How to align arthropod leg segments

Authors: Heather S. Bruce[1,2]*

Affiliations: 1. University of California Berkeley, Berkeley, CA. 2. Marine Biological Laboratory, Woods Hole, MA.
*Correspondence to: hbruce@mbl.edu

Abstract

How to align leg segments between the four groups of arthropods (insects, crustaceans, myriapods, and chelicerates) has tantalized researchers for over a century. By comparing the loss-of-function phenotypes of leg patterning genes in diverged arthropod taxa, including a crustacean, insects, and arachnids, arthropod legs can be aligned in a one-to-one fashion. By comparing the expression of *pannier* and *aurucan*, the proximal leg segments can be aligned. A model is proposed wherein insects and myriapods incorporated the proximal leg region into the body wall, which moved an ancestral exite (for example, a gill) on the proximal leg into the body wall, where it invaginated independently in each lineage to form tracheae. For chelicerates with seven leg segments, it appears that one proximal leg segment was incorporated into the body wall. According to this model, the chelicerate exopod and the crustacean exopod emerge from different leg segments, and are therefore proposed to have arisen independently. A framework for how to align arthropod appendages now opens up a powerful system for studying the origins of novel structures, the plasticity of developmental fields across vast phylogenetic distances, and the convergent evolution of shared ancestral developmental fields.

Introduction

Arthropods are the most successful animals on the planet, in part due to the diversity of their appendages. The vast diversity of arthropod appendage form is reflected in their diversity of

function: arthropod legs have been modified for walking, swimming, flying, chewing, grasping, hearing, breathing, sensing and chemoreception, osmoregulation, copulation, silk-production, venom delivery, and more(*1*, *2*). For over a century, researchers have debated the evolutionary relationships and trajectories of these structures, including the origin of insect wings, whether myriapod and insect respiratory systems are homologous or a result of convergent evolution(*3*), whether the exopods (the lateral leg branch when a leg is split) of fossil arthropods and crustaceans are homologous to each other (*4*, *5*), the composition of chelicerae (*6*), and the identity of an enigmatic lobe on the leg of horseshoe crabs, the flabellum. However, answering these questions requires knowledge of how individual leg segments correspond to each other between insects, crustaceans, myriapods, and chelicerates. A framework to homologize arthropod appendages (i.e. the one-by-one alignment of each leg segment of all major arthropod groups) would therefore open up a powerful system for studying these and many more long-standing questions.

Unfortunately, such a model has remained elusive for over a century. The leg segments of chelicerates, myriapods, crustaceans, and insects have different numbers, shapes, and names. Chelicerates can have either 7 or 8 leg segments, myriapods have either 6 or 7, insects have 6, and crustacean have 7 or 8 leg segments (Fig. 1)(*4*, *7–10*). Over the decades, researchers have proposed many different theories to account for this variation, invoking leg segment deletions, duplications, and fusions to account for the different numbers of leg segments between arthropod taxa (see for example (*4*, *9*, *11*, *12*)). These differences even contributed to some author's conclusions that the four arthropod groups arose independently, and therefore it is not possible to homologize and align their legs(*13*). However, with molecular tools confirming arthropod monophyly and mapping the topology of arthropod relationships with ever greater precision(*14*,

15), and as loss-of-function studies of leg patterning genes have been conducted in a wider swath of the arthropod tree of life, this long sought model can now be brought to light.

Arthropod leg terms

An understanding of arthropod appendage diversity requires a few arthropod terms (Fig. 1A). Using crustacean legs as an example, the proximal region ("protopod") of the leg often bears lobes of various shapes and functions (*4*). When emerging laterally, these lobes are called exites (for example, gills or plates Fig. 1B). When emerging medially, towards the midline of the body, they are called endites (for example, the lobes used for chewing and cleaning that emerge on legs modified into mouthparts (Fig. 1C).

Legs can be bifurcated or split ("biramous"), meaning that two distal leg branches ("rami") emerge from the same proximal leg base. In this case, the lateral leg branch is called the exopod, and the medial leg branch is called the endopod (Fig. 1A, D). While exopods and exites both emerge laterally, exopods are a continuation of the leg, so they have muscle insertions and are often segmented (Fig. 1D). In contrast, exites are lobe-like outgrowths on the leg that lack muscle insertions and segmentation (Fig. 1B). This difference is also reflected in their development: exopods and endopods are the result of the terminus of the developing limb bud becoming forked in two, while exites emerge later by budding off of the existing proximal leg (*16–18*).

Arthropod legs are divided into segments (Fig. 1E – H). Leg segments are delineated on either side by joints where muscles insert (Fig. 1B) (*9*, *19*). Leg segments sometimes have subdivisions within them where no muscle inserts, i.e. muscles pass through the subdivision without inserting. These subdivisions serve as points of flexion, such as the subdivisions in the

tarsus of insects, arachnids, and myriapods (Fig. 1F), but subdivisions without muscle insertions do not represent true segments.

Alignment of insect and crustacean legs suggests that insect wings evolved from leg exites

The origin of insect wings has been a contentious problem for over 130 years. Two competing theories have developed to explain their emergence. Given that insects evolved from crustaceans(*14*), one theory is that insect wings evolved from crustacean exites(*20*). The second theory proposes that wings are a novel structure on the body wall that may not be present in crustaceans(*21*). In this latter model, genes that pattern crustacean exites could have been co-opted and expressed by an unrelated tissue, the dorsal body wall, in order to form insect wings on the back(*22*). To test these two hypotheses, Bruce & Patel(*2*) used CRISPR-Cas9 gene editing to compare the function of five leg patterning genes, *Distalless (Dll)*, *dachshund (dac)*, *Sp6-9*, *extradenticle (exd)*, and *homothorax (hth)*, in the amphipod crustacean *Parhyale hawaiensis.* By comparing the leg segment deletion phenotypes in *Parhyale* to previously published results in insects, they found that the six distal leg segments of *Parhyale* and insects (leg segments 1 – 6, counting from the distal claw) could be aligned in a one-to-one fashion (Fig. 2). They then wanted to understand the proximal leg segments. To do so, they compared the expression of *pannier* (*pnr*) and the Iroquois complex gene *aurucan* (*ara*) in *Parhyale* and insects(*2*) with previous work on the expression and function of wing genes in *Parhyale* and insects(*23*, *24*). They found that, in both *Parhyale* and insects, the expression of *ara* distinguishes two proximal leg segments (leg segments 7 and 8; Fig. 2), while expression of *pnr* marks the true body wall. These data suggested that insects had incorporated two ancestral proximal leg segments, 7 and 8, into the body wall(*20*). Because each of these leg segments carried an exite, insects now have

two exites emerging from the apparent body wall: the dorsal wing, and in some insects, a lateral lobe. Therefore, insect wings are not the result of gene co-option by an unrelated tissue, but instead are derived from a structure that already existed in the crustacean ancestor.

This work demonstrated that crustacean and insect legs could be homologized in a straightforward, one-to-one relationship. No deletions, duplications or rearrangements were necessary to make sense of leg segment homologies: insects and crustaceans each have 8 leg segments. Notably, the possession of 8 leg segments appears to be the arthropod ground plan, because early-branching crustaceans have an 8th leg segment, the chelicerate ground plan has 8 leg segments(*9*), and early fossil arthropods prior to the split of chelicerates and mandibulates also have 8 leg segments(*25*). Thus, if insect and crustacean legs can be homologized, this model may extend to myriapods and chelicerates as well, in a generalizable model of appendages across all four groups of arthropods.

Homologizing chelicerate and pancrustacean legs

To align *Parhyale* and chelicerate legs, leg segment deletion phenotypes in *Parhyale* were compared to previously published results in chelicerates, as I had done for insects. Functional experiments in chelicerates have been performed for *Dll*, *Sp6-9*, *dac*, and *hth*. Based on the leg segment deletion phenotypes of these genes, the six distal leg segments of *Parhyale*, insects, and chelicerates (leg segments 1 – 6, counting from the distal claw) can be aligned in a one-to-one fashion, as follows.

In spiders, *Parhyale*, and insects, *Dll* is required for the development of leg segments 1 - 5, counting from the distal end of the leg (Fig. 3A-F)(*2*, *26–31*). In spiders, *Parhyale*, and insects, *Sp6-9* is required for the development of leg segments 1 – 6 (Fig. 3G – L)(*2*, *32–36*). In

spiders, harvestman, *Parhyale*, and insects, *dac* is required to pattern leg segments 3 – 5 (in insects, dac function extends partway into leg segment 2)(*2*, *37–40*). In spiders, harvestman, and *Parhyale*, a weak *dac* phenotype causes leg segment 4 to be truncated and fused onto leg segment 3 (Fig. 4A - F). In harvestman, *Parhyale*, and insects, a strong *dac* phenotype affects leg segments 3 – 5 (Fig. 4G - L). In *Parhyale* and insects, loss of *hth* deletes the proximal leg segments, leaving only the distal 2 leg segments intact. In harvestman, reduction of *hth* shortens and fuses the proximal leg segments, leaving the distal segments unaffected(*41*). It is not clear from the figures or text how many distal leg segments are unaffected - the most severely affected embryos did not survive to hatching and their cuticle was shriveled, thus obscuring what is due to loss of hth and what is due to the embryo not developing fully before hatching However, leg segment 1 and at least the distal half of leg segment 2 are clearly unaffected. Thus, in insects, crustaceans, and *Parhyale*, *hth* appears to function in all but the distal 2 leg segments. For the above comparisons, it is known that RNAi gives a range of partial knockdowns, but the above data focuses on what appear to be the null phenotypes.

To elucidate the composition of proximal leg segments in chelicerates, the expression of *pnr*, *ara,* and *Distalless* (*Dll*) was examined in embryos of the tarantula *Acanthoscurria geniculate* (Fig. 5), similar to what I had done previously in *Parhyale* and *Tribolium*(*2*). Three orthologs of *pnr* were identified in *Acanthoscurria* with closest homology to *Drosophila*, *Tribolium*, and *Parhyale pnr* (Fig. S1). However, only one of these was expressed at the stages examined, and was presumed to be *pnr*. An *Acanthoscurria Iroquois* gene was identified which was the reciprocal best BLAST hit to *Drosophila*, *Tribolium*, and *Parhyale ara*. An *Acanthoscurria Dll* gene was identified which was the reciprocal best BLAST hit to *Drosophila*, *Tribolium*, and *Parhyale Dll*. I*n Acanthoscurria*, *pnr* was expected to mark the “true” body wall,

ara to distinguish leg segments 7 and 8, and *Dll* to mark leg segments 1 – 5, as these genes do in *Parhyale* and *Tribolium.*

As expected, in *Acanthoscurria*, *Dll* is expressed in leg segment 1 – 5, and *pnr* is expressed in the most dorsal tissue, as it is in *Parhyale* and *Tribolium* (Fig. 5). Thus, it appears that *pnr* marks the "true" body wall in all arthropods. In *Parhyale* and *Tribolium*, *ara* is expressed in four domains: a dorsal armband on proximal leg segment 8 that is adjacent to the *pnr* domain; a second armband on proximal leg segment 7; a dot of expression on the medial side of leg segment 6; and some expression in the tip of the claw(*2*). In *Acanthoscurria*, at the embryonic stages examined, *ara* is expressed in three of these four domains: a dorsal armband abutting the *pnr* domain; a second armband on proximal leg segment 7; and some expression in the tip of the claw. The dot of *ara* expression in leg segment 6 was not observed. Perhaps this domain is expressed at embryonic stages that were not examined, or it is not expressed in the *Acanthoscurria* lineage. However, as predicted by the leg segment alignment model, the two armbands of *ara* expression in *Acanthoscurria* bracket a region proximal to the coxa (leg segment 7) and adjacent to *pnr*, just as it does in *Parhyale* and *Tribolium*. This suggests that *Acanthoscurria* also retains a proximal 8th leg segment.

