines are not well known. H4K16ac is the exception, and has been shown to be necessary for dosage compensation in flies[54] and hematopoietic differentiation in mammals[55]. However, we did not observe a significant correlation of this mark with switch gene transcription, or with commonly studied enhancer/promoter modifications on H3[27]. There is indirect evidence though for a role of H4K12ac in plasticity. For instance, H4K12ac has been linked to acetyl-CoA levels[35], suggesting that it may connect diet or metabolism to changes in gene expression. Furthermore, stimulating H4K12ac can promote memory formation, a paradigm of neuronal plasticity[37,56]. While the outcomes of neuronal and morphological plasticity are clearly different, the proximate mechanisms regulating plasticity may be shared.

Our time-resolved data also support a mechanistic role for H4K5/12ac in transcriptional elongation. This has been hinted at by ChIP-seq data showing an abundance of H4K12ac in gene bodies compared to other acetylated lysines[38,57]. Furthermore, the bromodomain and extraterminal domain (BET) family protein BRD4, which binds to acetylated histones, was recently shown to act as an elongation factor to facilitate RNAPII clearance through chromatin[58]. BRD4 promotes estrogen receptor-positive breast cancer by interacting with hyperacetylated H4K12[36], and the related protein BRD2 interacts specifically with H4K5/12ac in immortalized human cell lines[36,59,60]. It may be worth investigating whether H4K5/12ac has conserved roles in hormone-dependent processes by licensing transcription elongation at key switch genes[61,62]. Alternatively, TSA has also been shown to induce cell cycle arrest[63], which could explain its effect on development. Untangling these mechanistic possibilities may provide insight into both fundamental developmental biology and drug design.

The identified role of histone acetylation in plasticity also provides a potential avenue for epigenetic information storage. However, the half-life of histone acetylation is generally considered too short

https://doi.org/10.1038/s41467-023-37734-z

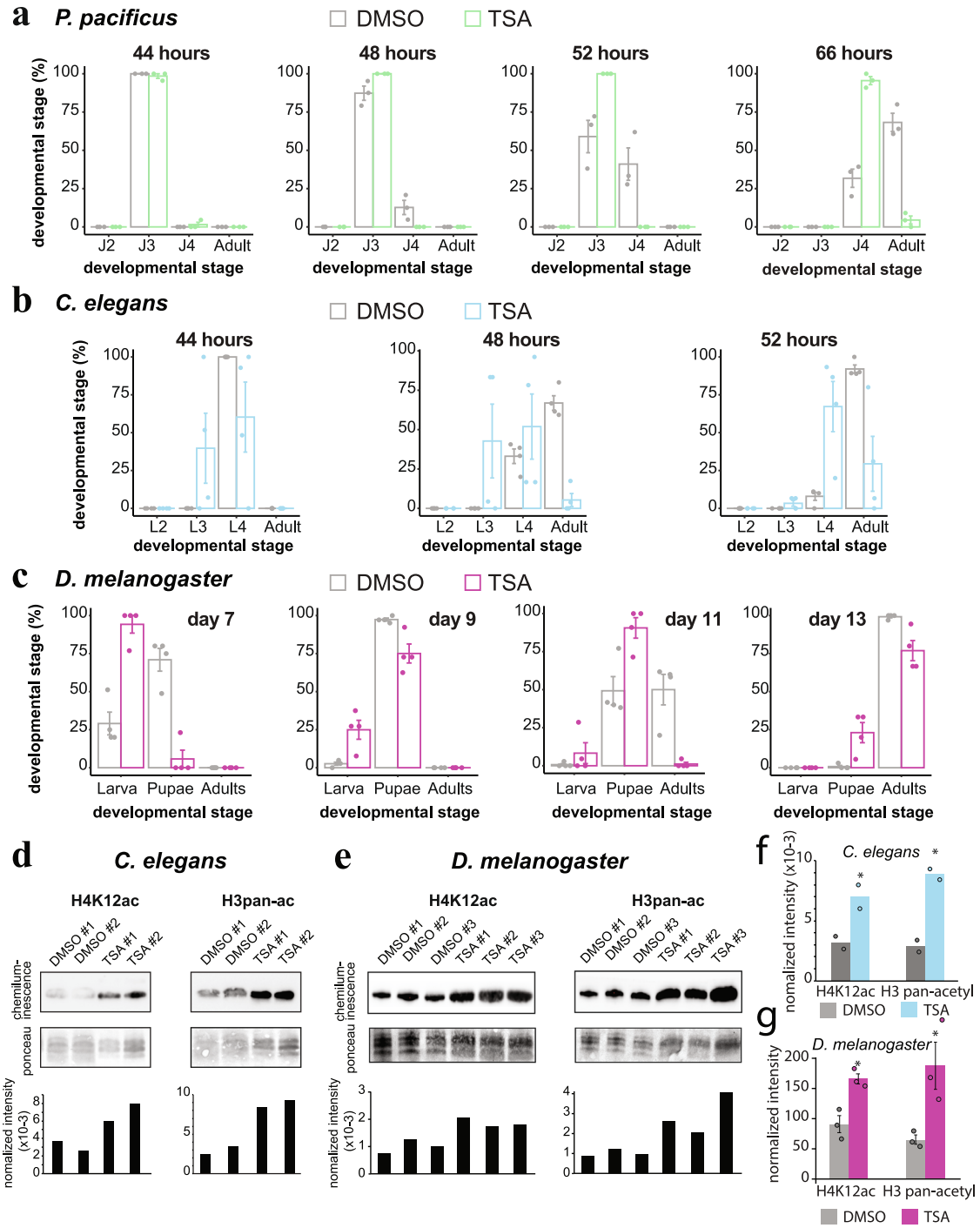


Fig. 5 | Histone acetylation links plasticity to developmental speed. **a** Developmental rate of *P. pacificus*, **b** *C. elegans*, and **c**, *D. melanogaster* +/- TSA, $n = 3$ biologically independent populations of *P. pacificus*, 4 for *C. elegans*, and 4 for *D. melanogaster*. Significance was assessed with a two-way ANOVA. Error bars represent S.E.M. **d** Representative WBs of acid-extracted histones from *C. elegans* and **e** *D. melanogaster*. Intensity of chemiluminescence was normalized to total histone levels detected by Ponceau S. **f** Mean normalized intensity of WBs in *C. elegans* ($n = 2$ biologically independent populations) and **g** *D. melanogaster* ($n = 3$ biologically independent populations). Statistical significance determined by 1-tailed students *t*-test. *P*-values for *C. elegans* H3pan-ac = 0.0061, H4K12ac = 0.038. *P*-values for *D. melanogaster* H3pan-ac = 0.019, H4K12ac = 0.0052. Error bars represent S.E.M. Source data are provided as a Source Data file.

(1–2 h) to provide long-term memory[64]. Nevertheless, more recent studies have shown that histone acetylation can persist through mitotic division;[59,65,66] reviving the question of their epigenetic potential. Our results add an important next layer by demonstrating that histone acetylation can provide long-term environmental memory in a multicellular organism. More broadly, our results also provide empirical and mechanistic evidence that developmental plasticity is channeled through histone modifications.

Finally, the confluence of developmental timing and plasticity has intriguing implications for the evolution of novel phenotypes. Regulating developmental timing, referred to as heterochrony, has been hypothesized for nearly a century to facilitate evolutionary novelty[67]. While ample morphological evidence supports this hypothesis, there are scant known molecular mechanisms. Our results suggest that acetylation/de-acetylation acts as a timer to progress through development, and as a lever to regulate switch gene (*eud-1*) expression. In

Pristionchus, prolonging deacetylation could delay development and extend switch gene expression into the critical window, which would initiate the Eu morph. If sufficiently selected upon, the Eu morph would ultimately canalize, as appears to have occurred in some *Pristionchus* species[68], potentially representing plasticity-first evolution[69]. Going forward, it will be important to evaluate how H4K5/12 acetylation affects both molecular and evolutionary mechanisms.

Methods

Environmental switching experiments

Gravid adult *P. pacificus* were bleach-synchronized[70] and eggs were seeded into agar or liquid media as described in Werner et al. 2017[23], and then switched to the other condition at the indicated time points. At 84 h post-synchronization, adult worms were phenotyped with a 100x/1.4 oil DIC objective (Zeiss), and percent Eu was plotted on the y-axis. Statistical significance was assessed by binomial logistic regression on Eu and St counts between a given time point and both the t'0 and the 84 h phenotype (see section below on 'Statistics on mouth-form frequencies' for more details). Error bars represent S.E.M. for n=5 independent biological replicates, with >20 worms counted per replicate.

Statistics on mouth-form frequencies

To test the statistical significance of the effect of small molecule inhibitors or switching experiments on mouth-form frequencies, while incorporating biological replicates, we performed a weighted binomial logistic regression in the R base statistical software package. Eu counts were modeled as 'successes' and St modeled as 'failures,' and replicates were included as repeated measurements: summary(glm(formula = cbind(Eu,St)~condition, data = results_table.txt, family = "binomial"))

Resulting *p*-values were adjusted for multiple testing using a *bonferroni*-correction (i.e., by the number of inhibitors tested), and are indicated in corresponding figure panels. $n \geq 3$ in all cases, with individual replicates shown as data points.

Fitting reciprocal transplant data

Raw data of transplantation experiments were normalized on the y-axes to '% plasticity' by subtracting the average minimum value (e.g., the average percent Eu when grown for the entire duration of development in liquid) from each time point. The normalized data were fit to a four-parameter logistic function, which is commonly used in pharmacology to fit a physiological response to drug concentration[71] or ligand binding to a protein in biochemistry (i.e., the Hill equation):

$$\%\text{plasticity} = a + \frac{(b-a)}{1+\left(\frac{K_T}{\text{time point}}\right)^{-n}}$$

Where a = the minimum asymptote (or plasticity after determination), b = maximum asymptote (or plasticity at $t' = 0$), K_T = the inflection point of the curve (or the time point when plasticity transitions to determination, equivalent to the effective concentration EC_{50} in pharmacological studies, or K_A in protein-binding experiments), and n = the slope at the steepest part of the curve (or the rate of determination, also known as the Hill coefficient). Note, K_T is not necessarily equivalent to 50% plasticity because the maximum asymptote or maximum plasticity b is not necessarily equal to 100%. To fit the data, a nonlinear least squares regression was performed in R using the 'nls' function with upper bound = 100, lower bound = 0, and starting estimates of $a = 0$, $b = 100$, $K_T = 60$, and $n = 10$, which reached convergence after four iterations. The fitted values were plotted using 'ggplot2,' and a line marking the inflection point K_T was added by 'geom_vline.'

Geometric morphometric (GMM) quantification

Two-dimensional mouth shapes were quantified using 13 fixed landmarks and 3 sliding semilandmarks. Landmark configurations from $n \geq 30$ animals per stage and condition were used as input for shape analysis by General Procrustes Alignment (GPA), and PCA was performed on the aligned Procrustes coordinates (i.e., shapes) as described in Theska et al.[72]. GPA was performed using 'gpagen' of geomorph (ver. 4.0.0); sliding of the semilandmarks was achieved by minimizing bending energy. Statistical testing for differences between group means of shape variation was performed on Procrustes coordinates of animals of the same stage in different conditions via permutational Multivariate Analysis of Variance (*PERMANOVA, or* 'Procrustes' ANOVA) with a randomized residual permutation procedure (RRPP). This was done using the 'procD.lm' function of geomorph with 100,000 iterations, seed = NULL. Resulting relative effect sizes (*Z*-scores) and FDR-adjusted *p*-values are presented adjacent to corresponding PCA plots. We considered a PERMANOVA effect as 'incompatible with the null hypothesis' if the relative effect size was greater than or equal to two times the standard deviation (i.e., $Z \geq 2.0$) and the associated *p*-value was below a type I error rate of 0.05 (i.e., $p < 0.05$).

Histone-modifying-enzyme inhibitor assays

Worm cultures were synchronized by adding bleach solution (1 ml bleach/0.5 ml 5M NaOH) to 3.5 ml of gravid adult worms, which were washed with M9 from 3 x 6 cm NGM agar plates per experimental condition. Eggs were then aliquoted to 10 ml S-medium liquid cultures or 6 cm NGM agar plates. For liquid experiments, inhibitors were added directly at the time of aliquoting eggs. For agar experiments, fresh 6 cm NGM plates were spotted with 500 µl of overnight cultures of *E. coli* OP50 (37 °C in LB, 180 rpm) plus inhibitors and spread evenly over the surface of the plate, then air-dried in a chemical hood before adding 100-500 bleached eggs. The following concentrations of inhibitors or control (DMSO) were added: 100 µl DMSO (Sigma Aldrich cat. #D8418) = 1% (v/v), 100 µl 6.6 mM TSA (Selleckchem cat. #S1045) dissolved in DMSO = 66 µM final conc., 100 µl 2.2 mM C646 (Sigma Aldrich cat. # 382113) dissolved in DMSO = 22 µM final conc., 100 µl 1.7 mM BIX (Sigma Aldrich cat. #B9311) dissolved in DMSO = 17 µM final conc., 100 µl 1.85 mM GSK343 (Sigma Aldrich cat. #SML0766) dissolved in DMSO = 18.5 µM final conc., and 30 µl 43 mM NU9056 (Tocris cat. #4903) dissolved in DMSO = 129 µM final conc.. For Fig. S2, inhibitors were resuspended at the following concentrations before titrating into LC: 1 M Na-butyrate (Sigma Aldrich cat. #B5887) dissolved in water, 1 mM apicidin (Sigma Aldrich cat. #A8851) dissolved in DMSO, 10 mM CI-944 (Sigma Aldrich cat. #C0621) dissolved in DMSO, 10 mM TMP195 (Selleckchem cat. #S8502) dissolved in DMSO, 10 mM Pyroxamide (Sigma Aldrich cat. #SML0296) dissolved in DMSO, 5 mM SAHA/Vorinostat (Sigma Aldrich cat. #SML0061) dissolved in DMSO, 4 mM Tubacin (Sigma Aldrich cat. #SML0065) dissolved in DMSO, 10 mM CUDC-101 (Selleckchem cat. #S1194) dissolved in DMSO, 10 mM Panobinostat (MedChemExpress cat. #HY-10224) dissolved in DMSO, and 10 mM Belinostat/PXD101 (MedChemExpress cat. #HY-10225) dissolved in DMSO.

Histone purification and western blot

For nematodes, crude nuclei were obtained as in Werner et al. 2018[40] but without sucrose cushion purification, with starting inputs of 200–500 µl worm pellets (10–20 × 10 cm plates of bleach-synchronized worms) collected at 72 h in the presence of 100 µl DMSO, 66 µM TSA, or 10 mM butyrate (see 'Histone-modifying-enzyme inhibitor assays' method section for resuspension and cat. #s). Histones were acid-extracted from crude nuclei, and precipitated in TCA (Sigma Aldrich cat. #T9159) following Shechter et al.[73], and re-suspended in 80 µl water. Histone yields from each purification were determined by using a calibration curve of recombinant H3 and H4 (NEB, cat. #M2503S, #M2504S, respectively). Absolute amounts of histone were determined by densitometry on coomassie-stained bands with Fiji[74] for *Pristionchus* data (conducted at the MPI in Tübingen) and Image Lab software (BioRad) for *C. elegans* and *D.*

melanogaster (conducted at the University of Utah). For Western Blot, 5 μg of histone sample was loaded per lane on a BioRad 'any Kd' Precast gel (cat. #4569033) alongside 5 μl PageRuler Prestained Protein Ladder (Thermo Scientific cat. #26616), and run 200V for 28 min. A Wet Transfer was performed with Bjerrum Schaffer-Nielsen Buffer + SDS (48 mM Tris, 39 mM glycine, 20% methanol, 0.0375% SDS, pH 9.2) in a Mini Protean Tetra box (BioRad) to 0.2 μM, 7 × 8.5 cm pre-cut nitrocellulose membranes (BioRad, cat. #162-0146), at 4 °C with magnetic stirring and an opposing cold-pack, 100V for 10 min, then 60V for 20 min. Total-protein transferred was visualized by 10-20 ml 0.1% ponceau/5% acetic acid for 5–10 min with rotation, washed with distilled water until bands became apparent, and then imaged on a Quantum gel imager (Vilber Lourmat) with white light using the 'preview' function (in Tübingen) or BioRad ChemiDoc (Utah). Nitrocellulose blots were then briefly washed in TBS (50 mM Tris-HCl, 150 mM NaCl, pH 7.5), then blocked with 20 ml 5% nonfat dry milk (BioRad, cat # 170-6404) in TBS for 1 h with rotation. Membranes were then washed 2 x 5–10 min in TBS plus 0.05% Tween 20 (TBS-T). Primary antibodies (Supplementary Data 1) were then incubated at 1:1000–2000 dilution in 5 ml 2.5% nonfat dry milk/TBS-T with membranes overnight (~12 h) at 4 °C with rotation. The next day, membranes were washed 4 x with TBS-T, 5 min each with rotation. Secondary antibodies corresponding to the animal immunoglobulin of the primary antibody, fused to horseradish peroxidase (Anti-rabbit IgG-HRP, Cell Signaling, cat. #7074S), were then incubated at 1:2,000 in 5 ml 2.5% nonfat dry milk/TBS-T for 1 h at room temperature. Membranes were then washed 4 x with TBS-T, plus one additional wash in TBS to remove residual Tween 20. To image, membranes were incubated in a 5 ml 1:1 A:B solution of Clarity Western ECL substrate (BioRad cat. #170-5060) for 5 min at room temperature with rotation, and chemiluminescence was detected on a Fusion SL imager (Vilber Lourmat; Tübingen) or ChemiDoc (BioRad; Utah) within 5 min.

For *Drosophila* experiments, flies were cultured on standard medium containing cornmeal, yeast, agar, and molasses, and maintained at 25 °C and 60% humidity on a 16:8 light:dark cycle. Additionally, the medium contained either TSA dissolved in DMSO, or DMSO only as a control. These were added to the medium at a concentration of 0.34% (v/v) once it had cooled to ~50 °C. This resulted in a final concentration of 10 μM TSA in the TSA-containing medium. The genotype of all flies used was w^{1118} (Bloomington Drosophila Stock Center #3605). For crude histone purification, between 100 and 150 wandering third-instar larvae were collected per replicate, and three replicates per treatment were used. Larvae were rinsed twice in PBS, and histone extraction, purification, and quantification were performed as described above. Western blots were performed as described above, except only 1 μg of histone sample was loaded per lane, and images were acquired using a ChemiDoc MP Imaging System (BioRad). Image and statistical analysis were performed as described above, except two-tailed F tests were conducted to ensure that variances did not differ significantly between treatments, and one-tailed Student's *t*-tests were used since evidence from nematodes indicated that TSA was expected to cause increased histone acetylation. Antibodies used: H3pan = Active Motif #39140, lot #34519009 (1:2,000); H4pan = Active Motif #39926, lot #18619005 (1:1,000); H4K12 = Millipore #04-119-S, lot #3766681 (1:1,000). Intensities of chemiluminescent bands were quantified and normalized to intensities of histone bands from Ponceau staining. Statistical testing on normalized intensities was performed by a two-tailed *student's t*-test in nematodes and one-tailed test in *Drosophila*, $n = 3$ independent biological replicates for all experiments.

Histone Liquid Chromatography Tandem-Mass Spectrometry (LC-MS/MS)

Histones were extracted from 72-h worm pellets as described for Western Blots (see above) for two independent biological replicates each of 1% DMSO, 10 mM Na-Butyrate, and 66 μM TSA. 50–100 μg of soluble histones were then reduced with 1 mM DTT/50 mM ammonium bicarbonate for 1 h at room temperature, and then alkylated for 1 h in the dark with 10 mM chloroacetamide/50 mM ammonium bicarbonate. The reaction was neutralized with 10 mM DTT for 30 min. A 10x Arg-C digestion buffer was then added to histones (1x = 5 mM CaCl2, 0.2 mM EDTA, 5 mM DTT), which were digested with 1:50 Arg-C protease (Promega, cat. #V1881): histone protein at 37 °C for 12–16 h. Digest completion was assessed by running an analytical SDS-PAGE of digested sample with undigested control sample. When complete, the reaction was stopped by adding 10% trifluoroacetic acid to a final concentration of 0.5%. Digested histones were then purified on homemade desalting C18 stage-tips[75], and run on an Easy-nLC 1200 system coupled to a QExactive HF-X mass spectrometer (both Thermo Fisher Scientific) in three technical replicates per biological sample as described elsewhere[76] with slight modifications: peptides were separated with a 127-min segmented gradient from to 10-33-50-90% of HPLC solvent B (80% acetonitrile in 0.1% formic acid) in HPLC solvent A (0.1% formic acid) at a flow rate of 200 nl/min. The 7 most intense precursor ions were sequentially fragmented in each scan cycle using higher energy collisional dissociation (HCD) fragmentation. In all measurements, sequenced precursor masses were excluded from further selection for 30 s. The target values for MS/MS fragmentation were 10^5 charges and 3×10^6 charges for the MS scan.

Data was analyzed by MaxQuant software version 1.5.2.8[77] with integrated Andromeda search engine[78]. Acetylation at lysine was specified as variable modification, and carbamidomethylation on cysteine was set as fixed modification. Endoprotease ArgC was defined as protease with a maximum of two missed cleavages and the minimum peptide length was set to five. Data was mapped to the 'El Paco' protein annotation version 1 and 286 commonly observed contaminants. Initial maximum allowed mass tolerance was set to 4.5 parts per million (ppm) for precursor ions and 20 ppm for fragment ions. Peptide, protein and modification site identifications were reported at a false discovery rate (FDR) of 0.01, estimated by the target/decoy approach[79]. The label-free algorithm was enabled, as was the "match between runs" option[80]. A spectrum quality control threshold score >100 and posterior error probability (PEP) <0.01 was defined. The averages of two biological replicates of acetylated-peptide intensities normalized to total H3 and H4 peptide intensities are presented in Fig. 2h, i, and statistical significance was assessed by a two-way *student's t*-test. For fold-change, relative ion intensities of TSA-treated samples were compared to both DMSO and Na-Butyrate.

Relative *eud-1* expression by Reverse Transcription-quantitative PCR (RT-qPCR)

Worm pellets (25–100 μl) collected at the indicated time points from each culture condition were freeze-thawed 3x between liquid nitrogen and a 37 °C heat block, then re-suspended in 500 μl Trizole (Ambion cat. #15596026). RNA was extracted following the manufacturer's protocol, then purified using a Zymo RNA clean & concentrator-25 columns. RNA was eluted in 50 μl water and quantified by nanodrop (A260/280). 1 μg RNA was used for reverse transcription with SuperScript II (Thermo Fisher cat. #18064071) following manufacturers protocols. Template RNA was degraded by the addition of 40 μl base (150 mM KOH, 20 mM Tris) for 10 min, 99 °C. cDNA pH was subsequently neutralized by 40 μl acid (150 mM HCl), then diluted by 100 μl TE (200 μl final volume, 1:100 dilution). Four μl cDNA was used as input for 10 μl qPCR reactions with 2x Luna Universal qPCR Master Mix (NEB cat. #M3003X) and 0.25 μM forward and reverse primers (Supplementary Data 1) on a LightCycler 480/II 384 (Roche, serial #6073). Four technical replicates were performed per biological replicate qPCR primer set. Standard quantitative PCR thermocycling conditions were used: 95 °C for 5 min to denature, followed by 45 cycles of 95 °C for 10 seconds (s) (4.8 °C/s), 60 °C for 10s (2.5 °C/s), and 72 °C for 10s

(4.8 °C/s). After qPCR, a thermal melting profile was also obtained for each primer set and verified for whether it was a single peak.

For all experiments quantifying *eud-1* expression by RT-qPCR, threshold Ct values were compared to housekeeping genes *Ppa-cdc-42* and *Ppa-Y45F10D.4*[81] to obtain $2^{\Delta Ct}$, and the geometric mean was calculated for each time point, then normalized to *t*'0 = eggs, representing $2^{-\Delta\Delta Ct}$, or fold change relative to *t*' = 0. Error bars represent S.E.M. from 3–5 biological replicates, and a one-tailed *student's t*-test between agar and liquid was performed for assessing statistical differences.

ChIP-qPCR

Nuclear fractionation, chromatin digestion, and immunoprecipitation (IP) were performed as previously described in Werner et al. 2018[40] with an additional pre-clear step prior to IP with washed, unconjugated beads for 30 min. This additional step was empirically determined to yield mildly greater enrichment *vs*. background (Input) compared to the original protocol, or an additional step of hydroxyapatite (hap) nucleosome purification. A 1M salt wash yielded greater signal *vs*. background, however it enriched for multivalent antibody binding, i.e. higher-order nucleosomes, compared to the other conditions. After pre-clear, ~10 μg of input (fully saturating beads) was used for 10-min incubations with antibody-conjugated Dynabeads (Thermo Fisher cat. #10004D) 5 μg Ab/20 μl beads, followed by five wash steps and three tube transfers. Co-precipitated DNA was purified by 0.4 mg/ml Proteinase K digestion for 2 h, 55 °C, and AMPure XP bead purification (cat. #A63881), and resuspended in 50 μl TE buffer.

Quantitative PCR (qPCR) was performed on a 1:100 dilution of co-precipitating DNA with four technical replicates, following the parameters discussed in *eud-1* qPCR (see above), and the average Ct was normalized according to the '% Input' method (Haring et al.) as follows:

$100*2^{((Ct\ 10\%Input - log2(10)) - Ct\ Ab\ sample)}$

Statistical significance was determined by a one-way *student's t*-test (based on phenotype, WB and LC-MS/MS results).

ChIP-Seq

ChIP-sequencing reads were aligned to the 'El Paco' genome assembly (Rödelsperger et al.) using Bowtie2 (version 2.3.4.1)(Langmead and Salzberg, 2012) with standard parameters. Mapped reads were filtered for quality (e.g. samtools view -b -q 10 'sample'.bam > 'sample'.filtered.bam), then sub-sampled to 6 million reads each per library. Metagene coverage plots were made using the bedGraph output files from MACS2 and a.bed file of the El Paco gene annotation using a custom 'awk' script to capture 5' to 3' ends of mRNA, and the Homer software package (Heinz et al.) function 'makeMetaGeneProfile.pl'. Peaks were called using Macs2 or Homer 'FindPeaks' with factor set to 'histone' on each replicate, normalized to input (both methods yielded similar peak numbers and locations). Replicate peaks within 100 bp of each other were merged with BedtoolsMerge. Tag counts at peak regions between replicates and between H4K12ac and H4K5ac were compared using 'AnnotatePeaks', and Pearson correlations were calculated in Excel. Chromatin state comparisons were done using BedTools 'IntersectBed' and previously determined genome-wide chromatin states (Werner et al. 2018[40]). A Fisher Exact test of H4K12ac peaks within regulatory regions (promoter, enhancer, and transcriptional transition) and repressed regions (repressive states 1-3) vs. expected genome-wide distributions was <0.01. Differences in H4K12ac abundance +/- TSA was performed using Macs2 bedgraph output (normalized to input) with Homer "AnnotatePeaks" with size 'given' against a peak file of gene coordinates (V3). Gene IDs with average replicate tag counts >1.5 fold-change between TSA and control were extracted. Protein sequences of these genes were obtained and homologs to *C. elegans* genes were identified with BLAST. Top hits with >25% identity, >50% positive amino acid matches and query length/sequence length >10% were input into WormBase Gene Set Enrichment Analysis[41].

STX64 experiments

To prepare STX64 plates, we spotted 6 cm NGM-agar plates with 200 μl of OP50 + 10 μl STX64 (1mg/ml in DMSO) for a final concentration of just under 1 μg/ml. For control DMSO plates, we added 200 μl OP50 +10 μl DMSO. Wild type PS312 worms or the *eud-1* constitutive overexpression line Ex[*eud-1*](RS2561) were bleached and eggs were aliquoted to STX64 or DMSO plates. Thirty worms from STX64 plates were picked to new DMSO plates ~48 h post bleach synchronization ($n = 3$).

Developmental delay

P. pacificus (PS312) and *C. elegans* (N2) were bleach-synchronized, then eggs were incubated 24 h in M9 buffer with rotation at room temperature, in order to further synchronize to the L2/J2 stage with starvation. Afterwards, animals were aliquoted into 10 ml standard S-Medium liquid cultures with 66 μM TSA (100 μl) or 100 μl DMSO, and observed at the indicated time points with a 40x oil immersion objective on an upright light microscope (Zeiss Axioscope 5). Stages were identified and scored based on vulva and mouth development, $n = 3$ for *P. pacificus* and 4 for *C. elegans*. Statistical significance was determined with a two-way ANOVA in R: aov(time ~ treatment*stage+ replicate)

Drosophila embryos were collected for 3h on standard grape juice agar egg-laying medium, and transferred individually to culture vials. The genotype of all flies used was w^{1118} (Bloomington Drosophila Stock Center #3605). Flies were cultured on standard medium containing cornmeal, yeast, agar, and molasses, plus 0.1% propionic acid and 0.1% Tegosept for mold inhibition. Additionally, the medium contained either TSA dissolved in DMSO, or DMSO only as a control. These were added to the medium at a concentration of 0.34% (v/v) once it had cooled to ~50 °C. This resulted in a final concentration of 10 μM TSA in the TSA-containing medium. For each condition (10 μM TSA and DMSO-only control), four replicates were established. Each replicate consisted of a vial with 50 embryos each, for a total of 200 embryos per treatment. Beginning 118h after transfer of embryos (118h–121h post-laying), pupae and adults were counted in each vial every 24h. Flies were maintained at 25 °C and 60% humidity on a 16:8 light:dark cycle, and counting occurred 3h after lights on. Pupae were counted through Day 13 (thereafter, any new pupae could conceivably represent progeny of the original flies). Flies were sexed and counted until all replicates of both treatments stopped producing new adults, which occurred on Day 16. Each day, all new adults were frozen at −80 °C for use in Western Blots.

Statistical analysis was performed in R. Differences in time to pupation and to eclosion between TSA-treated flies and DMSO-treated controls were analyzed with two-factor ANOVA, accounting for variation between replicates (Model: time ~ replicate + treatment).

In *Drosophila* raised in media containing 10 μM TSA, the average time to pupation was 8.68 days (95% CI: 8.33–9.03) and the average time to eclosion was 12.72 days (95% CI: 12.43–13.02). In control flies treated with DMSO only, the average time to pupation was 7.35 days (95% confidence interval: 7.22–7.48) and the average time to eclosion was 11.56 days (95% CI: 11.43–11.68). Thus, flies exposed to TSA took significantly longer to develop to pupation ($p < 0.00001$, two-way ANOVA) and to eclosion ($p < 0.00001$, two-way ANOVA). The average developmental delay in TSA-treated flies was 1.33 days over the period between laying and pupation and 1.16 days over the entire life cycle from laying to eclosion, indicating that TSA delays development in the larval stages of *Drosophila* but not the pupal stage. We found that flies raised in TSA were significantly less likely to survive from laying to pupation than the DMSO-only controls (56/200 survivors in TSA vs. 142/200 in DMSO, $p < 0.00001$, two-sample chi-square test with Yates correction). However, between pupation and eclosion, the survival rate did not differ significantly between the

treatments (43/56 survivors in TSA vs. 124/142 in DMSO, $p = 0.105$, two-sample chi-square test with Yates correction), further suggesting that the main developmental effects of TSA occur prior to pupation. Consistent with previous results (Pile et al.) we found that the surviving adults in the TSA-treated group were significantly more likely to be female than DMSO-treated controls (26/43 females in TSA vs. 51/124 in DMSO, $p = 0.044$, two-sample chi-square test with Yates correction).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All nematode strains used in this study are available from the corresponding author Ralf J. Sommer (ralf.sommer@tuebingen.mpg.de). All other materials (oligonucleotide primers and antibodies) used in this study were purchased from vendors. The ChIP-seq datasets generated during this study are available at the National Center for Biotechnology Information Sequence Read Archive (NCBI SRA) data base under the accession number PRJNA628502. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD018940. All microscopy images are available upon request. Source data are provided with this paper.

References

1. DeWitt, T. J. & Scheiner, S. M. *Phenotypic Plasticity: Functional and Conceptual Approaches* (OUP USA, 2004).
2. Pigliucci, M. *Phenotypic Plasticity: Beyond Nature and Nurture* (Johns Hopkins University Press, 2001).
3. Nijhout, H. F. Development and evolution of adaptive polyphenisms. *Evol. Dev.* **5**, 9–18 (2003).
4. Müller, G. B. Evo–devo: extending the evolutionary synthesis. *Nat. Rev. Genet* **8**, 943–949 (2007).
5. West-Eberhard, M. J. *Developmental Plasticity and Evolution* (Oxford University Press, 2003).
6. Laland, K. N. et al. The extended evolutionary synthesis: its structure, assumptions and predictions. *Proc. R. Soc. B Biol. Sci.* **282**, 20151019 (2015).
7. Yan, H. et al. Eusocial insects as emerging models for behavioural epigenetics. *Nat. Rev. Genet.* **15**, 677–688 (2014).
8. Valena, S. & Moczek, A. P. Epigenetic mechanisms underlying developmental plasticity in horned beetles. *Genetics Res. Int.* **2012**, 576303 (2012).
9. Sommer, R. J. Pristionchus *Pacificus: A Nematode Model for Comparative and Evolutionary Biology* (Brill Academic Pub, 2015).
10. Bento, G., Ogawa, A. & Sommer, R. J. Co-option of the hormone-signalling module dafachronic acid–DAF-12 in nematode evolution. *Nature* **466**, 494–497 (2010).
11. Werner, M. S., Claaßen, M. H., Renahan, T., Dardiry, M. & Sommer, R. J. Adult influence on juvenile phenotypes by stage-specific pheromone production. *Iscience* **10**, 123–134 (2018).
12. Lightfoot, J. W. et al. Small peptide–mediated self-recognition prevents cannibalism in predatory nematodes. *Science* **364**, 86–89 (2019).
13. Bose, N. et al. Complex small-molecule architectures regulate phenotypic plasticity in a nematode. *Angew Chem. Int. Ed. Eng.* **124**, 12606–12611 (2012).
14. Renahan, T. & Sommer, R. J. Nematode interactions on beetle hosts indicate a role of mouth-form plasticity in resource competition. *Front. Ecol. Evol.* **9**, 752695 (2021).
15. Renahan, T. et al. Nematode biphasic 'boom and bust' dynamics are dependent on host bacterial load while linking dauer and mouth-form polyphenisms. *Environ. Microbiol.* **23**, 5102–5113 (2021).
16. Quach, K. T. & Chalasani, S. H. Flexible reprogramming of Pristionchus pacificus motivation for attacking Caenorhabditis elegans in predator-prey competition. *Curr Biol* **32**, 1675–1688.e7 (2022).
17. Bui, L. T., Ivers, N. A. & Ragsdale, E. J. A sulfotransferase dosage-dependently regulates mouthpart polyphenism in the nematode Pristionchus pacificus. *Nat. Commun.* **9**, 4119 (2018).
18. Kieninger, M. R. et al. The nuclear hormone receptor NHR-40 Acts downstream of the sulfatase EUD-1 as part of a developmental plasticity switch in pristionchus. *Curr. Biol.* **26**, 2174–2179 (2016).
19. Ragsdale, E. J., Müller, M. R., Rödelsperger, C. & Sommer, R. J. A developmental switch coupled to the evolution of plasticity acts through a sulfatase. *Cell* **155**, 922–933 (2013).
20. Sommer, R. J. et al. The genetics of phenotypic plasticity in nematode feeding structures. *Open Biol.* **7**, 160332 (2017).
21. Bannister, A. J. & Kouzarides, T. Regulation of chromatin by histone modifications. *Cell Res.* **21**, 381–395 (2011).
22. Allis, C. D. & Jenuwein, T. The molecular hallmarks of epigenetic control. *Nat. Rev. Genet.* **17**, 487–500 (2016).
23. Werner, M. S. et al. Environmental influence on Pristionchus pacificus mouth form through different culture methods. *Sci. Rep-uk* **7**, 7207 (2017).
24. Mather, K. & Winton, D. D. Adaptation and counter-adaptation of the breeding system in primula: the nature of breeding systems. *Ann. Bot. London* **5**, 297–311 (1941).
25. Sommer, R. J. Phenotypic plasticity: from theory and genetics to current and future challenges. *Genetics* **215**, 1–13 (2020).
26. Sommer, R. J. *Pristionchus pacificus: A Nematode Model for Comparative and Evolutionary Biology* Vol. 11 (Brill Academic Pub, 2015).
27. Rada-Iglesias, A. et al. A unique chromatin signature uncovers early developmental enhancers in humans. *Nature* **470**, 279–283 (2011).
28. Simola, D. F. et al. Epigenetic (re)programming of caste-specific behavior in the ant Camponotus floridanus. *Science* **351**, *aac6633* (2016).
29. Ozawa, T. et al. Histone deacetylases control module-specific phenotypic plasticity in beetle weapons. *Proc. Natl. Acad. Sci. USA* **113**, 15042–15047 (2016).
30. Serobyan, V. et al. Chromatin remodelling and antisense-mediated up-regulation of the developmental switch gene eud-1 control predatory feeding plasticity. *Nat. Commun.* **7**, 12337 (2016).
31. Yoon, S. & Eom, G. H. HDAC and HDAC inhibitor: from cancer to cardiovascular diseases. *Chonnam Med. J.* **52**, 1–11 (2016).
32. Li, W. & Sun, Z. Mechanism of action for HDAC inhibitors—insights from Omics approaches. *Int. J. Mol. Sci.* **20**, 1616 (2019).
33. Bradner, J. E. et al. Chemical phylogenetics of histone deacetylases. *Nat. Chem. Biol.* **6**, 238–243 (2010).
34. Glastad, K. M. et al. Epigenetic regulator CoREST controls social behavior in ants. *Mol. Cell* **77**, 338–351.e6 (2019).
35. Peleg, S. et al. Life span extension by targeting a link between metabolism and histone acetylation in Drosophila. *Embo Rep.* **17**, 455–469 (2016).
36. Nagarajan, S., Benito, E., Fischer, A. & Johnsen, S. A. H4K12ac is regulated by estrogen receptor-alpha and is associated with BRD4 function and inducible transcription. *Oncotarget* **6**, 7305–7317 (2015).
37. Peleg, S. et al. Altered histone acetylation is associated with age-dependent memory impairment in mice. *Science* **328**, 753–756 (2010).
38. Wang, Z. et al. Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat. Genet.* **40**, 897–903 (2008).
39. Barnes, C. E., English, D. M. & Cowley, S. M. Acetylation & Co: an expanding repertoire of histone acylations regulates chromatin and transcription. *Essays Biochem.* **63**, 97–107 (2019).

40. Werner, M. S. et al. Young genes have distinct gene structure, epigenetic profiles, and transcriptional regulation. *Genome Res.* **28**, 1675–1687 (2018).
41. Davis, P. et al. WormBase in 2022—data, processes, and tools for analyzing Caenorhabditis elegans. *Genetics* **220**, iyac003 (2022).
42. Montani, M. S. G. et al. Histone deacetylase inhibitors VPA and TSA induce apoptosis and autophagy in pancreatic cancer cells. *Cell Oncol.* **40**, 167–180 (2017).
43. Lindemann, R. K. et al. Analysis of the apoptotic and therapeutic activities of histone deacetylase inhibitors by using a mouse model of B cell lymphoma. *Proc. Natl Acad. Sci. USA* **104**, 8071–8076 (2007).
44. Falkenberg, K. J. & Johnstone, R. W. Histone deacetylases and their inhibitors in cancer, neurological diseases and immune disorders. *Nat. Rev. Drug Discov.* **13**, 673–691 (2014).
45. Sieriebriennikov, B. et al. A Developmental switch generating phenotypic plasticity is part of a conserved multi-gene locus. *Cell Rep.* **23**, 2835–2843.e4 (2018).
46. Pérez-Jiménez, M. M. et al. Steroid hormones sulfatase inactivation extends lifespan and ameliorates age-related diseases. *Nat. Commun.* **12**, 49 (2021).
47. Ambros, V. Control of developmental timing in Caenorhabditis elegans. *Curr. Opin. Genet. Dev.* **10**, 428–433 (2000).
48. Serobyan, V., Ragsdale, E. J., Müller, M. R. & Sommer, R. J. Feeding plasticity in the nematode Pristionchus pacificus is influenced by sex and social context and is linked to developmental speed. *Evol. Dev.* **15**, 161–170 (2013).
49. Lenuzzi, M. et al. Influence of environmental temperature on mouth-form plasticity in Pristionchus pacificus acts through daf-11-dependent cGMP signaling. *J. Exp. Zool. Part B Mol. Dev. Evol.* https://doi.org/10.1002/jez.b.23094 (2022).
50. Howard, R. J. et al. The Ediacaran origin of Ecdysozoa: integrating fossil and phylogenomic data. *J. Geol. Soc. London* **179**, jgs2021–jgs2107 (2022).
51. Pile, L. A., Lee, F. W.-H. & Wassarman, D. A. The histone deacetylase inhibitor trichostatin A influences the development of Drosophila melanogaster. *Cell. Mol. Life Sci.* **58**, 1715–1718 (2001).
52. Simola, D. F. et al. A chromatin link to caste identity in the carpenter ant Camponotus floridanus. *Genome Res* **23**, 486–496 (2013).
53. Baroncelli, L. et al. Experience affects critical period plasticity in the visual cortex through an epigenetic regulation of histone post-translational modifications. *J. Neurosci.* **36**, 3430–3440 (2016).
54. Akhtar, A. & Becker, P. B. Activation of transcription through histone H4 acetylation by MOF, an acetyltransferase essential for dosage compensation in Drosophila. *Mol. Cell* **5**, 367–375 (2000).
55. Rodrigues, C. P., Shvedunova, M. & Akhtar, A. Epigenetic regulators as the gatekeepers of hematopoiesis. *Trends Genet.* **37**, 125–142 (2020).
56. Gräff, J. & Tsai, L.-H. Histone acetylation: molecular mnemonics on the chromatin. *Nat. Rev. Neurosci.* **14**, 97–111 (2013).
57. Heintzman, N. D. et al. Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* **459**, 108–112 (2009).
58. Kanno, T. et al. BRD4 assists elongation of both coding and enhancer RNAs by interacting with acetylated histones. *Nat. Struct. Mol. Biol.* **21**, 1047–1057 (2014).
59. Kanno, T. et al. Selective recognition of acetylated histones by bromodomain proteins visualized in lving cells. *Mol. Cell* **13**, 33–43 (2004).
60. LeRoy, G., Rickards, B. & Flint, S. J. The double bromodomain proteins Brd2 and Brd3 couple histone acetylation to transcription. *Mol. Cell* **30**, 51–60 (2008).
61. Ediriweera, M. K., Tennekoon, K. H. & Samarakoon, S. R. Emerging role of histone deacetylase inhibitors as anti-breast-cancer agents. *Drug Discov. Today* **24**, 685–702 (2019).
62. Vigushin, D. M. et al. Trichostatin A is a histone deacetylase inhibitor with potent antitumor activity against breast cancer in vivo. *Clin. Cancer Res. Official J. Am. Assoc. Cancer Res.* **7**, 971–976 (2001).
63. Hrgovic, I. et al. The histone deacetylase inhibitor trichostatin a decreases lymphangiogenesis by inducing apoptosis and cell cycle arrest via p21-dependent pathways. *BMC Cancer* **16**, 763 (2016).
64. Zheng, Y., Thomas, P. M. & Kelleher, N. L. Measurement of acetylation turnover at distinct lysines in human histones identifies long-lived acetylation sites. *Nat. Commu.n* **4**, 2203 (2013).
65. Behera, V. et al. Interrogating histone acetylation and BRD4 as mitotic bookmarks of transcription. *Cell. Rep.* **27**, 400–415.e5 (2019).
66. Samata, M. et al. Intergenerationally maintained histone H4 lysine 16 acetylation is instructive for future gene activation. *Cell* **182**, 127–144.e23 (2020).
67. McNamara, K. J. Heterochrony: the evolution of development. *Evol. Educ. Outreach* **5**, 203–218 (2012).
68. Susoy, V., Ragsdale, E. J., Kanzaki, N. & Sommer, R. J. Rapid diversification associated with a macroevolutionary pulse of developmental plasticity. *Elife* **4**, e05463 (2015).
69. Levis, N. A. & Pfennig, D. W. Evaluating 'plasticity-first' evolution in nature: key criteria and empirical approaches. *Trends Ecol. Evol.* **31**, 563–574 (2016).
70. Stiernagle, T. *Maintenance of C. elegans. Wormbook 1–11*. https://doi.org/10.1895/wormbook.1.101.1 (2006).
71. Goutelle, S. et al. The Hill equation: a review of its capabilities in pharmacological modelling. Fundam. *Clin. Pharm.* **22**, 633–648 (2008).
72. Theska, T., Sieriebriennikov, B., Wighard, S. S., Werner, M. S. & Sommer, R. J. Geometric morphometrics of microscopic animals as exemplified by model nematodes. *Nat. Protoc.* **15**, 2611–2644 (2020).
73. Shechter, D., Dormann, H. L., Allis, C. D. & Hake, S. B. Extraction, purification and analysis of histones. *Nat. Protoc.* **2**, 1445–1457 (2007).
74. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
75. Rappsilber, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* **2**, 1896–1906 (2007).
76. Kliza, K. et al. Internally tagged ubiquitin: a tool to identify linear polyubiquitin-modified proteins by mass spectrometry. *Nat. Methods* **14**, 504–512 (2017).
77. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372 (2008).
78. Cox, J. et al. Andromeda: A peptide search engine integrated into the maxQuant environment. *J. Proteome Res.* **10**, 1794–1805 (2011).
79. Elias, J. E. & Gygi, S. P. Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat. Methods* **4**, 207–214 (2007).
80. Luber, C. A. et al. Quantitative proteomics reveals subset-specific viral recognition in dendritic cells. *Immunity* **32**, 279–289 (2010).
81. Schuster, L. N. & Sommer, R. J. Expressional and functional variation of horizontally acquired cellulases in the nematode Pristionchus pacificus. *Gene* **506**, 274–282 (2012).

Acknowledgements

We would like to thank Silke Wahl for guidance on histone LC-MS/MS and loading digested peptides on C18 columns. We would like to acknowledge Hanh Witte and Bogdan Sieriebriennikov for creating and providing the *eud-1* CRISPR mutant. We would also like to thank Talia L. Karasov, James Lightfoot, and Tess Renahan for critical reading of our manuscript, and all members of the Sommer and Werner Laboratories. We would also like to thank WormBase. Funding was generously

provided by The Max Planck Society and the School of Biological Sciences at the University of Utah

Author contributions

M.S.W. and R.J.S. designed experiments; M.S.W. and T.L. performed reciprocal transplant and RT-qPCR experiments. M.S.W. and T.T. performed GMM. M.S.W. extracted histones and performed nematode WBs, and prepared digested peptides for LC-MS/MS. M.F.W. performed LC-MS/MS with supervision from B.M.. T.K. performed all fly experiments, and S.R. carried out the initial STX-64 experiments. R.J.S. and M.S.W. provided resources. Writing by M.S.W. and R.J.S. with input from all authors.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-023-37734-z.

Correspondence and requests for materials should be addressed to Ralf J. Sommer.

Peer review information *Nature Communications* thanks Susan Mango and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Author Correction: Histone 4 lysine 5/12 acetylation enables developmental plasticity of *Pristionchus* mouth form

Correction to: *Nature Communications*
https://doi.org/10.1038/s41467-023-37734-z,
published online 13 April 2023

https://doi.org/10.1038/s41467-023-38693-1

Published online: 18 May 2023

Check for updates

Michael S. Werner, Tobias Loschko, Thomas King, Shelley Reich, Tobias Theska, Mirita Franz-Wachtel, Boris Macek & Ralf J. Sommer

The original version of this Article contained an error in Figure 3a. The color key indicated bars representing DMSO treatment were coloured light purple rather than gray. The legend to Figure 3 inadvertently contains references to colors for panels h-j that are not present in the final figure.

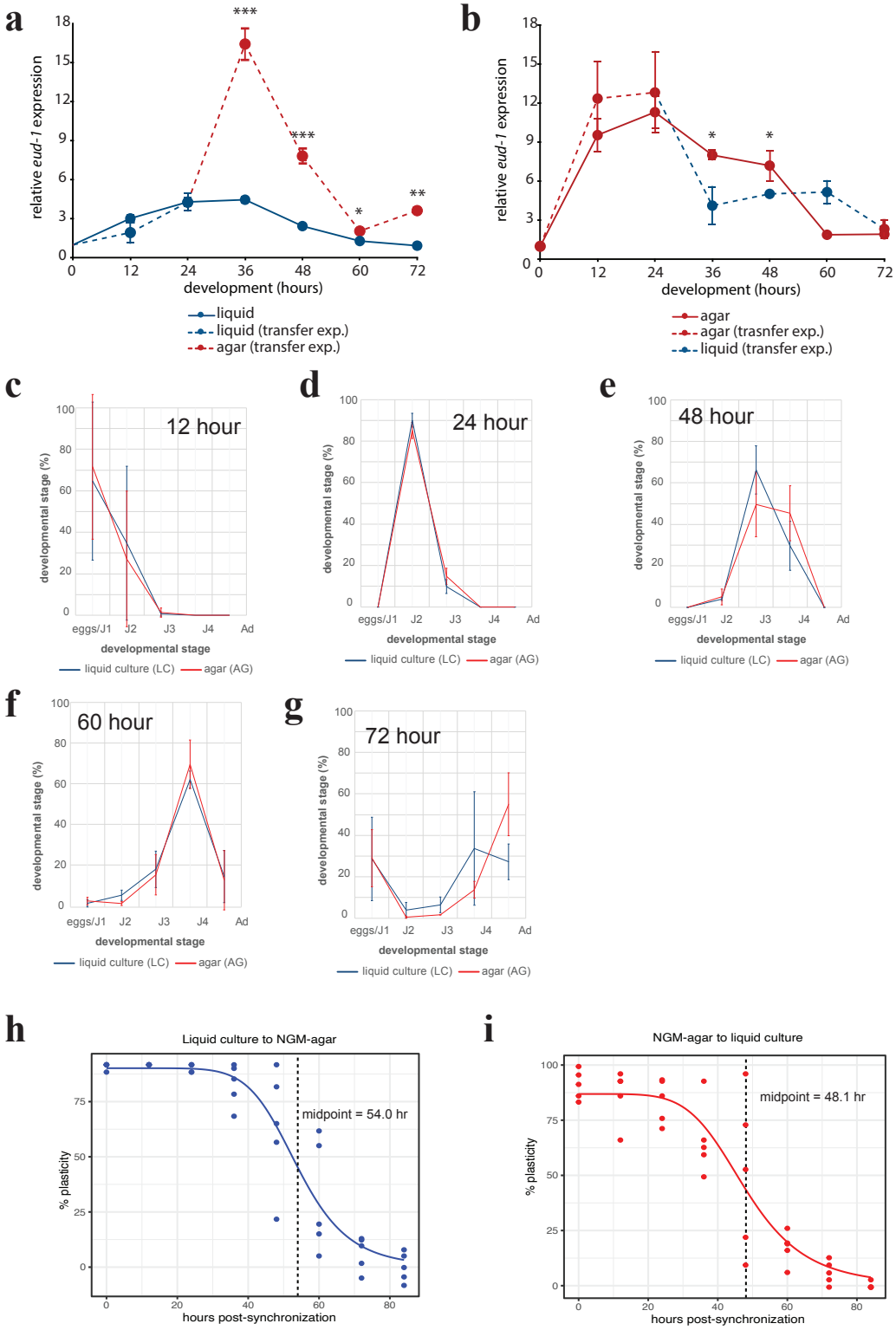
In the caption to Figure 2, the sentence "-values for H4K5ac…" should have read "*P*-values for H4K5ac…"

In the sentence on page 5 of this Article beginning "H4K12ac peaks were significantly enriched…" the term "TSA-induced genes" should have read "genes adjacent to hyper-acetylated regions." In the sentence on page 7 beginning "BRD4 promotes estrogen receptor-positive…" the text "human cell lines[59,60]. [36]" should have read "human cell lines[36,59,60]."

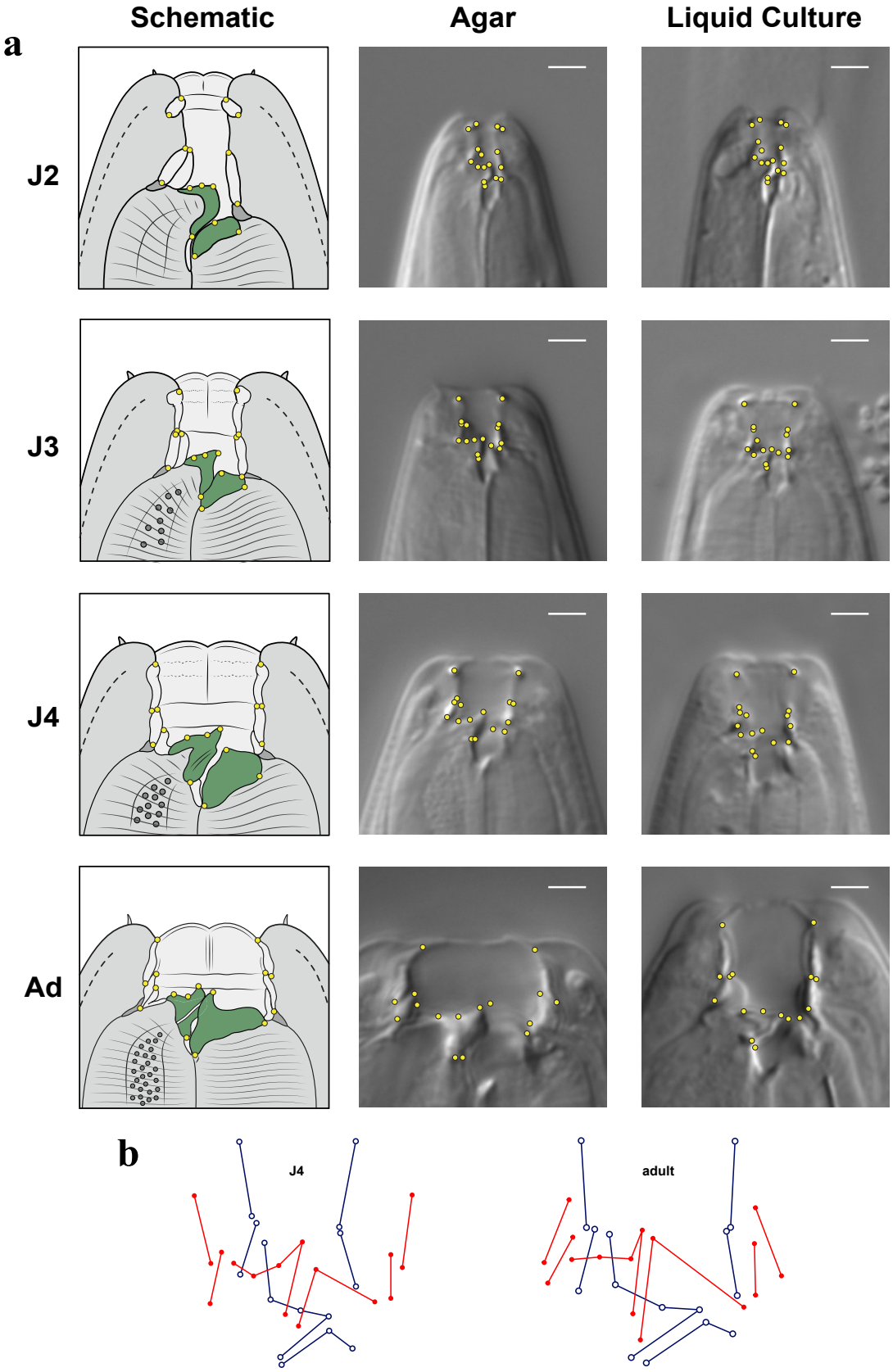
In the sentence on page 9 of this Article beginning "Gravid adult P. pacificus were…" the reference "Werner et al." should have read "Werner et al. 2017[23]."

In the sentences on pages 9 and 11 beginning "For nematodes, crude nuclei…" "Nuclear fractionation, chromatin…" and "Chromatin state comparisons were…" the references to "Werner et al." should have read "Werner et al. 2018[40]."

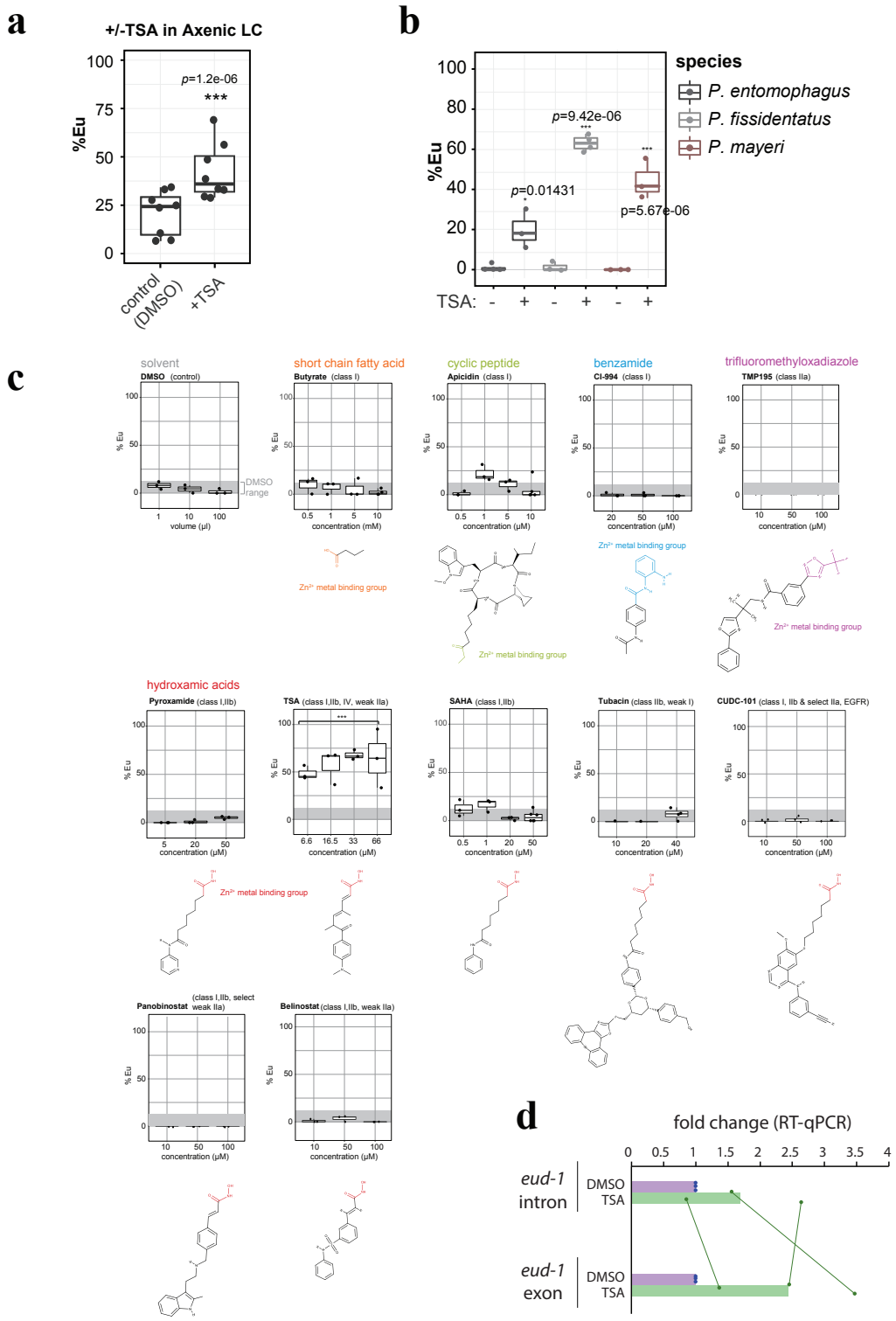
This has been corrected in the PDF and HTML versions of the Article.



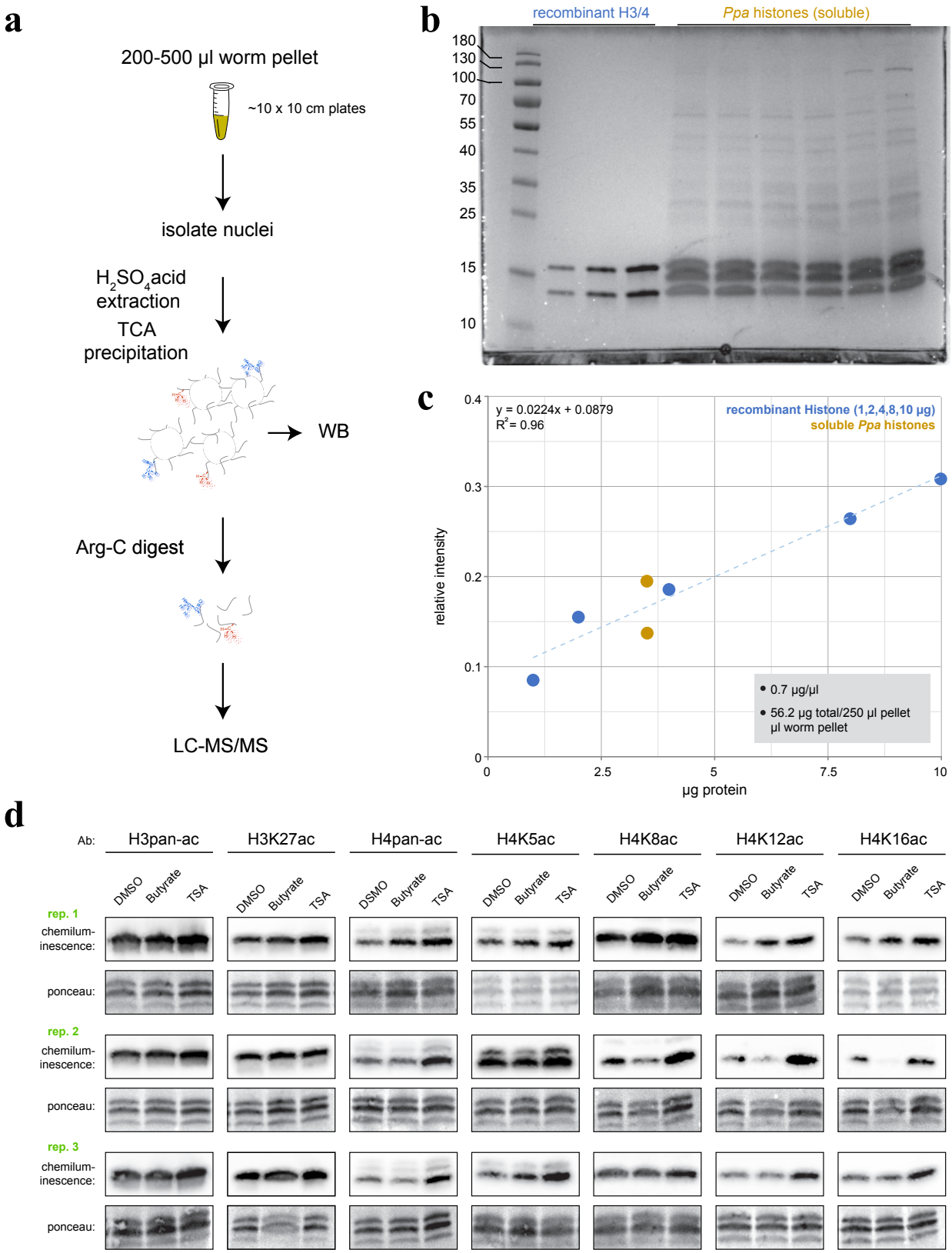
Supplementary Figure 1: Switch gene transcription corresponds to environmental sensitivity. a, Transcription of *eud-1* when worms are transferred from liquid culture to NGM-agar, or **b**, NGM-agar to liquid culture. Relative *eud-1* expression was measured by RT-qPCR (geometric mean of $2^{\Delta Ct}$ relative to *Ppa-cdc-42* and *Ppa-y45F10D.470*, normalized to t'=0). Error bars represent S.E.M. for 3 ind. worm populations, except 'liquid culture 12 hrs' and 'agar-liquid transfer 36 hrs' where n=4, and 'agar-liquid transfer 48 hrs' where n=5. Statistical significance was determined by a 1-sided student's t-test. p values for panel 'a': 36-hr=0.00041, 48-hr=0.00063, 60-hr=0.045, 72-hr=0.0011; p values for panel 'b': 36-hr=0.042, 48-hr=0.047. '*'=p<0.05, '**'=p<0.01, '***'=p<0.001. **c-g**, Developmental stages in agar (red) and liquid culture (blue) after hypochlorite synchronization, n=3 biological replicates, error bars reflect standard deviation centered around the mean. **h**, logistic fit of reciprocal transplant experiments (Fig. 1c) from liquid to NGM-agar and **i**, vice-versa. Source data are provided as a Source Data file.



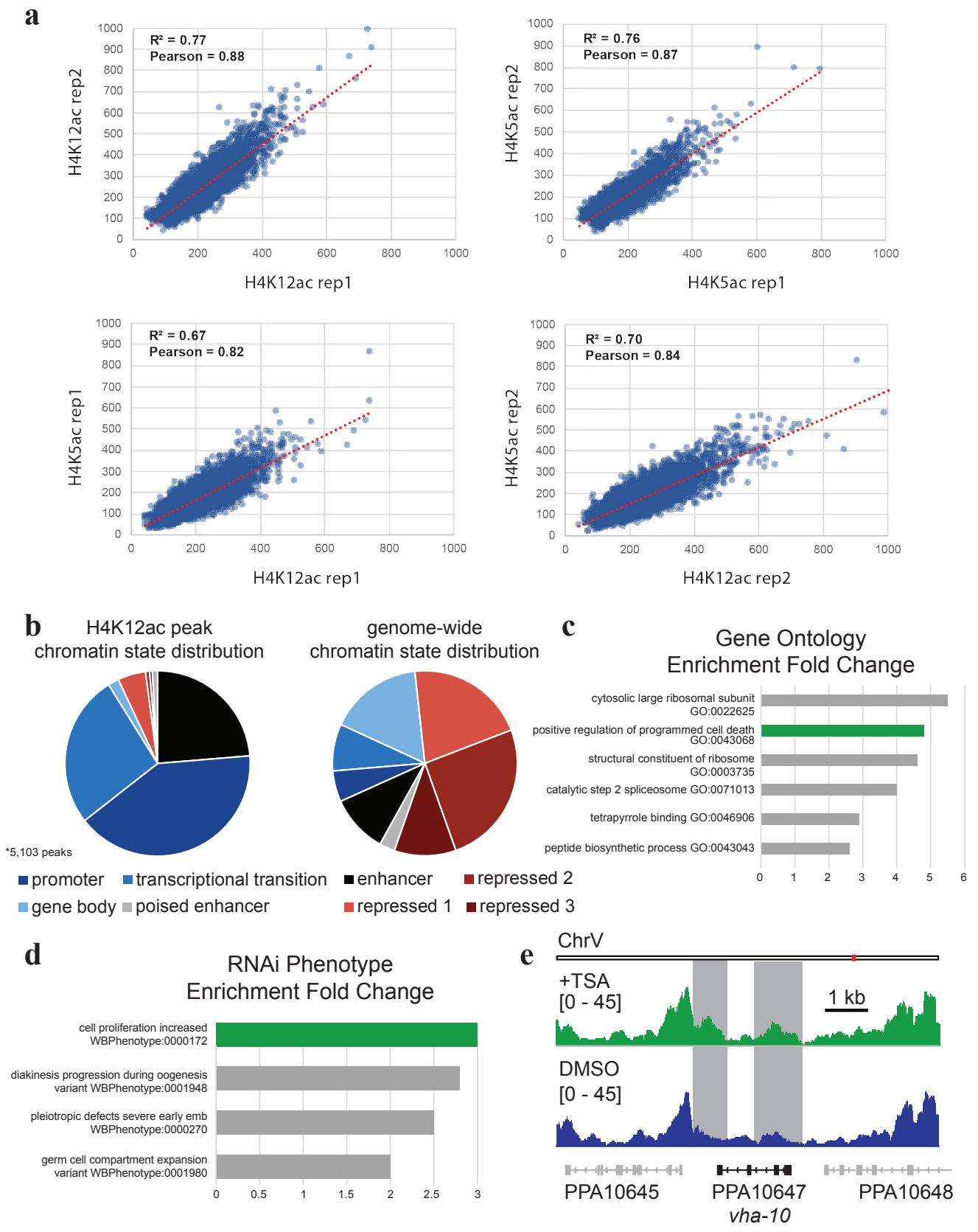
Supplementary Figure 2: Homologous landmarks for quantitative geometric morphometrics. a, Homologous structural landmarks used for GM in both NGM-agar and liquid culture at each developmental stage (except for J1). Note, only one morph (Eu) is shown for the adult illustration, and some landmarks in the microscopy images are out of the current focal plane. Scale bar = 5 µm. Schematics drawn in Adobe illustrator, and images taken using 100x 1.4 oil immersion objective with DIC prism on a Zeiss Axio Imager. **b**, Wire-frame plots of each morph (red=Eu, blue=St) in J4 and adults. Produced in R according to Theska et al. 2020. Source data are provided as a Source Data file.



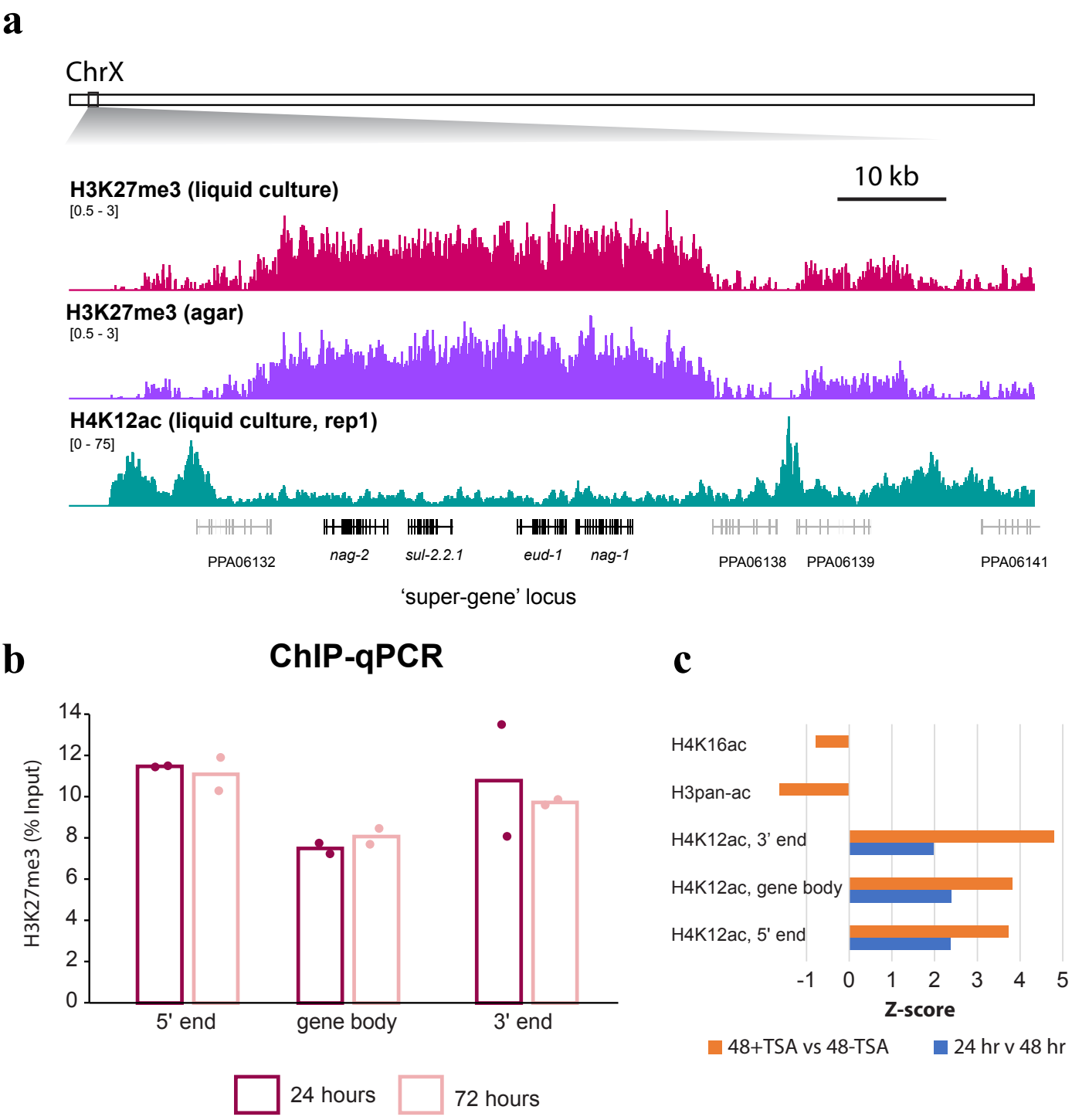
Supplementary Figure 3: TSA has a conserved and specific effect on mouth form. a, Phenotype of *P. pacificus* in axenic liquid culture +/- TSA, n=8 biologically independent worm populations. **b**, Different Pristionchus species in LC +/- TSA, n=7 biologically independent *P. fissidenatus* populations (3 DMSO, 4 TSA), 7 *P. entomophagus* populations (4 DMSO, 3 TSA), and 6 *P. mayeri* populations (3 each). **c**, Phenotype of *P. pacificus* animals grown in liquid culture with the indicated amounts of HDAC inhibitors. Samples without TSA were treated with 100 µl DMSO, n≥3 for all concentrations. Statistical significance in 'a-c' was calculated by glm relative to DMSO. Box-plot minima and maxima represent the 25% and 75% quantile, respectively, and middle bars represent the 50% quantile (median). Whiskers denote 1.5x the interquartile range. p value for TSA = 1.42e-09. '*'=p<0.05, '**'=p<0.01, '***'=p<0.001. **d**, RT-qPCR of *eud-1* intronic and exonic segments +/- TSA. qPCR was performed with distinct intronic and exonic primers and normalized to the geometric mean of *Ppa-cdc-42* and *Ppa-Y45F10D.470*. Bar plot indicates avg. fold-change between replicates, n=3. Lines connecting data points indicate paired replicates. Source data are provided as a Source Data file.