To test this hypothesis, the expression of *odd-skipped* was examined in *Acanthoscurria*. In *Drosophila*, the *odd-skipped* family of genes is expressed in the distal edge of each leg segment, where it induces cells to buckle and form the flexible joint(*42*). An *odd-skipped* gene was identified in *Acanthoscurria* which was the reciprocal best BLAST hit to *Cupiennius* (spider) *odd-related* 3 (*odd-r3*)(*43*). In *Acanthoscurria*, *odd-r3* is expressed in the distal region of leg segments 1 – 7 (Fig. 6). As predicted by the leg segment model presented here, there is an additional ring of *odd-r3* expression proximal to leg segment 7, which notably occurs in the

distal region of a leg-segment-like bulge. Given that *odd-skipped* is expressed in the distal region of leg segments(*42*), the ring of *odd-r3* expression in this leg-segment-like bulge suggests the presence of an additional proximal leg segment. Thus, the expression of *pnr*, *ara* and *odd-r3* and the presence of a proximal, leg-segment-like bulge together suggest that *Acanthoscurria* has an additional proximal 8th leg segment.

Discussion

The expression and embryological data shown here, in conjunction with the expression and functional data from previous publications, demonstrates that all arthropod legs can be aligned in a one-to-one fashion (Fig. 7). For example, the insect coxa, crustacean basis, and spider trochanter are homologous to each other, and are referred to as leg segment 6; the insect trochanter, crustacean ischium, and spider femur are homologous to each other, and are referred to as leg segment 5. In this model, arthropods ancestrally possessed a total of 8 leg segments. This fits well with the observation that arthropods in general have a maximum of 8 leg segments(*4*, *7*, *9*), as well as with the fossil data, where the ancestors of living arthropods also possessed 8 leg segments(*25*). Taxon-wide reductions of leg segment number are proposed to be due to the incorporation of proximal leg segments into the body wall, likely to support more of the body weight on the legs. Insects, which have 6 leg segments, have incorporated two proximal leg segments into the body wall (leg segments 7 and 8). Crustaceans with 7 leg segments, like *Parhyale*, have partially incorporated one leg segment into the body wall (leg segment 8). Myriapods with 6 or 7 leg segments incorporated 2 or 1 leg segments into the body wall, respectively (see below for further discussion).

Chelicerates with 8 leg segments, such as sea spiders, align one-to-one with the 8 leg segments of *Parhyale* and insects. However, in chelicerates with 7 leg segments, such as

Acanthoscurria, one of the two proximal leg segments is missing, and must be accounted for. One hypothesis is that one of the two proximal leg segments was simply deleted. However, the expression of *Acanthoscurria pnr*, *ara*, and *odd-r3* presented here suggests that the proximal-most 8th leg segment of these chelicerates was incorporated into the body wall, similar to how insects incorporated proximal leg segments into their body wall(*2*, *11*, *44*, *45*). In addition, the following observations from morphology, phylogenetics, paleontology, embryology, and molecular studies also support the hypothesis that the proximal-most leg segment of these chelicerates was incorporated into the body wall.

Early-branching(*15*) chelicerates often have eight leg segments like the arthropod ancestor, including sea spiders, trombidiform mites, hooded tick-spiders, and camel spiders, while later-branching chelicerates have seven leg segments, such as whip scorpions, vinegaroons, pseudoscorpions, scorpions, and spiders (Fig. 7b; (*9*)). An interesting exception is the horseshoe crab, *Limulus*(*9*). *Limulus* is an early-branching chelicerate(*15*) but has seven leg segments. However, it also has two proximal structures, a free endite and a pleurite (Fig. 1E), which may represent the remnant of the missing 8th leg segment. The free endite (also "epicoxite") is a mound of spiny exoskeleton ventral-medial to the coxa (leg segment 7) of each walking leg. The free endites of the walking legs encircle the mouth to aid in food processing (*46*). While endites are traditionally thought of as emerging from a leg segment, the *Limulus* free endite does not belong to any obvious leg segment, but is nevertheless moveable by muscles(*46*). Dorsal to the coxa of each walking leg is a conspicuous Y-shaped pleurite (body wall exoskeleton plate), which articulates with the coxa(*7*, *46*). As moveable endites and pleurites are characteristic of leg segments(*2*, *4*, *11*, *44*, *45*), these may represent the remnant of a proximal 8th leg segment in *Limulus* (*47*).

This hypothesis may explain a striking difference between living and extinct horseshoe crabs. In stark contrast to *Limulus*, fossil xiphosurids such as *Offacolus*(*48*) and *Dibasterium*(*49*) have a large, segmented exopod emerging from the base of the limb (Fig. 8a). In *Offacolus*, the limb base is even more elongated than in *Limulus*, such that it occupies the entire lateral region of the body. This proximal leg region is not well preserved in these fossils, and thus Sutton et. al were uncertain how many leg segments the limb base contained. However, a comparison to *Limulus* suggests that the limb base of *Offacolus* may actually be two leg segments: the elongate coxa like that seen in *Limulus* (*7*, *46*), and a smaller proximal leg segment from which the fossil exopod emerged. If this smaller proximal leg segment degenerated into the body wall in the lineage leading to living horseshoe crabs like *Limulus*, this would explain where the ancestral eighth leg segment went, the loss of the ancestral exopod, and the origin of the free endite and pleurite.

Embryological evidence of this proximal 8th leg segment can be observed in modern chelicerates with seven leg segments. In embryos of the tarantula spider *Acanthoscurria*, which is in the outgroup to other spiders, there is an additional leg segment-like structure proximal to the coxa in all walking legs (Fig. 6b; observable in *Acanthoscurria* embryos in (*31*, *50*); in *Cupiennius* embryos in (*51*, *52*); and in *Parasteatoda* embryos in (*53*, *54*)). Furthermore, even though there is no apparent exopod or outgrowth on this proximal fused leg segment, genes associated with outgrowing appendages, *wingless* (*wg*) and *Distalless* (*Dll*), are expressed here. In spiders, a dot of *wg* is expressed above each walking leg (Fig. 8c; (*55*, *56*)), and in *Limulus*, a dot of *Dll* is expressed above each walking leg(*57*). This is reminiscent of insects, which have incorporated two leg segments into the body wall(*2*): *wg* is expressed in two domains above the leg, one domain on each incorporated leg segment (Fig. 8d; (*58*, *59*)). In insects, these *wg*

domains pattern exites. In chelicerates, this *wg* domain may also be patterning an exite, for example a gland, like those found in the protonymphon larvae of sea spiders (*60*). Alternatively, the *wg* dot may be patterning the remnant of the exopod (see below for further discussion).

Dorsoventral patterning of arthropod appendages

The proximal-distal alignment model of arthropod legs presented here allows the previous molecular work on arthropod dorsal-ventral patterning to be reinterpreted into a simple and coherent model of arthropod leg development. *wg* is known to pattern body segments first, and then later, appendages. Previous studies in all four arthropod clades found that *wg* indeed patterns the ventral side of appendages. However, there seemed to be no obvious correspondence between the various lines and dots of *wg* expression in each clade. However, if insect and spider lateral body wall are interpreted as incorporated proximal leg segments, *wg* expression across all four arthropod clades agrees quite well.

The crustacean *Triops* demonstrates this interpretation. In crustaceans(*61*), *wg* is initially expressed in a solid stripe in each body segment, just as it is in insects(*59*, *62*, *63*), myriapods (*64*, *65*), and spiders (*55*) (Fig. 9a) . The *Triops* leg grows out, not like a cylinder, but like a shelf that wraps around dorsoventrally. Endites emerge near the midline, exites emerge laterally, and endopod and exopod emerge between these. As these outgrowths develop, the line of *wg* expression is broken up and becomes restricted to the ventral region of each of these outgrowths (Fig. 9b).

If insect wing and lateral body wall are interpreted as proximal leg, *wg* expression in insects mirrors that of crustaceans (Fig. 9c). In *Tribolium* (*58*, *63*) and cricket (*59*) embryos, *wg* is expressed in the initial body segment stripe, and as the legs grow out, this line of *wg* expression is broken up such that *wg* is expressed in a ventral stripe in each leg. In addition to the

ventral leg stripe, there are two regions of *wg* expression dorsal to the leg (Fig. 6D), corresponding to the insect wing and the lobe. These two additional regions of *wg* expression are precisely what one would expect if insects incorporated two leg segments into the body wall, and each leg segment produces an exite that is patterned ventrally by *wg*. Insect and myriapod endites are also accounted for in this model, as the single body stripe of *wg* is interrupted and becomes restricted to the ventral side of the developing mouthparts of insects and myriapods (*66–68*).

A similar sequence is observed in spiders. According to the model presented here, arachnids with seven leg segments should have a cryptic ancestral proximal leg segment incorporated into the body wall, from which the ancestral chelicerate exopod used to emerge. In spider embryos, the initial body segment-patterning stripe of *wg* resolves into two domains, a ventral stripe on each appendage, and a dot dorsal to each coxa (Fig. 9c; (*55*, *56*)). Thus, arachnids have one *wg* domain in the lateral body wall corresponding to one fused leg segment where the exopod once emerged, and insects have two *wg* domains in the lateral body wall corresponding to two fused leg segments, each with an exite (Fig. 9c). The expression domains of *dpp* are also consistent with this reinterpretation.

Independent origins of exopods

On the proximal-most leg segment of many ancient arthropods including trilobites, *Leanchoilia*, and *Offacolus*, there emerges a structure that many authors believe is an exopod (although see (*7*, *12*)). The exopods of these fossil arthropods are proposed by some authors to be homologous to the exopod of crustaceans (*4*, *8*, *69*), meaning that exopods had a single origin inherited by the common ancestor of all arthropods. This is a reasonable hypothesis when morphology alone is considered. However, when the molecular evidence is considered together

with morphology, exopod homology becomes doubtful: the crustacean exopod emerges from leg segment 6, while the chelicerate and early arthropod exopod emerges from leg segment 8 (Fig. 10a). Thus, unexpectedly, the crustacean exopod appears to have a separate origin from the exopod of early-branching arthropods and chelicerates.

As discussed above, arthropod legs almost universally have a maximum of 8 leg segments. Exopods are essentially a bifurcation or splitting of the leg axis (*16*, *17*), such that two distal leg branches continue from the same proximal leg segment. Therefore, in a biramous leg, one would expect the maximum number of leg segments, counting down either axis, to always be 8 (except when legs are annulated). If the chelicerate leg bifurcates at the proximal-most leg segment, there should be 7 remaining leg segments along the endopod and exopod, mirroring each other, for a total of 8 segments (Fig. 10a). This 1 + 7 configuration indeed appears to be the case for fossil arthropods near or within chelicerates, such as trilobites(*7*, *70*), *Emeraldella* (*71*), *Sidneyia* (*72*). In contrast, crustacean legs have a 3 + 5 configuration, where the leg bifurcates at the third-most proximal leg segment, and therefore both endopod and exopod each have up to 5 segments (*4*, *8*), mirroring each other, for a total of 8 segments.

From a molecular standpoint, the alignment of the expression and function of insect, crustacean, and chelicerate leg patterning genes results in a crustacean exopod that emerges from leg segment 6, and a chelicerate and ancient arthropod exopod that emerges from leg segment 8 (Fig. 10a). One possible interpretation of this alignment is that exopods are evolved once in the common ancestor of chelicerates and crustaceans, but in the lineage leading to crustaceans, the exopod moved from the 8th to the 6th leg segment position. It is unlikely that this could have been achieved on a molecular level from one or two simple mutations, as these would not be

sufficient to rewire a complex leg patterning pathway to repress bifurcation in the 8th leg segment while also activating bifurcation in the 6th leg segment.

Moving the exopod from the 8th to the 6th leg segment also appears unlikely from a morphological standpoint, as it would require a dubiously complex set of leg segment rearrangements in order to keep the total number of leg segments constant between chelicerates and crustaceans (Fig. 10b): four deletions and two additions of leg segments. If the lineage leading to crustaceans deleted four leg segments and added two leg segments, then the configuration of leg patterning genes between chelicerates and crustaceans would be different (Fig. 10c). However, the proximal expression and distal functional domains are essentially the same (Fig. 10d). Given that leg segment additions are rare (*4*), and that leg genes have the same configuration in both chelicerates and crustaceans, this is not a parsimonious hypothesis. Therefore, two independent origins for chelicerate and crustacean exopods is more plausible.

Several interesting and potentially useful implications emerge if chelicerate and crustacean exopods are not homologous. First, given that chelicerate and crustacean legs bifurcate at different points along the axis, and these two regions express different patterning genes, it is likely that different genetic mechanisms led to the generation of the exopod in these two groups. Thus, chelicerate and crustacean exopods likely represent independent evolutionary gains of a bifurcated leg axis, and could be used as a dyad system to compare mechanisms of convergent evolution.