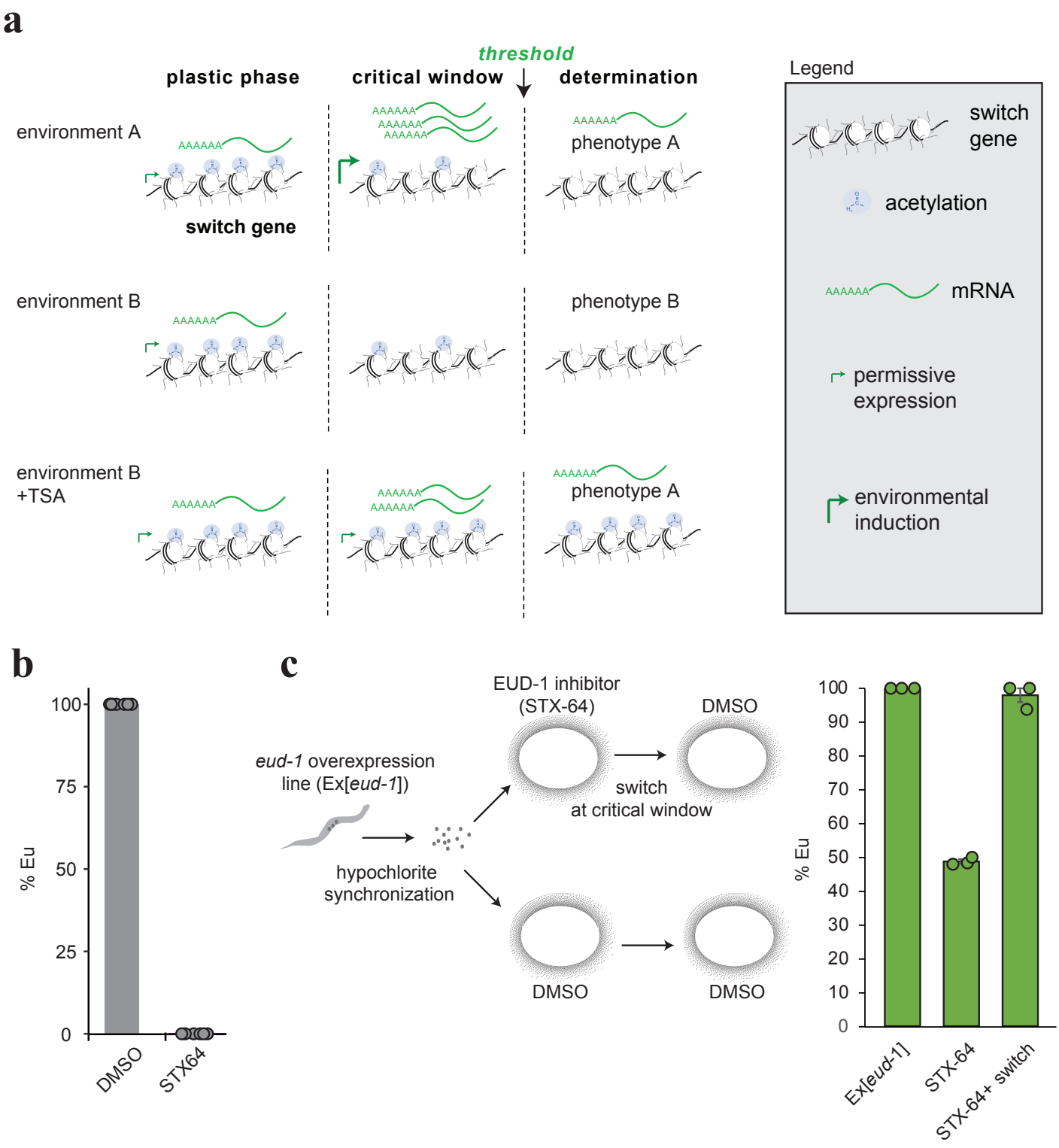
Supplementary Figure 4: **Histone Purification and Raw Data for Western Blots. a**, Schematic of histone acid extraction for Western Blot and LC-MS/MS after digestion with Arg-C protease. **b**, Example SDS-PAGE of histone extraction and a recombinant H3/4 calibration curve. Proteins were visualized with Coomassie Brilliant Blue R-250. **c**, Example calibration curve with recombinant histone used to calculate *P. pacificus* histone quantity for WB and LC-MS/MS, n=2 technical replicates. **d**, Raw data of chemiluminescent signal from three independent biological replicate WBs for each antibody, and total histone amount transferred to nitrocellulose membranes visualized with Ponceau Red. Source data are provided as a Source Data file.



Supplementary Figure 5: Analysis of H4K5/12ac +/- TSA. a, ChIP-seq tag correlations between H4K5 and 12 acetylation at H4K12ac peaks. **b**, Distribution of H4K12ac peaks in each chromatin state, compared to genome-wide distributions, from Werner et al., 2018. Peaks were significantly enriched in regulatory states (promoter, enhancer and transcriptional transition; Fisher's Exact Test). **c-d**, Gene Set Enrichment Analysis of *C. elegans* best-hit homologs (WormBase). Bar plots represent significant observed vs. expected genes in Gene Ontology (c) and RNAi phenotypes (d). **e**, Example gene (*Ppa-vha-10*) with 'programed cell death' ontology. Source data are provided as a Source Data file.

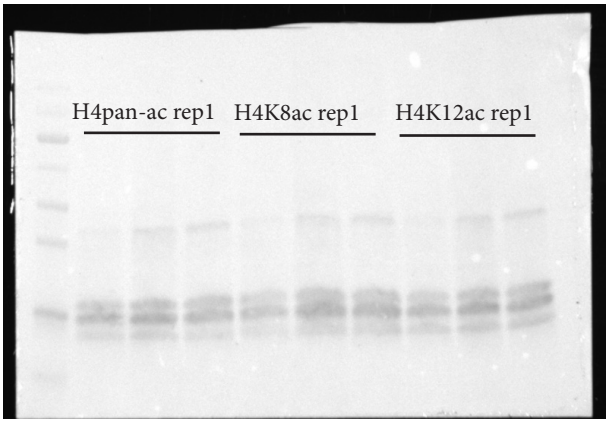


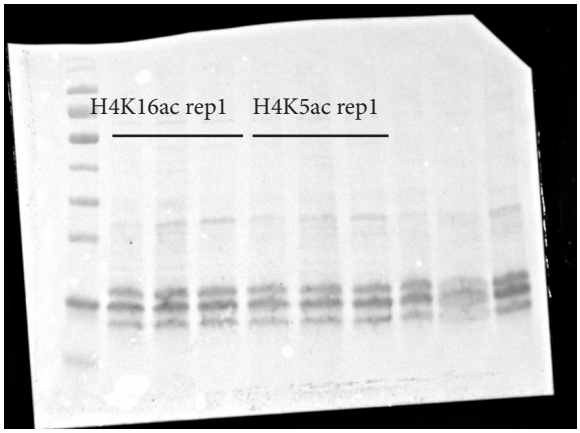
Supplementary Figure 6: H3K27me3 chromatin immunoprecipitation. a, ChIP-seq of H3K27me3 from 48-hour worms grown in liquid culture, compared to H4K12ac in the same conditions. **b**, ChIP-qPCR of H3K27me3 at 24 hours and 72 hours across *eud-1* (same loci probed in main text figure 3), n = 2 biological replicates. Differences were not significant, Student's t-test. **c**, Z-scores of ChIP-qPCR data from Fig. 3h-k. Source data are provided as a Source Data file.

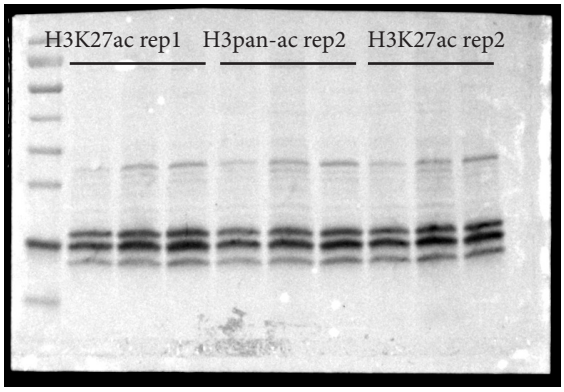


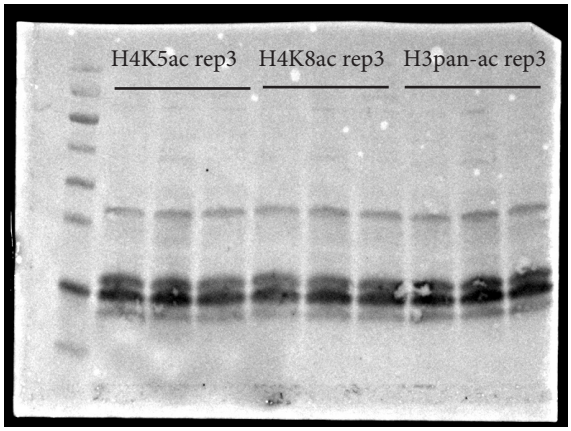
Supplementary Figure 7: Model for establishment of plasticity during the critical window. a, Plasticity is established by environment-independent H4K5/12ac and permissive transcription during the critical window. Ultimately H4K5/12 is deacetylated leading toward repression of switch genes and the end of the critical window. When TSA is added in liquid culture, the fixed juvenile chromatin state of H4K5/12ac prevents switch gene repression. **b**, STX-64 induces the St phenotype at 1 µg/ml. n=6 STX64, 7 DMSO, with at least 20 worms per replicate. **c**, Switching between STX-64 plates and DMSO plates at 48 hours 'rescues' the Ex[eud-1] phenotype=3 biologically independent worm populations. Error bars represent S.E.M. Source data are provided as a Source Data file.

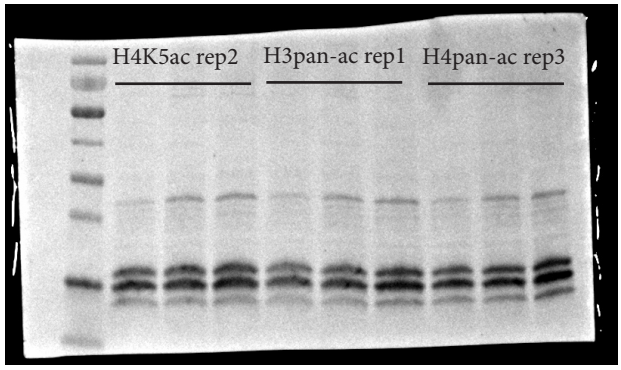
Source Data: Ponceau Red Stains in western blots (Supplementary Fig. 4d)

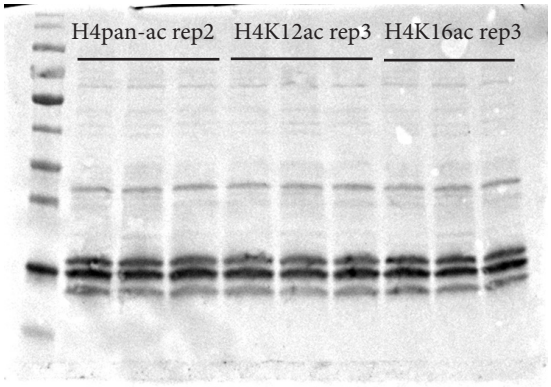


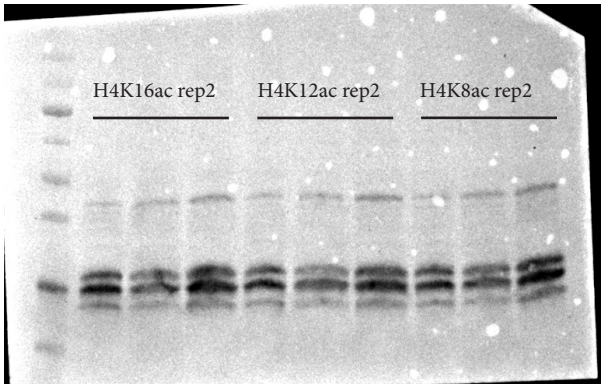


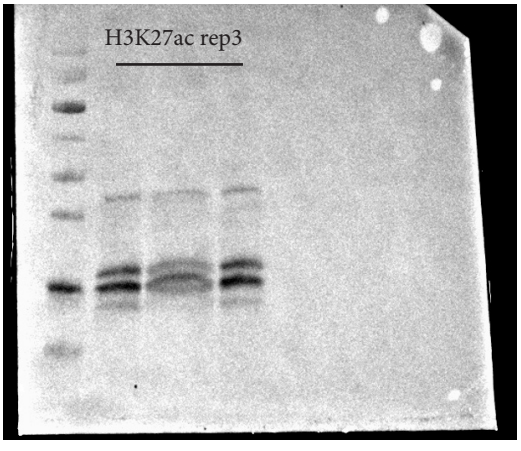




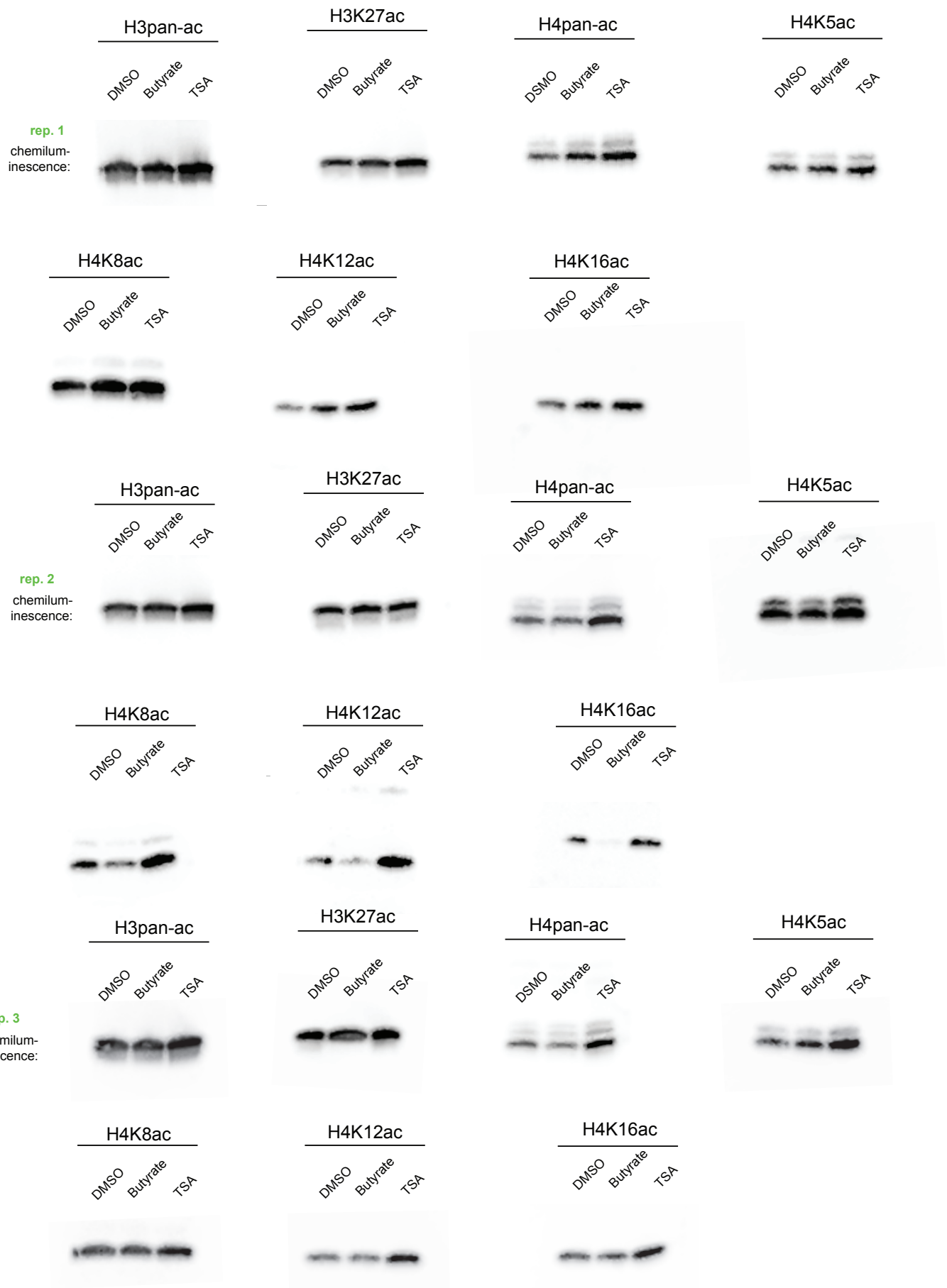








Source Data: Western Blots (Supplementary Fig. 4d). Note, strips of nitrocellulose membrane corresponding to biological replicates of DMSO, Butyrate and TSA were cut after ponceau staining and probed separately for each antibody to reduce primary-Ab incubation volume.



Starvation resistance in the nematode *Pristionchus pacificus* requires a conserved supplementary nuclear receptor

Tobias Theska, Tess Renahan, & Ralf J. Sommer*

Max Planck Institute for Biology, Tübingen,
Department for Integrative Evolutionary Biology,
Max-Planck-Ring 9,
72076 Tübingen, Germany

***Corresponding author**

Ralf J. Sommer (ralf.sommer@tuebingen.mpg.de)

ORCIDs

Tobias Theska - https://orcid.org/0000-0002-0920-4427
Tess Renahan - https://orcid.org/0000-0002-4326-0254
Ralf J. Sommer - https://orcid.org/0000-0003-1503-7749

Keywords

nuclear receptor, phenotypic plasticity, starvation resistance, nematode, *Pristionchus pacificus*

Abstract

Nuclear hormone receptors (NHRs) are a deeply-conserved superfamily of metazoan transcription factors, which fine-tune the expression of their regulatory target genes in response to a plethora of sensory inputs. In nematodes, NHRs underwent an explosive expansion and many species have hundreds of *nhr* genes, most of which remain functionally uncharacterized. However, recent studies elucidated that two sister receptors, *Ppa*-NHR-1 and *Ppa*-NHR-40, are crucial regulators of feeding-structure morphogenesis in the diplogastrid model nematode *Pristionchus pacificus*. In this study, we functionally characterize *Ppa*-NHR-10, the sister paralog of *Ppa*-NHR-1 and *Ppa*-NHR-40, aiming to reveal whether it too regulates aspects of feeding-structure development. We used CRISPR/CAS9-mediated mutagenesis to create knock-out mutations of this receptor and applied a combination of geometric morphometrics and unsupervised clustering to characterize potential mutant phenotypes. However, we found that *Ppa*-NHR-10 does not affect feeding-structures morphogenesis. Instead, multiple RNA-seq experiments revealed that many of the target genes of this receptor are involved in lipid catabolic processes. We hypothesized that their mis-regulation could affect the survival of mutant worms during starvation, where lipid catabolism is often essential. Indeed, using novel survival assays, we found that mutant worms show drastically decreased starvation resistance, both as young adults and as dauer larvae. We also characterized genome-wide changes to the transcriptional landscape in *P. pacificus* when exposed to 24hrs of acute starvation, and found that *Ppa*-NHR-10 partially regulates some of these responses. Taken together, we were able to demonstrate that *Ppa*-NHR-10 is broadly required for starvation resistance and regulates different biological processes than its closest paralogs *Ppa*-NHR-1 and *Ppa*-NHR-40.

Background

Phenotypic plasticity is the ability of a single genotype to produce different phenotypes depending on the environmental cues that the organism receives during development. As such, phenotypic plasticity has re-gained wide-spread attention as a driving force in organismal evolution [1], especially in the context of contemporary debates surrounding a proposed extension of standard evolutionary theory [1-4].

Diplogastrid nematodes such as *Pristionchus pacificus* and its conspecifics are model organisms used to study the molecular mechanisms that govern phenotypic plasticity and the role it plays in phenotypic evolution [4-7]. These worms display a morphological novelty in their mouths that renders them unique amongst roundworms: they possess movable cuticular teeth which allow them to predate on other nematodes (Fig. 1A). This phenotype is plastic; depending on the environment, one of two alternative adult morphologies can be adopted: the predatory "eurystomatous" morph which has two cuticular teeth, or the strictly bacterial-feeding "stenostomatous" morph that has only a single tooth (Fig. 1A) [5,7]. Recent studies utilizing *P. pacificus* began to elucidate the complex genetics behind this polyphenism [4,8-11] and identified a modular gene regulatory network (GRN) which controls the development of plastic feeding structures found in Diplogastridae (Fig. 1B). An upstream module perceives environmental information and relays it to a central plasticity-switch module that controls which of the two alternative feeding morphologies will be developed [4,10,12]. Downstream of the plasticity-switch module is a phenotype-execution module that governs the morphogenesis of the cuticular feeding structures [4]. At the core of this module are two transcription factors, *Ppa*-NHR-1 and *Ppa*-NHR-40, which co-regulate the expression of genes involved in the modification and degradation of the cuticular material that makes up the worm's feeding structures (Fig. 1A,B) [11]. Additionally, recent studies identified mucin-type protein DPY-6 and chitin synthase CHS-1 as members of the phenotype-execution module which participate in the synthesis of the feeding structures [13,14]. However, the transcription factors regulating their expression have not yet been identified.

Another plastic phenotype maintained in *P. pacificus,* and other nematodes such as distantly related "rhabditid" nematode *Caenorhabditis elegans*, is related to their life cycles. Under suboptimal environmental conditions, e.g. overcrowding or food depletion, young larvae can enter an indirect life cycle (Fig. 1C) as dauers, an arrested developmental stage [15,16]. Dauers are specialized non-feeding larvae adapted for stress survival [15,17] and dispersal [18]. Characteristic morphological features of dauers include a thickened body wall cuticle and a closed buccal cavity with a plug clogging the alimentary channel (Fig. 1D) [15]. Thus, dauers cannot feed and they have to generate energy from stored neutral lipids via lipolysis and fatty acid β-oxidation [19-22]. Once environmental conditions turn more favorable, worms can exit dauer and resume development into reproducing adults

(Fig. 1C). Interestingly, like diplogastrid feeding plasticity, various aspects of dauer plasticity are regulated by nuclear receptors (NHRs) [16,23,24], which constitute one of the most important transcription factors classes in animals [25]. They generally share a deeply conserved protein organization and contain a DNA-binding domain (DBD), which usually recognizes specific DNA motifs in the *cis*-regulatory regions of their target genes, and a ligand-binding domain (LBD), which mediates the activity state of the receptor in response to specific ligands [25,26]. Besides their roles in phenotypic plasticity [27], NHRs are known to regulate copious metabolic and physiological functions in nematodes, including immune reactions, detoxification, fatty-acids synthesis and β-oxidation, as well as responses to acute starvation [28-32]. Note that the majority of NHRs in invertebrates have not been deorphanized, leaving many aspects of their molecular function unknown [25-28].

Recent advances in genomics also revealed that, in contrast to other animals, nematodes often possess hundreds of *nhr* genes [11,33]. The overwhelming majority of these, called "supplementary nuclear receptors" (*supnrs*), originated from an explosive burst of duplications of the nuclear receptor HNF4 [34,35]. Another striking observation among distantly-related nematodes with similarly high numbers of NHRs (e.g. *C. elegans* [n=266] and *P. pacificus* [n=254]) showed that the vast majority of NHRs is completely lineage-specific with no one-to-one orthologs in other species (Fig. 1E) [11]. Additionally, the expression of *supnr* genes in the nervous system is highly biased toward sensory neurons, suggesting that the expansion of NHRs, which is particularly blatant in free-living nematodes, could be linked to the worm's complex sensory environments filled with entangled microbial foods [33]. Thus, *supnrs* might act as sensors of environmental food signals or internal metabolic signals, and may adjust the worm's physiological states or behavioral responses according to perceived environmental information [28,30,32,33].

Interestingly, *Ppa*-NHR-1 and *Ppa*-NHR-40, the aforementioned regulators of feeding-structure morphogenesis in *P. pacificus*, are representatives of these nematode-specific *supnrs* (Fig. 1E) [11] and, specifically, two of the very few *supnrs* which are actually conserved across distantly-related species. Although previous studies demonstrated that they are not only deeply conserved, but also co-express in the same pharyngeal tissues that produce the feeding structures of both *P. pacificus* and *C. elegans*, they were found to control feeding-structure development only in diplogastrids [11,36]. This can be explained by the fact that NHR-1 and NHR-40 regulate vastly different pools of regulatory target genes between these two species, and that they control very young (and rapidly evolving) cuticle-modifying enzymes only in *P. pacificus*. Together, these data suggest that these two *supnrs* were likely co-opted into the modular feeding-plasticity GRN of diplogastrids during the evolution of their tooth-like feeding structures [36].

Intriguingly, phylogenetic analyses of all *nhr* genes in *C. elegans* and *P. pacificus* found that another *supnr*, *Ppa*-NHR-10, is the sister paralog of *Ppa*-NHR-1, and that these receptors, together with *Ppa*-NHR-40, form a small cluster of deeply conserved *supnrs* in the otherwise highly species-specific canopy of the *nhr* phylogenetic tree (Fig. 1E) [11,36].

Therefore, we speculated that *Ppa*-NHR-10, too, could have been co-opted as another important regulator of feeding-structure development in *P. pacificus*, possibly even the hitherto unidentified transcription factor of those genes in the phenotype-execution module which are involved in cuticle synthesis (Fig. 1B). Thus, we set out to investigate the biological function of P*pa*-NHR-10, using an integrative approach that combines CRISPR-mediated knockouts of this receptor, geometric morphometric analyses of mutant mouth morphologies, and regulatory-target identification via transcriptomics. This study demonstrates that *Ppa*-NHR-10 is not involved in regulating mouth-form plasticity. Instead, *Ppa*-NHR-10 is required for the survival of acute starvation in young adults and long-term starvation in dauer. We also elucidated genome-wide patterns of transcriptional responses to acute starvation in *P. pacificus* and found that NHR-10's regulatory targets include multiple enzymes involved in lipid catabolism and transport, some of which are starvation-response genes (SRGs).

Results

Ppa-nhr-10 is expressed in the skin, pharyngeal glands, uteri, spermathecae, and rectal glands

Previous studies demonstrated that *Ppa-nhr-1* and *Ppa-nhr-40* are co-expressed in the pharyngeal muscle cells, the tissues that produce the cuticular teeth of *P. pacificus* [11]. Additionally, all of their joint regulatory targets (metallopeptidases, chitinases, and secreted proteins [11,36]) are expressed in a single cell: g1D. This cell is a gland cell of *P. pacificus*, which is morphologically integrated with the dorsal pharynx, where it runs through the dorsal tooth and opens into the mouth lumen [37]. We therefore wanted to determine if *Ppa-nhr-10* is also expressed in these cell types.

We generated two transcriptional reporter lines and found that *Ppa-nhr-10* is expressed in a variety of tissues (Fig. 2). These included the skin of the worms (hypodermis) (Fig. 2A, 2B, and 2C), the anterior and posterior uteri and spermathecae (Fig. 2A, 2B, and 2D), the rectal glands (Fig. 2A), and, expectedly, the pharyngeal glands (Fig. 2A and 2C). While worms consistently showed an expression in all three of the pharyngeal gland cells (g1D, g1VL, g1VR [37]), the relative intensity of the expression signal varied across the two independently acquired reporter lines: in one reporter line the ventral pharyngeal glands (g1V) dominate the signal (Fig. 2A), while in the other the dorsal pharyngeal gland (g1D) reports more strongly (Fig. 2B and 2C). Additionally, one of the lines shows only a weak expression in the rectal glands (Fig. 2A *vs.* 2B). Thus, *Ppa-nhr-10* is expressed in a variety of tissues and is not restricted to the head region.

The fact that *Ppa-nhr-10* is the deeply conserved paralog of *Ppa-nhr-1* and *Ppa-nhr-40*, and that the expression patterns of all three receptors overlap in pharyngeal tissues associated with the morphologically novel teeth, would be consistent with the idea that this *supnr* may be another member of the GRN controlling feeding-structure development in *P. pacificus*. Thus, we set out to characterize whether the knockouts of *Ppa-nhr-10* would lead to any mouth-morphogenesis related mutant phenotypes.

Loss of *Ppa-nhr-10* does not affect stoma morphology and only weakly affects mouth-form plasticity

Using CRISPR/CAS9, we aimed to create mutants of *Ppa-nhr-10* (genome annotation: El Paco v3, gene_ID: PPA01780, chromosome: ChrIII) by introducing small frameshift-causing insertion-deletion (indel) mutations that truncate the nuclear receptor either at the site of the DNA-binding domain (DBD) or the ligand-binding domain (LBD). While homozygous frameshift mutants in the DBD were lethal, we succeeded in generating homozygous mutant lines carrying frameshifts that truncate *Ppa*-NHR-10's LBD (crRNA

sequence: GGCGCGTGGGTTCGGGCGTA). These results indicate that DBD mutants represent lethal *loss-of-function (lof)* alleles of *Ppa-nhr-10*, and that frameshift mutations in its LBD represent *reduction-of-function (rof)* alleles that are viable. These findings are similar to previous observations in *P. pacificus* which found that certain frameshift mutations can escape non-sense mediated decay [14,38], a phenomenon that is currently under investigation in our lab. We chose two of these *reduction-of-function* mutants to establish reference alleles, *tu1654* (strain RS3920) and *tu1655* (strain RS3921), which we used for all subsequent experiments in this study (Fig 3A).

First, we found that both *Ppa-nhr-10* mutants show a detectable decrease in the number of predatory morphs when compared to wild-type worms (Fig. 3B). While this effect is well supported by our data (wild type *vs. Ppa-nhr-10(tu1654)*: $Z = 2.1$, $P = 0.0058$; wild type *vs. Ppa-nhr-10(tu1655)*: $Z = 2.6$, $P = 0.0004$) and consistent across two alleles (*Ppa-nhr-10(tu1654) vs. Ppa-nhr-10(tu1655)*: $Z = 0.05$, $P = 0.4839$), it is only marginal in size (~ 4-5%). This is in stark contrast to the effects observed for *loss-of-function* and *gain-of-function* mutations in *Ppa-nhr-40,* which lead to a complete loss of the predatory or bacterivorous morphs, respectively [9,11]. Given the small effect that *reduction-of-function* mutations of *Ppa-nhr-10* exert on the worm's morph ratio, and taking into account the diversity of factors influencing the mouth-form decision [4,5,12], we conclude that *Ppa*-NHR-10 is unlikely to be a major regulator of genes involved in phenotype execution.

This idea was further supported by our geometric morphometric analysis of 2-dimensional landmark data from 144 individual worms (48 per strain), which showed homogenous and highly overlapping ranges of morphological variation in mouth form across wild-type worms and *Ppa-nhr-10* mutants (Fig. 3C). Subsequent model-based clustering of all individuals in the form dataset revealed that mutant worms could not be classified as such, or differentiated from wild-type worms, based on their mouth morphologies alone (Fig. 3D). Together, these data support the absence of any discernible mutant phenotypes in the mouth of *Ppa-nhr-10* mutants. These findings are in contrast to the phenotypes caused by frameshift mutations in *Ppa-nhr-1,* which lead to structural alterations and gives rise to intermediary mouth morphology combining anatomical features of predators and bacterivores [11].

Thus, the results of our morphological analyses do not provide support for the hypothesis that *Ppa*-NHR-10 is another constituent of the phenotype-execution module of the GRN which governs feeding-structure development in *P. pacificus*. This also indicates that not the entire NHR-1/-10/-40 cluster was co-opted into said functional context during diplogastrid evolution. Intrigued by this observation, we set out to elucidate the biological function of NHR-10 in *P. pacificus,* starting with exploratory RNA-seq experiments.

Exploratory transcriptomics suggests *Ppa*-NHR-10 might be a regulator of lipid catabolism

Since nuclear receptors are transcription factors, it is possible to identify their putative regulatory target genes, and thus infer their potential biological functions, via comparative transcriptomics. We analyzed the transcriptomes of wild-type worms (PS312), and both *Ppa-nhr-10* frame-shift mutant alleles (*tu1654* and *tu1655*). Specifically, we performed RNAseq at two developmental stages (larvae and adults), aiming to identify potential "core targets" of *Ppa*-NHR-10 which we considered to be genes that are consistently mis-regulated in both mutant strains, in at least one of the two developmental stages.

In doing so, we were able to identify 32 potential regulatory target genes, of which 25 were down-regulated and seven up-regulated (Fig. 4). We found 11 genes consistently differentially expressed in mutant larvae and 12 in mutant adults (compared to their wild-type counterparts). The larval targets include multiple down-regulated collagens and up-regulated proteins which are known to be involved in stress responses, like C-type lectins or the UDP-glycosyltransferases (Fig. 4). Interestingly, we also found two proteins, both of them strongly down-regulated, which are predicted to be involved in the binding and transport of lipids: PPA13553 (an ortholog of the *C. elegans* vitellogenin *vit-6*) and PPA05630 (a predicted apolipoprotein). On the other hand, all of the 12 targets we identified in adult mutants were down-regulated and they include a more diverse array of proteins, ranging from a membrane-trafficking protein (PPA08133/*Ppa-scm-1*), to a component of the RNA polymerase II complex (PPA14053/*Ppa-lin-25*), to a chromatin organization modifier (ppa_stranded_DN31598_c1_g4_i2), to one of the TGF-β signaling receptors (PPA15180/*Ppa-daf-4*).

Additionally, we found nine putative targets which are strongly and constitutively mis-regulated across all developmental stages of mutant animals, and the majority of them have predicted functions in lipid catabolism. Briefly, canonical lipid catabolism involves three processes: (I) lipolysis - the cleavage of neutral storage lipids (TAGs) into glycerol and free fatty acids by lipases, (II) activation - the ligation of CoA to free fatty acids by fatty acyl-CoA synthetases (ACSs), and (III) β-oxidation - the breakdown of activated fatty acids into acetyl-CoA molecules. The latter process is jointly performed by four enzyme classes: acyl-CoA dehydrogenases (ACDHs), enoyl-CoA hydratases (ECHs), 3-hydroxyacyl CoA dehydrogenases (3-HACDs), and thiolases [29-31]. Interestingly, the stage-independent putative targets of *Ppa*-NHR-10 include two predicted lipases (PPA07082 and Iso_D.954.1), two predicted ACSs (PPA05769 and PPA13813), as well as a predicted ECH (PPA40903). Furthermore, these genes are among the most strongly affected by the mutations of *Ppa-nhr-10*, which is reflected by the fact that they show some of the largest log2-fold changes (L2FC) among all putative regulatory targets (Fig. 4).

Taken together, our exploratory transcriptome and differential-expression analyses suggest that *Ppa*-NHR-10 might be a regulator of body fat in *P. pacificus*. This is evident in the observation that the expression of genes predicted to be involved in the major

processes of lipid catabolism and transport is negatively affected in *Ppa-nhr-10* mutants. Given that lipid metabolism and life span are tightly linked in all animals [39], and knowing that functional interferences with any of the lipid catabolic processes in nematodes can impair their survival capabilities across various environmental conditions [29,30], we decided to quantify the survival capabilities of *Ppa-nhr-10* mutants in different conditions.

Survival analyses reveal that *Ppa-nhr-10* is required for survival of acute starvation in young adults

In *C. elegans*, interference with the major lipid catabolic pathways can lead to altered life spans and/or rates of survival over time. While mutations in some genes of these pathways constitutively impair the worm's survival capabilities, irrespective of environmental factors, others are known to cause such impairments only in certain conditions like starvation [29-31,39]. Therefore, we examined the survival capabilities of wild-type *P. pacificus* and *Ppa-nhr-10* mutants in two different conditions: well-fed and acutely-starved. To achieve this, we transferred young adult hermaphrodites into small cell culture flasks which either contained a standard liquid culture (LC) medium rich in OP50 or a LC medium that contained no food, and scored their survival probabilities over time until all worms in the flasks were dead.