Second, the position of the exopod, on either the 6th or 8th leg segment, could be a powerful morphological character for determining the phylogenetic position of otherwise ambiguous arthropod fossils. This in turn might reconfigure existing arthropod phylogenies and necessitate a reinterpretation of ground states and character evolution in different arthropod taxa.

For example, the problematic fossil arthropod *Agnostus*, which has an exopod that emerges from leg segment 8, would be more closely allied with chelicerates, rather than a stem crustacean. For many fossil arthropods, the number of segments in the exopod will not be informative as they number fewer than 5, which would be equally valid for either a chelicerate or crustacean. However, the maximum number of segments in the endopod should be seven for chelicerates and five for crustaceans (except when legs are annulated), which should be an informative polarizing character.

If chelicerate and crustacean exopods are not homologous, when did the crustacean exopod evolve? If it evolved in Mandibulata, then we might expect as-yet-unknown stem myriapods to have an exopod on the 6th leg segment. However, if the stem myriapod retained the chelicerate exopod, it would be on the 8th leg segment. Alternatively, perhaps the stem myriapod had already lost the chelicerate exopod, but did not evolve their own exopod like crustaceans, so we would expect an animal without an exopod. This new model of exopods may help to uncover the yet-to-be-identified stem myriapod. Perhaps the stem myriapod has already been identified, but due to an exopod, was misclassified as either a crustacean or closer to chelicerates.

A third interesting outcome of this model is that the *Limulus* flabellum is not the remnant of the chelicerate exopod. Some authors have interpreted the flabellum as the remnant of the chelicerate exopod (*4*, *69*), because it emerges from what appears to be the most proximal leg segment, and because the flabellum expresses *Dll* (*57*) . While *Dll* indeed patterns leg segments, including the exopod, it is also expressed in exites (*2*, *73*), where it patterns sensory structures (*74*). Thus, *Dll* expression alone cannot determine exopod identity.

If the flabellum is not an exopod, perhaps it is an exite. Consistent with this hypothesis is the fact that, like crustacean exites, the flabellum is an unsegmented lobe that lacks muscle

insertions(*46*, *75*), it has a sensory function (*76*), and it develops by budding off of the proximal leg ((*57*, *70*), rather than bifurcation of the terminus of the leg as exopods do. These features of the flabellum have led several authors to interpret it as an exite (epipod) (*7*, *70*, *77*). The data presented here is also consistent with the exite hypothesis, because it suggests that the proximal-most leg segment, which carried the chelicerate exopod, was reduced and incorporated into the body wall in *Limulus*. In this model, the leg segment that carries the flabellum in *Limulus* is not the same leg segment that carries the exopod in fossil chelicerates. This is unexpected, given that exites are proposed to have evolved in crustaceans (*4*, *78*). However, this proposal is based on morphology and fossils. To determine more definitively whether the flabellum is an exopod or an exite, the function of *Dll* should be examined in *Limulus*. In *Parhyale*, *Sp-69* knockout deletes the entire exopod and endopod up to leg segment 7, but leaves the exites unaffected (Fig. 11)(*2*). If the *Limulus* flabellum is an exite, then *Sp6-9* knockout will truncate the distal leg up to leg segment 7 (coxa), but leave the flabellum unaffected. If the *Limulus* flabellum is a one-segmented exopod composed of leg segment 6, then *Sp6-9* knockout will delete the flabellum and the distal leg up to leg segment 7. It should be noted that *Dll* knockout will not be informative here, as *Dll* knockout only deletes leg segments 1 – 5, so would be unlikely to affect the flabellum if it is an exopod. If the *Limulus* flabellum is indeed an exite, this means that exites evolved well before crustaceans, and may be part of the ground pattern of euarthropods. This may allow the lateral lobes in fossil arthropods, such as *Agnostus*(*78*) and *Sanctacaris* (*48*), *Sinoburius* to be reinterpreted as exites.

Are myriapod and insect respiratory systems homologous?

No functional data for leg patterning genes is available for myriapods. However, morphological and embryological evidence suggests that myriapods, like insects, have incorporated proximal leg segment(s) into the body wall. Incorporation of proximal leg segments into the myriapod body wall would bring their leg segment count to 8, in agreement with other arthropods.

Morphologically, all groups of myriapods have pleurites - regions of hardened exoskeleton in the membranous body wall surrounding the base of the free leg. In insects, pleurites often articulate with the first free leg segment and serve as attachment points for muscles that move the legs. The pleurites of insects and wingless hexapods, which incorporated leg segment 7 and 8 into the body wall(*2*, *20*), are arrayed as the remnants of two concentric circles around the first free leg segment, leg segment 6 (*11*, *45*). In hexapods that branched prior to the evolution of insect wings, the pleurites are variously degenerated. However, in winged insects, the pleurites reinforce the thorax for flight, and have thus become robust (*11*). Consistent with this, the pleurites of myriapods, which did not evolve wings like insects, are variously degenerated as they are in pre-wing hexapods(*79*).

In the embryos of many insects, the proximal part of the developing leg ("subcoxa") broadens and flattens to form the adult pleurites (*44*, *80–84*). Thus, what was proximal leg now functions as body wall(*85*). In myriapod embryos, there is also a broad leg region ("limb-base") on which the rest of the leg is mounted that will form the adult pleurites (*45*, *86*, *87*), and from which the eversible sac of symphylans emerges (*87*). This "limb-base" may represent the remnant of the proximal leg that was incorporated into the myriapod body wall.

Notably, the myriapod respiratory spiracle and trachea are carried on the pleurite (*88*), which is here hypothesized to be the remnant of the incorporated 8th leg segment (*88*). In insects,

the spiracle and tracheae also occur on the incorporated 8th leg segment(*2*, *89*). Notably, *trachealess (trh)* and *ventral veins lacking (vvl)*, which are expressed in and required for insect tracheal formation, are also expressed in the crustacean gill(*90*), an exite on the leg. Thus, insect tracheae may be an invaginated exite on the incorporated 8th leg segment. If insect tracheae are an invaginated exite on the 8th leg segment, then perhaps the morphologically and functionally similar myriapod tracheae are as well. This would mean that an ancestral respiratory exite on the leg of the mandibulate ancestor was independently invaginated in two phylogenetically distant arthropod clades and in two separate terrestrialization events. In this model, insect and myriapod trachea are homologous and convergent structures: they both evolved from a homologous structure, the exite of the 8th leg segment, but invaginated convergently in separate lineages and at separate timepoints in response to terrestrialization. Myriapod-insect trachea would therefore be a useful dyad system for investigating the genetic basis of convergent evolution and developmental plasticity of the same ancestral structure.

To test this hypothesis, the expression of *pnr*, *Iroquois* genes like *ara*, joint-markers like *odd-skipped*, exite-patterning genes like *vestigial*, *scalloped*, and trachea-patterning genes like *ventral veins lacking*, and *trachealess* should be examined in myriapod embryos. If the spiracle is bracketed by *Iroquois* genes and abutted dorsally by *pnr* and ventrally by *odd-skipped*, then it is homologous to this same region in *Parhyale* and *Tribolium*, and may be carried on the incorporated 8th leg segment. If myriapod trachea/spiracles express *vestigial*, *scalloped*, *ventral veins lacking*, and/or *trachealess*, then they are likely invaginated exites.

Lateral lobes throughout Arthropoda may be exites of the incorporated 8th leg segment

Interestingly, it has been a matter of debate whether exites exist outside of crustaceans. Many fossil arthropods have lateral lobes on the legs, which have been proposed to be exites (*77*). Many groups of living arthropods have lobes or plates that project from the dorsal-lateral body wall. These include the paranotal lobes of wingless hexapods like silverfish, the helmets of hemipterans, the carapace of *Daphnia* and ostracods, the paratergites of polydesmid millipedes, the tergites of ricinulei and trigonotarbids, and the pteromorph "wings" of oribatid mites. All of these arthropods have fewer than 8 leg segments, and in the model presented here, have incorporated the missing proximal leg segment(s) into the body wall. This ancestral proximal leg segment(s) may have had an exite, which in the above groups would now emerge from what appears to be the dorsal-lateral body wall. Curiously, a systematic examination of the expression of wing/exite genes in taxa outside of crustacea has not been undertaken to help resolve this question (but see Sharma 2017 (*91*) for expression of *vvl* and *trh* in chelicerates).

To test if these lateral lobes emerge from incorporated proximal leg, the expression of *pnr*, *Iroquois* genes like *ara*, and joint-markers like *odd-skipped* should be examined in embryos in these groups. If these lateral lobes are bracketed by *Iroquois* genes and abutted dorsally by *pnr* and ventrally by *odd-skipped*, then they are homologous to this same region in *Parhyale* and *Tribolium*, and may represent the exite of the incorporated 8th leg segment. Exite-patterning genes like *vestigial*, *scalloped*, *ventral veins lacking*, and *trachealess* might be expressed, which would support an exite identity for these structures. However, if they are not expressed, this does not necessarily disprove an exite identity, as exite-patterning may have diverged between arthropod groups. If exites evolved prior to crustaceans, this may necessitate a reappraisal of the accepted trajectories of character acquisition and evolution, and thus the affinities of many fossil arthropods. Exites may have even been part of the arthropod ground plan.

Future experiments

In all arthropod species in which gene function has been examined, the proximal 8th leg segment is not free, having been incorporated to varying degrees into the body wall. Thus, a necessary test of the model presented here is to perform functional studies in an arthropod in which the proximal 8th leg segment is free, such as sea spiders, camel spiders, trombidiid mites, ostracods, or mantis shrimp. These would preferentially be done using CRISPR-Cas9 in order to achieve easily interpretable null or nearly null phenotypes, wherein all affected segments are completely deleted, and adjacent leg segments are unaffected wild-type.

Materials and Methods

In situ HCR version 3.0

Acanthoscurria cDNA sequences were submitted to Molecular Instruments(*92*), and the probe sets are available from the company. In situs were performed using their whole mount *Drosophila* embryo protocol, with the following exceptions. Embryos were left on bench, not rocked back and forth. Embryos were permeabilized in SDS detergent solution (1% SDS, 0.5% Tween 20, 50mM Tris-HCL pH7.5, 1mM EDTA pH8.0, 150mM NaCl) for 30min instead of ProK, which improved morphology in our hands. Embryos were in pre-hyb and pre-amp for 1-2 hours to allow complete settling. Embryos were hybridized for 2 days instead of 1, and probes were used at 5x concentration to boost signal to noise.

Data and materials availability: All data is available in the main text or the supplementary materials.

Competing interests: Authors declare no competing interests.

Acknowledgments: I thank Nipam H. Patel for reagents. *Acanthoscurria* embryos and transcriptome were kindly provided by Mathias Pechmann.

References

1. R. C. Brusca, G. J. Brusca, *Invertebrates* (Sinauer Associates Incorporated, 2003; http://books.google.com/books?id=lSggPQAACAAJ&dq=inauthor:brusca+(intitle:invertebrates+2003)&hl=&cd=1&source=gbs_api).

2. H. S. Bruce, N. H. Patel, Knockout of crustacean leg patterning genes suggests that insect wings and body walls evolved from ancient leg segments. *Nat Ecol Evol*. **4**, 1703–1712 (2020).

3. J. W. Wagele, T. Bartolomaeus, Deep Metazoan Phylogeny: The Backbone of the Tree of Life, 216.

4. G. A. Boxshall, The evolution of arthropod limbs. *Biol. Rev.* **79**, 253–300 (2004).

5. J. T. Haug, A. Maas, C. Haug, D. Waloszek, in *Functional Morphology and Diversity*, Les Watling, M. Thiel, Eds. (Oxford, 2013; https://bioone.org/journals/Palaeodiversity/volume-12/issue-1/pale.v12.a1/Cretaceous-chimera--an-unusual-100-million-year-old-neuropteran/10.18476/pale.v12.a1.full).

6. P. P. Sharma, E. E. Schwager, C. G. Extavour, G. Giribet, Evolution of the chelicera: a dachshunddomain is retained in the deutocerebral appendage of Opiliones (Arthropoda, Chelicerata). *Evol. Dev.* **14**, 522–533 (2012).

7. R. E. Snodgrass, *A textbook of arthropod anatomy*. (Ithaca, N.Y., Comstock Pub. Associates, 1952).

8. F. R. Schram, *Crustacea* (Oxford University Press, USA, 1986).

9. J. W. Shultz, Morphology of locomotor appendages in Arachnida: evolutionary trends and phylogenetic implications. *Zool J Linn Soc*. **97:1–56** (1989) (available at http://onlinelibrary.wiley.com/doi/10.1111/j.1096-3642.1989.tb00552.x/full).

10. D. Grimaldi, M. S. Engel, *Evolution of the Insects* (Cambridge University Press, 2005; http://books.google.com/books?id=odQmAAAAQBAJ&pg=PR4&dq=inauthor:grimaldi+(2005+inauthor:engel)&hl=&cd=1&source=gbs_api).

11. R. E. Snodgrass, *Morphology and mechanism of the insect thorax* (City of Washington, Smithsonian institution, 1927), vol. 80.

12. L. Størmer, *On the relationships and phylogeny of fossil and recent Arachnomorpha* (Oslo, 1944; http://scholar.google.comjavascript:void(0)), *A comparative study on Arachnida, Xiphosura, Eurypterida, Tribolita, and other fossil arthropoda*.

13. S. M. Manton, Habits, functional morphology and the evolution of pycnogonids. *Zool J Linn Soc*. **63: 1–21** (1978) (available at http://onlinelibrary.wiley.com/doi/10.1111/j.1096-3642.1978.tb02087.x/full).

14. J. Lozano-Fernandez, M. Giacomelli, J. F. Fleming, A. Chen, J. Vinther, P. F. Thomsen, H. Glenner, F. Palero, D. A. Legg, T. M. Iliffe, D. Pisani, J. Olesen, Pancrustacean Evolution Illuminated by Taxon-Rich Genomic-Scale Data Sets with an Expanded Remipede Sampling. *Genome Biol Evol*. **11**, 2055–2070 (2019).

15. J. Lozano-Fernandez, A. R. Tanner, M. Giacomelli, R. Carton, J. Vinther, G. D. Edgecombe, D. Pisani, Increasing species sampling in chelicerate genomic-scale datasets provides support for monophyly of Acari and Arachnida. *Nat Commun*, 1–8 (2019).

16. C. Wolff, G. Scholtz, The clonal composition of biramous and uniramous arthropod limbs. *Proceedings of the Royal Society B: Biological Sciences*. **275**, 1023–1028 (2008).

17. A. Hejnol, G. Scholtz, Clonal analysis of Distal-less and engrailed expression patterns during early morphogenesis of uniramous and biramous crustacean limbs. *Dev Genes Evol*. **214**, 473–485 (2004).

18. G. A. Boxshall, D. Jaume, Exopodites, epipodites and gills in Crustaceans. *Arthropod Systematics & Phylogeny*, 1–27 (2009).

19. G. Boxshall, *Arthropod Limbs and their Development* (2013), *Arthropod Biology and Evolution*.

20. J. Kukalová-Peck, Origin of the insect wing and wing articulation from the arthropodan leg. *Can. J. Zool.* **61**, 1618–1669 (1983).

21. G. Crampton, The phylogenetic origin and the nature of the wings of insects according to the paranotal theory. *Journal of the New York Entomological Society*. **24**, 1–39 (1916).

22. E. L. Jockusch, L. M. Nagy, Insect evolution: How did insect wings originate? *Current Biology*. **7**, R358–R361 (1997).

23. C. M. Clark-Hachtel, Y. Tomoyasu, Two sets of candidate crustacean wing homologues and their implication for the origin of insect wings. *Nat Ecol Evol* (2020), doi:10.1038/s41559-020-1257-8.

24. C. M. Clark-Hachtel, D. M. Linz, Y. Tomoyasu, Insights into insect wing origin provided by functional analysis of vestigial in the red flour beetle, Tribolium castaneum. **110**, 16951–16956 (2013).

25. J. Yang, J. Ortega-Hernández, D. A. Legg, T. Lan, J. Hou, X. Zhang, Early Cambrian fuxianhuiids from China reveal origin of the gnathobasic protopodite in euarthropods. *Nat Commun*. **9**, 1–9 (2018).

26. S. M. Cohen, G. Jürgens, Proximal-distal pattern formation in Drosophila: cell autonomous requirement for Distal-less gene activity in limb development. *EMBO J*. **8**, 2045–2055 (1989).

27. B. Cohen, A. A. Simcox, S. M. Cohen, Allocation of the thoracic imaginal primordia in the Drosophila embryo. *Development*. **117**, 597–608 (1993).

28. G. Campbell, A. Tomlinson, The roles of the homeobox genes aristaless and Distal-less in patterning the legs and wings of Drosophila. *Development*. **125**, 4483–4493 (1998).

29. A. Beermann, D. G. Jay, R. W. Beeman, M. Hülskamp, D. Tautz, G. Jürgens, The Short antennae gene of Tribolium is required for limb development and encodes the orthologue of the Drosophila Distal-less protein. *Development*. **128**, 287–297 (2001).

30. D. R. Angelini, T. C. Kaufman, Functional analyses in the hemipteran Oncopeltus fasciatus reveal conserved and derived aspects of appendage patterning in insects. *Developmental Biology*. **271**, 306–321 (2004).

31. M. Pechmann, S. Khadjeh, N. Turetzek, A. P. McGregor, W. G. M. Damen, N.-M. Prpic, Novel Function of Distal-less as a Gap Gene during Spider Segmentation. *PLoS Genet*. **7**, e1002342 (2011).

32. A. Beermann, M. Aranda, R. Schröder, The Sp8 zinc-finger transcription factor is involved in allometric growth of the limbs in the beetle Tribolium castaneum. *Development*. **131**, 733–742 (2004).

33. N. D. Schaeper, N.-M. Prpic, E. A. Wimmer, A conserved function of the zinc finger transcription factor Sp8/9 in allometric appendage growth in the milkweed bug Oncopeltus fasciatus. *Dev Genes Evol*. **219**, 427–435 (2009).

34. C. Estella, R. S. Mann, Non-Redundant Selector and Growth-Promoting Functions of Two Sister Genes, buttonhead and Sp1, in Drosophila Leg Development. *PLoS Genet*. **6**, e1001001 (2010).

35. T. Königsmann, N. Turetzek, M. Pechmann, N.-M. Prpic, Expression and function of the zinc finger transcription factor Sp6–9 in the spider Parasteatoda tepidariorum. *Dev Genes Evol*. **130**, 1–12 (2017).

36. E. V. W. Setton, P. P. Sharma, Cooption of an appendage-patterning gene cassette in the head segmentation of arachnids. **128**, 201720193–10 (2018).

37. G. Mardon, N. M. Solomon, G. M. Rubin, dachshund encodes a nuclear protein required for normal eye and leg development in Drosophila. *Development*. **120**, 3473–3486 (1994).

38. B. C. Tavsanli, E. J. Ostrin, H. K. Burgess, B. W. Middlebrooks, T. A. Pham, G. Mardon, Structure–function analysis of the Drosophila retinal determination protein Dachshund. *Developmental Biology*. **272**, 231–247 (2004).

39. N. Turetzek, M. Pechmann, C. Schomburg, J. Schneider, N.-M. Prpic, Neofunctionalization of a Duplicate dachshundGene Underlies the Evolution of a Novel Leg Segment in Arachnids. *Mol. Biol. Evol.* **33**, 109–121 (2015).

40. P. P. Sharma, E. E. Schwager, G. Giribet, E. L. Jockusch, C. G. Extavour, Distal-lessand dachshund pattern both plesiomorphic and apomorphic structures in chelicerates: RNA interference in the harvestman Phalangium opilio(Opiliones). *Evol. Dev*. **15**, 228–242 (2013).

41. P. P. Sharma, O. A. Tarazona, D. H. Lopez, E. E. Schwager, M. J. Cohn, W. C. Wheeler, C. G. Extavour, A conserved genetic mechanism specifies deutocerebral appendage identity in insects and arachnids. *Proceedings of the Royal Society B: Biological Sciences*. **282**, 20150698–20150698 (2015).

42. C. Mirth, M. Akam, Joint development in the Drosophila leg: cell movements and cell populations. *Developmental Biology*. **246**, 391–406 (2002).

43. N.-M. Prpic, W. G. M. Damen, Notch-mediated segmentation of the appendages is a molecular phylotypic trait of the arthropods. *Developmental Biology*. **326**, 262–271 (2009).

44. Y. Kobayashi, Formation of Subcoxae-1 and 2 in Insect Embryos: The Subcoxal Theory Revisited. *Proc Arthropod Embryol Soc Jpn*. **48**, 33–38 (2017).

45. R. Matsuda, Morphology and evolution of the insect thorax. *Memoirs of the Entomological Society of Canada*. **Volume 102**, 5–431 (1970).

46. S. M. Manton, J. P. Harding, Mandibular mechanisms and the evolution of arthropods. … *of the Royal* …. **247** (1964) (available at http://rstb.royalsocietypublishing.org/content/247/737/1.short).

47. L. Størmer, *Studies on trilobite morphology. Part III. The ventral cephalic structures with remarks on the zoological position of the trilobites* (Norsk Geol. Tidssk, 1951; http://scholar.google.com/scholar?q=related:WtkO_2OEme0J:scholar.google.com/&hl=en &num=20&as_sdt=0,5).

48. M. D. Sutton, D. E. G. Briggs, D. J. Siveter, D. J. Siveter, P. J. Orr, The arthropod Offacolus kingi (Chelicerata) from the Silurian of Herefordshire, England: computer based morphological reconstructions and phylogenetic affinities. *Proceedings of the Royal Society B: Biological Sciences*. **269**, 1195–1203 (2002).

49. D. E. G. Briggs, D. J. Siveter, D. J. Siveter, M. D. Sutton, R. J. Garwood, D. Legg, Silurian horseshoe crab illuminates the evolution of arthropod limbs. **109**, 15702–15705 (2012).

50. M. Pechmann, Embryonic development and secondary axis induction in the Brazilian white knee tarantula Acanthoscurria geniculata, C. L. Koch, 1841 (Araneae; Mygalomorphae; Theraphosidae). *Dev Genes Evol*. **130**, 1735–20 (2020).

51. M. Pechmann, S. Khadjeh, F. Sprenger, N.-M. Prpic, Patterning mechanisms and morphological diversity of spider appendages and their importance for spider evolution. *Arthropod Structure and Development*. **39**, 453–467 (2010).

52. C. Wolff, M. Hilbrant, The embryonic development of the central American wandering spider Cupiennius salei. *Front. Zool.* **8**, 15–35 (2011).

53. M. Pechmann, E. E. Schwager, N. Turetzek, N.-M. Prpic, Regressive evolution of the arthropod tritocerebral segment linked to functional divergence of the Hox gene labial. *Proceedings of the Royal Society B: Biological Sciences*. **282**, 20151162 (2015).

54. B. Mittmann, C. Wolff, Embryonic development and staging of the cobweb spider Parasteatoda tepidariorum C. L. Koch, 1841 (syn.: Achaearanea tepidariorum; Araneomorphae; Theridiidae). *Dev Genes Evol*. **222**, 189–216 (2012).

55. N.-M. Prpic, R. Janssen, B. Wigand, M. Klingler, W. G. M. Damen, Gene expression in spider appendages reveals reversal of exd/hth spatial specificity, altered leg gap gene dynamics, and suggests divergent distal morphogen signaling. *Developmental Biology*. **264**, 119–140 (2003).

56. W. G. M. Damen, Parasegmental organization of the spider embryo implies that the parasegment is an evolutionary conserved entity in arthropod embryogenesis. *Development*. **129**, 1239–1250 (2002).

57. B. Mittmann, G. Scholtz, Distal-less expression in embryos of Limulus polyphemus (Chelicerata, Xiphosura) and Lepisma saccharina (Insecta, Zygentoma) suggests a role in the development of mechanoreceptors, chemoreceptors, and the CNS. *Dev Genes Evol*. **211**, 232–243 (2001).

58. K. A. Ober, E. L. Jockusch, The roles of wingless and decapentaplegic in axis and appendage development in the red flour beetle, Tribolium castaneum. *Developmental Biology*. **294**, 391–405 (2006).