First, we found that the obtained time-to-death data from the well-fed condition is compatible with the hypothesis that there are no differences in the overall survival capabilities of *Ppa-nhr-10* mutants and wild-type worms (Fig. 5A and 5C). Neither the global Kaplan-Meier (K-M) survival analysis of well-fed worms (Fig. 5C: $\chi^2 = 5.7$, $P = 0.060431$), nor the pairwise comparisons of strain-specific survival curves via a log-rank test, revealed any significant differences in survival (Fig. 5A). Yet, hazard ratios were estimated to be slightly elevated in the mutant strains (1.15 for *Ppa-nhr-10(tu1654)* and 1.13 for *Ppa-nhr-10(tu1655)*). These hazard ratios describe the risk-of-death that these mutants face at any given point of time during the experiment, relative to the wild-type level of risk. A hazard ratio of 1.0 thus indicates that mutant risks-of-death are equivalent to the wild type, and hazard ratios higher or lower than 1.0 indicate relatively higher or lower risks-of-death in mutant animals, respectively. While the hazard ratios mentioned above indicate that *Ppa-nhr-10* mutants may show a ~15% increase in the relative risk-of-death at any given timepoint in the well-fed condition, these estimates were still found to be compatible with the null hypothesis given their associated uncertainty (Fig. 5C).

Notably, we found the exact opposite to be true for the starvation experiments: the global K-M analysis indicated stark differences in the overall survival probabilities among starving worms (Fig. 5C: $\chi^2 = 175$, $P = 1.51e\text{-}38$) and the log-rank pairwise comparisons revealed that the survival curves of both *Ppa-nhr-10* mutants, while being virtually identical to each other (Fig. 5B: $P = 1.0$), differed markedly from the wild-type survival

curve (Fig. 5B: P < 2e-16). Additionally, we found highly elevated hazard ratios for *Ppa-nhr-10* mutants compared to wild-type *P. pacificus* (Fig. 5C: 2.23 for *Ppa-nhr-10(tu1654)* and 2.12 for *Ppa-nhr-10(tu1655)*). These estimates indicate that for every wild-type worm that dies at any given point of time during the starvation experiment, more than two mutant worms will face the same fate. These findings strongly support decreased survival rates in mutant worms (P[*Ppa-nhr-10(tu1654)*]= 4.79e-30 and P[*Ppa-nhr-10(tu1655)*]= 3.91e-27) (Fig. 5C). While our condition-specific survival experiments revealed that acute starvation induces highly elevated rates of premature deaths in adult *Ppa-nhr-10* mutants, it is important to note that the original exploratory RNA-seq data that inspired these survival experiments were derived from well-fed worms. Thus, we wondered whether any of *Ppa*-NHR-10's targets are starvation sensitive and therefore performed condition-specific RNA-seq experiments on well-fed and acutely-starved wild-type and mutant worms in the LC set-up.

Starvation induces substantial transcriptional responses in genes predicted to be involved in metabolism, stress responses, and translation

To begin, we aimed to elucidate genome-wide changes of the transcriptional landscape in wild-type worms which faced 24hrs of acute starvation. Using the same differential expression pipeline as described above, we were able to identify a total of 1,007 genes showing a response to starvation (Fig. 6A). Of these, 294 genes were notably up-regulated (cluster I in Fig. 6) and 713 notably down-regulated (cluster II in Fig. 6). Thus, most of the starvation-responsive genes (SRGs) strongly decrease expression levels under starvation compared to when food is abundantly available.

We found that phylogenetically old genes, rather than diplogastrid-specific orphans, are overrepresented in both SRG clusters (Fig. 6B). Yet, while they do constitute the minority of SRGs, there are still hundreds of orphans responding to starvation (Fig. 6B). Additionally, we asked whether recently identified co-expression modules [40] were overrepresented among the SRGs. Indeed, genes belonging to various neuronal and oogenesis-related co-expression modules were strongly enriched amongst the up-regulated SRGs. On the other hand, members of intestinal and hypodermal co-expression modules were strongly enriched in the down-regulated SRG cluster (Fig. 6B).

Amongst the up-regulated SRGs, there was a clear bias towards genes encoding proteins that are critical for translational processes and ribosome assembly (Fig. 6D), including large ribosomal-subunit proteins, elongation factors, and tRNA synthetases (Fig. 6C). In contrast, genes which are down-regulated in response to acute starvation seem to control genes that predominantly encode proteins involved in diverse metabolic pathways and stress responses (Fig. 6D). The strongest detected metabolic signals originated from genes encoding proteins involved in lipid metabolism, such as fatty-acid elongases and

desaturases, lipases, and vitellogenins (Fig. 6C,D). Genes involved in the synthesis and breakdown of amino acids were overrepresented among down-regulated SRGs, too (Fig. 6D). The detected stress-response signals in this cluster originated from C-type lectins, p450 cytochromes, and lysozymes (Fig. 6C and 6D).

Condition-specific transcriptome analyses indicate that *Ppa*-NHR-10 contributes to the regulation of starvation responses

With the reference knowledge of genome-wide patterns of starvation responses as described above, we next set out to analyze the transcriptomes of starved *Ppa-nhr-10* mutants to see whether any of them are members of the identified SRGs. The condition-specific transcriptome analysis revealed an updated list of 37 putative *Ppa*-NHR-10 targets, of which 32 were down-regulated and five up-regulated (Fig. 7).

The comparison of gene expression in well-fed and starved *Ppa-nhr-10* mutant animals indicates that the original exploratory RNA-seq experiments missed a large number of putative targets (28 out of 37). Many of the newly identified targets have a wide variety of predicted functions which are not easily grouped into overarching processes (Fig. 7). Yet, we recovered nine of the putative target genes that were found in the exploratory analysis (Fig. 4 and Fig. 7), including the TGF-β receptor PPA15180/*Ppa-daf-4*, the p450 cytochrome PPA25752/*Ppa-cyp-33B1*, and the two lipid catabolic enzymes PPA05769/*Ppa-acs-14* and PPA40903 (Fig. 4 and Fig. 7). This confirms that, while the initial RNA-seq approach missed out on starvation-specific effects, it still identified some of the targets which likely convey starvation resistance.

More importantly, the condition-specific RNA-seq approach added to the initial observation that the putative targets tend to be ones that are involved in the transport and breakdown of lipids and fatty acids (Fig. 7). PPA33234/*Ppa-strl-1* is a lipid-binding protein predicted to be mitochondrial and involved in the transport of cholesterol for steroid biosynthesis. PPA27650 is predicted to be involved in intermembrane transport of lipids as part of glycerolipid synthesis. Genes that encode important players in the β-oxidation of fatty acids were recovered, too: PPA05769/*Ppa-acs-14* is a fatty acyl-CoA synthetase, the type of enzyme that "activates" free fatty acids for β-oxidation. Additionally, the gene *ppa_stranded_DN29959_c0_g2_i3* encodes an acyl-CoA dehydrogenase (ACDH) and PPA40903 an enoyl-CoA hydratase (ECH). These two enzymes catalyze the first two of the four steps of mitochondrial β-oxidation [29-31].

Lastly, we wondered what the gene regulatory basis of the starvation-survival phenotype in *Ppa-nhr-10* mutants looks like and anticipated two different scenarios. One, that NHR-10's regulatory targets would be members of genetic networks which are not part of the starvation response machinery (Fig. 6), but might, when mis-regulated, lead to decreased survival capabilities in mutant worms that face starvation. Alternatively, *Ppa-*

NHR-10's regulatory targets could actually constitute members of the SRGs. In this latter scenario, the impaired survival capabilities of starving mutants could be explained by malfunctions in the starvation-response itself. Our analyses suggested a combination of both: while most of *Ppa*-NHR-10's putative targets do not respond to starvation, we found four SRGs among them (Fig. 7). These included the enoyl-CoA hydratase and p450 cytochrome which were already found in our exploratory RNA-seq analysis (Fig. 7). Taken together, these findings suggest that *Ppa*-NHR-10 controls some of the worm's transcriptional responses to acute starvation.

Dauer assays demonstrate that mutations in *Ppa-nhr-10* also affect dauer biology

Finally, we wondered whether *Ppa*-NHR-10 might play a more general role in starvation survival (beyond its requirement for starvation resistance in young adults) and investigated the effects of *Ppa-nhr-10* mutations on the development of dauer larvae (Fig. 1B). For that, we quantified three different aspects of dauer biology using a robust combination of assays: i) the number of worms which entered the dauer stage due to overcrowding, ii) the time it took the worms to exit dauer and resume development when re-introduced to food, and iii) the survival of dauers over extended periods of time (Fig. 8A).

Interestingly, for dauer entry, we found a mutant effect in only one of our *Ppa-nhr-10* alleles (*tu1654*). Across culture conditions, worms carrying this allele showed a significantly decreased number of dauer larvae in the overall population when compared to the wild-type animals and members of the second mutant strain (Fig. 8B). However, such allele-specific differences could not be observed in the remainder of our dauer assays (Fig. 8C and 8D). The dauer-exit data demonstrated that mutant animals consistently needed more time to exit the dauer stage and resume normal development. While all worms eventually completed the dauer exit (and developed into fully viable adults), it took *Ppa-nhr-10* mutants twice as long as wild-type worms to do so (Fig. 8C). Lastly, we found impaired long-term survival capabilities in mutant dauer larvae (Fig. 8D). For example, after spending one week in dauer, the estimated overall survival rates of wild-type and mutant dauers were indistinguishable. In contrast, after a second week, wild-type worms showed survival rates which were indistinguishable from the previous week, whereas mutant worms died off rapidly, leaving almost no survivors (Fig. 8D). This trend continued into the third week, where mutant dauers were essentially absent, while a fifth of wild-type dauers were still alive. Thus, these assays revealed that all investigated aspects of dauer biology were negatively affected in *Ppa-nhr-10* mutants. This observation confirmed the hypothesis we conceived based on the combined results from our survival analysis and our condition-specific RNA-seq data, and demonstrates that *Ppa*-NHR-10 is

indeed required in the broader context of starvation resistances throughout the *P. pacificus* life cycle, including the ecologically relevant dauer stage [18,41,42].

Discussion

In this study, we aimed to reveal whether *Ppa*-NHR-10 is a regulator of feeding-structure development in *P. pacificus* and, thus, address whether a small but complete cluster of three closely-related and deeply-conserved supplementary nuclear receptors was functionally co-opted into the gene regulatory network that governs feeding-structure plasticity in *Pristionchus* nematodes [11,36]. We were able to demonstrate that *Ppa*-NHR-10 is indeed, amongst other tissues like the skin, expressed in the relevant cell types which are part of the morphological novelties of diplogastrid mouths (e.g., the g1D pharyngeal gland cell associated with the dorsal tooth). This is intriguing, as *Ppa*-NHR-1 itself and virtually all of the regulatory targets it shares with *Ppa*-NHR-40 are co-expressed in the same cell [11]. Yet, unlike mutations in the other two receptors, frameshift mutations in *Ppa*-NHR-10 neither disrupted the worm's ability to develop into predators, nor did it lead to aberrant mouth morphologies. These results indicate that only *Ppa*-NHR-1 and *Ppa*-NHR-40, but not the complete NHR-1/-10/-40 cluster, were co-opted when the plasticity GRN was assembled during the evolution of mouth morphological novelty in diplogastrids [36]. This also demonstrates that the genetic basis for the evolution of plastic traits and novel structures is more complicated than often presumed, and that paralogs - at least when it comes to nematode *supnrs* - might not necessarily be adequate predictors of the biological functions of genes.

Through a combination of exploratory and condition-specific mRNA-seq experiments, we were able to identify *Ppa*-NHR-10's putative regulatory targets. These encompass a broad variety of proteins including many metabolic enzymes and multiple signaling elements (Figs. 4, 7). Interestingly, this pattern is very different from the regulatory target patterns of its paralogs *Ppa*-NHR-1 and *Ppa*-NHR-40: most of their targets belong to a single class of metallopeptidases (astacins) that is crucial for cuticle degradation and the vast majority of them are located on the X-chromosome [11,36]. This is the same chromosome that also harbors *Ppa-nhr-1* and *Ppa-nhr-40*, as well as the multi-gene locus that regulates the plastic switch decision [8,10,43]. This pattern, again, does not hold true for the loci harboring *Ppa*-NHR-10's targets; it does not display a bias towards one particular chromosome, let alone chromosome X (Sup. Tab. 1). In contrast to *Ppa*-NHR-1's and *Ppa*-NHR-40's biases towards regulating classes of specific enzymes [11,36], *Ppa*-NHR-10 seems to regulate aspects of the activation and β-oxidation of fatty acids (Fig. 4 and Fig. 7), processes which are known to be regulated by a variety of NHRs in *C. elegans*, too [28-31]. The differences in the regulatory "behaviors" among these closely-related NHRs demonstrate why, in this particular case, *Ppa*-NHR-10's paralogs were inadequate predictors of its function.

We also elucidated genome-wide transcriptional responses to 24hrs of acute starvation in young adult hermaphrodites of *P. pacificus*. When focusing on the broad-scale

patterns in this data, three interesting observations emerge: First, during starvation, *P. pacificus* increases the expression of numerous genes which are critical elements of the translational machinery, and it decreases the expression of metabolic pathway genes, especially those affecting lipid transport, and the desaturation and elongation of fatty acids for fat storage purposes (Fig. 6C and 6D). Thus, the biological pathways affected by acute starvation are consistent between *P. pacificus* and *C. elegans* [17,44-46], indicating that the underlying starvation-response mechanisms are deeply conserved; an observation that is also reflected in the fact that most of the SRGs are phylogenetically old (Fig. 6B). Yet differences exist in the details: when *C. elegans* starves, it does not change the expression levels of ribosome-biogenesis genes, but it decreases the translation of their transcripts [46]. We speculate that *P. pacificus* might employ a slightly different strategy, in which it increases the expression of translation-related genes without increasing the translation of their transcripts. This way, the worms would not invest large amounts of energy into the protein synthesis [17], while maintaining the ability to quickly do so if food becomes available again. Future proteomic studies will be able to directly test this hypothesis and study the evolution of these molecular starvation-response strategies.

Second, the majority of up-regulated genes are associated with the nervous system and oogenesis, while the vast majority of the down-regulated genes are associated with the metabolic hub of nematodes, the intestine (Fig. 6B). Additionally, we found numerous down-regulated genes to be members of a co-expression module which is associated with the hypodermis (Fig. 6B "cuticle 1" [40]), the site of long-term lipid storage [30]. Together with the observation that lipid anabolism is down-regulated during starvation, these data could suggest that *P. pacificus*, when facing acute starvation as adults, might undergo drastic changes to their neural regulatory landscape which may, in turn, lead the worms to switch strategies from synthesizing and storing lipids to maintain their soma, to investing into egg production and ensuring population survival. This would be in line with recent predictions concerning neural regulation of body fat during starvation in *C. elegans* [30].

Third, a variety of stress responses are down-regulated in the absence of bacterial food. This includes multiple proteins predicted to be involved in the detoxification of xenobiotic compounds (p450 cytochromes), the immune response (C-type lectins), and the defense against microbial pathogens (lysozymes) [47] (Fig. 6D). This corroborates an interesting observation: the uracil-auxotrophic OP50 strain of *E. coli*, the standard laboratory food source for nematodes [48], might be slightly toxic to *P. pacificus* [49]. Upon ingesting these bacteria, the worms may have to activate some of these stress response pathways in order to be able to sustainably live off of them as a diet. Taken together, this newly-established reference knowledge of responses to acute starvation complements recent nutrigenomic advances [40] in *P. pacificus*, and it will inform future studies that may aim to evaluate the effects of different diets on the transcriptional landscape of the worm.

The present study also demonstrates that *Ppa*-NHR-10 is broadly required for starvation resistance in *P. pacificus* (Fig. 5 and Fig. 8) and that some of its regulatory targets are starvation-responsive (Fig. 7). This indicates that *Ppa*-NHR-10 is one of the probably many transcriptional regulators that contribute to the wild-type starvation response. Interestingly, the expression levels of the *Ppa-nhr-10* gene itself do not change during acute starvation (L2FC = -0.10894 and s = 0.79898), suggesting that *Ppa*-NHR-10 is already present in the worm's cells before the starvation signal is sensed, and not expressed in response to starvation. Future studies may identify the ligand of this receptor and ChIPseq experiments could be used to determine its target DNA motif. This will lead to a better understanding of *Ppa*-NHR-10's role in the starvation response, and elucidate how the food-depletion signal is detected and translated into a biological response via this *supnr*.

Lastly, we found that the loss of *Ppa*-NHR-10 negatively affected multiple aspects of dauer biology, using a unique method combining several assays. This allowed us to investigate three pertinent dauer characteristics - entry, exit, and survival - from the same initial population of dauers in conditions similar to their ecology. In its natural environment, *Pristionchus* is found primarily as dauers in soil and on its invertebrate hosts, predominantly beetles [7,18]. The ability to enter dauer is vital to ensure success in the wild, as diminishing food resources are a common occurrence. Worms that do not enter dauer under unfavorable conditions risk preserving their populations. In addition to persevering through unpredictable food availability, *Pristionchus* faces interspecific competition [41,42]. The speed of dauer exit may determine the success of a population if cohabitating with another nematode species, as first to the feast may establish, repopulate, and consume all accessible bacteria [42]. Long-term survival in dauer provides protection during extended periods of stressful conditions [18-22,24], and may play a role in evading competition while waiting for microbial populations to reestablish [42]. Therefore, starvation resistance is a trait that is crucial for the fitness of these worms in the wild: adults need to adjust to starvation and invest into the production of viable eggs, allowing future generations of worms to persist; dauers need to be able to survive over extended periods of time, allowing them to disperse and establish new populations of reproducing worms under suitable conditions.

A current limitation of our understanding of *Ppa-nhr-10* is rooted in the fact that we were unable to obtain viable homozygous *loss-of-function (lof)* alleles via frameshifts in its DBD (indicating lethality), and that the LBD-domain mutants likely represent *reduction-of-function (rof)* alleles. It is currently unclear how these two types of mutations compare in terms of their functional impact on nematode *supnrs*. Yet, frameshift mutations in the DBD and LBD of *Ppa-nhr-1* both had *loss-of-function* effects [11]. Regardless, a *reduction-of-function* could explain why we found inconsistent (that is allele-specific) effects of *Ppa-nhr-10* mutations on dauer entry. Alternatively, these could also be explained by failure to statistically detect the effect on dauer entry in *Ppa-nhr-10(tu1655)* due to a lack of power or

sampling biases. This would be in line with the fact that we did not observe any other allele-specific effects in the remainder of experiments and assays throughout our study.

In summary, we demonstrated that the supplementary nuclear receptor *Ppa*-NHR-10 contributes to the transcriptional regulation of starvation responses in *P. pacificus*, and that it is required for starvation resistance in adults and dauer larvae. Based on our expression pattern analysis and transcriptome data, we speculate that this is dependent on fatty acid β-oxidaton in the worms' skin - the major site of lipid metabolism [30,31]. Future molecular genetic investigations will be able to directly test this hypothesis based on the candidate targets we identified in our experiments, ultimately expanding our understanding of the physiological mechanisms that underlie starvation resistance in these nematodes.

Conclusion

In this study, we investigated the biological function of the nuclear receptor *Ppa*-NHR-10, the sister receptor of the two known regulators of feeding-structure morphogenesis in the diplogastrid nematodes (*Ppa*-NHR-1 and *Ppa*-NHR-40). We found that *Ppa*-NHR-10 is not involved in feeding-structure development and that it regulates entirely different types of targets genes than its sister paralogs [11,36]. We showed that it contributes to the genome-wide transcriptional responses to acute starvation, and revealed that both dauer larvae (a specialized dispersal stage) and adults of *P. pacificus* are less resistant to starvation if this receptor is experimentally rendered non-functional. Thus, our findings demonstrate that these worms require *Ppa*-NHR-10 to survive some of the harsh conditions they frequently encounter in the wild [15,18-22,41,42].

Material & methods

Nematode husbandry

Standard protocols for the maintenance of laboratory cultures of rhabditid nematodes [48] were followed. Worms were cultured on 6cm Petri dishes with nematode growth medium (NGM). 300µl of *Escherichia coli* (OP50) were provided as a food source. Culture plates were stored at 20°C.

Light microscopy with differential interference contrast (DIC) and morph ratios

For microscopy, all specimens were mounted on object slides with 5% Noble Agar pads which contained the sedative sodium azide (0.5% NaN_3) and subsequently examined using a Zeiss Axio Imager.Z1 microscope with a Zeiss Plan-Apochromate 100 x 1.4 DIC objective. Image stacks were taken using a monochromatic Zeiss Axiocam 506 CCD camera. The Zen 2 Pro Software (Carl Zeiss Microscopy GmbH, 2011; version 2.0.14283.302; 64-bit) for digital microscopic analysis and image acquisition.

Relative ratios of eurystomaous (Eu) and stenostomatous (St) morphologies in each strain were scored using the aforementioned microscopy settings. Worms that had a large right ventrosublateral tooth, a hook-shaped dorsal tooth, and a promesostegostom whose anterior tips were clearly posterior to those of the gymnostom were classified as Eu morphs; animals that did not meet these combined characteristics were scored as St morphs. Statistical differences in morph ratios were assessed via a PERMANOVA (see "Dauer assays" section of the methods for details).

CRISPR/CAS9 mutagenesis

We used CRISPR/CAS9 to create knock-out mutations of *Ppa-nhr-10*, specifically targeting their DNA- and ligand-binding domains. Standard mutagenesis protocols for model nematodes [50,51] were followed. Target-specific CRISPR RNAs (crRNAs) were generated

based on the publicly available genome sequence of *P. pacificus* (www.pristionchus.org; El Paco ver. 3). CAS9 protein, universal trans-acting CRISPR RNA (tracrRNA), and target-specific crRNAs were ordered from IDT. Ribonucleoprotein complexes (RNPs) were created by combining 0.5μl of CAS9 nuclease (10μg/ml stock) with 5μl of tracrRNA (0.4μg/μl stock) and 2.8μl of crRNA (0.4μg/μl stock). This mix was incubated at 37°C for 15min and subsequently cooled down to room temperature. The RNP mix was diluted by adding enough Tris-EDTA buffer to bring the total volume of the injection mix to 20μl. In order to avoid clogging in the injection needles, the final mix was centrifuged at 14,000rpm for 2min. Injections were conducted with an Eppendorf FemtoJet microinjector at 400x magnification under an inverse Zeiss AxioVert microscope that was equipped with a Zeiss Plan-Apochromat 40x DIC objective and coupled to an Eppendorf TransferMan micromanipulator. Injections were aimed at the distal tip of the hermaphrodite gonad in young adults; animals were punctured in a 45° angle to the anterior-posterior body axis and RNP complexes were injected into the gonadal rachis. Injected P0 animals were isolated onto separate NGM plates and allowed to lay eggs for 24h at 20°C. After that, P0s were removed from the plates and F1 animals were left to grow. Upon reaching maturity, F1s were singled out to fresh plates, allowed to lay eggs for 24h and subsequently used for genotyping. DNA of candidate F1 animals was recovered via single worm lysis (SWL), target sites were PCR amplified (using the Taq PCR Master Mix produced by Qiagen), and candidate mutations were identified using Sanger sequencing (performed by Azenta Life Sciences). The F2 offspring of candidate F1-heterozygotes were genotyped in a similar fashion. Multiple homozygous mutant lines were isolated, including frameshift mutants with premature stop codons. Backups of the original homozygous mutant strains were frozen and stored at -80°C.

Transgenesis

To visualize which cells actively express *Ppa-nhr-10*, we generated transcriptional reporter lines. We PCR-amplified a 1.371 bp long putative promoter sequence, ending right before the start codon of *Ppa-nhr-10*, and subsequently cloned it into a pUC19 vector that contained a codon-optimized GFP with *P. pacificus* introns and the *Ppa-rpl-23* 3'UTR [52]. We generally used the NEBuilder HiFi DNA Assembly Kit for Gibson assemblies, *E. coli* (NEB 5-alpha) for bacterial transformation, and InVitrogen's PureLink HQ Mini Plasmid kit for plasmid extraction from 3ml bacterial overnight cultures. The *Ppa-nhr-10p::GFP* construct was linearized with SphI-HF (New England BioLabs) according to manufacturer's instructions. 10 ng/μl of our linearized plasmid were injected into the gonad of the wild-type strain of *P. pacificus* (PS312), together with 60 ng/μl of SphI-linearized genomic carrier DNA and 10 ng/μl of SphI-linearized co-injection marker *Ppa-egl-20p::RFP* [52,53]. We confirmed that expression patterns were consistent across two

independent reporter lines (Fig. 2A and 2B). Fluorescent worms were imaged using a Leica SP8 confocal laser-scanning microscope (cLSM).

Exploratory RNA-seq experiments - sequencing mutant larvae and adults

To facilitate exploratory RNA-seq analysis on our *Ppa-nhr-10* mutants comparable to the RNA-seq experiments our lab previously conducted on *Ppa-nhr-1* and *Ppa-nhr-40* mutants, we adopted (with minor modifications) the sample collection and sequencing approaches that were used by Sieriebriennikov *et al.* [11]. Animals were collected for RNA extraction at 24h (mostly J2 and a few J3 larvae), 50h (J3 and J4 larvae), and 70h (young adults) after bleach-isolation of eggs. The presence and expected distribution of the desired stages at each time point was verified prior to sample collection by screening 50 specimens on each plate under a dissecting microscope. Total RNA was extracted using the Direct-Zol RNA Mini prep kit (Zymo Research) according to the instructions provided by the manufacturer. Following Sieriebriennikov *et al.*'s approach [11], we combined 500ng of RNA that was extracted at the 24h time point with 500ng of RNA extracted at the 50h time point. This sample type thus contained 1μg RNA extracted from J2-4 larvae. On the other hand, we gathered 1μg of RNA extracted at the 70h time point (adult worms) and did not mix it with RNA extracted from any other time point. Thus, we ended up with two sample types ("larval" and "adult"), which we shipped to Novogene for library preparation and mRNA sequencing. Libraries were sequenced as 150bp paired end reads on Illumina's NovaSeq6000 platform. We sequenced the mRNA of three strains: the wild-type *P. pacificus* strain (PS312), and two independently acquired frameshift mutants of *Ppa-nhr-10* (*tu1654* and *tu1655*). Two biological replicates were collected for the larval and adult stages of each of these three strains.

Condition-specific RNA-seq experiments - sequencing starving worms

We also performed condition-specific RNA-seq experiments, in order to identify *Ppa*-NHR-10's regulatory targets in the same conditions in which we identified the survival phenotype of mutant animals. Additionally, we aimed to investigate the wild-type starvation response of *P. pacificus*. Therefore, using our starvation assay set-up (see "Starvation assays and survival analysis"), we sequenced mRNA extracted from young adult hermaphrodites which have been well-fed or acutely-starved for 24 hours (in liquid cultures). Just like for the exploratory RNA-seq experiments, we used the wild-type strain (PS312) of *P. pacificus* and the two reference *Ppa-nhr-10* alleles (*tu1654* and *tu1655*) for these condition-specific RNA-seq experiments. After the liquid culture treatments, we generally washed the worms three times with 0.5% Triton X-100/PBS and three times with PBS through a 20μm nylon filter net, and subsequently preformed total-RNA extractions with the Direct-Zol RNA Mini prep kit (Zymo Research), following manufacturer instructions. For each sample, ~1μg of total RNA was extracted and shipped to Novogene for library

preparation and mRNA sequencing. As before, libraries were sequenced as 150bp paired end reads on Illumina's NovaSeq6000 platform. This time, we sequenced three biological replicates for each the three strains in each of the two conditions.

Quantification of transcripts and differential expression analysis (DEA)

Transcript abundances were estimated from raw read files relative to the *P. pacificus* reference transcriptome using SALMON (ver. 1.5.2) [54]. Using the entire *P. pacificus* reference genome as a decoy, we built a decoy-aware transcriptome and indexed it for subsequent read mapping using an auxiliary *k*-mer hash (k=31). Reads were quantified against this index using Salmon's *quant* command by running the program in mapping-based mode with selective alignment as mapping strategy [55]. We allowed Salmon to infer the library type, and to learn and correct for fragment-level GC-biases and sequence-specific biases. Differential expression analyses (DEA) were carried out in R using BIOCONDUCTOR (ver. 3.17) [56], TXIMPORT (ver. 1.28.0) [57], and DESEQ2 (ver. 1.40.1) [58]. Transcript-level abundance estimates generated by SALMON were summarized into gene-level count matrices using the *tximport* function. A DESeq data set was created with the *DESeqDataSetFromTximport* function and pre-filtered for transcripts which had at least ten counts across all samples. Stage-specific (exploratory analysis) and condition-specific mutation effects (condition-specific analysis) on transcript abundances were modeled with the *DESeq* function [57,58]. Obtained maximum likelihood estimates (MLE) of the log2-fold changes (L2FC) for each gene were shrunken with the *lfcShrink* function using the adaptive shrinkage estimator from the ASHR package [59]. Resulting (shrunken) minimum mean squared error (MMSE) estimates of the L2FC were reported and used to rank and visualize candidate targets. Wondering whether there is sufficient evidence in the data that a given gene is indeed at least 50% up- or down-regulated, and whether the direction of this effect was correctly estimated, we always included a L2FC-threshold of > 0.585 (absolute value) and a *s*-value threshold of < 0.005 into our hypothesis tests.

Quantification of morphological differences in mutant animals

We quantified morphological differences in the mouths of wild-type and mutant worms, using our recently published protocol [60] which combines 2-dimensional landmark-based geometric morphometrics (GMM) with unsupervised clustering.

For this study, we obtained image stacks of the nematode mouths in lateral position and recorded the X and Y coordinates of 18 fixed landmarks which capture the overall form of the mouth using FIJI (ver. 2.1.0) [61]. All steps of the GMM analysis were performed in R (ver. 4.3.0) [62] using a combination of the GEOMORPH (ver. 4.0.5) [63-65], MORPHO (ver. 2.11) [66,67], and MCLUST (ver. 6.0.0.) packages [68,69]. We originally generated landmark data for 150 animals (50 per strain; three strains: PS312, *Ppa-nhr-10(tu1654)*, *Ppa-nhr-10(tu1655)*), performed a GPA with the *gpagen* function of

GEOMORPH, and subsequently checked for Procrustes distance outliers in our original shape data. Outliers were identified separately for each of the strains of interest, using the *plotOutliers* function. Six extreme outliers were detected and removed. The outlier-corrected data set (n=144) contained landmark configurations of 48 animals per strain. GPA was performed on the outlier-corrected data to obtain a shape matrix, to which we appended the log-transformed centroid size in order to generate a form data set *sensu* Mitteroecker *et al.* [70]. Subsequently, we ran a principal component analysis (PCA) on the form data set to visualize potential morphological differences among wild-type and mutant worms in a form space. Lastly, we assessed whether mutant worms could be classified as such (and distinguished from wild-type worms) based on their mouth morphology. For this, we used model-based clustering via the *Mclust* function. Only "meaningful" principal components (mPCs) of form variation (identified with MORPHO's *getMeaningfulPCs* function) were used as input variables for clustering, in order to avoid overparameterization [60]. Clustering results were visualized by coloring each specimen according to cluster membership.

Starvation assays and survival analysis

Young adult hermaphrodites (day one of adulthood) were washed off of multiple agar plates with 0.5% Triton-X100/PBS and a lazy-L scraper. The collected worm solutions were filtered through a 20 μm nylon net filter and washed three times with 1ml 0.5% Triton-X100/PBS and, subsequently, three times with 1ml PBS. The bacteria-free worm pellets were resuspended in 250μl S-medium without cholesterol and added to empty agarose plates. The S-medium was allowed to be absorbed by the agarose and the clean worms were allowed to recover for one hour. During the recovery time, liquid cultures were set up for the experiments. We used small cell culture flasks (50ml) for our assays and added a total amount of 2.5ml liquid culture (LC) to each of them. The LCs for the acute starvation assays consisted only of 2.5ml S-medium (without cholesterol). The control LC with plenty of food were set up as follows: for each 2.5ml of final LC, we centrifuged 10ml of fresh bacterial stock solution (*E. coli* OP50, OD_{600} = 0.5) for 30min at 4250rpm (4°C). The bacterial pellets were resuspended in the according amount (2.5ml per assay flask) of S-medium without liquid culture and added to the cell culture flasks. To prevent unwanted microbial contaminations, we added 1μl of nystatin and 1.2μl of streptomycin to each assay flask, irrespective of whether it contained OP50 or not. Finally, we picked the young hermaphrodites which recovered from the washing step and added them to the assay flasks (~50 worms per flask). All assay flasks were kept in a shaking incubator set to 20°C and 130rpm and the ratio of alive worms was scored every 24hrs until all worms in the flask were dead. Media were replaced every third day in order to limit bacterial contaminations and remove offspring animals produced in the well-fed conditions. Overall differences in the survival curves within each of the two treatments were

estimated using the Kaplan-Meier (KM) approach [71]. Pairwise comparisons among individual survival curves were based on the KM estimator and performed using the log-rank test [72]. Additionally, we estimated hazard rates for each of the strains in both experiments via Cox proportional hazard regression modeling [71,72]. Effect sizes were calculated as hazard ratios which compare the hazard rates of *Ppa-nhr-10* mutants (*tu1654* and *tu1655*) to the wild-type hazard rate. *P*-values were corrected for false discoveries using the Bonferroni method. All survival analyses were carried out in R (ver. 4.3.0) [62] using the SURVIVAL (ver. 3.5-5), GGSURVFIT (ver. 0.3.0), SURVMINER (ver. 0.4.9), and GTSUMMARY (ver. 1.7.1) packages [73-76]. All survival experiments were repeated four times (N=4) independently with all three strains.

Pfam domain prediction, WormCat annotation, and overrepresentation analyses

Protein domains encoded in the proteome of *P. pacificus* were predicted using HMMER (ver. 3.3.2) [77] in conjunction with the Pfam-A.hmm database (ver. 3.1b2) [78] (hmmscan, e-value cutoff: < 0.001). The ORTHOFINDER software (default mode; ver. 2.5.4) [79] was used to define orthogroups of *P. pacificus* and *C. elegans* proteins; WormCat (ver. 2) annotations [47] for *P. pacificus* were created based on the obtained orthogroups. In a few cases, we found multiple *C. elegans* orthologs with different WormCat terms for an individual *P. pacificus* gene. In such cases, we transferred all unique WormCat terms among the orthologs to the given gene in *P. pacificus*. All overrepresentation analyses (ORAs) were based on Fisher's Exact test; obtained *P*-values were FDR-corrected.

Dauer assays

We performed three different assays regarding dauer biology: dauer entry, dauer exit, and dauer survival. The workflow of these assays is depicted in Fig. 8A. Starting population of 500 young adults were added to 10ml liquid cultures (LCs) with OP50 as a food source (S-medium with or without cholesterol; 4μl nystatin and 5μl Streptomycin). Worms were propagated for two weeks at 22°C and 180rpm. After that we extracted 10μl from each LC and counted the number of dauers versus the number of feeding stages in the sample to quantify dauer entry (in %). The remaining volume of the LCs was treated with SDS following standard protocols [80] to isolate dauers. Some of the obtained dauers were added to agarose plates spotted with 300μl OP50 (30 worms per plate) and followed for multiple days to quantify the speed of dauer exit. Plates were kept at 20°C and worms were scored every 24 hrs until every individual exited dauer. The rest of the SDS-isolated dauers was added to LCs for the survival assays. These LCs did not contain food nor cholesterol and were again kept at 22°C and 180rpm for up to three weeks. Every week, we extracted 75μl of worm solution and added 25μl to three OP50-spotted agarose plates. We let the worms recover on these plates for 24 hrs and then scored alive versus dead worms to obtain the ratio of survivors (in %). All experiments were repeated four times independently. For the dauer exit and dauer survival assays we exclusively used dauers

isolated from LCs without cholesterol. Mutant-specific phenotypes were assessed via a PERMANOVA [81,82] with null model residual randomization (RRPP) using the *lm.rrpp* function of the RRPP package (ver. 1.3.1) [83,84]. Relative effect sizes (*Z*-scores) and *P*-values are reported in Fig. 8B-8D. We considered our dauer data as incompatible with the null hypothesis (i.e., "statistically significant") if the relative effect size was at least twice as large as the standard deviation ($Z \geq 2.0$) and the *P*-value was below a type I error rate of 5% ($P < 0.05$). If at least one of these criteria was not met, we considered the data compatible with the null hypothesis (i.e., "not statistically significant"). Null model residuals were randomized and permuted 10,000 times. Pairwise comparisons of mutant and wild-type strains were performed with the *pairwise* function of RRPP [83,84].