59. N. Niwa, Y. Inoue, A. Nozawa, M. Saito, Y. Misumi, H. Ohuchi, H. Yoshioka, S. Noji, Correlation of diversity of leg morphology in Gryllus bimaculatus (cricket) with divergence in dpp expression pattern during leg development. *Development*. **127**, 4373–4381 (2000).

60. G. Brenneis, E. V. Bogomolova, C. P. Arango, F. Krapp, From egg to “no-body”: an overview and revision of developmental pathways in the ancient arthropod lineage Pycnogonida. *Front Zool*. **14**, 6 (2017).

61. C. Nulsen, L. M. Nagy, The role of wingless in the development of multibranched crustacean limbs. *Dev Genes Evol*. **209**, 340–348 (1999).

62. S. M. Cohen, G. Jürgens, Mediation of Drosophila head development by gap-like segmentation genes. *Nature*. **346**, 482–485 (1990).

63. L. M. Nagy, S. Carroll, Conservation of wingless patterning functions in the short-germ embryos of Tribolium castaneum. *Nature*. **367**, 460–463 (1994).

64. C. L. Hughes, T. C. Kaufman, Exploring Myriapod Segmentation: The Expression Patterns of even-skipped, engrailed, and wingless in a Centipede. *Developmental Biology*. **247**, 47–61 (2002).

65. R. Janssen, N.-M. Prpic, W. G. M. Damen, Gene expression suggests decoupled dorsal and ventral segmentation in the millipede Glomeris marginata (Myriapoda: Diplopoda). *Developmental Biology*. **268**, 89–104 (2004).

66. N.-M. Prpic, Homologs of wingless and decapentaplegic display a complex and dynamic expression profile during appendage development in the millipede Glomeris marginata (Myriapoda: Diplopoda). *Front. Zool.* **1**, 6 (2004).

67. J. F. Coulcher, M. J. Telford, Comparative gene expression supports the origin of the incisor and molar process from a single endite in the mandible of the red flour beetle Tribolium castaneum. *EvoDevo*. **4**, 1–12 (2013).

68. E. L. Jockusch, C. Nulsen, S. J. Newfeld, L. M. Nagy, Leg development in flies versus grasshoppers: differences in dpp expression do not lead to differences in the expression of downstream components of the leg patterning pathway. *Development*. **127**, 1617–1626 (2000).

69. D. Walossek, K. J. Müller, Cambrian 'Orsten'-type arthropods and the phylogeny of Crustacea. *Arthropod Relationships* (1997) (available at http://link.springer.com/chapter/10.1007/978-94-011-4904-4_12).

70. L. Størmer, *Studies on Trilobite Morphology. Part I: The Thoracic Appendages and Their Phylogenetic Significance* (1939; http://scholar.google.com/scholar?q=related:Zlt52K7FEQgJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5).

71. M. Stein, P. A. Selden, A restudy of the Burgess Shale (Cambrian) arthropod *Emeraldella brocki* and reassessment of its affinities. *Journal of Systematic Palaeontology*. **10**, 361–383 (2012).

72. M. Stein, Cephalic and appendage morphology of the Cambrian arthropod Sidneyia inexpectans Walcott, 1911. *Zoologischer Anzeiger*, 15 (2013).

73. T. A. Williams, C. Nulsen, L. M. Nagy, A complex role for distal-less in crustacean appendage development. *Developmental Biology*. **241**, 302–312 (2002).

74. N. Gorfinkiel, G. Morata, I. Guerrero, The homeobox gene Distal-less induces ventral appendage development in Drosophila. *Genes & Development*. **11**, 2259–2271 (1997).

75. J. W. Shultz, Gross muscular anatomy of Limulus polyphemus (Xiphosura, Chelicerata) and its bearing on evolution in the Arachnida. *Journal of Arachnology*. **29(3):283-303** (2001), doi:10.1636/0161-8202(2001)029%5B0283:GMAOLP%5D2.0.CO;2.

76. W. F. Hayes, Fine Structure of the Chemoreceptor Sensillum in Limulus. *J. Morphol.*, 1–35 (1971).

77. D. Legg, thesis (2014).

78. A. Maas, C. Haug, J. T. Haug, J. Olesen, Early crustacean evolution and the appearance of epipodites and gills. *Arthropod* …. **67 (2) 255-273** (2009) (available at http://www.senckenberg.de/files/content/forschung/publikationen/arthropodsystematics/asp_67_2/asp_67_2_maas_255-273.pdf).

79. H. Bäcker, M. Fanenbruck, J. W. Wägele, A forgotten homology supporting the monophyly of Tracheata: The subcoxa of insects and myriapods re-visited. *Zoologischer Anzeiger - A Journal of Comparative Zoology*. **247**, 185–207 (2008).

80. J. F. Coulcher, G. D. Edgecombe, M. J. Telford, Molecular developmental evidence for a subcoxal origin of pleurites in insects and identity of the subcoxa in the gnathal appendages. *Scientific reports*. **5**, 1–8 (2015).

81. Y. Kobayashi, K. Niikura, Y. Oosawa, Y. Takami, Embryonic development of Carabus insulicola (Insecta, Coleoptera, Carabidae) with special reference to external morphology and tangible evidence for the subcoxal theory. *Journal of Morphology*. **274**, 1323–1352 (2013).

82. M. L. Roonwal, Studies on the embryology of the African migratory locust, Locusta migratoria migratorioides Reiche and Frm.(Orthoptera, Acrididae). II. Organogeny. *Philosophical Transactions of the Royal Society of \ldots*. **CCXXVII-B543** (1937).

83. R. Heymons, *Beiträge zur Morphologie und Entwicklungsgeschichte der Rhynchoten* (1899).

84. T. Uchifune, R. Machida, Embryonic development of Galloisiana yuasai Asahina, with special reference to external morphology (insecta: Grylloblattodea). *Journal of Morphology*. **266**, 182–207 (2005).

85. J. Bitsch, The morphological groundplan of hexapoda: critical review of recent concepts. *Annales de la Société entomologique de France*. **30**, 103–129 (1994).

86. K. W. Verhoeff, Über die Entwicklungsstufen der Steinläufer, Lithobiiden, und Beiträge zur Kenntnis der Chilopoden. *Zoologische Jahrbücher*. **8** (1905).

87. O. W. Tiegs, The embryology and affinities of the symphylas based on a study of Hanseniella agilis. *Journal of Cell Science* (1940) (available at http://jcs.biologists.org/content/s2-82/325/1.short).

88. T. Wesener, Sternites and spiracles - The unclear homology of ventral sclerites in the basal millipede order Glomeridesmida (Myriapoda, Diplopoda). *Arthropod Structure*, 9 (2014).

89. Y. Kobayashi, K. Niikura, Y. Oosawa, Y. Takami, Embryonic development of Carabus insulicola (Insecta, Coleoptera, Carabidae) with special reference to external morphology and tangible evidence for the subcoxal theory. *J. Morphol.* **274**, 1323–1352 (2013).

90. X. Franch-Marro, N. Martín, M. Averof, J. Casanova, Association of tracheal placodes with leg primordia in Drosophila and implications for the origin of insect tracheal systems. *Development*. **133**, 785–790 (2006).

91. P. P. Sharma, Chelicerates and the Conquest of Land: A View of Arachnid Origins Through an Evo-Devo Spyglass. *Integrative and Comparative Biology*. **57**, 510–522 (2017).

92. H. M. T. Choi, M. Schwarzkopf, M. E. Fornace, A. Acharya, G. Artavanis, J. Stegmaier, A. Cunha, N. A. Pierce, Third-generation in situhybridization chain reaction: multiplexed, quantitative, sensitive, versatile, robust. *Development*. **145**, dev165753-122 (2018).

93. M. Pechmann, N.-M. Prpic, Appendage patterning in the South American bird spider Acanthoscurria geniculata (Araneae: Mygalomorphae). *Dev Genes Evol*. **219**, 189–198 (2009).

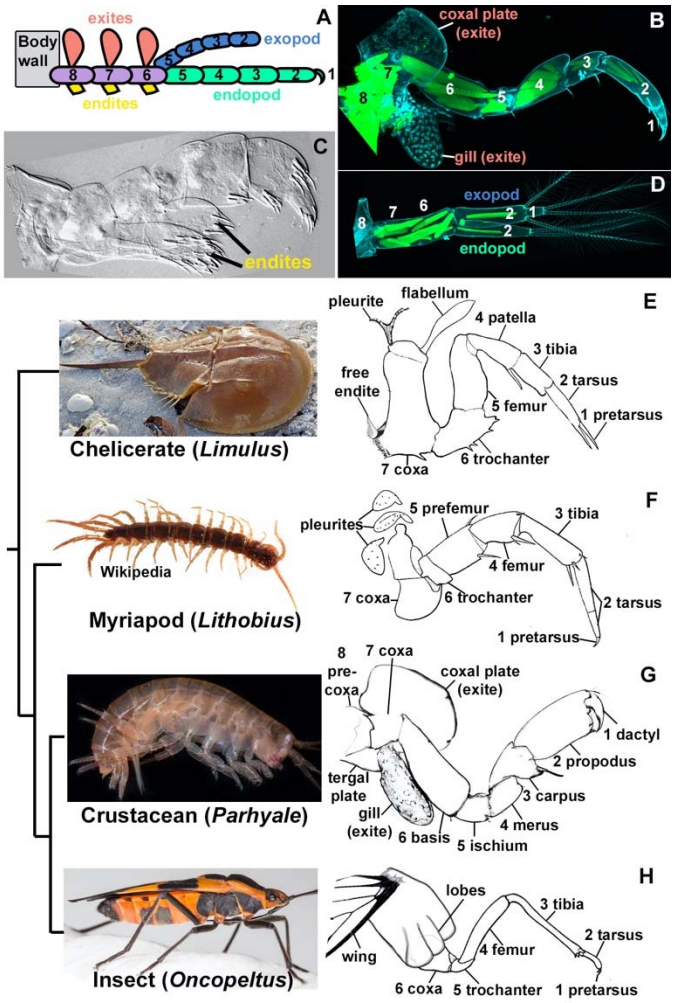


Fig. 1. Arthropod legs. (A) Generalized crustacean leg based on {Holthuis:1993wl}. Proximal leg segments (purple), exopod (blue), endopod (teal), exites (peach), endites (yellow). (B) Confocal image of thoracic walking leg of a crustacean (*Parhyale*). Muscle (green phalloidin), autofluorescent cuticle. Exites (here coxal plate and gill) do not have muscles inside ("intrinsic"). Leg segments have muscles that insert via a tendon on joints. Muscles are braced around the walls of one segment and insert on the rim of a distal segment to move it. (C) Mouthpart (maxilliped) of a crustacean (*Parhyale*). Endites emerge medially. (D) Biramous leg of a crustacean (*Parhyale*). Endopod and exopod are segmented with internal musculature. Compare to exites in B which are not segmented and lacks muscles inside. (E – H) Chelicerates, myriapods, crustaceans, and insects have different numbers, shapes, and names for their leg segments. Phylogeny based on {LozanoFernandez:2019ko}