Data processing, data visualization, and illustration

All data was processed, formatted, and wrangled in R (ver. 4.3.0) [62] using the TIDYVERSE [85]. All scientific plots depicted in this study were created using the GGPLOT2 package (ver. 3.4.2) [85]. Heatmaps were created using the PHEATMAP package (ver. 1.0.12) [86]. In order to make our plots accessible to people with color-vision deficiencies or color blindness [87], we exclusively used the scientifically-derived colormaps provided by the SCICO (ver. 1.4.0) [88] and VIRIDIS (ver. 0.6.3) packages [89]. Microscopic images were edited and adjusted for levels, lightness, and contrast, in Affinity Photo (1.10.5). Scientific illustrations and final figures were created in Affinity Designer (ver. 1.10.5).

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

All data generated or analyzed during this study are included in this published article and its supplementary information files. The raw-read data from the mRNA-seq experiments are available in the European Nucleotide Archive (ENA) repository under the accession number PRJEB65081. Upon request, all worm strains generated in this study can be provided by the Sommer lab.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was funded by the Max Planck Society. T.T. was supported by the International Max Planck Research School (IMPRS) "From Molecules to Organisms".

Author contributions

T.T. conceptualized the study, designed and performed most experiments, obtained and analyzed the data, created visualizations, and wrote the manuscript. T.R. and T.T. planed and conducted the dauer experiments. R.J.S. supervised the project and provided resources. All authors edited and reviewed the original draft of the manuscript.

Acknowledgements

We thank the Light Microscopy facility at the Max Planck Institute for Biology, especially Christian Feldhaus and Aurora Panzera, for their assistance with confocal laser-scanning microscopy. We are grateful to Heike Haussmann for freezing samples of all the worm strains produced for this study. Additionally, we thank Christian Rödelsperger, Adrian Streit, and Devansh R. Sharma for helpful discussions of our work.

References

1. West-Eberhard MJ. Developmental plasticity and evolution. Oxford; New York: Oxford University Press; 2003.
2. Laland KN, Uller T, Feldman MW, Sterelny K, Müller GB, Moczek A, et al. The extended evolutionary synthesis: its structure, assumptions and predictions. Proceedings of the Royal Society B: Biological Sciences. 2015;282:20151019.
3. Uller T, Feiner N, Radersma R, Jackson IS, Rago A. Developmental plasticity and evolutionary explanations. Evolution & Development. 2020;22:47–55.
4. Sommer RJ. Phenotypic plasticity: from theory and genetics to current and future challenges. Genetics. 2020;215:1–13.
5. Sommer RJ, Dardiry M, Lenuzzi M, Namdeo S, Renahan T, Sieriebriennikov B, et al. The genetics of phenotypic plasticity in nematode feeding structures. Open biology. 2017;7:160332.
6. Hong RL, Sommer RJ. *Pristionchus pacificus*: a well-rounded nematode. Bioessays. 2006;28:651–659.
7. Sommer RJ. *Pristionchus pacificus*: a nematode model for comparative and evolutionary biology. Vol. 11. Leiden; The Netherlands: Brill; 2015.
8. Ragsdale EJ, Müller MR, Rödelsperger C, Sommer RJ. A developmental switch coupled to the evolution of plasticity acts through a sulfatase. Cell. 2013;155:922–933.
9. Kieninger MR, Ivers NA, Rödelsperger C, Markov GV, Sommer RJ, Ragsdale EJ. The nuclear hormone receptor NHR-40 acts downstream of the sulfatase EUD-1 as part of a developmental plasticity switch in *Pristionchus*. Current Biology. 2016;26:2174–2179.
10. Sieriebriennikov B, Prabh N, Dardiry M, Witte H, Röseler W, Kieninger MR, et al. A developmental switch generating phenotypic plasticity is part of a conserved multi-gene locus. Cell Reports. 2018;23:2835–2843.
11. Sieriebriennikov B, Sun S, Lightfoot JW, Witte H, Moreno E, Rödelsperger C, et al. Conserved nuclear hormone receptors controlling a novel plastic trait target fast-evolving genes expressed in a single cell. PLOS Genetics. 2020;16:e1008687.
12. Werner MS, Loschko T, King T, Reich S, Theska T, Franz-Wachtel M, et al. Histone 4 lysine 5/12 acetylation enables developmental plasticity of *Pristionchus* mouth form. Nature Communications. 2023;14:2095.
13. Sun S, Theska T, Witte H, Ragsdale EJ, Sommer RJ. The oscillating Mucin-type protein DPY-6 has a conserved role in nematode mouth and cuticle formation. Genetics. 2022;220:iyab233.
14. Sun S, Witte H, Sommer RJ. Chitin contributes to the formation of a feeding structure in a predatory nematode. Current Biology. 2023;33:15–27.
15. Cassada RC, Russell RL. The dauerlarva, a post-embryonic developmental variant of the nematode *Caenorhabditis elegans*. Developmental biology. 1975;46:326–342.
16. Ogawa A, Streit A, Antebi A, Sommer RJ. A conserved endocrine mechanism controls the formation of dauer and infective larvae in nematodes. Current Biology. 2009;19:67–71.
17. Baugh LR, Hu PJ. Starvation responses throughout the *Caenorhabditis elegans* life cycle. Genetics. 2020;216:837–878.
18. Renahan T, Lo W, Werner MS, Rochat J, Herrmann M, Sommer RJ. Nematode biphasic 'boom and bust' dynamics are dependent on host bacterial load while linking dauer and mouth-form polyphenisms. Environmental Microbiology. 2021;23:5102–5113.

19. O'Riordan VB, Burnell AM. Intermediary metabolism in the dauer larva of the nematode *Caenorhabditis elegans* II. The glyoxylate cycle and fatty-acid oxidation. Comparative Biochemistry and Physiology B, Comparative Biochemistry. 1990;95:125–130.

20. Braeckman B, Houthoofd K, Vanfleteren J. Intermediary metabolism. In: The *C. elegans* research community, editor. WormBook: The Online Review of *C. elegans* Biology. 2009.

21. Erkut C, Gade VR, Laxman S, Kurzchalia TV. The glyoxylate shunt is essential for desiccation tolerance in *C. elegans* and budding yeast. Elife. 2016;5:e13614.

22. Hibshman JD, Doan AE, Moore BT, Kaplan RE, Hung A, Webster AK, et al. daf-16/FoxO promotes gluconeogenesis and trehalose synthesis during starvation to support survival. Elife. 2017;6:e30057.

23. Motola DL, Cummins CL, Rottiers V, Sharma KK, Li T, Li Y, et al. Identification of ligands for DAF-12 that govern dauer formation and reproduction in *C. elegans*. Cell. 2006;124:1209–1223.

24. Hu PJ. Dauer. In: The *C. elegans* research community, editor. WormBook: The Online Review of *C. elegans* Biology. 2007.

25. Miglioli A, Canesi L, Gomes ID, Schubert M, Dumollard R. Nuclear receptors and development of marine invertebrates. Genes. 2021;12:83.

26. Tao LJ, Seo DE, Jackson B, Ivanova NB, Santori FR. Nuclear hormone receptors and their ligands: metabolites in control of transcription. Cells. 2020;9:2606.

27. Sommer RJ, Ogawa A. Hormone signaling and phenotypic plasticity in nematode development and evolution. Current Biology. 2011;21:R758–766.

28. Antebi A. Nuclear receptor signal transduction in *C. elegans*. In: The *C. elegans* research community, editor. WormBook: The Online Review of *C. elegans* Biology. 2015.

29. Srinivasan S. Regulation of body fat in *Caenorhabditis elegans*. Annual Review of Physiology. 2015;77:161–178.

30. Lemieux GA, Ashrafi K. Neural regulatory pathways of feeding and fat in *Caenorhabditis elegans*. Annual Review of Genetics. 2015;49:413–438.

31. Watts JL, Ristow M. Lipid and carbohydrate metabolism in *Caenorhabditis elegans*. Genetics. 2017;207:413–446.

32. Peterson ND, Samantha YT, Huang QJ, Wani KA, Schiffer CA, Pukkila-Worley R. Non-canonical pattern recognition of a pathogen-derived metabolite by a nuclear hormone receptor identifies virulent bacteria in *C. elegans*. Immunity. 2023;56:768–782.

33. Sural S, Hobert O. Nematode nuclear receptors as integrators of sensory information. Current Biology. 2021;31:4361–4366.

34. Sluder AE, Mathews SW, Hough D, Yin VP, Maina CV. The nuclear receptor superfamily has undergone extensive proliferation and diversification in nematodes. Genome Research. 1999;9:103–120.

35. Robinson-Rechavi M, Maina CV, Gissendanner CR, Laudet V, Sluder A. Explosive lineage-specific expansion of the orphan nuclear receptor HNF4 in nematodes. Journal of Molecular Evolution. 2005;60:577–586.

36. Theska T, Sommer RJ. Feeding-structure morphogenesis in "rhabditid" and diplogastrid nematodes is not controlled by a conserved genetic module. bioRxiv. 2023. https://doi.org/10.1101/2023.11.07.565949

37. Riebesell M, Sommer RJ. Three-dimensional reconstruction of the pharyngeal gland cells in the predatory nematode *Pristionchus pacificus*. Journal of Morphology. 2017;278:1656–1666.

38. Moreno E, Sieriebriennikov B, Witte H, Rödelsperger C, Lightfoot JW, Sommer RJ. Regulation of hyperoxia-induced social behaviour in *Pristionchus pacificus* nematodes requires a novel cilia-mediated environmental input. Scientific Reports. 7:17550.

39. Hansen M, Flatt T, Aguilaniu H. Reproduction, fat metabolism, and life span: what is the connection? Cell Metabolism. 2013;17:10–19.

40. Athanasouli M, Akduman N, Röseler W, Theam P, Rödelsperger C. Thousands of *Pristionchus pacificus* orphan genes were integrated into developmental networks that respond to diverse environmental microbiota. PLOS Genetics. 2023;19:e1010832.

41. Renahan T, Sommer RJ. Nematode interactions on beetle hosts indicate a role of mouth-form plasticity in resource competition. Frontiers in Ecology and Evolution. 2021;9:752695.

42. Renahan T, Sommer RJ. Multidimensional competition of nematodes affects plastic traits in a beetle ecosystem. Frontiers in Cell and Developmental Biology. 2022;10:985831.

43. Dardiry M, Eberhard G, Witte H, Rödelsperger C, Lightfoot JW, Sommer RJ. Divergent combinations of cis-regulatory elements control the evolution of phenotypic plasticity. PLOS Biology. 2023;21:e3002270.

44. Stadler M, Fire A. Conserved translatome remodeling in nematode species executing a shared developmental transition. PLOS Genetics. 2013;9:e1003739.

45. Larance M, Pourkarimi E, Wang B, Murillo AB, Kent R, Lamond AI, et al. Global proteomics analysis of the response to starvation in *C. elegans*. Molecular & Cellular Proteomics. 2015;14:1989–2001.

46. Harvald EB, Sprenger RR, Dall KB, Ejsing CS, Nielsen R, Mandrup S, et al. Multi-omics analyses of starvation responses reveal a central role for lipoprotein metabolism in acute starvation survival in *C. elegans*. Cell Systems. 2017;5:38–52.

47. Holdorf AD, Higgins DP, Hart AC, Boag PR, Pazour GJ, Walhout AJ, et al. WormCat: an online tool for annotation and visualization of *Caenorhabditis elegans* genome-scale data. Genetics. 2020;214:279–294.

48. Stiernagle T. Maintenance of *C. elegans*. In: The *C. elegans* research community, editor. WormBook: The Online Review of *C. elegans* Biology. 2006.

49. Sanghvi GV, Baskaran P, Röseler W, Sieriebriennikov B, Rödelsperger C, Sommer RJ. Life history responses and gene expression profiles of the nematode *Pristionchus pacificus* cultured on *Cryptococcus* yeasts. PLOS One. 2016;11:e0164881.

50. Witte H, Moreno E, Rödelsperger C, Kim J, Kim JS, Streit A, et al. Gene inactivation using the CRISPR/Cas9 system in the nematode *Pristionchus pacificus*. Development Genes and Evolution. 2015;225:55–62.

51. Ghanta KS, Ishidate T, Mello CC. Microinjection for precision genome editing in *Caenorhabditis elegans*. STAR Protocols. 2021;2:100748.

52. Han Z, Lo WS, Lightfoot JW, Witte H, Sun S, Sommer RJ. Improving transgenesis efficiency and CRISPR-associated tools through codon optimization and native intron addition in *Pristionchus* nematodes. Genetics. 2020;216:947–956.

53. Eren GG, Roca M, Han Z, Lightfoot JW. Genomic integration of transgenes using UV irradiation in *Pristionchus pacificus*. microPublication Biology. 2022;29.

54. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. Nature Methods. 2017;14:417–419.

55. Srivastava A, Malik L, Sarkar H, Zakeri M, Almodaresi F, Soneson C, et al. Alignment and mapping methodology influence transcript abundance estimation. Genome Biology. 2020;21:1–29.

56. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, et al. Bioconductor: open software development for computational biology and bioinformatics. Genome Biology. 2004;5:1–16.

57. Love MI, Anders S, Kim V, Huber W. RNA-Seq workflow: gene-level exploratory analysis and differential expression. F1000Research. 2015;4:1070.

58. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology. 2014;15:1–21.

59. Stephens M. False discovery rates: a new deal. Biostatistics. 2017;18:275–294.

60. Theska T, Sieriebriennikov B, Wighard SS, Werner MS, Sommer RJ. Geometric morphometrics of microscopic animals as exemplified by model nematodes. Nature Protocols. 2020;15:2611–2644.

61. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: an open-source platform for biological-image analysis. Nature Methods. 2012;9:676–682.

62. R Core Team R. R: A language and environment for statistical computing. R Foundation for Statistical Computing. 2023.

63. Adams DC, Otárola-Castillo E. geomorph: an R package for the collection and analysis of geometric morphometric shape data. Methods in Ecology and Evolution. 2013;4:393–399.

64. Adams, D. C., Collyer, M. & Kaliontzopoulou, A. Geomorph: Software for geometric morphometric analyses. R package version 4.0.5. 2023. Available from: https://cran.r-project.org/web/packages/geomorph/index.html

65. Baken EK, Collyer ML, Kaliontzopoulou A, Adams DC. geomorph v4. 0 and gmShiny: Enhanced analytics and a new graphical interface for a comprehensive morphometric experience. Methods in Ecology and Evolution. 2021;12:2355–2363.

66. Schlager S. Morpho and Rvcg–Shape Analysis in R: R-Packages for geometric morphometrics, shape analysis and surface manipulations. In: Zheng G, Li S, Szekely G, editors. Statistical Shape and Deformation Analysis. Academic Press; 2017. p. 217–256.

67. Schlager S, Jefferis G. Morpho: calculations and visualisations related to geometric morphometrics. R package version 2.11. 2023. Available from: https://cran.r-project.org/web/packages/Morpho/index.html

68. Scrucca L, Fop M, Murphy TB, Raftery AE. mclust 5: clustering, classification and density estimation using Gaussian finite mixture models. The R journal. 2016;8:289.

69. Fraley C, Raftery A, Scrucca L. mclust: Normal mixture modeling for model-based clustering, classification, and density estimation. R package version 6.0.0. 2023. Available from: https://cran.r-project.org/web/packages/mclust/index.html

70. Mitteroecker P, Gunz P, Bernhard M, Schaefer K, Bookstein FL. Comparison of cranial ontogenetic trajectories among great apes and humans. Journal of Human Evolution. 2004;46:679–698.

71. Rich JT, Neely JG, Paniello RC, Voelker CC, Nussenbaum B, Wang EW. A practical guide to understanding Kaplan-Meier curves. Otolaryngology—Head and Neck Surgery. 2010;143:331–336.

72. Dey T, Lipsitz SR, Cooper Z, Trinh QD, Krzywinski M, Altman N. Regression modeling of time-to-event data with censoring. Nature Methods. 2022;19:1513–1515.

73. Therneau T. A package for survival analysis in R. R package version 3.5-5. 2023. Available from: https://cran.r-project.org/web/packages/survival/index.html

74. Sjoberg D, Baillie M, Haesendonckx S, Treis T. ggsurvfit: Flexible Time-to-Event Figures. R package version 0.3.0. 2023. Available from: http://www.danieldsjoberg.com/ggsurvfit/

75. Kassambara A, Kosinski M, Biecek P, Fabian S. survminer: Drawing Survival Curves using 'ggplot2.' R package version 0.4.9. 2023. Available from: https://cran.r-project.org/web/packages/survminer/index.html

76. Daniel DS, Whiting K, Curry M, Lavery JA, Larmarange J. Reproducible summary tables with the gtsummary package. The R Journal. 2021;13:570–580.

77. Mistry J, Finn RD, Eddy SR, Bateman A, Punta M. Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. Nucleic Acids Research. 2013;41:e121.

78. Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, et al. The Pfam protein families database: towards a more sustainable future. Nucleic Acids Research. 2016;44:D279–285.

79. Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. Genome Biology. 2019;20:1–14.

80. Karp X. Working with dauer larvae. In: The *C. elegans* research community, editor. WormBook: The Online Review of *C. elegans* Biology. 2018.

81. Anderson MJ. A new method for non-parametric multivariate analysis of variance. Austral Ecology. 2001;26:32–46.

82. Anderson MJ. Permutational multivariate analysis of variance (PERMANOVA). In: Balakrishnan N Colton T, Everitt B, Piegorsch W, Ruggeri F, Teugels JL, editors. Wiley Statsref: Statistics Reference Online. 2017. pp. 1–15.

83. Collyer ML, Adams DC. RRPP: An r package for fitting linear models to high-dimensional data using residual randomization. Methods in Ecology and Evolution. 2018;9:1772–1779.

84. Collyer M, Adams D. RRPP: linear model evaluation with randomized residuals in a permutation procedure. R package version 1.3.1. 2023. Available from: https://cran.r-project.org/web/packages/RRPP/index.html

85. Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, et al. Welcome to the Tidyverse. Journal of Open Source Software. 2019;4:1686.

86. Kolde R, Kolde MR. Package 'pheatmap.' R package version 1.0.12. 2023. Available from: https://cran.r-project.org/web/packages/pheatmap/index.html

87. Crameri F, Shephard GE, Heron PJ. The misuse of colour in science communication. Nature Communications. 2020;11:5444.

88. Pedersen TL, Crameri F. Scico: Colour palettes based on the scientific colour-maps. R package version 1.4.0. 2023. Available from: https://cran.r-project.org/web/packages/scico/index.html

89. Garnier S, Ross N, Rudis R, Camargo PA, Sciaini M, Scherer C. viridis—colorblind-friendly color maps for R. R package version 0.6.3. 2023. Available from: https://sjmgarnier.github.io/viridis/

Figures

Fig. 1. Phenotypic plasticity throughout the *Pristionchus pacificus* life cycle and the gene regulatory network of feeding-structure development. (A) Alternative adult mouth forms in *P. pacificus*. Eu = eurystomatous, St = stenostomatous. (B) *P. pacificus* life cycle, including the alternative "indirect life cycle" via dauer development. (C) Closed mouth morphology of dauer larvae. bp = buccal plug. (D) modular gene regulatory network (GRN) controlling feeding-structure development in *P. pacificus*. (E) Phylogeny of all *nhr* genes encoded in the genomes of *C. elegans* and *P. pacificus*. Tree was re-drawn based on the data of Sieriebriennikov *et al.* [11]. The asterisk (*) indicates the location of the branch that harbors the three deeply conserved sister receptors NHR-1, NHR-10, and NHR-40.

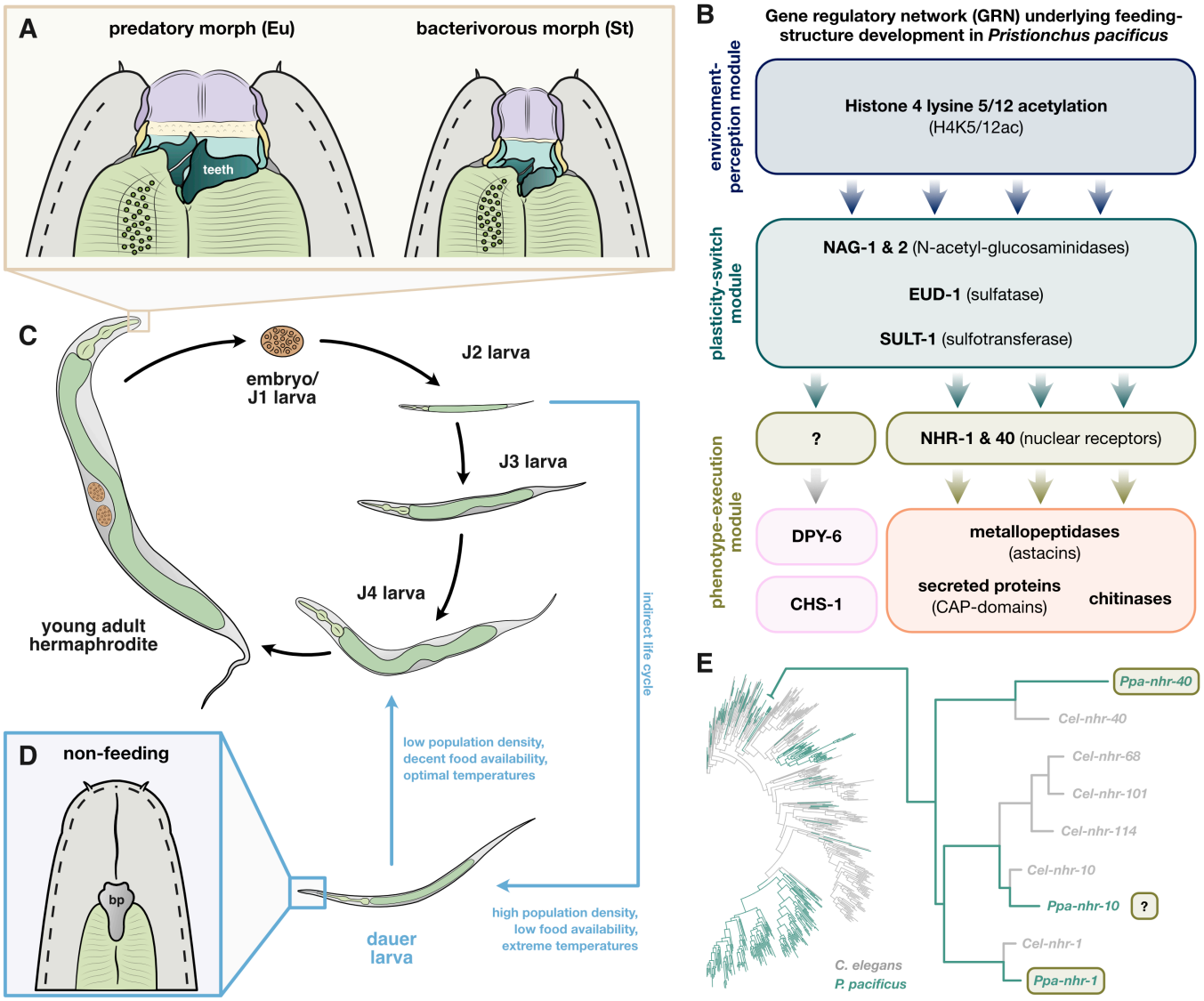


Fig. 2. Expression pattern of *Ppa-nhr-10* across tissues. A 1.5kb long promoter sequence of *Ppa-nhr-10* (ending right before the start codon) was fused to *P. pacificus*-optimzed *GFP.* Relative strength of fluorescent-protein expression is indicated via a relative intensity scale. (A) and (B) each show a representative worm from two independent reporter lines in different orientations: (A) right body side facing reader, (B) left body side facing reader. Asterisk (*) in (A) indicates the position of anterior gonad arm. Note that both, anterior and posterior gonad arms, are always reporting; absence of the anterior arm in (A) is due to image stacking only. The crossed double-arrows with letters in (A) and (B) indicate anatomical orientation of the worm: A = anterior, P = posterior, D = dorsal, V = ventral. (C) Zoom-in on the head of the worm depicted in (B), merged with DIC image. (D) Zoom-in on the anterior gonad arm (outlined with white dots) of the worm depicted in (B). Note that neither the sperm cells nor the eggs express *Ppa-nhr-10*; only the somatic gonad tissues do so. amp = ampulla of the gland cell(s), g1D = dorsal pharyngeal gland, g1V = ventral pharyngeal glands, sp = sperm.

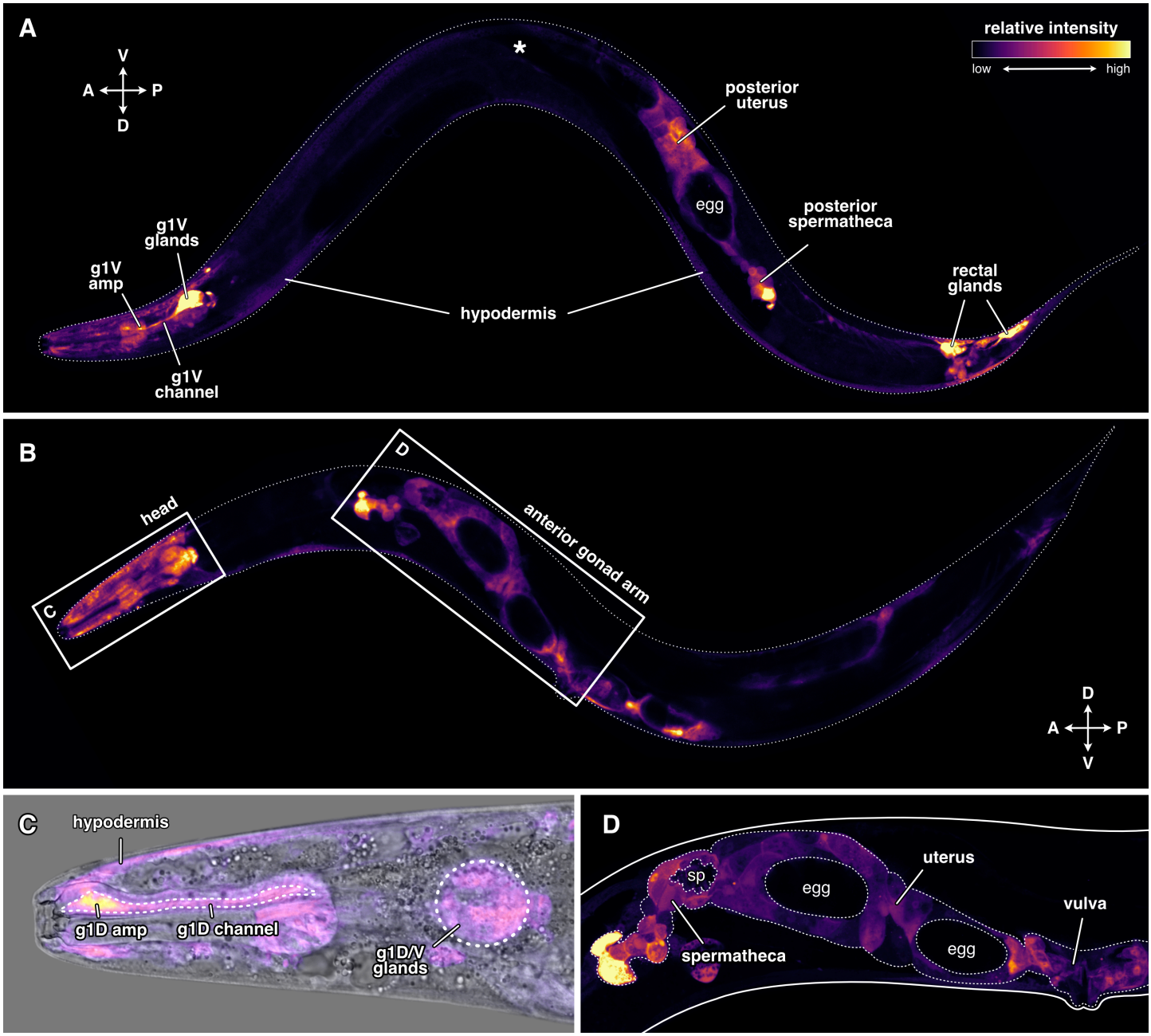


Fig. 3. *Ppa-nhr-10* mutations do not affect mouth morphology and only weakly affects mouth-form plasticity. (A) Gene model of *Ppa-nhr-10* and frameshift alleles which lead to early truncation of the ligand-binding domain. e = exon. (B) Ratio of predatory (Eu) and bacterivorous (St) animals in the population. Worms were raised and scored on agar plates with *E. coli* OP50 as a food source. Three biological replicates were obtained (N=3) and 150 worms of each strain were scored per replicate. Significant differences are indicated in black, non-significant ones in gray. *Z*-scores and *P*-values were derived from a two-tailed PERMANOVA with null model residual randomization (RRPP). Bars represent mean values across all replicates; error bars indicate ±1 standard error of the mean. (C) Geometric morphometric analysis of form differences in the nematode mouths. The PCA plot shows the distribution of mouth morphologies of individual worms in a form space. Note that the ranges of morphological variation (outlined by convex hulls) in *Ppa-nhr-10* mutant strains completely overlap with the wild-type range of morphological variation. (D) Unsupervised (model-based) clustering reveals that all 144 individuals belong to a single morphological cluster, demonstrating that mutant worms cannot be differentiated from wild-type worms based on mouth morphology. Only predatory (Eu) animals were used for the analyses in (C) and (D).

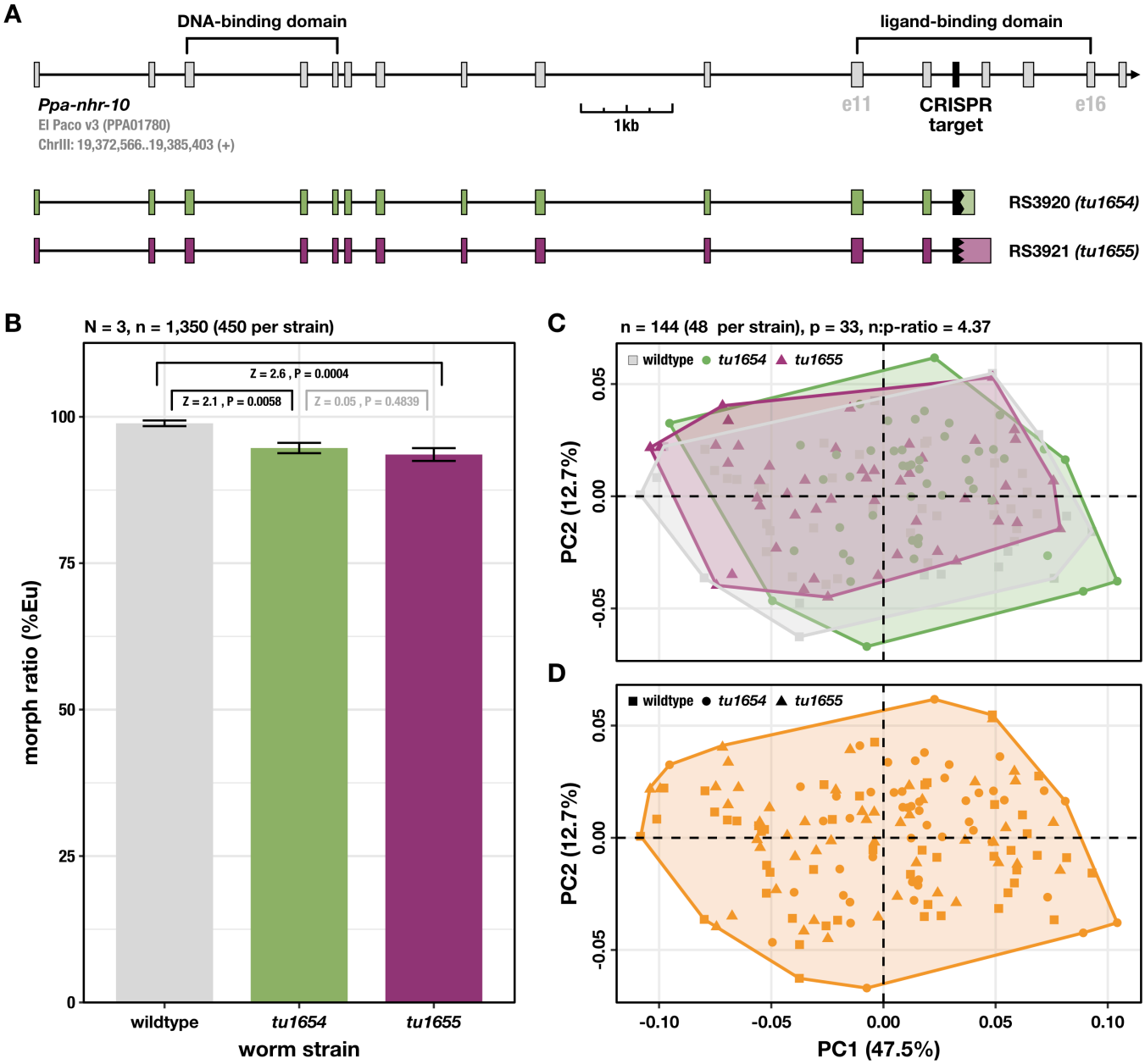


Fig. 4. *Ppa-nhr-10* mutations affect expression levels of 32 potential regulatory target genes. Differential gene expression analysis between wild-type and *Ppa-nhr-10* mutant animals (*tu1654* and *tu1655*). Bold names along the y-axis correspond to names of the mis-regulated putative targets (El_Paco_v3). Target genes which have an exact *C. elegans* ortholog were additionally given the public gene name of this ortholog (indicated in brackets). Panels indicate in which life history stage the putative targets genes were found. Labels within the panels indicate function information for the gene in form of WormCat terms (black) and/or predicted PFAM domains (gray). WormCat annotations were adapted from *C. elegans*, based on *Cel-Ppa*-orthogroups identified with OrthoFinder. PFAM domains were only indicated for genes we could not assign a WormCat term to, or genes whose WormCat terms ended in "unassigned". Cutoffs used in the differential gene expression analysis with DESeq2 and for PFAM domain prediction with HMMER are indicated in the lower left. DEGs = differentially expressed genes, MMSE = minimum mean squared error, FSOS = false sign or smaller.