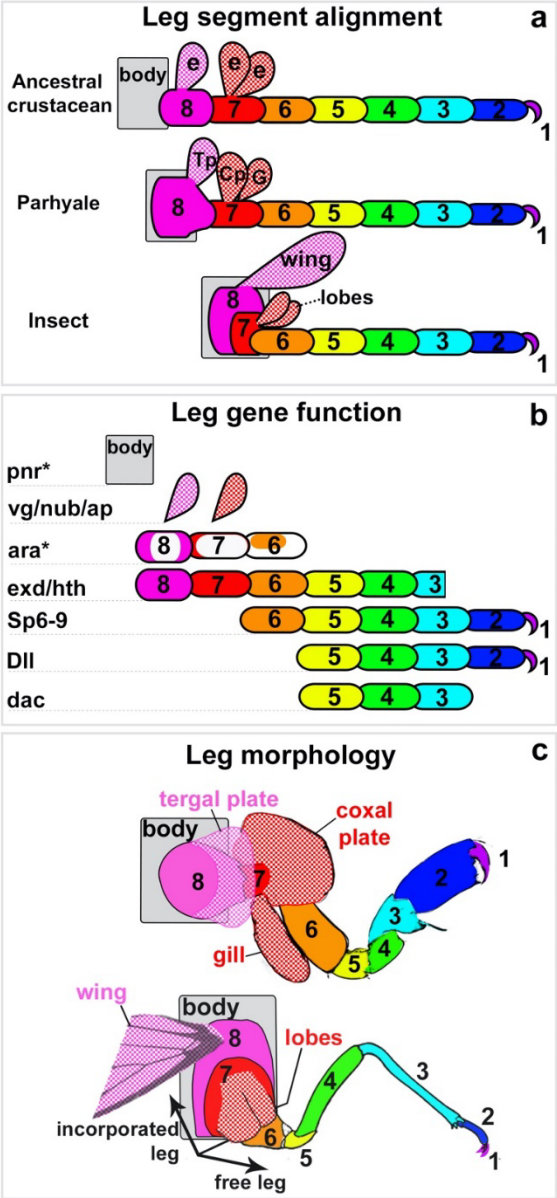


Fig. 2. Leg segment homologies (colours) between insects, *Parhyale*, and a hypothetical ancestral crustacean (from Bruce & Patel 2020 (*2*)). (a) based on leg gene function alignment. (b) A schematic of each leg segment that requires each leg patterning gene in crustaceans and insects. Based on the function of *exd*, *hth*, *Dll*, *Sp6-9*, and *dac*, the six distal leg segments of crustaceans and insects (leg segment 1 through leg segment 6) correspond with each other in a one-to-one fashion. Expression (*) of *pnr* and *ara*, as well as expression and function of wing genes, suggests that insects retain two additional proximal leg segments (7 and 8), each with an exite. In this model, the exites of pink leg segment 8 are homologous: the ancestral crustacean precoxa exite (pink, e), *Parhyale* tergal plate (Tp), and insect wing; and the exites of red leg segment 7 are homologous: the ancestral coxa exite (red, e), *Parhyale* coxal plate (Cp) and gill (G), and insect supracoxal lobes. (c) Leg segment morphologies in *Parhyale* and insect.

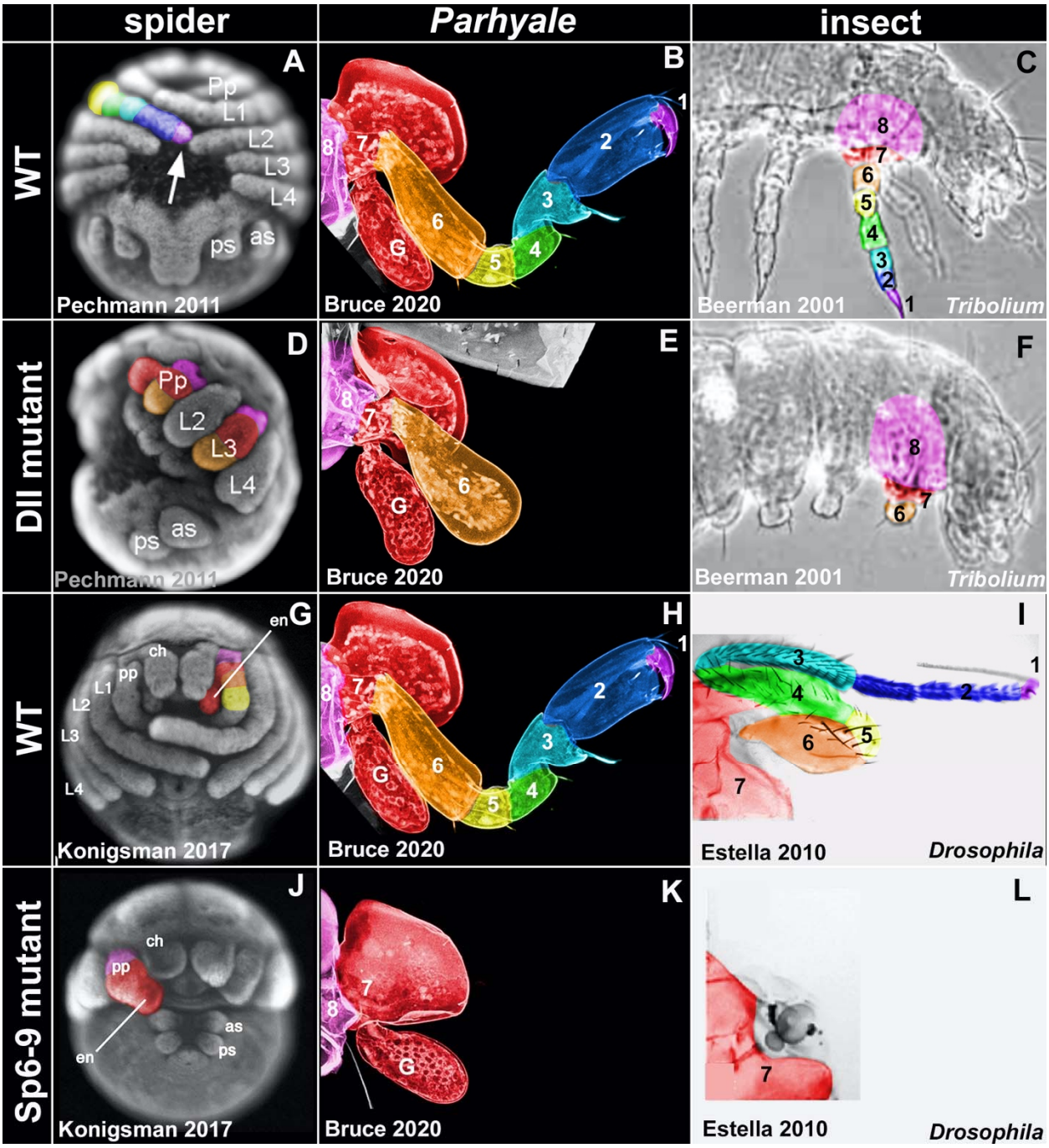


Fig. 3. In spiders, *Parhyale*, and insects, *Dll* is required for the development of leg segments 1 - 5, counting from the distal end of the leg. A, D from Pechmann 2011. B, E, H, K from Bruce&Patel (?) 2020. C, F From Beerman 2001. G, J from Konigsman 2017. I, L from Estella 2010.

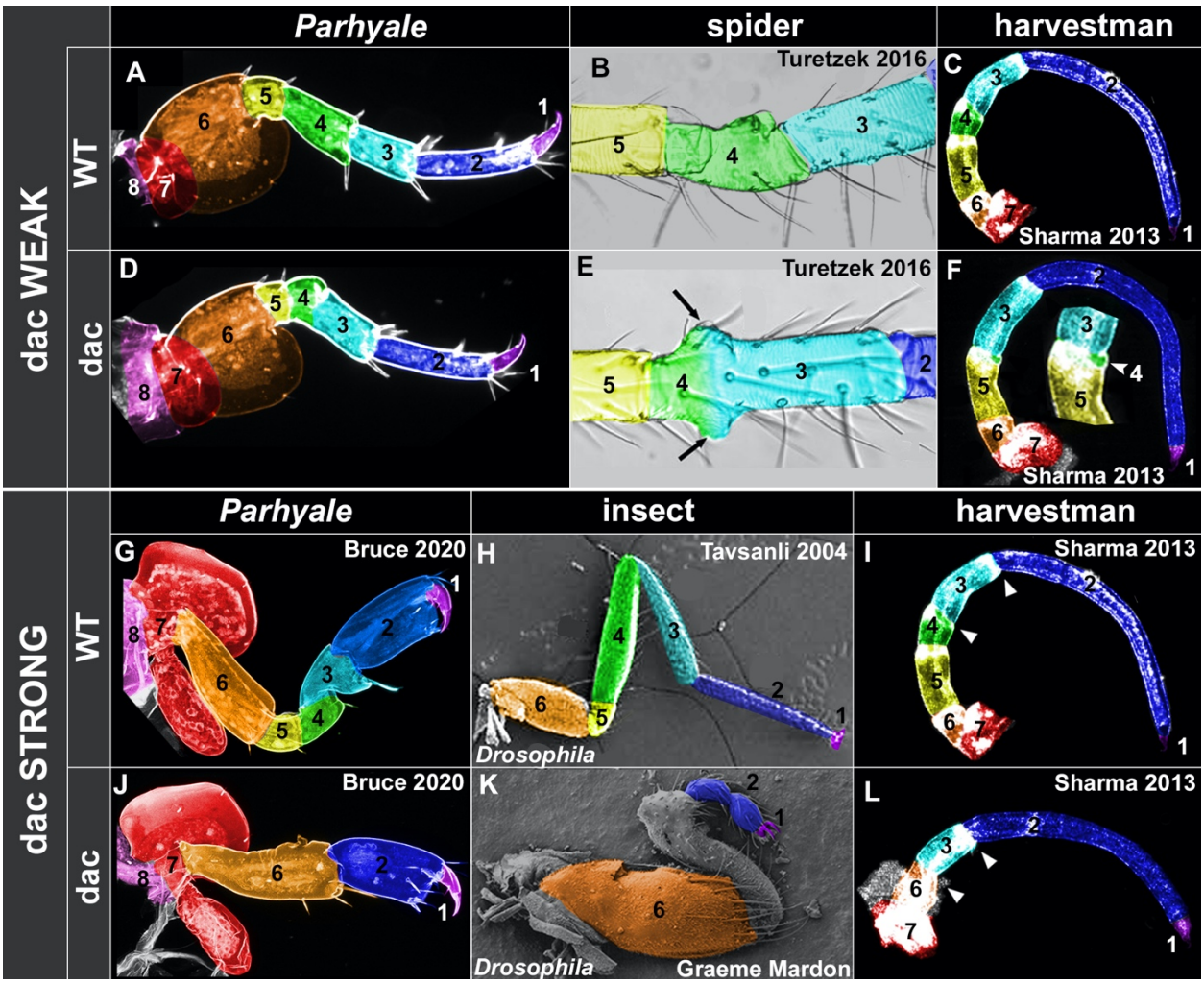


Fig. 4. *Dac* function in *Parhyale*, spider, harvestman, and insect (*Drosophila*). In spiders, harvestman, and *Parhyale*, a weak *dac2* phenotype causes green leg segment 4 to be truncated and fused onto cyan leg segment 3. In harvestman, *Parhyale*, and *Drosophila*, a strong *dac2* phenotype affects leg segments 3 – 5.

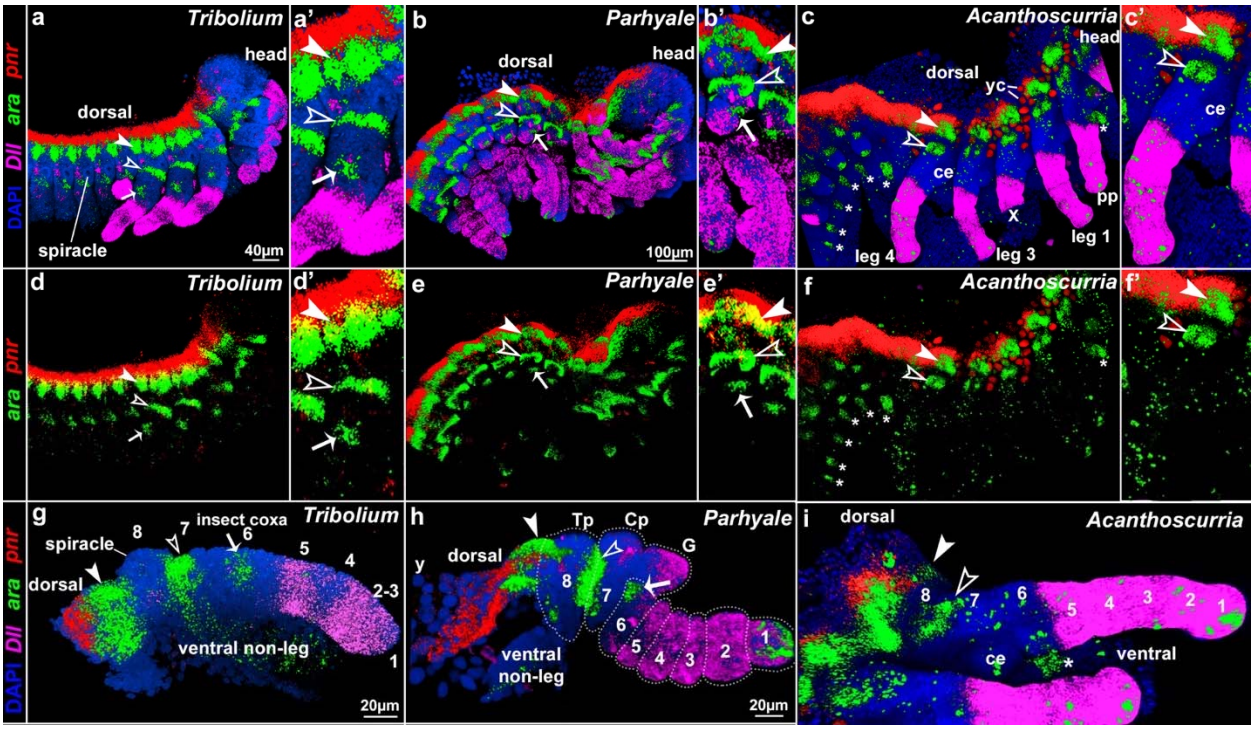


Fig. 5 Elucidating proximal leg segments in arthropods. (a, a', d, d') dissected right half of *Tribolium* embryo. a' and d' are zoomed in images of a and d respectively. (b, b', e, e'), dissected right half of *Parhyale* embryo. b' and e' are zoomed in images of b and e respectively. (c, c', f, f') dissected right half of *Acanthoscurria* embryo. c' and f' are zoomed in images of c and f respectively. (g) dissected leg of *Tribolium* embryo. (h) dissected T3 leg of *Parhyale* embryo. Large cells dorsal to *pnr* expression in b and f are extra-embryonic yolk (y) that exist prior to dorsal closure. (i) dissected leg of *Acanthoscurria* embryo. In all three arthropods, leg segments 1 through 5 are identified by *Dll* expression (magenta)(*93*). In all three arthropods, the two *ara* (green) armband domains bracket a region proximal to leg segment 7. In all three arthropods, *pnr* (red) marks the most dorsal domain and is adjacent and partially overlapping the dorsal *ara* domain. In g, h, i, *ara* is expressed in a smattering of ventral non-leg cells. *Tribolium* larvae have a fused tibia and tarsus, the tibiotarsus, here labelled 2-3(*80*). In *Acanthoscurria*, leg segment 7 is easily identified by the coxal endite (ce) that bulges medially. In both *Tribolium* (e) and *Parhyale* (f) legs, a muscle expressing *pnr* and *ara* that extends the length of leg segments 7 and 8 was masked to clearly show the ectodermal domains. Tp, tergal plate. Cp, coxal plate. G, gill. In *Acanthoscurria*, auto-fluorescent yolk cells (yc); ventral expression domain of ara (*). X in c indicates distal leg that was accidentally removed.

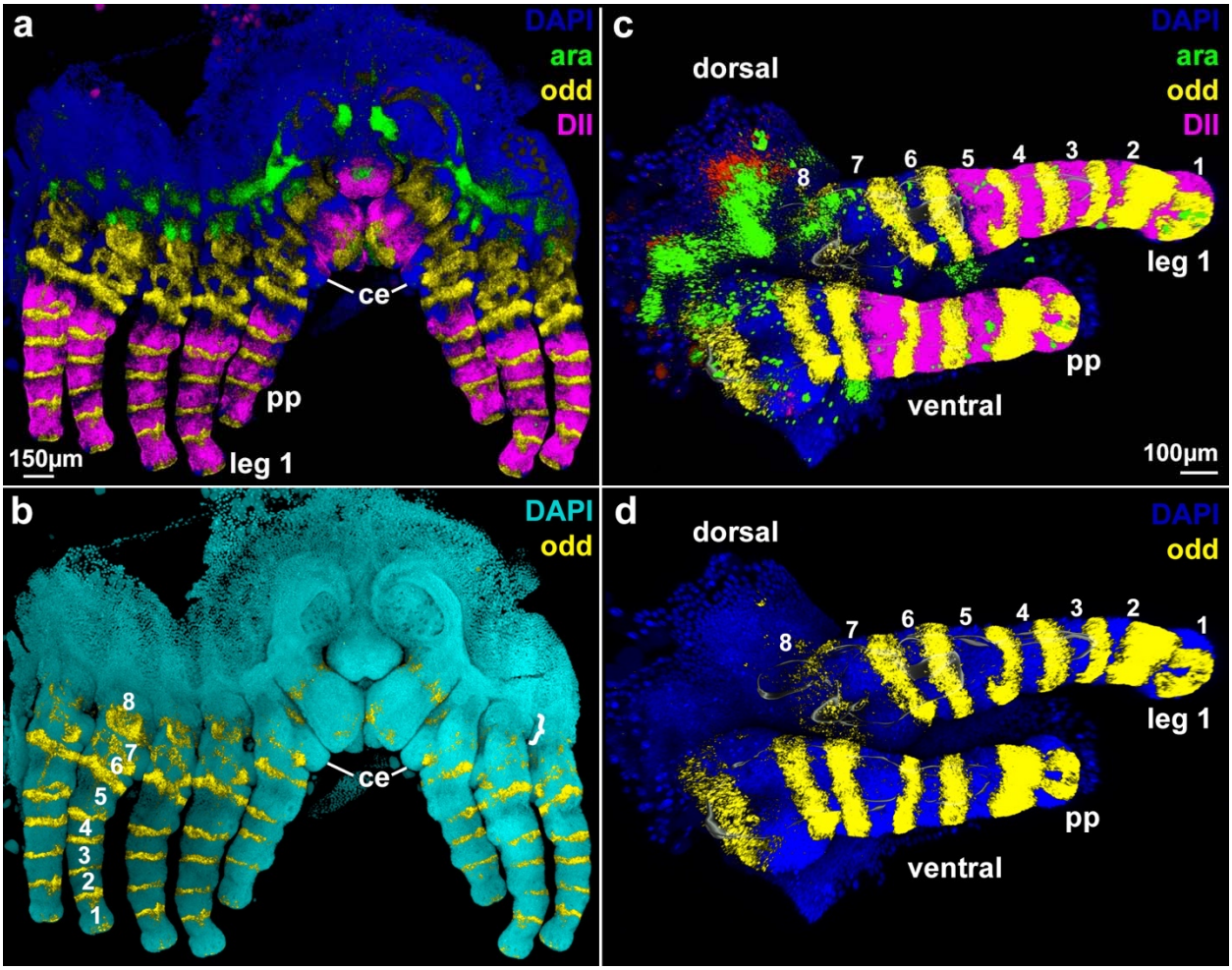


Fig. 6. Expression of *odd-r3* in *Acanthoscurria*. *odd-r3* is expressed in the distal region of each leg segment where the joint will later form. a, b Stage 12.5 *Acanthoscurria* embryos dissected away from yolk mass. Leg segment 7 is readily identified by the presence of a medial bulge, the coxal endite (ce). Proximal to leg segment 7, there is a leg-segment-like bulge (white curly brace), which expresses *odd-r3* in the distal region. *pnr* expression in this late stage embryo is reduced. DAPI in c is shown in cyan to better observe morphology. c, d dissected pedipalp (pp) and walking leg 1 from Stage 11.5 embryo where leg segments are less developed. Internal muscle that expresses *odd-r3* is masked to make the ectodermal *odd-r3* expression clearer, but this muscle can still be observed in a and b.

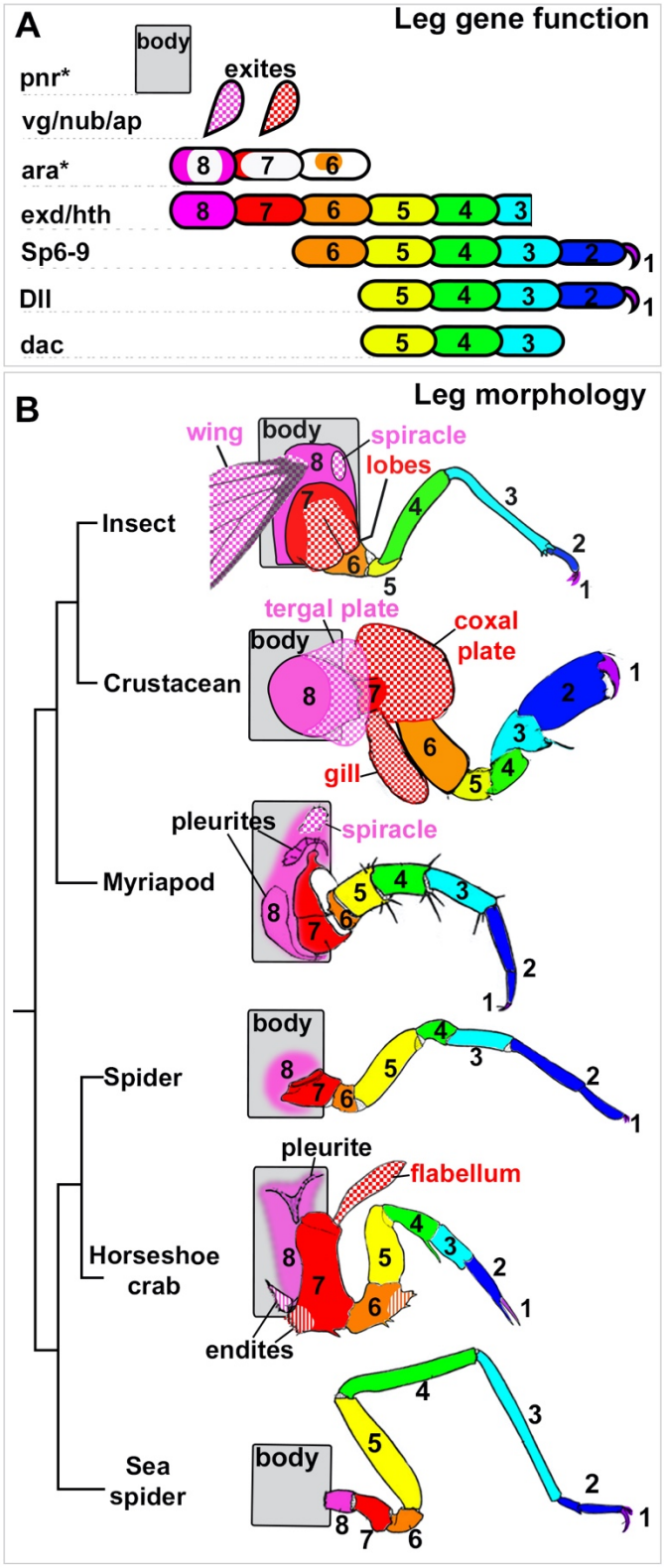

A
Leg gene function
body
pnr*
exites
vg/nub/ap
ara*
8
7
6
exd/hth
8
7
6
5
4
3
Sp6-9
6
5
4
3
2
1
Dll
5
4
3
2
1
dac
5
4
3
B
Leg morphology
Insect
wing
body
spiracle
lobes
Crustacean
tergal plate
body
coxal plate
gill
Myriapod
pleurites
spiracle
Spider
body
Horseshoe crab
pleurite
flabellum
endites
Sea spider
body

Fig. 7. Model of how to align all arthropod legs. A. Schematic of which genes function is related to (specific) leg segments. B. Morphology and homologies of arthropod leg segments based on leg gene function in insects, *Parhyale*, and chelicerates. Colors and patterns indicate proposed homologies. Exites (checker pattern); endites (stripe pattern). Drawings in B modified from Snodgrass 1952.