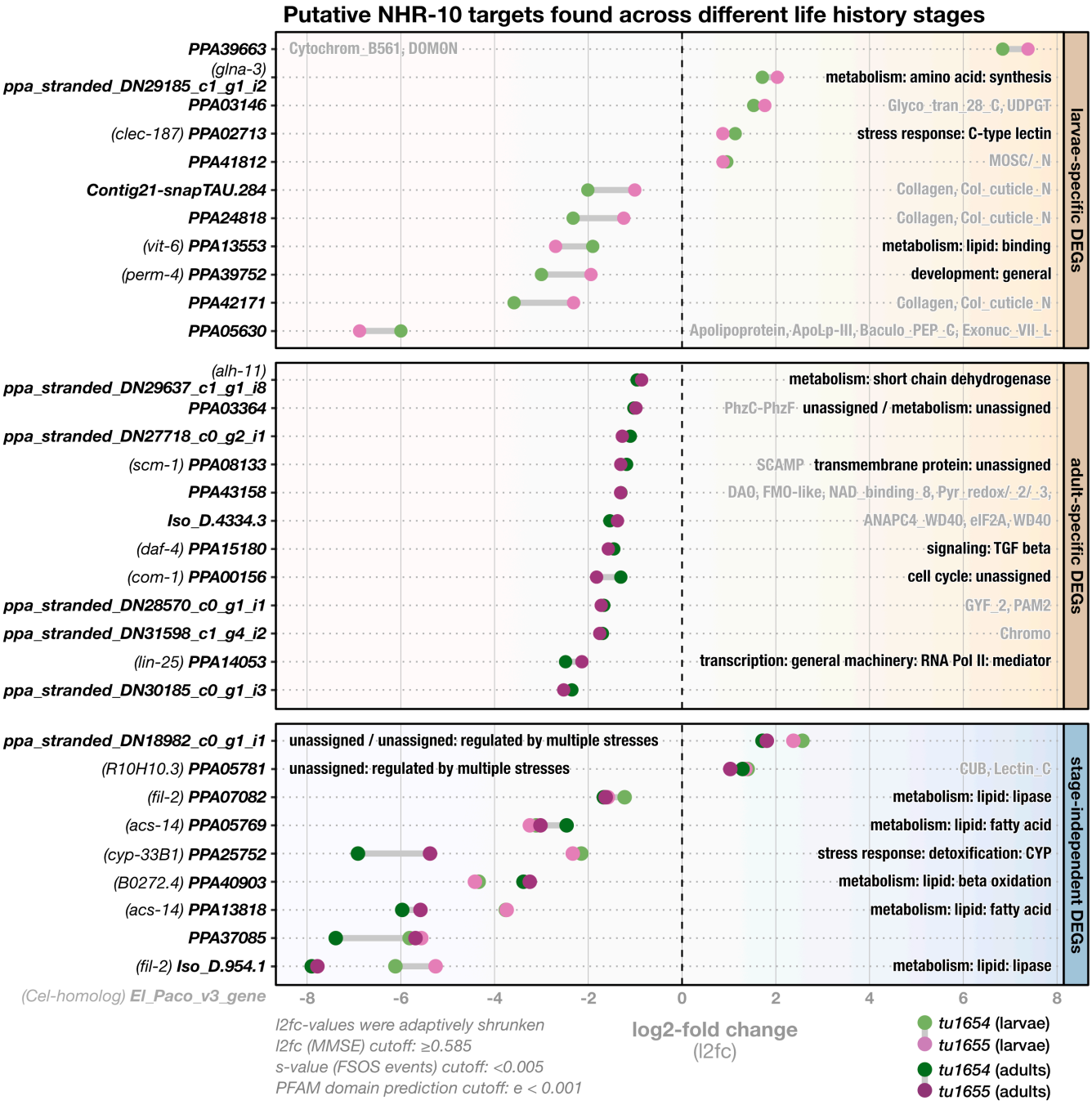


Fig. 5. *Ppa-nhr-10* mutations cause premature death during acute starvation. Comparison of overall survival curves of wild-type *P. pacificus* and two *Ppa-nhr-10* mutants that were well-fed (A) and completely starved (B). Survival curves show the mean survival, across all replicates, over time. Transparent ribbons behind the curves indicate the associated uncertainty in form of ±1 standard errors of the mean. Each of the two experiments (A and B) was repeated four times independently (N=4); the total numbers of worms (n) that were tracked across all replicates are indicated on top of both plots. Indicated *P*-values for pairwise comparisons resulted from a log-rank test that followed a Kaplan-Meier survival probability estimation; bold and black font indicates significant differences in the pairwise comparisons of overall survival. (C) Summary of the global Kaplan-Maier statistics for both experiments (in gray boxes) next to results from Cox proportional hazard regressions (white boxes). Points indicate the estimated hazard ratios for both mutant strains within each experiment (relative to the wild-type reference level of 1.0). Intervals indicate associated uncertainty in form of 95% confidence intervals. *P*-values in bold black indicate significant differences between the mutant strain and the wild type.

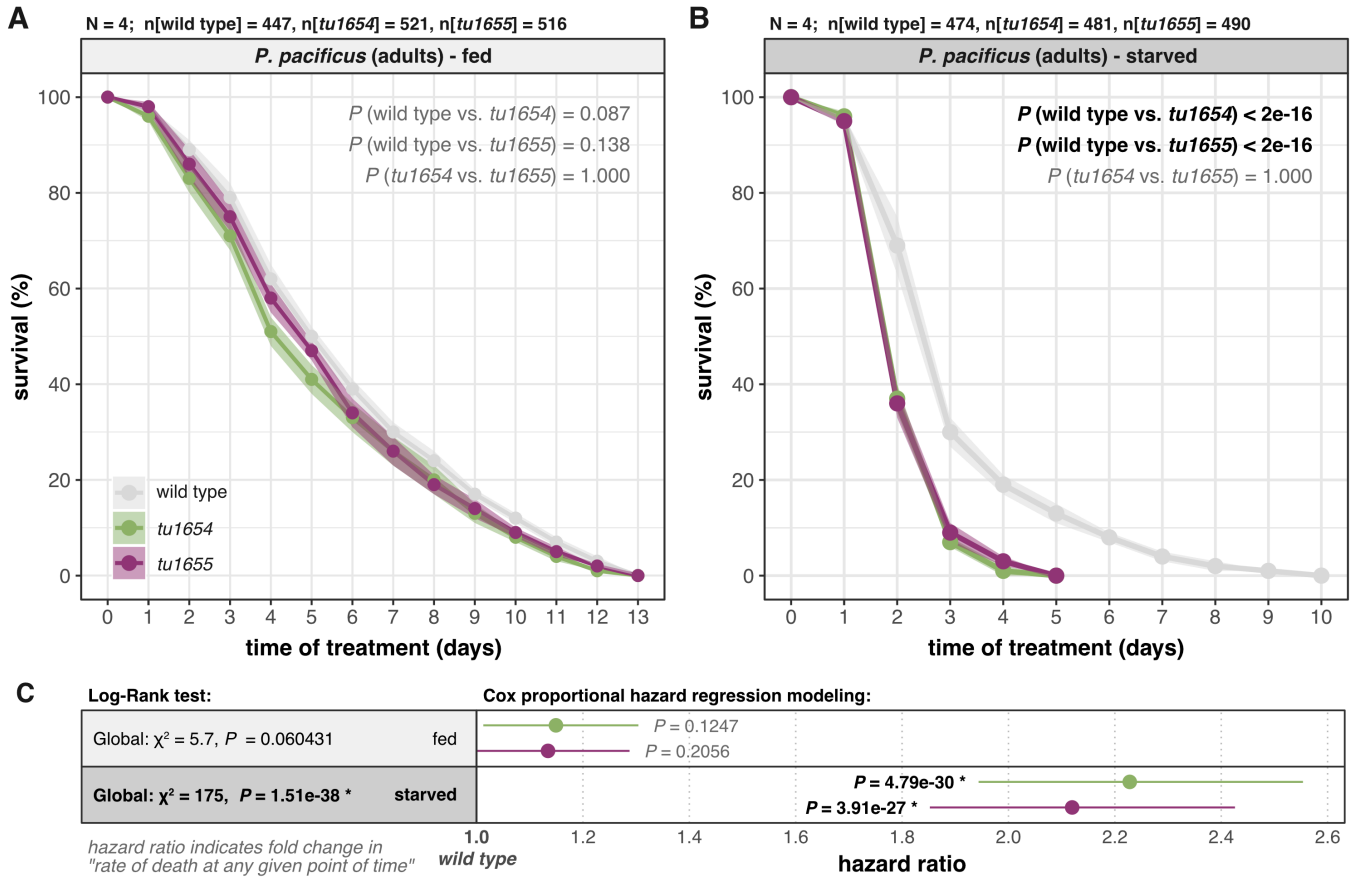


Fig. 6. Genome-wide transcriptional responses to acute starvation in wild-type *P. pacificus*. (A) Hierarchically clustered heat map showing the 1,007 genes which were identified as differentially expressed in response to 24h of acute starvation (SRGs). Trees on the left and the top show clustering results based on similarity in expression levels. Numbers on the gene tree indicate the number of genes in each of the two response clusters. WT = wild type. (B) Bar chart showing the overrepresented gene-age classes and co-expression networks (according to [40]) identified the two response clusters, relative to the genome. "Phylogenetically old genes" refers to those genes in the genome of *P. pacificus* which are not diplogastrid-specific orphan genes. (C) Lollipop plot showing overrepresented PFAM domains in each response cluster. Y-axis gives PFAM domain names. X-axis gives the percentage of all unique genes encoding this PFAM domain, which were found to be present in the response cluster (e.g 60% means that 60% all unique genes in the genome that encode this domain are present in that cluster). (D) Bubble chart showing overrepresented WormCat terms in each of the response clusters. Enriched categories were broken down hierarchically (see WormCat legend in the upper right). Bubble size scales with the gene count indicated to the right of the bubble. The significance scale (q-value) depicted in (C) also applies to the bubbles depicted in (D).

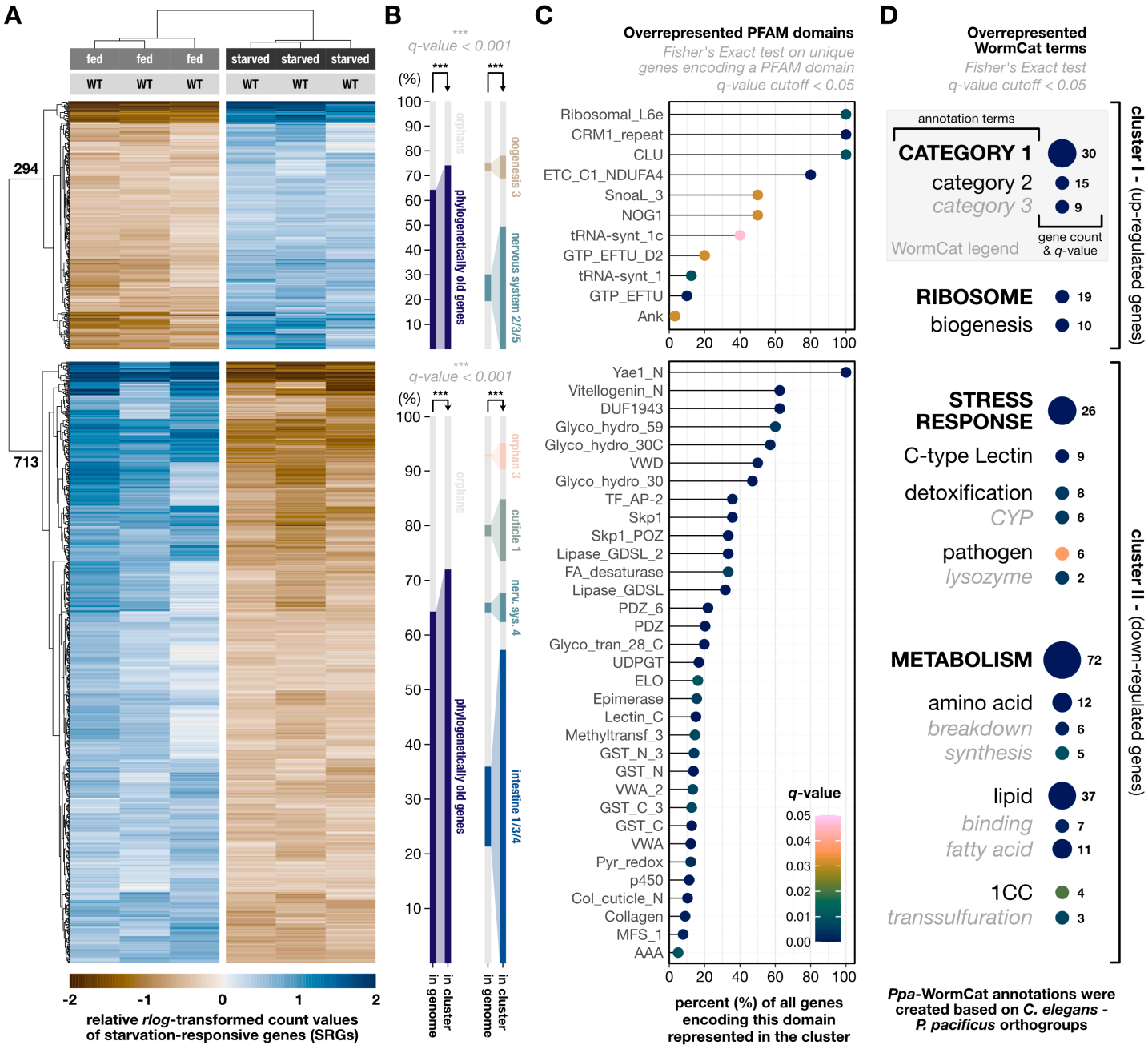


Fig. 7. *Ppa-nhr-10* mutations disturb the expression levels of 37 putative regulatory targets in starving *P. pacificus*. Bold names along the y-axis correspond to names of the misregulated putative targets (El Paco v3) which are likely to cause the observed starvation-survival phenotype. Genes which have an exact *C. elegans* ortholog were additionally given the public gene name of this ortholog (indicated in brackets). Panels indicate whether these targets are members of the SRGs or not. Labels within the panels indicate functional information for the gene in form of WormCat terms (black) and/or predicted PFAM domains (gray). WormCat annotations were adapted from *C. elegans*, based on *Cel-Ppa*-orthogroups identified with OrthoFinder. PFAM domains were only indicated for genes we could not assign a WormCat term to, or genes whose WormCat terms ended in "unassigned". Cutoffs used in the differential gene expression analysis with DESeq2 and for PFAM domain prediction with HMMER are indicated in the lower left. Note that, in the current El Paco v3 annotations, PPA25752 (in SRG targets) seems to be a gene model that artificially fuses two separate genes. This is supported by multiple available RNA datasets and our PFAM domain predictions, which revealed that the first half of the hypothetical protein encodes a 7TM_GPCR, while the second half encodes a p450 cytochrome. As it seems unlikely that a single protein would act as both, a GPCR and a cytochrome, we repeated our differential expression analysis after manually splitting PPA25752 into two separate genes. We found no evidence of differential expression for the resulting GPCR-gene, but could perfectly replicate the differential expression signature we originally found for PPA25752, for the cytochrome-encoding gene. FSOS = false sign or smaller, MMSE = minimum mean squared error, SRG = starvation-response gene.

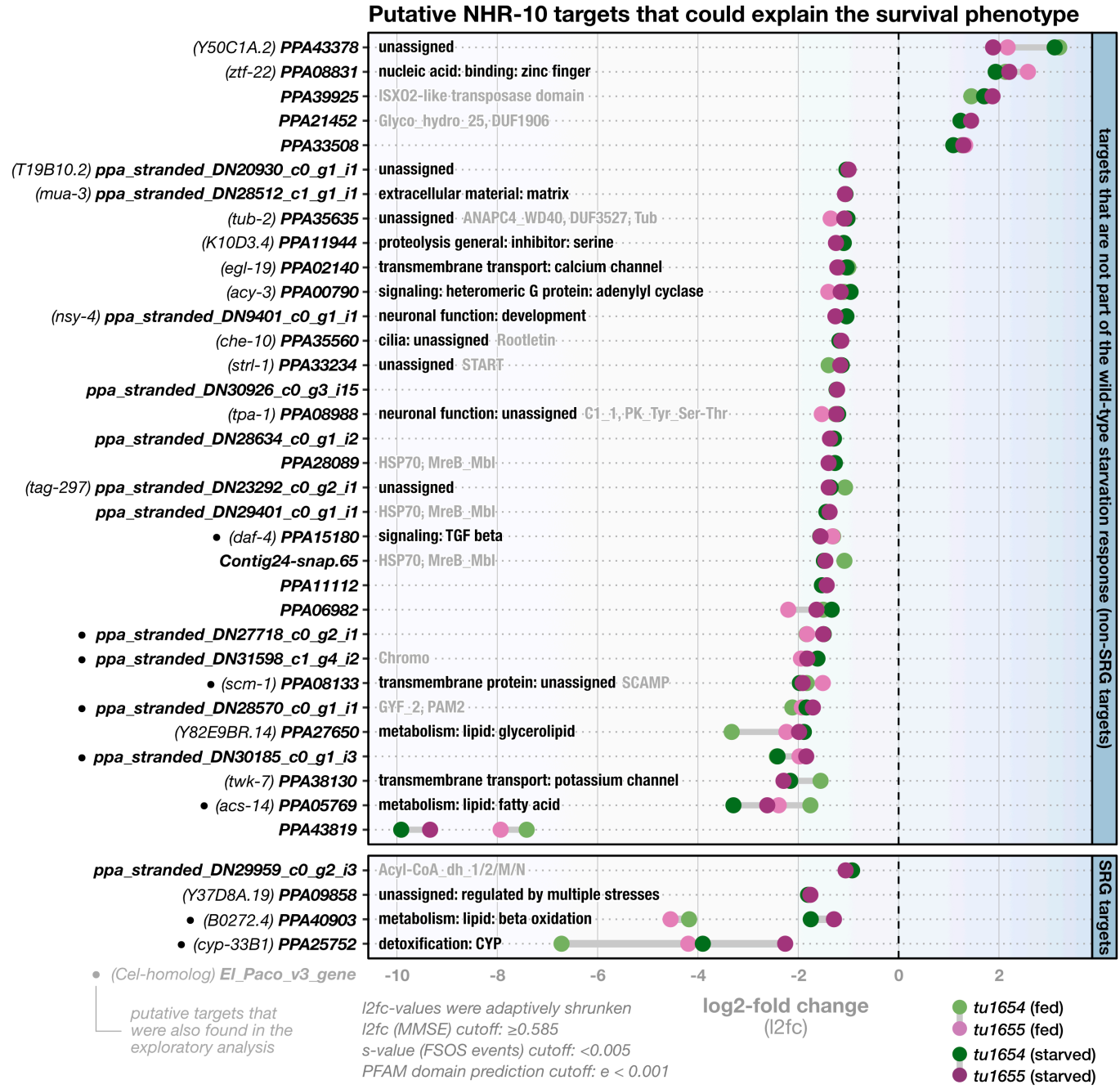
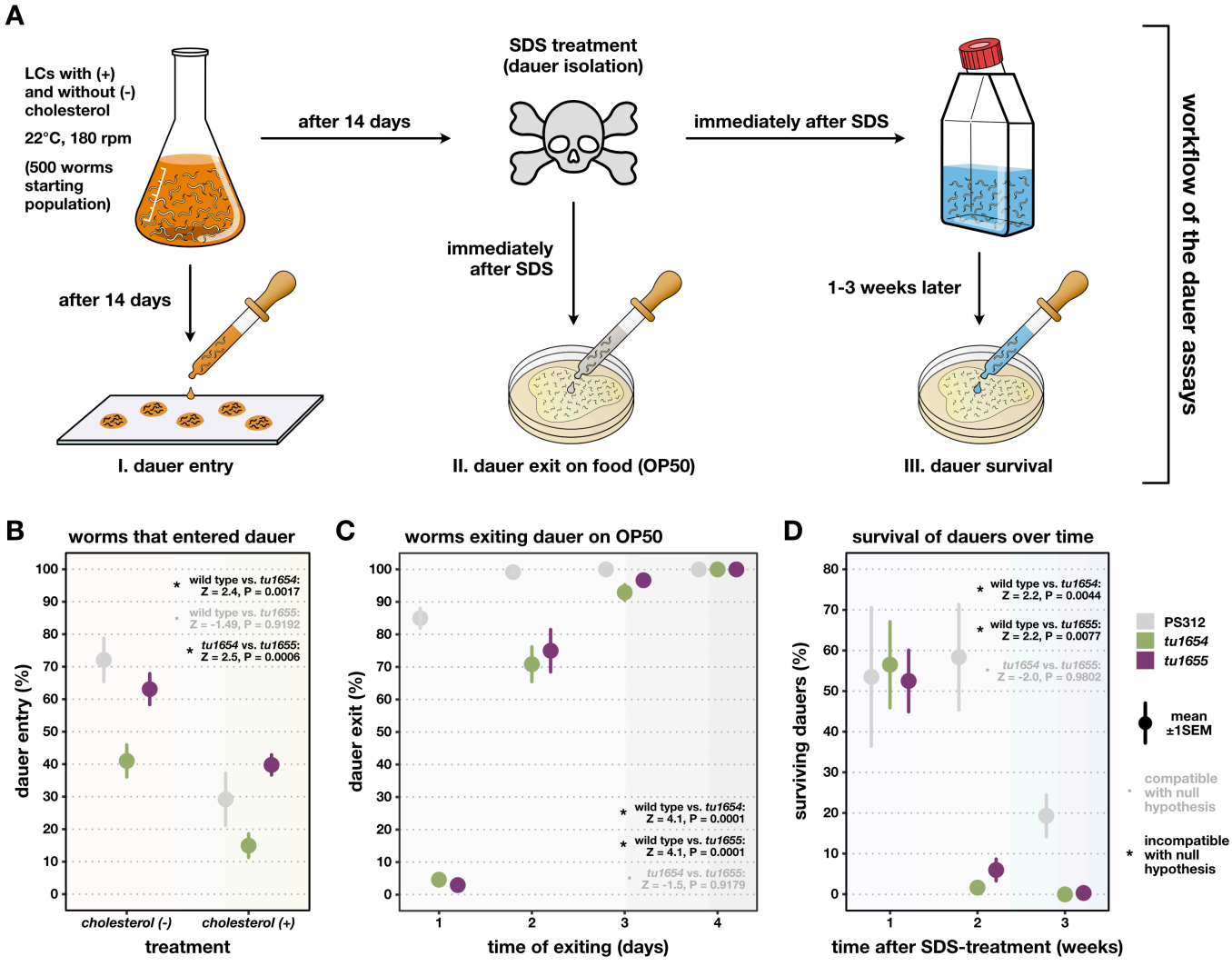


Fig. 8. *Ppa-nhr-10* mutants have various dauer-related phenotypes, including impaired long-term survival. (A). Workflow of the three dauer assays. (B-D) Results of the individual phenotypic assays indicated above each plot in the workflow chart. Point ranges indicate the mean estimates across all replicates ±1 standard error of the mean. Note that the dauers used in the exit and survival assays (C-D) were obtained from LCs without cholesterol. Additionally, all of the worms in (C) managed to exit dauer and fully resumed developments into viable adults. Hypothesis tests used a two-tailed PERMANOVA with null model residual randomization (RRPP); relative effect sizes (*Z*-scores) and *P*-values for pairwise comparisons among the three strains are reported in the upper right of each panel. Incompatibility with the null hypothesis ($Z \geq 2$ and $P < 0.05$) indicates the presence of a strain-specific phenotype; compatibility with the null hypothesis ($Z < 2$ or $P \geq 0.05$) supports the absence of a strain-specific phenotype. LCs = liquid cultures, rpm = rotations per minute, SDS = sodium dodecyl sulfate, SEM = standard error of the mean. All experiments were repeated multiple times independently (N=4). Number of worms in (B): n[PS312, (-)] = 460; n[*Ppa-nhr-10(tu1654)*, (-)] = 537; n[*Ppa-nhr-10(tu1655)*, (-)] = 432; n[PS312, (+)] = 750; n[*Ppa-nhr-10(tu1654)*, (+)] = 877; n[*Ppa-nhr-10(tu1655)*, (+)] = 879. Number of worms in (C): n[PS312] = 240; n[*Ppa-nhr-10(tu1654)*] = 240; n[*Ppa-nhr-10(tu1655)*] = 240. Number of worms in (D): n[PS312] = 395; n[*Ppa-nhr-10(tu1654)*] = 175; n[*Ppa-nhr-10(tu1655)*] = 177.



Supplementary Information

Sup. Tab. 1. Overrepresentation analysis shows no bias towards any given chromosomal location among *Ppa*-NHR-10's targets. The 37 candidate regulatory targets (DEGs) of *Ppa*-NHR-10, which were identified in the condition-specific RNA-seq experiments, were used as input for the analysis. No significant bias (FDR-corrected *P*-value < 0.05) towards any particular chromosome was found for *Ppa*-NHR-10's target genes. DEG = differentially expressed gene, FDR = false discovery rate.

chromosome	number_of_genes	number_of_DEGs	fold_enrichment	FDR_corrected_pvalue
Chr1	6869	5	0.5684764688986556	0.35319808452624085
Chr2	4462	11	1.9253031606236448	0.13589186241137188
Chr3	3736	5	1.045199374963829	0.9702253905888654
Chr4	6594	10	1.1843690824582545	0.8351614949834785
Chr5	4097	5	0.9531034573748757	1
ChrX	3002	1	0.26015089039739275	0.35319808452624085

Feeding-structure morphogenesis in "rhabditid" and diplogastrid nematodes is not controlled by a conserved genetic module

Tobias Theska & Ralf J. Sommer*

Max Planck Institute for Biology Tübingen (MPI-B), Department for Integrative Evolutionary Biology, Max-Planck-Ring 9, 72076 Tübingen, Germany

***Corresponding author**

Ralf J. Sommer (ralf.sommer@tuebingen.mpg.de)

ORCIDs

Tobias Theska - https://orcid.org/0000-0002-0920-4427
Ralf J. Sommer - https://orcid.org/0000-0003-1503-7749

Keywords

morphological novelty, stoma development, evo-devo, *Caenorhabditis elegans*, *Pristionchus pacificus*

Abstract

Disentangling the evolution of the molecular processes and genetic networks that facilitate the emergence of morphological novelties is one of the main objectives in evolutionary developmental biology. Here, we investigated the evolutionary history of a gene regulatory network controlling the development of novel tooth-like feeding-structures in diplogastrid nematodes. Focusing on NHR-1 and NHR-40, the two transcription factors that regulate the morphogenesis of these feeding structures in *Pristionchus pacificus*, we sought to determine whether they have a similar function in out-group nematode *Caenorhabditis elegans*, which has typical "rhabditid" flaps instead of teeth. Contrary to our initial expectations, we found that they do not have a similar function. While both receptors are co-expressed in the tissues that produce the feeding structures in the two nematodes, genetic inactivation of either receptor had no impact on feeding-structure morphogenesis in *C. elegans*. Transcriptomic experiments revealed that NHR-1 and NHR-40 have highly species-specific regulatory targets. These results suggest two possible evolutionary scenarios: either the genetic module responsible for feeding-structure morphogenesis in Diplogastridae already existed in the last common ancestor of *C. elegans* and *P. pacificus*, and subsequently disintegrated in the former as NHR-1 and NHR-40 acquired new targets, or it evolved in conjunction with teeth in Diplogastridae. These findings indicate that feeding-structure morphogenesis is regulated by different genetic programs in *P. pacificus* and *C. elegans*, hinting at developmental systems drift during the flap-to-tooth transformation. Further research in other "rhabditid" species is needed to fully reconstruct the developmental genetic changes which facilitated the evolution of novel feeding structures in Diplogastridae.

Research Highlights

Combining CRISPR-based mutagenesis, geometric morphometrics, and transcriptomics, we found that the genetic module governing the morphogenesis of novel feeding structures in diplogastrid nematodes is not conserved in the "rhabditid" *C. elegans*.

Introduction

Evolutionary developmental biology (evo-devo) investigates how organismal development evolves and in which ways developmental processes guide organismal evolution, with the ultimate aim of mechanistically understanding phenotypic transformations and the emergence of new traits (Minelli, 2015). In evo-devo, morphological novelties are defined based on structural homology and relative to ancestral character states (Wagner, 2014). Currently, two different types of novelties are broadly recognized: discrete novelties and individualized novelties. Discrete novelties are new structures which have no homologous counterpart in an ancestral species; individualized novelties constitute major deviations from one or more pre-existing traits which progressively evolved a new qualitative dimension or functional capacity (Wagner, 2014; Müller, 2021).

Nematodes (roundworms) have historically been used as model organisms to study myriad facets of biology and evolution (Corsi *et al.*, 2015; Sommer, 2015), given the invaluable characteristics of some species for functional genetics. Indeed, selected nematode model organisms mainly reproduce as self-fertilizing hermaphrodites and therefore produce isogenic populations. In addition, many nematodes have a stringent mode of development in which a small and invariable number of cells generate the body, a phenomenon known as eutely. These features made it possible to homologize the worm's anatomical traits down to the level of individual cells (Haag *et al.*, 2018) rendering eutelic model nematodes, such as the "rhabditid" *Caenorhabditis elegans* and the diplogastrid *Pristionchus pacificus*, promising subjects for comparative evo-devo studies of morphological character transformations.

Despite some superficial similarities in their simplified cylindrical body shape, nematodes are far from morphologically uniform. A closer look reveals a stunning diversity in cuticular feeding structures (Wright, 1976; Malakhov, 1994; Susoy *et al.*, 2015), which arguably explains how these microscopic animals were able to conquer nearly all ecosystems. When referring to the roundworm's mouth (stoma), nematologists traditionally distinguish between the buccal cavity (the lumen of the mouth) and the buccal capsule (the inner cuticular lining of the mouth) (Fürst von Lieven & Sudhaus, 2000). Based on the underlying cells which secrete it, the buccal capsule can be subdivided into individual cuticular feeding structures (from anterior to posterior): cheilostom, gymnostom, prostegostom, mesostegostom, metastegostom, and telostegostom (De Ley *et al.*, 1995). The cheilostom is secreted by skin cells (hypodermis) and is continuous with the outer body-wall cuticle of the worm. The gymnostom, which forms the middle part of the buccal capsule, was

found to be associated with the anterior and posterior arcade syncytia (De Ley *et al.*, 1995; Burr & Baldwin, 2016) - nematode-specific cells with enigmatic functions. The four posterior-most feeding structures are collectively referred to as the stegostom - the part of the buccal capsule that is made by pharyngeal cells. The prostegostom and mesostegostom are each produced by a set of "e" cells (pharyngeal epithelia); the metastegostom and telostegostom are each secreted by a set of "pm" cells (pharyngeal muscles) (De Ley *et al.*, 1995; Burr & Baldwin, 2016; Harry *et al.*, 2022). Interestingly, all of these elements are cuticular; they, just like the body-wall cuticle, constantly have to be remodeled, degraded and re-secreted, during the worm's periodic larval molts.

Some of these feeding structures differ drastically between the closely related "rhabditid" and diplogastrid nematodes (Figure 1a,b,d). "Rhabditids" represent a paraphyletic group (Figure 1e) whose members typically have elongated buccal capsules in which cheilo-, gymno-, pro-, and mesostegostoms form a rigid cuticular tube (Figure 1a). Their metastegostoms are shaped into simple cuticular triangles called flaps (Figure 1a,e) (Fürst von Lieven & Sudhaus, 2000). These structures are not actively moveable and likely function as valves, which prevent regurgitation of bacterial foods from the pharynx to the buccal cavity. In contrast, in Diplogastridae, a monophyletic clade that emerged from the paraphyletic "rhabditids" (Figure 1e), buccal capsules seldomly come in the form of rigid tubes and flaps do not exist (Fürst von Lieven & Sudhaus, 2000; Sudhaus, 2014). Instead, cheilostom and gymnostom are developed as cuticular rings which are telescoped into each other, giving the buccal capsule a barrel-like shape (Figure 1b,d). Strikingly, the metastegostom of diplogastrids evolved into large hook-shaped teeth (Figure 1b,d,e) that can actively be moved via pharyngeal muscle contractions and used for predatory feeding (Figure 1c) (Fürst von Lieven & Sudhaus, 2000; Susoy *et al.*, 2015; Sommer, 2015). Their pro- and mesostegostom evolved into slender and flexible hinges, which facilitate these movements relative to the remainder of the buccal capsule (compare Figure 1a and 1b) (Harry *et al.*, 2022).

Given nematode eutely, the individual cells which secrete the buccal capsule can be homologized. Thus, one can also directly homologize the individual feeding structures they produce, including the metastegostomatal teeth of diplogastrids and the flaps of the paraphyletic "rhabditids" (Figure 1a,b,e). A recent study reconstructed the buccal cavity of *P. pacificus* and showed that their feeding structures are indeed homologous to the ones in *C. elegans* (Harry *et al.*, 2022). Nonetheless, the movable teeth of *P. pacificus* and other diplogastrids are unlike any other feeding structure in nematodes. While they evolved from a pre-existing structure (the metastegostom), they have reached an extreme point of phenotypic

deviation from its ancestral state (flaps); and thereby developed into uniquely specialized traits that gained the quality of raptorial structures (hook-shaped teeth), which now serve a function their precursors did not, predation. Hence, the diplogastrid's teeth are an example of individualized novelty. Together, the fact that the evolutionary history of this trait can be reconstructed unambiguously (as robust phylogenetic backbones are available), and that these nematodes readily lend themselves for functional genetic experiments, renders the flap-to-tooth transformation that accompanied the "rhabditid"-to-diplogastrid transition a model system for studies on the molecular and developmental underpinnings of morphological novelties.

Revealing the molecular components of the gene regulatory network (GRN) which controls the development of diplogastrid feeding structures was a major focus of many recent genetic studies (for reviews see, Sommer *et al.*, 2017; Sommer, 2020). This is particularly due to the fact that most diplogastrids display a feeding polyphenism: in response to various environmental inputs such as pheromone signals (Werner *et al.*, 2018) or compositional differences in the media surrounding them (Werner *et al.*, 2017), these worms can develop either into narrow-mouthed bacterivores (stenostomatous morph; "St") or wide-mouthed predators (eurystomatous morph; "Eu") (Susoy *et al.*, 2015). In *P. pacificus*, both morphs carry a moveable metastegostomatal tooth at the anterior tip of their dorsal pharynx (Figure 1b,d). However, Eu morphs possess a second cuticular tooth at the anterior tip of the right subventral pharynx (Figure 1b), where St animals only possess a flat cuticle ridge (Fürst von Lieven & Sudhaus, 2000; Sommer, 2015; Theska *et al.*, 2020, Harry *et al.* 2022). By moving these two teeth against each other, Eu worms can rupture the cuticle of their prey (Figure 1c).

Work over the last decade has established that a modular GRN governs the development of these phenotypically plastic feeding structures (Figure 1f). In the center of this network sits a plasticity-switch module that controls which of the two alternative feeding morphologies will be developed (Sommer, 2020). It consists of an X-chromosomal multi-gene locus that is expressed in the central nervous system and the pharynx of the worm (Sieriebriennikov *et al.*, 2018). Upstream of the plasticity-switch module is an environment-perception module, which governs the activity of the plasticity-switch module in correspondence with environmental signals perceived during early larval development. This module works largely through the epigenetic mechanism of histone 4 lysine K 5/12 acetylation (H4K5/12ac) (Werner *et al.*, 2023). The GRN is completed by a phenotype-execution module that is downstream of the plasticity switch and controls the morphogenesis of the feeding structures (Sommer, 2020). The core of this module are two transcription factors

(TFs), the nuclear receptors NHR-1 and NHR-40, which control the expression of multiple protein classes involved in the modification and degradation of extracellular matrix and cuticle - the materials that make up the worm's feeding structures (Kieninger *et* al., 2016; Sieriebriennikov *et al.*, 2020) (Figure 1f). Recent studies also identified the mucin-type protein DPY-6 and the chitin synthase CHS-1 as members of this module (Sun & Theska *et al.*, 2022; Sun *et al.*, 2023). DPY-6 is the first known protein constituent of the buccal capsule and located mostly in the cheilostom. CHS-1 synthesizes chitin, and *Ppa-chs-1* mutants are teethless, showing for the first time that chitin is a structural constituent of nematode feeding structures (Sun *et al.*, 2023). However, the transcription factors, which regulate the expression of these cuticle-synthesis related components of the phenotype-execution module, have not yet been identified (Figure 1f).

Interestingly, the evolutionary histories of the genes establishing the *P. pacificus* mouth-form GRN are diverse. All known components of the plasticity-switch module are derived from recent gene-duplication events, which accompanied the flap-to-tooth transformation during the "rhabditid"-to-diplogastrid transition (Figure 1e,f) (Sieriebriennikov *et al.*, 2018; Sommer, 2020). As such, these young genes seemed to have gained their functions related to feeding-structure development immediately or early in their evolutionary history (Figure 1e). This is in stark contrast to some of the components of the phenotype-execution module. Both DPY-6 and CHS-1 are deeply-conserved proteins with functions related to buccal-capsule development, which predate the "rhabditid"-to-diplogastrid transition (Figure 1e) (Sun & Theska *et al.*, 2022; Sun *et al.*, 2023). Additionally, both NHR-1 and NHR-40 exist as one-to-one orthologs in *C. elegans* and *P. pacificus* (Figure 1g) and therefore already existed in their last common ancestor (LCA) (Figure 1e) (Sieriebriennikov *et al.*, 2020). Yet, the target genes that these receptors co-regulate to mediate feeding-structure development in diplogastrids are fast evolving (Sieriebriennikov *et al.*, 2020), even though the classes of cuticle-modifying proteins to which they belong (astacins, chitinases, and cysteine-rich secreted proteins) are ancient (Figure 1e). This contrast is even more stunning given that, in nematodes, NHRs are rapidly evolving, too (Robinson-Rechavi *et al.*, 2005; Sural & Hobert, 2021). Most of the *nhr* genes in the genomes of *C. elegans* and *P. pacificus* are entirely lineage-specific and only a small handful of them (like *nhr-1* and *nhr-40*) is conserved between these worms (Figure 1g) (Sieriebriennikov *et al.*, 2020). In *P. pacificus*, both of these receptors are expressed in the pharyngeal muscles, including the specific cells that secrete the metastegostomatal teeth (Figure 1h). Additionally, *nhr-1* is expressed in the dorsal pharyngeal gland (Sieriebriennikov *et al.*, 2020), which, in diplogastrids, is a massively enlarged gland cell that runs through the pharyngeal muscles, connects to

the metastegostom, and opens into the buccal cavity via a hole in the dorsal tooth (Figure 1h) (Riebesell & Sommer, 2017). Interestingly, all shared regulatory targets of NHR-1 and NHR-40 in *P. pacificus* are co-expressed in that precise cell, too (Figure 1i) (Sieriebriennikov *et al.*, 2020).

In light of the fact that the teeth of diplogastrids are individualized novelties that evolved from "rhabditid" flaps, we were eager to ask: does the entire phenotype-execution module - including the NHRs regulating cuticle-modifying enzymes - represent an ancient genetic cassette that is also functionally conserved as a mouth-development program in "rhabditids", such as *C. elegans*? This would imply that the feeding-structure GRN of diplogastrids evolved by building a new plasticity-switch module (which is made of evolutionarily young genes) upon an old mouth-development program. Alternatively, the two old NHRs may control a different biological function in "rhabditids". In this case, they might have been co-opted into a new regulatory context - that is borrowed for stoma morphogenesis - in diplogastrids. To address these possibilities, we set out to study whether these nuclear receptors are expressed in the stomatal tissues of "rhabditid" model nematode *C. elegans* and whether their functional loss would cause developmental aberrations in the mouth of these worms.

Material & methods

Nematode husbandry

Standard protocols for the maintenance of laboratory cultures of rhabditid nematodes (Stiernagle, 2006) were followed. Worms were cultured on 6cm Petri dishes with nematode growth medium (NGM). 300μl of *Escherichia coli* (OP50) were provided as a food source. Culture plates were stored at 20°C.

Light microscopy with differential interference contrast (DIC)

For microscopy, all specimens were mounted on object slides with 5% Noble Agar pads containing sodium azide (0.5% NaN_3) and subsequently examined using a Zeiss Axio Imager.Z1 microscope with a Zeiss Plan-Apochromate 100 x 1.4 DIC objective. Image stacks (Grayscale, 16-bit) were taken using a monochromatic Zeiss Axiocam 506 CCD camera. The Zen 2 Pro Software (version 2.0.14283.302) was used for digital microscopy and image acquisition.

CRISPR/CAS9 mutagenesis

We used the CRISPR/CAS9 system to create mutations in the coding sequences of *Cel-nhr-1* and *Cel-nhr-40*, specifically targeting their DNA-binding domains. We followed standard mutagenesis protocols for nematodes (Ghanta *et al.*, 2021). Target-specific CRISPR RNAs (crRNAs) were generated based on the genome sequence of *C. elegans* (www.wormbase.org). The CAS9 protein, the universal trans-acting CRISPR RNA (tracrRNA), and the crRNAs were ordered from IDT (Alt-R product line). Ribonucleoprotein complexes (RNPs) were created by combining 0.5μl of CAS9 nuclease (10μg/ml stock) with 5μl of tracrRNA (0.4μg/μl stock) and 2.8μl of crRNA (0.4μg/μl stock). This mix was incubated at 37°C for 15min and subsequently cooled to room temperature. The RNP mix was diluted by adding enough Tris-EDTA buffer to bring the total volume of the mix to 20μl. In order to avoid clogging in the injection needles, the final mix was centrifuged at 14,000rpm for 2min. Injections were conducted with an Eppendorf FemtoJet microinjector and an inverse Zeiss AxioVert microscope coupled to an Eppendorf TransferMan micromanipulator. RNP complexes were injected into the distal gonadal rachis of young adult hermaphrodites. Injected P0 animals were isolated and allowed to lay eggs for 24h at 20°C. After that, P0s were removed from the plates and F1 animals were left to develop. Upon reaching maturity, F1s were singled to fresh plates, allowed to lay eggs for 24h, and subsequently genotyped. DNA of candidate F1 animals was recovered via single worm lysis (SWL), target sites were PCR amplified (using the

Taq PCR Master Mix produced by Qiagen), and candidate mutations were identified using Sanger sequencing (performed by Azenta Life Sciences). If heterozygous F1 mutants were found, their F2 offspring were genotyped in a similar fashion to identify homozygous mutant carriers. Multiple homozygous mutant lines were isolated, including frameshift mutants with premature stop codons. Backups of these homozygous mutant strains were stored at -80°C and can be provided by the Sommer lab upon request.

Transgenesis and SEC-based protein tags

Our CRISPR-based transgenesis experiments largely followed the well-established SEC approach (Dickinson *et al.*, 2015) with minor modifications. In short, we used the original pDD268 plasmid to create a transcriptional reporter for *Cel-nhr-1* (pre-heatshock) and an N-terminally mNeonGreen(mNG)-tagged version of *Cel*-NHR-1 (post-heatshock). pDD268 was digested overnight with ClaI and SpeI in a thermocylcer set to 37°C. We excluded the native start codon of *Cel-nhr-1* from the 3'-homology arm in the donor plasmid in order to reduce the likelihood of internal splicing initiation in the mutant strain that expresses the mNG-tagged version of *Cel*-NHR-1 (RS3881). We used the InVitrogen HiFi Miniprep Kit (including the optional wash step with Qiagen PB buffer) to extract plasmids from 4ml bacterial overnight cultures. Digested SEC-plasmids and PCR-amplified homology arms were cleaned with the Qiaquick MinElute kit, assembled into donor plasmids with NEB's DNA HiFi Assembly Kit, and subsequently transformed into *E. coli* (NEB 5-alpha). We injected donor plasmids (20 ng/μl) together with prepared RNP complexes. Fluorescent worms were imaged using a Leica SP8 confocal laser-scanning microscope (cLSM).

Quantification of mouth-form differences among wild-type and mutant worms

We quantified morphological differences in the mouths of wild-type and mutant worms, using our recently published protocol (Theska et al., 2020) which combines two-dimensional landmark-based geometric morphometrics (GMM) with model-based clustering. We obtained image stacks of the nematode mouths in lateral position and recorded the X and Y coordinates of 15 fixed landmarks which capture all feeding structures using FIJI (ver. 2.1.0) (Schindelin et al., 2012). All steps of the GMM analysis were performed in R (ver. 4.3.0) (R Core Team, 2023) using a combination of the GEOMORPH (ver. 4.0.5) (Adams *et al.*, 2013; Baken *et al.*, 2021), MORPHO (ver. 2.11) (Schlager, 2017), and MCLUST (ver. 6.0.0.) packages (Scrucca *et al.*, 2016). General Procrustes Analysis (GPA) was performed with the *gpagen* function of

GEOMORPH, by minimizing Procrustes distances. We generated landmark data for 150 animals (50 per strain; three strains: N2, RS3635, RS3682), performed GPA, and subsequently checked for Procrustes distance outliers, using the *plotOutliers* function of GEOMORPH. Seven outliers were removed from the data for all subsequent steps of analysis. The outlier-corrected data set (n=143) contained landmark configurations from 44 wild-type animals (N2), 49 *Cel-nhr-1(tu1424)* mutant animals (RS3635), and 50 *Cel-nhr-40(tu1464)* mutant animals (RS3682). GPA was performed on the outlier-corrected data to obtain a shape matrix, to which we appended the log-transformed centroid size in order to generate a form data set (*sensu* Mitteroecker *et al.*, 2004). We subsequently performed a principal component analysis (PCA) on the form data set to visualize potential morphological differences among wild-type and mutant worms in a form space. Lastly, we assessed whether mutant worms could be classified as such and distinguished from wild-type worms, based on their mouth morphology. For this, we performed model-based clustering with the *Mclust* function of MCLUST. We only used "meaningful" principal components (mPCs) of form variation (identified with MORPHO's *getMeaningfulPCs* function) as input variables for clustering, in order to avoid overparameterization (Theska *et al.*, 2020). Clustering results were visualized by coloring each specimen according to cluster membership.

RNA-seq experiments

To facilitate RNA-seq experiments on our *Cel-nhr-1(tu1424)* and *Cel-nhr-40(tu1464)* mutants being directly comparable to the RNA-seq experiments our lab previously conducted on *Ppa-nhr-1* and *Ppa-nhr-40* mutants, we adopted the sample collection approach of Sieriebriennikov *et al.* (2020) with minor modifications. Animals were collected for RNA extraction at 33h (mostly L2 and a few L3 larvae), 44h (L3 and L4 larvae), and 58h (young adults) after bleaching; the presence and expected distribution of the desired stages at each time point was verified prior to sample collection by screening 50 specimens on each plate under a dissecting microscope. Total RNA was extracted using the Direct-Zol RNA Mini prep kit (Zymo Research) according to the instructions provided by the manufacturer. We combined 500ng of RNA extracted at the 33h time point with 500ng of RNA extracted at the 44h time point. This sample type thus contains a total of 1μg RNA extracted from L2-4 larvae. We gathered 1μg of RNA extracted at the 58h time point (adult worms) and did not mix it with RNA extracted from any other time point. Therefore, we ended up with two sample types ("larval" and "adult"), which we shipped to Novogene for library preparation and mRNA sequencing. Libraries were sequenced as 150bp paired end

reads on Illumina's NovaSeq6000 platform. We sequenced the mRNA of three strains: the wild-type *C. elegans* strain N2, the *Cel-nhr-1* frameshift mutant RS3635(*tu1424*), and the *Cel-nhr-40* frameshift mutant RS3682(*tu1464*). Two biological replicates were collected for each of the larval and adult samples of each of these three strains. The obtained raw reads have been deposited in the European Nucleotide Archive (ENA) under the study accession number PRJEB68220.

Quantification of transcripts and differential gene expression (DGE) analysis

Transcript abundances were estimated from raw read files relative to the *C. elegans* transcriptome using SALMON (ver. 1.5.2) (Patro *et al.*, 2017; Srivastava *et al.*, 2020) [84,85]. Using the entire *C. elegans* reference genome as a decoy, we built a decoy-aware transcriptome and indexed it for subsequent read mapping using an auxiliary *k*-mer hash (k=31). Reads were quantified against this index using Salmon's *quant* command by running the program in mapping-based mode with selective alignment as mapping strategy. We allowed Salmon to infer the library type, and to learn and correct for fragment-level GC-biases and sequence-specific biases. Differential expression analyses (DEA) were carried out in R using BIOCONDUCTOR (ver. 3.17) (Gentleman *et al.*, 2004), TXIMPORT (ver. 1.28.0) (Love *et al.*, 2015), and DESEQ2 (ver. 1.40.1) (Love *et al.*, 2014). Transcript-level abundance estimates generated by SALMON were summarized into gene-level count matrices using the *tximport* function. A DESeq data set was created with the *DESeqDataSetFromTximport* function and pre-filtered for transcripts which had at least ten counts across all samples. Stage-specific mutation effects on transcript abundances were modeled with the *DESeq* function. Obtained maximum likelihood estimates (MLE) of the log2-fold changes (L2FC) were shrunken with the *lfcShrink* function using the adaptive shrinkage estimator from the ASHR package (Stephens, 2017). Resulting minimum mean squared error (MMSE) estimates of the L2FC were reported for this study and used to visualize candidate target genes. We always included a L2FC-threshold of > 0.585 (absolute value) and an *s*-value threshold of < 0.001 into our hypothesis tests for differential expression. This is equivalent to asking whether there is sufficient evidence in the data that a given gene shows at least a 50% increase or decrease in expression levels, and that the direction of the observed effect was correctly estimated. Both as proof-of-concept and to facilitate comparability of datasets, we also used this pipeline to reanalyze the transcriptome data of wild-type *P. pacificus* (PS312), as well as *Ppa-nhr-1(-)* and *Ppa-nhr-40(-)* mutants obtained by Sieriebriennikov *et al.* (2020). Results of the DEAs can be found in Sup. Data 2 and 3.

Pfam domain prediction, WormCat annotation, and overrepresentation analyses (ORA)

Protein domains encoded in the proteome of *C. elegans* were predicted using HMMER (ver. 3.3.2) (Mistry *et al.*, 2013) in conjunction with the Pfam-A.hmm database (ver. 3.1b2) (Finn *et al.*, 2016) (hmmscan, e-value cutoff: < 0.001). WormCat (ver. 2) annotations were used to explore biological functions of target genes (Holdorf *et al.*, 2020). ORTHOFINDER (default mode; ver. 2.5.4) (Emms & Kelly, 2019) was used to define orthogroups between *P. pacificus* and *C. elegans* proteins, based on which WormCat annotations for *P. pacificus* were created. ORAs were based on Fisher's Exact test; WormCat terms were considered overrepresented if FDR-corrected *P*-values were below a type I error rate of 5% ($P < 0.05$). Results of the ORAs can be found in Sup. Data 4 and 5.

Data processing, data visualization, and illustration

All data was processed, formatted, and wrangled in R (ver. 4.3.0) (R Core Team, 2023) using the TIDYVERSE (Wickham *et al.*, 2019). All scientific plots were created using GGPLOT2 (ver. 3.4.2) (Wickham *et al.*, 2019). In order to make our plots accessible to people with color-vision deficiencies or color blindness, we used the scientifically-derived colormaps provided by the SCICO (ver. 1.4.0) and VIRIDIS (ver. 0.6.3) packages (Pedersen & Crameri, 2023; Garnier *et al.*, 2023). Microscopic images were edited and adjusted for color, levels, lightness, and contrast in Affinity Photo (ver. 1.10.5). Scientific illustrations and final figures were created in Affinity Designer (ver. 1.10.5).

Results

Cel-nhr-1 is broadly expressed in somatic tissues including pharyngeal muscles

First, we wondered whether *C. elegans's* copies of *nhr*-1 and *nhr*-40 are expressed in the same tissues in which they are expressed in *P. pacificus*. Previous studies have shown that *Cel-nhr-40* is indeed expressed in the anterior pharyngeal muscles of *C. elegans*, as well as other tissues including head neurons, skin, and body wall muscles (Brožová *et al.*, 2006; Taylor *et al.*, 2021). Thus, the currently available data suggests that the expression of *nhr-40* in the pharynx - the tissue which secrets the metastegostom - is evolutionarily conserved across "rhabditids" and diplogastrids. We speculated that the same holds true for *nhr-1*, which would indicate that both receptors are co-expressed in the mouth-associated tissues in flap-bearing "rhabditid" nematodes. To verify this, we studied the expression of *nhr-1* in *C. elegans* via the creation of transcriptional and translational reporter lines.

Our transcriptional reporter lines revealed that *Cel-nhr-1* is broadly expressed in multiple somatic tissues at varying intensities. We found high expression levels in the entire skin (hypodermis) and body wall musculature from head to tail, as well as in the majority of the worm's pharyngeal muscles (Figure 2a). The latter includes pm3-5, the cells which make up most of the muscular portions of the procorpus, metacorpus, and isthmus (Figure 2b). Relatively lower expression levels were found in the stomatal tissues (aa, ap, e1, e3, pm1, pm2), the terminal bulb of the pharynx (pm6-8), the intestine, and some ventral aspects of the somatic gonad (Figure 2a,b,c). Furthermore, animals carrying the translational reporter express a fluorescently labeled version of NHR-1 that is restricted to the nuclei of the cells in which *Cel-nhr-1* is expressed (Figure 2d). Specifically, using these localization patterns, we were able to identify the nuclei of the stomatal and pharyngeal muscle cells for which we already have seen expression signals in the transcriptional reporter line (Figure 2b,c). These results confirm the presence of *Cel*-NHR-1 protein in exactly those cells that secrete the posterior buccal capsule (Figure 2d). Additionally, we found *Cel*-NHR-1 in the nuclei of multiple nerve-ring neurons, possibly including amphid-associated neurons or glia cells (Figure 2d).

Taken together, the expression and protein-localization patterns of *Cel*-NHR-1 are highly consistent with the ones previously reported for *Cel*-NHR-40. Most importantly, both receptors are co-expressed in the anterior pharyngeal muscle cells in both *P. pacificus* and *C. elegans*, suggesting that this expression pattern was indeed already established in their last common ancestor. This, together with the fact that

NHR-1 and NHR-40 co-regulate mouth development in *P. pacificus*, suggests that these receptors may in fact govern similar processes in *C. elegans*. To investigate this possibility, we created knock-out alleles for both of these receptors to see if that causes altered mouth morphologies reminiscent of those induced by the same kind of mutations in *P. pacificus* (Kieninger *et al.*, 2016; Sieriebriennikov *et al.*, 2020).

Loss of *nhr-1* or *nhr-40* does not disrupt mouth morphogenesis in *C. elegans*

We obtained multiple independent homozygous mutant alleles of *Cel-nhr-1* and *Cel-nhr-40* by using CRISPR/CAS9 to introduce frameshift mutations into their gene bodies, which truncate the receptor proteins at the site of their DNA-binding domain (DBD) and render them non-functional. All of these homozygous mutants were viable, fertile, and showed no obvious morphological alterations. However, as the mouths of these nematodes are only a few micrometers long and wide (Figure 1a), which makes the identification of potential structural alterations difficult, we used landmark-based geometric morphometrics to quantify potential mutant phenotypes in mouth morphology (Figure 3c) (Theska *et al.*, 2020). Two frameshift alleles, *Cel-nhr-1(tu1424)* and *Cel-nhr-40(tu1464)*, were chosen as reference alleles and used for all subsequent experiments and analyses (Figure 3a,b).

Surprisingly, we found that the loss of neither of these two NHRs affects the overall morphology of the adult mouth. A principal component analysis (PCA) of mouth-form data from 143 individuals revealed that, in the morphospace produced by PC1 and PC2 (which capture the majority of form variation in the data), wild-type worms and *nhr* mutants show homogenous and strongly overlapping ranges of morphological variation in their mouths (Figure 3d). This finding corroborated the superficial observation that wild-type and mutant worms were indistinguishable based on their morphology. We were able to estimate that PC1 is the only "meaningful" principal component (*sensu* Schlager, 2017) and that the other PCs are likely to contain noise rather than signal. Subsequent model-based clustering of all individuals based on the morphological variation contained in PC1 revealed that neither *Cel-nhr-1(tu1424)* nor *Cel-nhr-40(tu1464)* mutants could be classified as such or differentiated from wild-type worms in terms of their mouth morphology (Figure 3e) (note that this also holds true if clustering is perform based on the variation in PC2, or PC1 and PC2 together). Thus, geometric morphometric and model-based clustering analyses coherently suggest the absence of discernible malformations in the mouths of the two *nhr*-mutant strains when compared to wild-type animals (Figure 3d,e).

These findings are in contrast with the phenotypes caused by the loss of the same receptors in *P. pacificus,* which either lead to blatant structural malformations in

the adult mouth or affect which of the worm's alternative mouth morphologies is developmentally executed (Sieriebriennikov *et al.*, 2020). Thus, the results of our morphological analyses in *C. elegans* did not support the hypothesis that NHR-1 and NHR-40 have a conserved biological function in the regulation of mouth morphogenesis across diplogastrids and "rhabditids". Given that both receptors are deeply conserved between these worms and actively expressed in the same tissues, we wondered whether the absence of mouth-form related phenotypes in *Cel-nhr-1(tu1424)* and *Cel-nhr-40(tu1464)* mutants might be explained by differences in the regulatory targets of these receptors between the two species.

NHR-1 and NHR-40 have largely different regulatory targets between *P. pacificus* and *C. elegans*

First, we wondered whether NHR-1 and NHR-40 have any common regulatory targets in *C. elegans*, just as they do in *P. pacificus*. To determine this, we first reanalyzed the transcriptomes of the *Ppa-nhr-1(tu1163)*, *Ppa-nhr-1(tu1164)*, *Ppa-nhr-40(tu1418)*, and *Ppa-nhr-40(tu1423)* mutants obtained by Sieriebriennikov *et al.* (2020) alongside the transcriptomes of our *Cel-nhr-1(tu1424)* and *Cel-nhr-40(tu1464)* mutants. The purpose of the reanalysis of the *P. pacificus* data was two-fold: we wanted to analyze *nhr*-mutant transcriptomes in both species with the exact same pipeline and apply more stringent criteria in our differential expression analyses than in previous studies (e.g. including minimum effect sizes) (Figure 4 and Supplementary Figure 1).

The reanalysis of the *P. pacificus* mutant data largely corroborated the original findings and removed only a few putative targets. Sieriebriennikov *et al.* (2020) found, amongst a few others, 12 metallopeptidases (astacins), five secreted proteins (CAP-domain proteins), and two chitinases to be the major common regulatory targets of *Ppa*-NHR-1 and *Ppa*-NHR-40. Our reanalysis also recovered 12 metallopeptidases (Supplementary Figure 1) of which 10 were among the previously reported; all of them were found to be down-regulated in larval and adult worms. We also recovered the two chitinases and one of the secreted proteins among the common targets that are consistently mis-regulated throughout development (Supplementary Fig 1). Thus, this reanalysis of the *P. pacificus* transcriptome data both validated our computational pipeline and demonstrated that the original results by Sieriebriennikov *et al.* (2020) were robust against more stringent analytical cutoffs.

By performing similar analyses on the *C. elegans* transcriptome data we found that, just like in *P. pacificus*, *Cel*-NHR-1 and *Cel*-NHR-40 do indeed share a group of regulatory targets (Figure 4). A total of 17 genes are constitutively mis-regulated in

both larvae and adults of *Cel-nhr-1(tu1424)* and *Cel-nhr-40(tu1464)* mutants, of which nine were down-regulated and eight up-regulated compared to wild-type worms. Strikingly, we found that the compositional patterns of the shared regulatory targets of these receptors is highly species-specific (Figure 4 and Supplementary Figure 1).

First, the comparative analysis revealed that targets which are shared by NHR-1 and NHR-40 in a stage-independent manner were biased towards a specific chromosome in either species. Where they were biased towards the X chromosome in *P. pacificus* (P = 1.5e-07, Supplementary Fig 1), they were biased towards chromosome V in *C. elegans* (P = 0.00011, Figure 4). These chromosomes are non-homologous (Yoshida *et al.*, 2023).

Second, we did not find a single metallopeptidase or chitinase among the stage-independently shared targets of NHR-1 and NHR-40 in *C. elegans*. Thus, with the exception of a single CAP-domain-containing secreted protein (*vap-1*), none of the protein classes which were found to constitute the major regulatory targets of both receptors in *P. pacificus* could be found in a similar pool of targets for same receptors in *C. elegans*. Intriguingly, we found that 12 of the 17 shared regulatory targets of NHR-1 and NHR-40 in the latter species currently have no assigned biological roles, although six of them are at least known to be regulated by various stress factors (Figure 4). We wondered whether any of these biologically uncharacterized genes encode known protein family (pfam) domains (Finn *et al.*, 2016), which could at least hint at their biochemical functions. We were able to predict domains in the amino acid sequences encoded by only five of the biologically uncharacterized genes, and just a single one of them - a thrombospondin (*F11C7.2*) - contains a domain that is somewhat understood. In contrast, the other four domain-containing proteins possess seemingly nematode-specific domains of unknown function (DUF19 and DUF4473). Therefore, most of NHR-1's and NHR-40's shared regulatory targets in *C. elegans* currently remain uncharacterized both at the biological and biochemical levels. Amongst the few shared targets for which a biological function is already known, we found two lipid-binding proteins with metabolic roles (*far-3* and *far-8*), a neuropeptide-like protein (*nlp-34*), and a serine protease inhibitor (*Y49G5A.1*) (Figure 4). Seemingly, these genes are not easily grouped into any overarching biological process that might be co-regulated by NHR-1 and NHR-40 in this species. Yet, all but one of their shared regulatory targets are known to be expressed predominantly in glia cells, sensory neurons, interneurons, motoneurons, or combinations thereof (Figure 4) (Taylor *et al.*, 2021). Therefore, we speculate that, by co-regulating the neuronal expression of a pool of shared targets with hitherto unassigned biological roles, NHR-1 and NHR-40 might

jointly control the execution of a neuronal (and possibly sensory) response in *C. elegans*.

Third, stage-specific overrepresentation analyses for functional categories amongst the identified target genes of NHR-1 and NHR-40 revealed that these receptors control essentially the same kind of regulatory targets (metallopeptidases, chitinases, and secreted proteins) throughout the development of *P. pacificus* and that only a few targets are specific to either a single receptor or developmental stage (Figure 5). Interestingly, the same does not hold true for *C. elegans*. In fact, the genes with currently unassigned biological functions were the only category that was overrepresented among the targets of both mutants in a stage-independent manner (Figure 5). These genes aside, we found multiple stage-specific common targets of NHR-1 and NHR-40 in *C. elegans*. For example, approximately half of all genes which encode major sperm proteins are down-regulated in larvae, and more than a fifth of all collagen-encoding genes were down-regulated in adults of both *nhr* mutants. Other adult-specific common targets include hedgehog-like proteins, acid phosphatases involved in lysosomal protein breakdown, and functionally uncharacterized prions. Almost all of these targets were consistently down-regulated in both mutant strains. The only adult-specific group of shared putative targets that is consistently up-regulated are CUB-domain proteins (Figure 5), which play important roles in the worm's innate immune responses to pathogenic bacteria and fungi (Holdorf *et al.*, 2020). Furthermore, our analyses revealed multiple receptor-specific target groups. In adult *Cel-nhr-1(tu1424)* mutants, broad aspects of amino acid and lipid metabolism are mis-regulated. This includes amino-acid breakdown, lipolysis, fatty-acid synthesis, and β-oxidation. Similarly affected are genes encoding enzymes with broad metabolic roles like flavin-containing monooxygenases and short chain dehydrogenases (Figure 5). Additionally, various stress reactions such as detoxification via p450 cytochromes and UDP-glucuronossyltransferases, immune responses involving C-type lectins, and unfolded protein responses in the endoplasmic reticulum seem to require NHR-1, too (Figure 5). NHR-40, on the other hand, seems to regulate some proteolytic processes in adult worms, including (surprisingly) a small number of metallopeptidases - the main targets of both NHR-1 and NHR-40 in *P. pacificus* (Figure 5).

Discussion

In this study, we took a comparative evo-devo approach to elucidate the evolutionary history of a GRN that underlies the development of novel feeding structures in diplogastrid nematodes. Using *C. elegans* as a representative member of "rhabditids" (the paraphyletic outgroup of flap-carrying nematodes from which the Diplogastridae emerged), we investigated whether the mouth-morphogenesis related functions of NHR-1 and NHR-40 (Figure 1f) predated the emergence of diplogastrid teeth, or whether they evolved in conjunction with them (Figure 1e).

We anticipated to find that these two transcription factors have a conserved function in feeding-structure development between *C. elegans* and *P. pacificus*, based on a combination of observations: First, the teeth of diplogastrids are individualized novelties that evolved from "rhabditid" flaps. As such, both of these feeding structures can be homologized and traced back to the metastegostom of their last common ancestor (LCA). Second, the two molecular players of the phenotype-execution module, which are involved in cuticle synthesis, DPY-6 and CHS-1, have already been found to exert conserved functions in *C. elegans* and *P. pacificus* (Sun & Theska *et al.*, 2022; Sun *et al.*, 2023), further corroborating the hypothesis that this module was established before the flap-to-tooth transformation. Lastly, the two NHRs, which control the aspect of the phenotype-execution module that governs cuticle modification and degradation, are deeply conserved between *C. elegans* and *P. pacificus*, too (Sieriebriennikov *et al.*, 2020), and thus likely to regulate similar developmental processes in both species.

However, the results of the integrative analysis described in this study do not support our initial expectations. Despite the fact that the expression studies on *Cel*-NHR-1, together with the available expression data on *Cel*-NHR-40, indicate that both receptors are co-expressed in the tissues which give rise to the feeding structures of both *C. elegans* and *P. pacificus*, genetic inactivation of either receptor does not disrupt or otherwise affect feeding-structure development in the "rhabditid" species. Therefore, feeding-structure development in *C. elegans* and *P. pacificus* is not controlled by the exact same genetic module. This observation was further supported by the comparative transcriptomic experiments, which revealed that NHR-1 and NHR-40 have highly species-specific types of regulatory targets and thus control different biological processes in *C. elegans* and *P. pacificus*.

Throughout *P. pacificus's* ontogeny, both receptors co-regulate the expression of a narrow range of "core targets", which is essentially restricted to the chitinases, metallopeptidases, and secreted proteins involved in the cuticle-modifying and

-degrading processes of feeding-structure development. The same receptors also share numerous regulatory targets in *C. elegans*, but, as opposed to their *P. pacificus* counterparts, they represent functionally more diverse groups of genes that are often co-regulated in a stage-specific manner (Figure 5). Additionally, in *C. elegans*, both receptors regulate various non-overlapping biological processes. Interestingly, none of *C. elegans's* chitinase genes were affected by the loss of either receptor at any stage of development. Yet, we found that five secreted proteins displayed altered expression levels in adult *Cel-nhr-1(tu1424)* mutants (of which one is an ortholog to a core target of *Ppa*-NHR-1 and *Ppa*-NHR-40), and that five metallopeptidases were misregulated in adult *Cel-nhr-40(tu1464)* mutants (of which one is an ortholog to two of *Ppa*-NHR-1's and *Ppa*-NHR-40's core targets). Besides the differences between the NHR-1 and NHR-40 targets in these two species, there are some correlations of currently unknown significance. For example, independent of the developmental stage, both receptors co-regulate the expression of small pools of genes that tend to be located on a particular chromosome (X in *P. pacificus* vs. V in *C. elegans*) and co-expressed in specific cell types (pharyngeal gland in *P. pacificus* vs. glia and neurons in *C. elegans*). Furthermore, and adding to previous reports (Sieriebriennikov *et al.*, 2020), we found that the expression levels of either receptor remained unaffected by the inactivation of the other, at any stage of development in both species. This suggests that, within their respective GRNs, NHR-1 and NHR-40 do not activate or repress each other transcriptionally, but that they might co-regulate species-specific pools of shared target genes via post-transcriptional interactions, possibly by binding to the same promotor regions as heterodimers.

Taken together, the findings presented in this study could support two different evolutionary scenarios. Either, the phenotype-execution module which governs feeding-structure morphogenesis in diplogastrids indeed already existed in the LCA of *C. elegans* and *P. pacificus*. In this scenario, NHR-1 and NHR-40 used to control the expression of many metallopeptidases, chitinases, and secreted proteins in order to modify feeding structures throughout the ontogeny of "rhabditids" and diplogastrids. This, in turn, would mean that the entire phenotype-execution module (in its diplogastrid-like form) already existed before the Diplogastridae and their teeth evolved, and that *C. elegans* happens to be a "rhabditid" species in which this module secondarily disintegrated as NHR-1 and NHR-40 acquired new regulatory targets. In this case, the small number of metallopeptidases and secreted proteins we found to be misregulated in adults of either *Cel-nhr-1(tu1424)* or *Cel-nhr-40(tu1464)* mutants may reflect minor remnants of this ancestral genetic module. Alternatively, the phenotype-execution module of *P. pacificus's* feeding-structure GRN assembled

and gained its function in conjunction with the emergence of Diplogastridae. In this scenario, the absence of discernible mouth-morphogenesis related functions for the two NHRs in *C. elegans* would be explained by the fact these receptors used to control different biological processes in the LCA of "rhabditids" and diplogastrids (probably ones akin to the metabolic and physiological processes we found to be controlled by them in *C. elegans*). However, final conclusions regarding the ancestral regulatory functions of NHR-1 and NHR-40 cannot yet be drawn, since our study only addressed their roles in *C. elegans.* In order to established character polarity for the regulatory functions of these receptors (as they differ so clearly between the two species for which they are now known), future studies will have to determine the targets of NHR-1 and NHR-40 in additional "rhabditid" species more closely and more distantly related to the Diplogastridae than *C. elegans*. This way, we will learn where the phenotype-execution module of the diplogastrids emerged, and whether *Caenorhabditis* nematodes secondarily lost this module or never possessed it in the first place. This, however, is complicated by the fact that it is currently unclear which of the "rhabditids" constitutes the sister taxon of Diplogastridae, even though current studies suggest that it is *Rhabditoides* (van Megen *et al.*, 2009; Sudhaus, 2014).

In summary, the currently available data show that the phenotype-execution module that underlies feeding-structure development in diplogastrids is not functionally conserved between *P. pacificus* and *C. elegans*, suggesting that it was subject to developmental systems drift (Sommer, 2012; Haag *et al.*, 2018) during the flap-to-tooth transformation that accompanied the "rhabditid"-to-diplogastrid transition. We speculate that the two ancient transcription factors NHR-1 and NHR-40 were likely co-opted into a new regulatory context when the diplogastrid teeth evolved, given that most of the genes in the feeding-structure GRN (including many of NHRs' shared targets) derived from diplogastrid-specific gene duplication events (Sieriebriennikov *et al.*, 2020). In any case, the numerous details surrounding the evolutionary history of the flap-to-tooth transformation in diplogastrids highlight the complexity of the developmental genetic processes which underlie morphological evolution (Carroll, 2008), and are in line with similarly complex evolutionary histories identified in other evo-devo systems, such as the evolution of novel extraembryonic tissues in insects (van der Zee *et al.*, 2005; Panfilio *et al.*, 2013).

Acknowledgements

The plasmid pDD268 was a gift from Daniel Dickinson and Bob Goldstein (Addgene plasmid #132523; http://n2t.net/addgene:132523; RRID:Addgene_132523). We thank the Light Microscopy facility at the Max Planck Institute for Biology, for their assistance with cLSM. We are grateful to Heike Haussmann for freezing samples of all the worm strains produced for this study. Ziduan Han and Bogdan Sieriebriennikov provided expertise on molecular biological approaches in early stages of this project. Additionally, we thank Christian Rödelsperger, Devansh R. Sharma, Michael S. Werner, and James Lightfoot for helpful discussions of our work. We are grateful to Tess Renahan for critically reading this manuscript and for valuable suggestions throughout all stages of the project. T.T. was supported by the International Max Planck Research School (IMPRS) 'From Molecules to Organisms'.

Competing interests

The authors declare no competing interests.

Data availability

Raw reads from the mRNA-seq experiments have been deposited in the European Nucleotide Archive (ENA) under the accession number PRJEB68220. Landmark data for geometric morphometrics, results of the differential expression analyses, and results of the overrepresentation analyses are provided in the supplementary material (Supplementary Data 1-5).

References

Adams, D. C., & Otárola-Castillo, E. (2013). geomorph: An R package for the collection and analysis of geometric morphometric shape data. *Methods in Ecology and Evolution*, *4*(4), 393–399.

Baken, E. K., Collyer, M. L., Kaliontzopoulou, A., & Adams, D. C. (2021). Geomorph v4. 0 and gmShiny: Enhanced analytics and a new graphical interface for a comprehensive morphometric experience. *Methods in Ecology and Evolution*, *12*(12), 2355–2363.

Brožová, E., Šimečková, K., Kostrouch, Z., Rall, J. E., & Kostrouchová, M. (2006). NHR-40, a *Caenorhabditis elegans* supplementary nuclear receptor, regulates embryonic and early larval development. *Mechanisms of Development*, *123*(9), 689–701.

Burr, A., & Baldwin, J. G. (2016). The nematode stoma: Homology of cell architecture with improved understanding by confocal microscopy of labeled cell boundaries. *Journal of Morphology*, *277*(9), 1168–1186.

Carroll, S.B. (2008). Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell, 134,* 25–36.

Corsi, A. K., Wightman, B., & Chalfie, M. (2015). A transparent window into biology: A primer on *Caenorhabditis elegans*. *Genetics*, *200*(2), 387–407.

De Ley, P., Van de Velde, M.C., Mounport, D., Baujard, P. & Cooman, A. (1995) Ultrastructure of the stoma in Cephalobidae, Panagrolaimidae and Rhabditidae, with a proposal for a revised stoma terminology in Rhabditida (Nematoda). *Nematologica, 41*(1-4), 153–182.

Dickinson, D. J., Pani, A. M., Heppert, J. K., Higgins, C. D., & Goldstein, B. (2015). Streamlined genome engineering with a self-excising drug selection cassette. *Genetics*, *200*(4), 1035–1049.

Emms, D. M., & Kelly, S. (2019). OrthoFinder: Phylogenetic orthology inference for comparative genomics. *Genome Biology*, *20*, 1–14.

Finn, R. D., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Mistry, J., Mitchell, A. L., Potter, S. C., Punta, M., Qureshi, M., & Sangrador-Vegas, A. (2016). The Pfam protein families database: Towards a more sustainable future. *Nucleic Acids Research*, *44*(D1), D279–D285.

Fürst von Lieven, A., & Sudhaus, W. (2000). Comparative and functional morphology of the buccal cavity of Diplogastrina (Nematoda) and a first outline of the phylogeny of this taxon. *Journal of Zoological Systematics and Evolutionary Research*, *38*(1), 37–63.

Garnier, S., Ross, N., Rudis, R., Camargo, P. A., Sciaini, M., & Scherer, C. (2021). viridis—Colorblind-friendly color maps for R. *R Package Version 0.6*, *1*.

Gentleman, R. C., Carey, V. J., Bates, D. M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., & Gentry, J. (2004). Bioconductor: Open software development for computational biology and bioinformatics. *Genome Biology*, *5*(10), 1–16.

Ghanta, K. S., Ishidate, T., & Mello, C. C. (2021). Microinjection for precision genome editing in Caenorhabditis elegans. *STAR Protocols*, *2*(3), 100748.

Haag, E. S., Fitch, D. H., & Delattre, M. (2018). From "the worm" to "the worms" and back again: The evolutionary developmental biology of nematodes. *Genetics*, *210*(2), 397–433.

Harry, C. J., Messar, S. M., & Ragsdale, E. J. (2022). Comparative reconstruction of the predatory feeding structures of the polyphenic nematode Pristionchus pacificus. *Evolution & Development*, *24*(1–2), 16–36.

Holdorf, A. D., Higgins, D. P., Hart, A. C., Boag, P. R., Pazour, G. J., Walhout, A. J., & Walker, A. K. (2020). WormCat: An online tool for annotation and visualization of *Caenorhabditis elegans* genome-scale data. *Genetics*, *214*(2), 279–294.

Howard, R.J., Giacomelli, M., Lozano-Fernandez, J., Edgecombe, G.D., Fleming, J.F., Kristensen, R.M., … & Pisani, D. (2022). The Ediacaran origin of Ecdysozoa: integrating fossil and phylogenomic data. J*ournal of the Geological Society*, *179*(4):jgs2021–2107.

Kieninger, M. R., Ivers, N. A., Rödelsperger, C., Markov, G. V., Sommer, R. J., & Ragsdale, E. J. (2016). The nuclear hormone receptor NHR-40 acts downstream of the sulfatase EUD-1 as part of a developmental plasticity switch in *Pristionchus*. *Current Biology*, *26*(16), 2174–2179.

Love, M. I., Anders, S., Kim, V., & Huber, W. (2015). RNA-Seq workflow: Gene-level exploratory analysis and differential expression. *F1000Research*, *4*, 1070.

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12), 1–21.

Malakhov, V.V. (1994). *Nematodes: Structure, development, classification and phylogeny*. Smithsonian Books.

Minelli, A. (2015). Grand challenges in evolutionary developmental biology. *Frontiers in Ecology and Evolution*, *2*, 85, 1–11.

Mistry, J., Finn, R. D., Eddy, S. R., Bateman, A., & Punta, M. (2013). Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. *Nucleic Acids Research*, *41*(12), e121–e121.

Mitteroecker, P., Gunz, P., Bernhard, M., Schaefer, K., & Bookstein, F. L. (2004). Comparison of cranial ontogenetic trajectories among great apes and humans. *Journal of Human Evolution*, *46*(6), 679–698.

Müller, G. B. (2021). Developmental Innovation and Phenotypic Novelty. *Evolutionary Developmental Biology: A Reference Guide,* Nuño de la Rosa, L., Müller, G.B. (Ed.) (pp.69–84). Springer.

Panfilio, K.A., Oberhofer, G., & Roth, S. (2013). High plasticity in epithelial morphogenesis durin insect dorsal closure. *Biology Open, 2*(11), 1108–1118.

Patro, R., Duggal, G., Love, M. I., Irizarry, R. A., & Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. *Nature Methods*, *14*(4), 417–419.

Pedersen, T. L., & Crameri, F. (2023). Scico: Colour palettes based on the scientific colour-maps. *R Package Version 1.4.0*.

R Core Team, R. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Riebesell, M., & Sommer, R.J. (2017). Three-dimensional reconstruction of the pharyngeal gland cells in the predatory nematode *Pristionchus pacificus*. *Journal of Morphology, 278*(12), 1656–1666.

Robinson-Rechavi, M., Maina, C. V., Gissendanner, C. R., Laudet, V., & Sluder, A. (2005). Explosive lineage-specific expansion of the orphan nuclear receptor HNF4 in nematodes. *Journal of Molecular Evolution*, *60*, 577–586.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., & Schmid, B. (2012). Fiji: An open-source platform for biological-image analysis. *Nature Methods*, *9*(7), 676–682.

Schlager, S. (2017). Morpho and Rvcg–Shape Analysis in R: R-Packages for geometric morphometrics, shape analysis and surface manipulations. In Zheng, G., Li, S., Szekely, G. (Ed.), *Statistical Shape and Deformation Analysis* (pp. 217–256). Elsevier.

Scrucca, L., Fop, M., Murphy, T. B., & Raftery, A. E. (2016). mclust 5: Clustering, classification and density estimation using Gaussian finite mixture models. *The R Journal*, *8*(1), 289.

Sieriebriennikov, B., Prabh, N., Dardiry, M., Witte, H., Röseler, W., Kieninger, M. R., Rödelsperger, C., & Sommer, R. J. (2018). A developmental switch generating phenotypic plasticity is part of a conserved multi-gene locus. *Cell Reports*, *23*(10), 2835–2843.

Sieriebriennikov, B., Sun, S., Lightfoot, J. W., Witte, H., Moreno, E., Rödelsperger, C., & Sommer, R. J. (2020). Conserved nuclear hormone receptors controlling a

novel plastic trait target fast-evolving genes expressed in a single cell. *PLoS Genetics*, *16*(4), e1008687.

Sommer, R. J. (2012). Evolution of regulatory networks: Nematode vulva induction as an example of developmental systems drift. In Soyer, O. (Ed.), *Evolutionary Systems Biology* (pp. 79–91). Springer.

Sommer, R. J. (Ed.) (2015). *Pristionchus pacificus: A nematode model for comparative and evolutionary biology* (Vol. 11). Leiden, The Netherland: Brill.

Sommer, R. J. (2020). Phenotypic plasticity: From theory and genetics to current and future challenges. *Genetics*, *215*(1), 1–13.

Sommer, R. J., Dardiry, M., Lenuzzi, M., Namdeo, S., Renahan, T., Sieriebriennikov, B., & Werner, M. S. (2017). The genetics of phenotypic plasticity in nematode feeding structures. *Open Biology*, *7*(3), 160332.

Srivastava, A., Malik, L., Sarkar, H., Zakeri, M., Almodaresi, F., Soneson, C., Love, M. I., Kingsford, C., & Patro, R. (2020). Alignment and mapping methodology influence transcript abundance estimation. *Genome Biology*, *21*(1), 1–29.

Stephens, M. (2017). False discovery rates: A new deal. *Biostatistics*, *18*(2), 275–294.

Stiernagle, T. (2006). Maintenance of *C. elegans*. In the *C. elegans* research community (Ed.), *WormBook: The Online Review of C. elegans Biology*.

Sudhaus, W. (2014). 7.17 Order Rhabditina: "Rhabditidae". In Schmidt-Rhaesa, A. (Ed.), *Volume 2 Nematoda* (pp. 537-556). Berlin, Boston: De Gruyter.

Sun, S., Theska, T., Witte, H., Ragsdale, E. J., & Sommer, R. J. (2022). The oscillating Mucin-type protein DPY-6 has a conserved role in nematode mouth and cuticle formation. *Genetics*, *220*(3), iyab233.

Sun, S., Witte, H., & Sommer, R. J. (2023). Chitin contributes to the formation of a feeding structure in a predatory nematode. *Current Biology*, *33*(1), 15–27.

Sural, S., & Hobert, O. (2021). Nematode nuclear receptors as integrators of sensory information. *Current Biology*, *31*(19), 4361–4366.

Susoy, V., Ragsdale, E. J., Kanzaki, N., & Sommer, R. J. (2015). Rapid diversification associated with a macroevolutionary pulse of developmental plasticity. *Elife*, *4*, e05463.

Taylor, S. R., Santpere, G., Weinreb, A., Barrett, A., Reilly, M. B., Xu, C., Varol, E., Oikonomou, P., Glenwinkel, L., & McWhirter, R. (2021). Molecular topography of an entire nervous system. *Cell*, *184*(16), 4329–4347.

Theska, T., Sieriebriennikov, B., Wighard, S. S., Werner, M. S., & Sommer, R. J. (2020). Geometric morphometrics of microscopic animals as exemplified by model nematodes. *Nature Protocols*, *15*(8), 2611–2644.

van der Zee, M., Berns, N., & Roth, S. (2005). Distinct functions of the *Tribolium zerknüllt* genes in serosa specification and dorsal closure. *Current Biology*, *15*(7), 624–636.

van Megen, H., van den Elsen, S., Holterman, M., Karssen, G., Mooyman, P., Bongers, T., ... & Helder, J. (2009). A phylogenetic tree of nematodes based on about 1200 full-length small subunit ribosomal DNA sequences. *Nematology*, *11*(6), 927–950.

Wagner, G. P. (2014). *Homology, Genes, and Evolutionary Innovation*. Princeton University Press.

Werner, M. S., Claaßen, M. H., Renahan, T., Dardiry, M., & Sommer, R. J. (2018). Adult influence on juvenile phenotypes by stage-specific pheromone production. *Iscience*, *10*, 123–134.

Werner, M. S., Loschko, T., King, T., Reich, S., Theska, T., Franz-Wachtel, M., Macek, B., & Sommer, R. J. (2023). Histone 4 lysine 5/12 acetylation enables developmental plasticity of *Pristionchus* mouth form. *Nature Communications*, *14*(1), 2095.

Werner, M. S., Sieriebriennikov, B., Loschko, T., Namdeo, S., Lenuzzi, M., Dardiry, M., Renahan, T., Sharma, D. R., & Sommer, R. J. (2017). Environmental influence on *Pristionchus pacificus* mouth form through different culture methods. *Scientific Reports*, *7*(1), 7207.

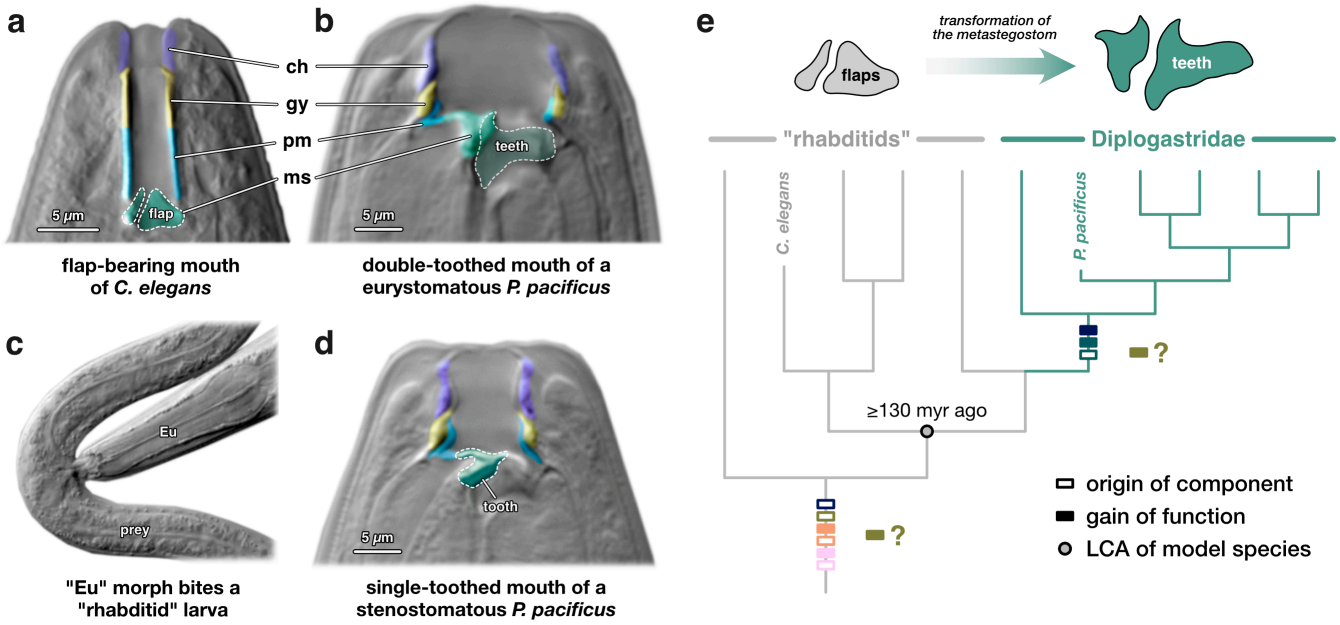
Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., & Hester, J. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, *4*(43), 1686.

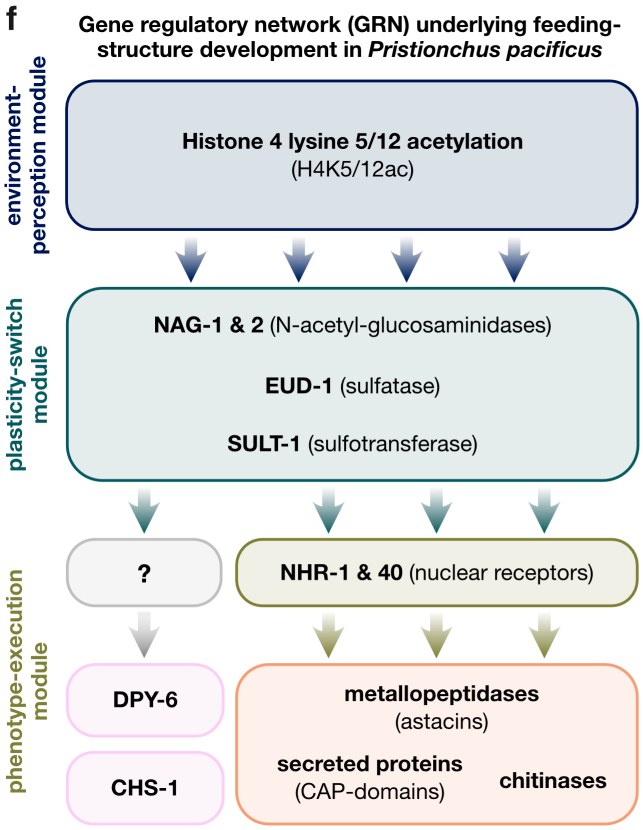
Wright, K. (1976). Functional organization of the nematode's head. *The Organization of Nematodes, Neil A. Croll (Ed.). Academic Press, London, UK*, 71–105.

Yoshida, K., Rödelsperger, C., Röseler, W., Riebesell, M., Sun, S., Kikuchi, T., & Sommer, R. J. (2023). Chromosome fusions repatterned recombination rate and facilitated reproductive isolation during *Pristionchus* nematode speciation. *Nature Ecology & Evolution*, *7*(3), 424–439.

Figure legends

Figure 1. Evolutionary history and developmental genetics of feeding structures in "rhabditids" and Diplogastridae. Microscopic images (with DIC) of the mouths of (a) *C. elegans*, (b) eurystomatous *P. pacificus*, (d) stenostomatous *P. pacificus*. Homologous feeding structures are correspondingly color-coded (ch = cheilostom, gy = gymnostom, pm = promesostegostom, ms = metastegostom). (c) Microscopic images (with DIC) of a eurystomatous *P. pacificus* preying on a "rhabditid" larva. (e) Phylogeny depicting the flap-to-tooth transformation of the metastegostom during the "rhabditid"-to-diplogastrid transition. Time estimate for LCA is based on Howard *et al.* (2022). Color coding of the boxes on the branches of the phylogeny indicates the origins of components of (f) the modular GRN controlling feeding-structure development in *P. pacificus*. Empty boxes in (e) correspond to the origin of the molecular players; filled boxes indicate the gain of their functions related to feeding-structure development (LCA = last common ancestor). (g) Phylogeny of all *nhr* genes encoded in the genomes of *C. elegans* and *P. pacificus* recreated from the data in Sieriebriennikov *et al.* (2020). The asterisk (*) indicates the location of the branch that harbors the one-to-one orthologs of NHR-1 and NHR-40. (h, i) Schematic representation of the expression patterns of NHR-1 and NHR-40 and their shared regulatory target genes in *P. pacificus*. Depicted is the head region in lateral position; the buccal cavity is on the left.





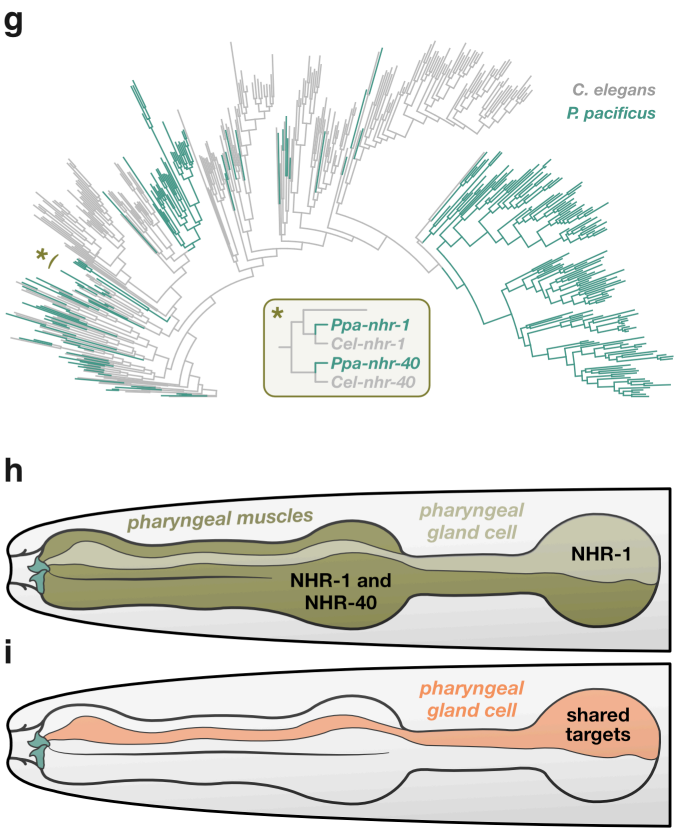


Figure 2. Expression and protein localization patterns of NHR-1 in *C. elegans*. (a) Pattern of *Cel-nhr-1* transcription in a young adult hermaphrodite. The same worm is depicted twice: top image shows its left lateral body surface; bottom image shows its mid-sagittal plane (A = anterior, P = posterior, D = dorsal, V = ventral). A zoom-in on the head region of the worm is depicted in (a) is provided in (b), showing the expression of *Cel-nhr-1* in the major pharyngeal muscle cells, as well as the head hypodermis. A close-up of the dorsal half of the stomatal tissues in another young hermaphrodite of the transcriptional reporter strain (c) reveals *Cel-nhr-1* expression in all of the cells which secrete the cuticular feeding structures (hyp = hypodermis, aa = anterior arcade syncytium, ap = posterior arcade syncytium, e1/3 = pharyngeal epithelial cells 1 and 3, pm = pharyngeal muscle cell, ch = cheilostom, gy = gymnostom, pm = promesostegostom, ms = metastegostom). A translational reporter strain (d) reveals that the *Cel*-NHR-1 protein localizes to the nuclei of the stomatal cells (D = dorsal, V = ventral, nr = nerve ring). Strength of fluorescence is indicated via a relative-intensity scale (a).

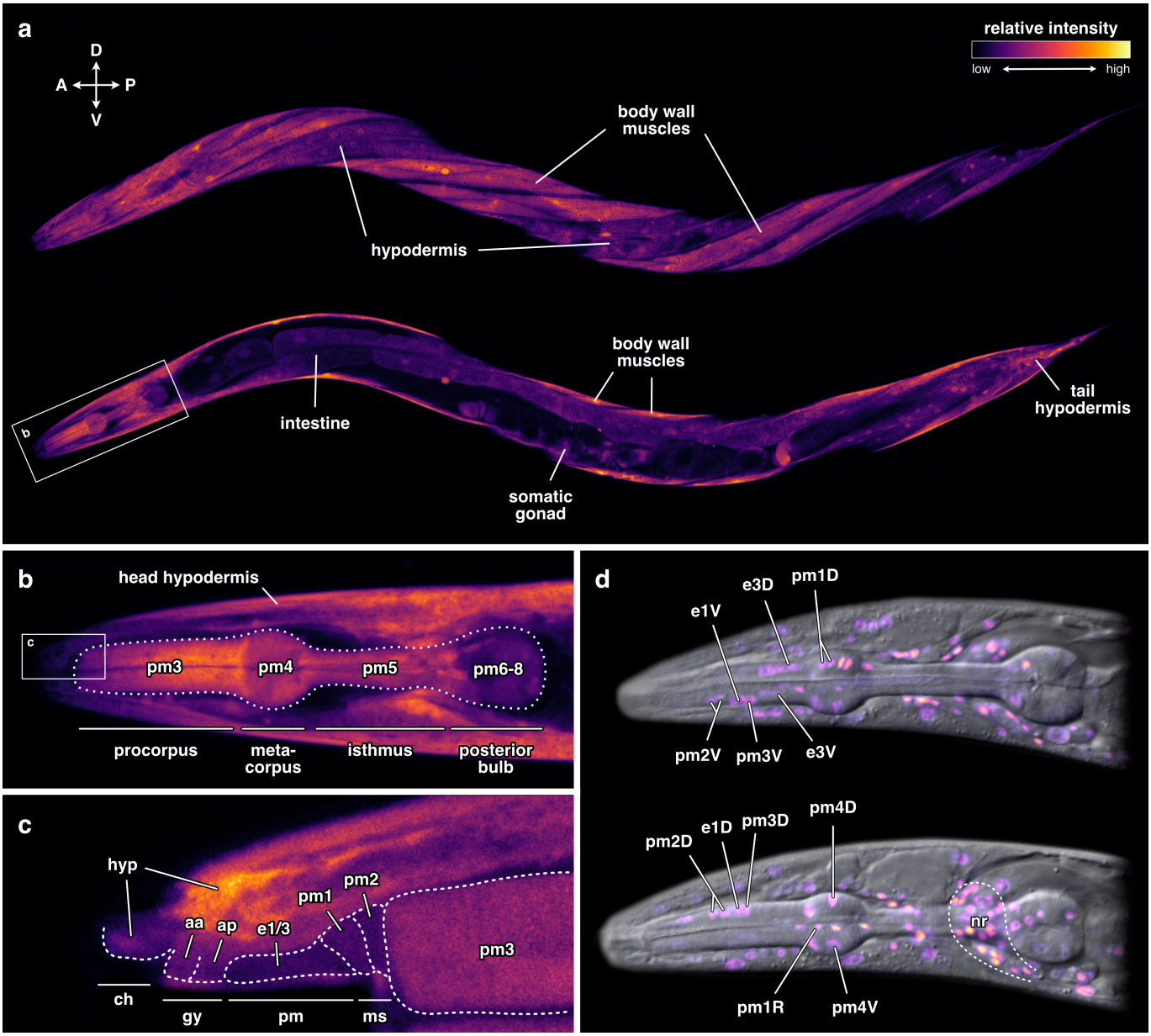


Figure 3. Neither the loss of NHR-1 or NHR-40 causes malformations in the *C. elegans* buccal capsule. (a, b) Gene models of *Cel-nhr-1* and *Cel-nhr-40* and their reference frameshift alleles which led to an early truncation of their DNA-binding domains (DBD). (c) Schematic representation of the adult *C. elegans* mouth depicting the landmark configuration used for geometric morphometric and clustering analysis (d, e). (d) PCA of form differences in wild-type *C. elegans*, as well as *Cel-nhr-1(tu1424)* and *Cel-nhr-40(1464)* mutants. Note that the ranges of morphological variation (outlined by convex hulls) in mutant strains strongly overlap with the wild-type range of morphological variation. (e) Unsupervised (model-based) clustering reveals that all 143 individuals belong to a single morphological cluster, demonstrating that mutant worms cannot be differentiated from wild-type worms based on their mouth morphology.

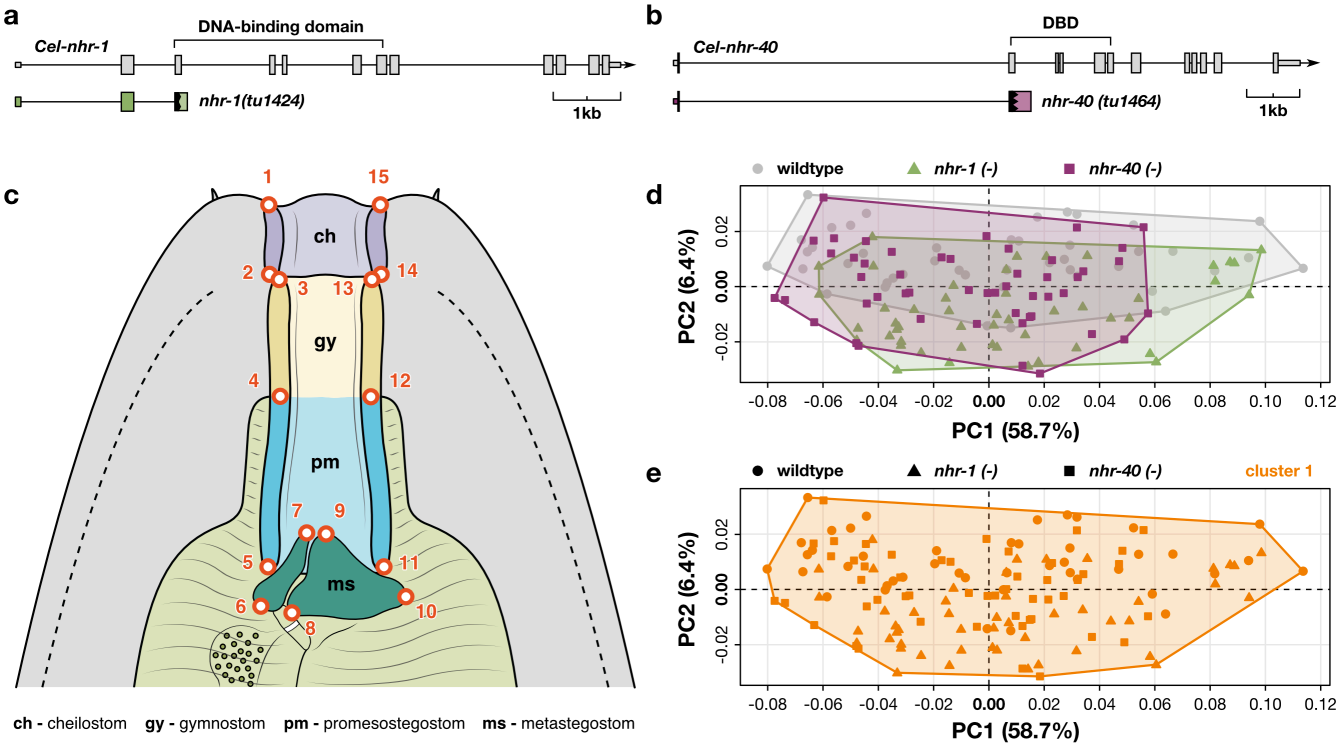


Figure 4. NHR-1 and NHR-40 share a pool of regulatory targets with currently unassigned biological functions. Bold names along the y-axis correspond to names of the shared regulatory targets of *Cel*-NHR-1 and *Cel*-NHR-40 which are misregulated throughout all developmental stages. Labels within the plot indicate functional information for the gene in form of WormCat terms (black) and/or predicted PFAM domains (gray). Roman numerals to the right of the plot indicate the chromosomal location of the shared targets. Note that the majority of target genes is located on chromosome V (bold letters indicate statistically overrepresented chromosomal locations). Dots to the right of the roman numerals indicate tissues in which the genes are known to be expressed (gray dot = neurons, black circle = glia cells, gray dot with black outline = neurons and glia cells). Cutoffs used in the differential gene expression analysis with DESeq2 and for PFAM domain prediction with HMMER are indicated in the lower left. MMSE = minimum mean squared error, FSOS = false sign or smaller.

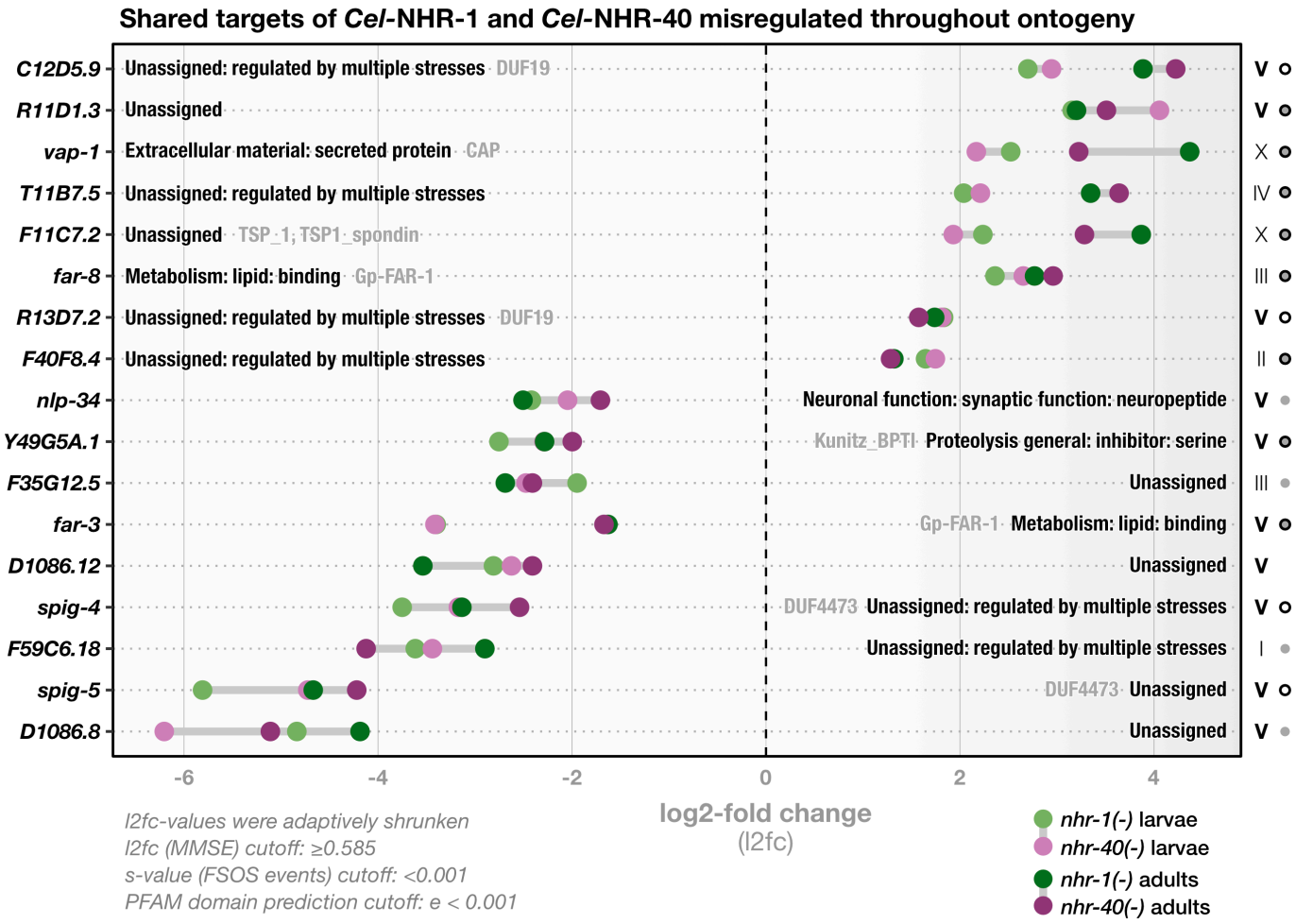
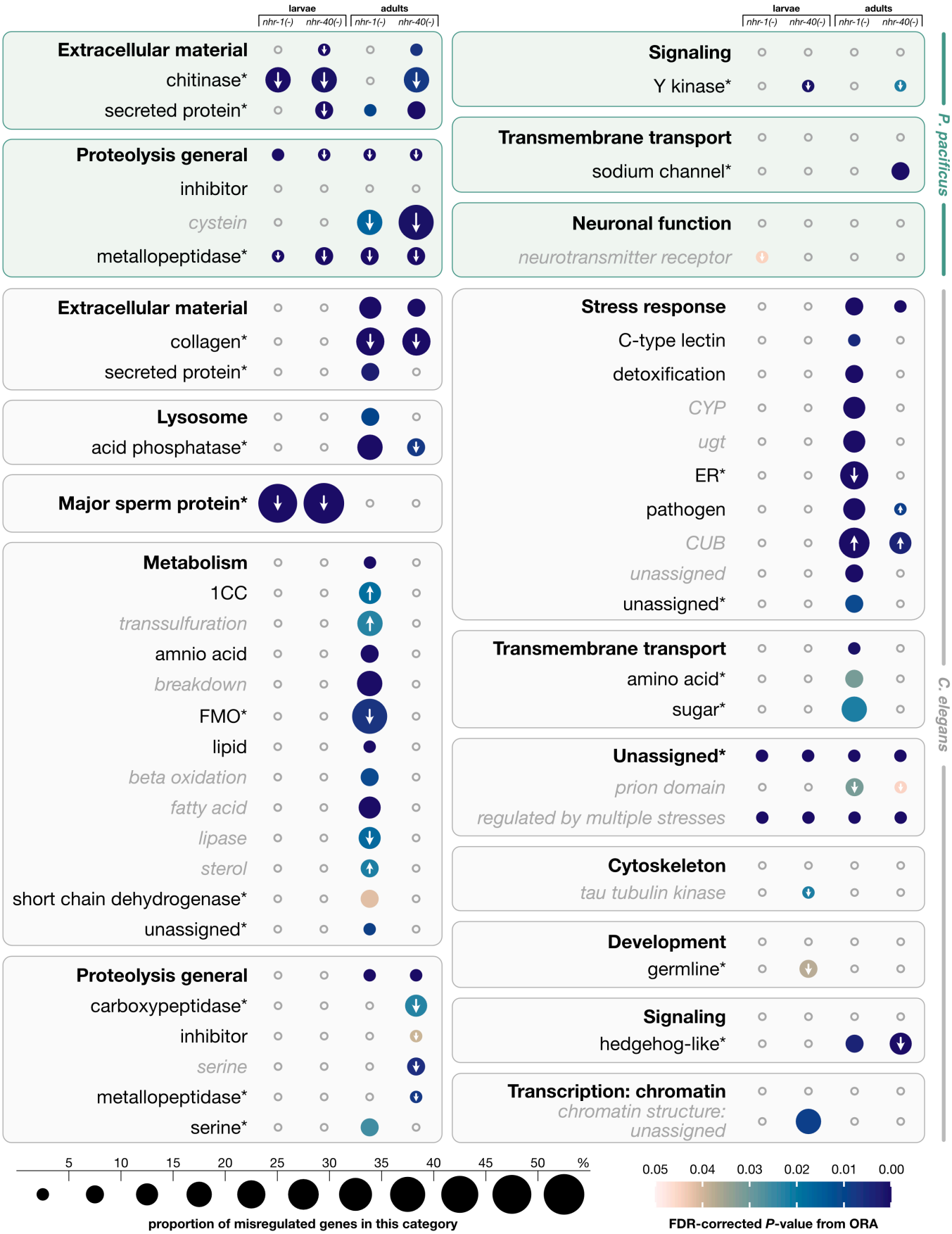


Figure 5. Comparative transcriptomics of *nhr-1* and *nhr-40* mutants in *P. pacificus* and *C. elegans* reveals species-specific composition of regulatory targets. Bubble chart showing overrepresented functional categories among the targets of each receptor in larvae and/or adults of *C. elegans* and *P. pacificus*. Enriched categories were broken down hierarchically. Boxes represent functional categories for which any overrepresentation signal could be detected. Functional categories were broken down into three hierarchical levels of specificity. Bold black font indicates the highest functional level (i.e., the least specific) and the name of the functional category represented by the box. Non-bold black font indicates the second category level (i.e., a more specific aspect of the overarching category). Gray italicized font indicates the most specific (and least inclusive) functional terms within the overarching category, for which overrepresentation signals could be detected. Asterisks (*) next to a category term indicate that no more-specific lower-level categories are defined within this term. The presence of bubbles indicates overrepresentation of the given functional category among target genes; absence of bubbles (depicted as gray circles) indicates no overrepresentation. Color bar indicates the statistical support for the overrepresentation signal (FDR-corrected *P*-values were derived from Fisher's Exact test). Bubble size indicates the percentage of genes in the given functional category which are misregulated in that particular sample type (out of all genes in the genome with that particular functional annotation). White arrows indicate that all of these targets are consistently up- or down-regulated. FDR = false discovery rate, ORA = overrepresentation analysis.



Supplementary Figure 1. Re-analysis of target genes that are constitutively misregulated in *Ppa-nhr-1* and *Ppa-nhr-40* mutants. Bold names along the y-axis correspond to names of the shared regulatory target genes of *Ppa*-NHR-1 and *Ppa*-NHR-40 which are misregulated throughout ontogeny. Labels within the plot indicate functional information for the gene in form of WormCat terms. Roman numerals to the right of the plot indicate the chromosomal location of the shared targets. Note that the majority of the targets is located on the X chromosome (bold letters indicate statistically overrepresented chromosomal location). Cutoffs used in the differential gene expression analysis with DESeq2 are indicated in the lower left. MMSE = minimum mean squared error, FSOS = false sign or smaller.