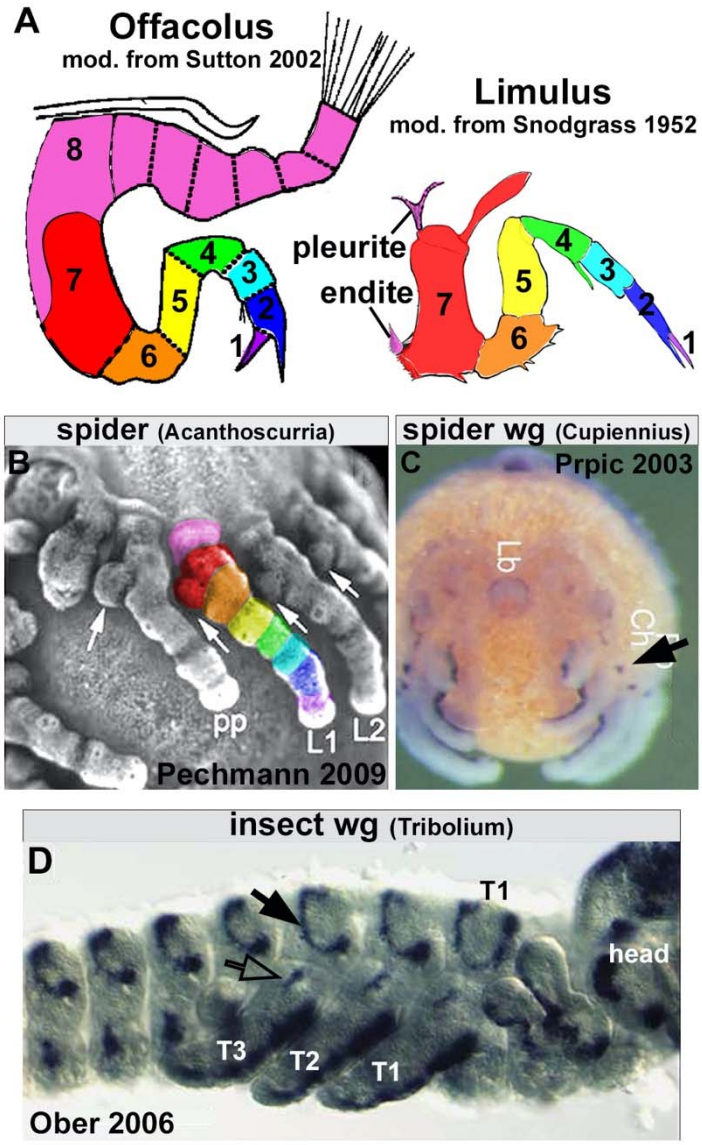


Fig. 8. The ancestral leg segment 8 may have been incorporated into the body wall of chelicerates with seven leg segments. A. Comparison of limbs in *Limulus* and the fossil xiphosurid *Offacolus*. Drawings of legs of *Limulus* (after Snodgrass 1952) and *Offacolus* (after Sutton 2002) were scaled to the same size, then red leg segment 7 in *Limulus* was superimposed on *Offacolus* to draw an approximation of this leg segment. If red leg segment 7 is the same size and shape in *Offacolus* and *Limulus*, then the exopod of *Offacolus* would emerge from a proximal 8th leg segment, here in pink. B. Embryo of bird spider *Acanthoscurria* with leg segments colored in. The embryonic spider coxa is readily identified by the conspicuous endite (arrows). An additional leg segment-like structure (pink) can be observed proximal to the spider coxa on all leg segments. C. *wg* is expressed in a ventral stripe on each leg, as expected, but also in a dot on the dorsal-most region of the leg (arrow). D. *wg* is expressed in two regions (closed arrow and open arrow) above the insect coxa in *Tribolium*

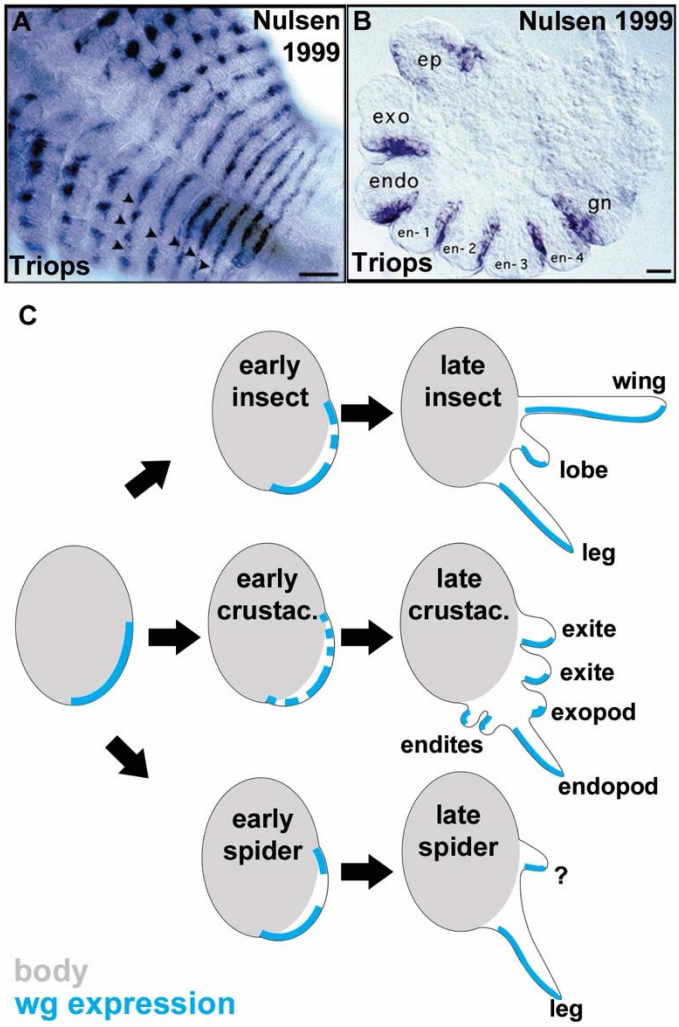
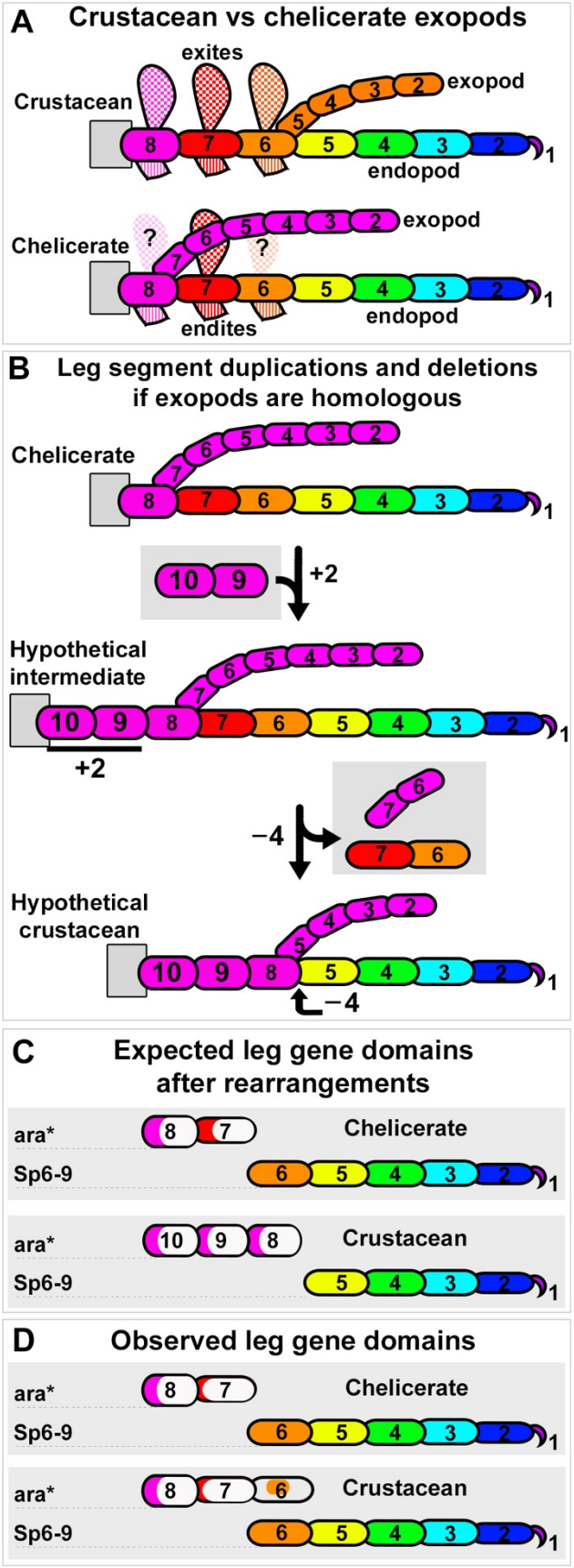


Fig. 9. *wg* expression across all arthropods makes sense from the standpoint of the model. A, B. *Triops* crustacean, from Nulsen and Nagy 1999. C. In all arthropods, *wg* is initially expressed in a solid stripe in each body segment. The crustacean leg grows out like a shelf that wraps around dorsoventrally. As endites, endopod, exopod, and exites develop, the line of *wg* expression is broken up and becomes restricted to the ventral region of each. If insects incorporated two leg segments into the body wall, and each with an exite (wing and lobe), and spiders incorporated one segment into the body wall (perhaps patterning the exopod remnant?), then *wg* expression in insects and spiders mirrors *wg* expression in crustaceans.


A
Crustacean vs chelicerate exopods
exites
Crustacean
exopod
endopod
Chelicerate
exopod
endites
endopod
B
Leg segment duplications and deletions if exopods are homologous
Chelicerate
+2
Hypothetical intermediate
+2
−4
Hypothetical crustacean
−4
C
Expected leg gene domains after rearrangements
ara*
Sp6-9
Chelicerate
ara*
Sp6-9
Crustacean
D
Observed leg gene domains
ara*
Sp6-9
Chelicerate
ara*
Sp6-9
Crustacean

835

Fig. 10. Exopods of chelicerates and crustaceans emerge from different leg segments. A. Chelicerate and crustacean exopods emerge from different leg segments. B. Illustration of leg segment deletions and duplications that would be necessary in order for chelicerate and crustacean exopods to be homologous but still maintain the observed number of 8 leg segments. C. expected leg gene expression and function in crustaceans if legs underwent the duplications and deletions in B and exopods are homologous. D. Observed leg gene expression and function is not consistent with the duplications and deletions that would be required if crustacean exopods evolved from chelicerate exopods.

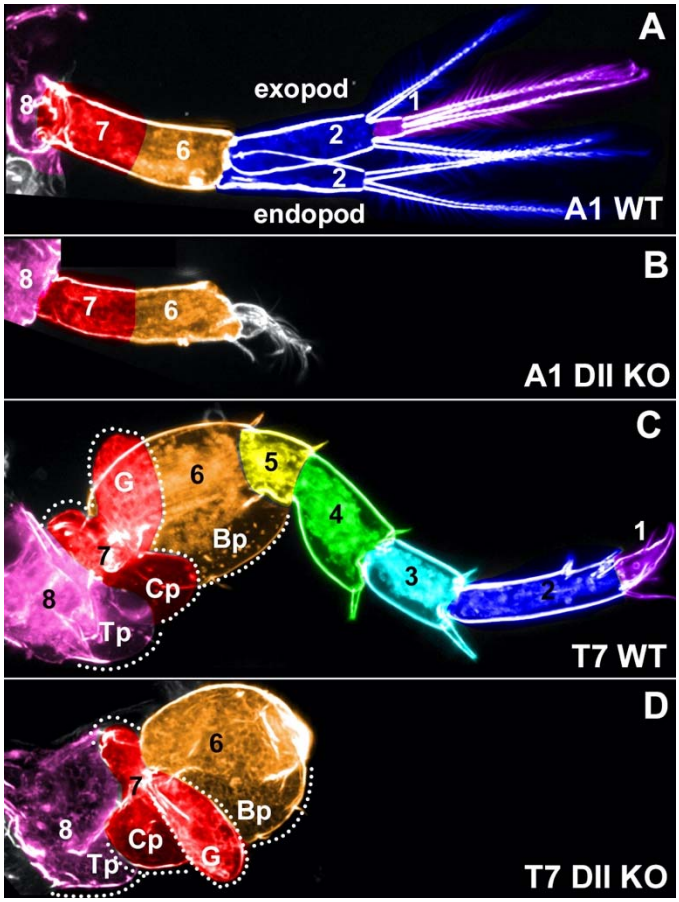


Fig. 11. *Dll* KO in *Parhyale* deletes the entire exopod and endopod, but leaves the exites unaffected. A, B first abdominal appendage (A1, biramous swimmeret), exopod and endopod are deleted. Patch of bristles on leg segment 6 that velcros left and right appendages together. C, D. seventh thoracic appendage (T7, jumping leg), endopod (telopod) is deleted, exites are unaffected. Gills point in different directions due to sample mounting.

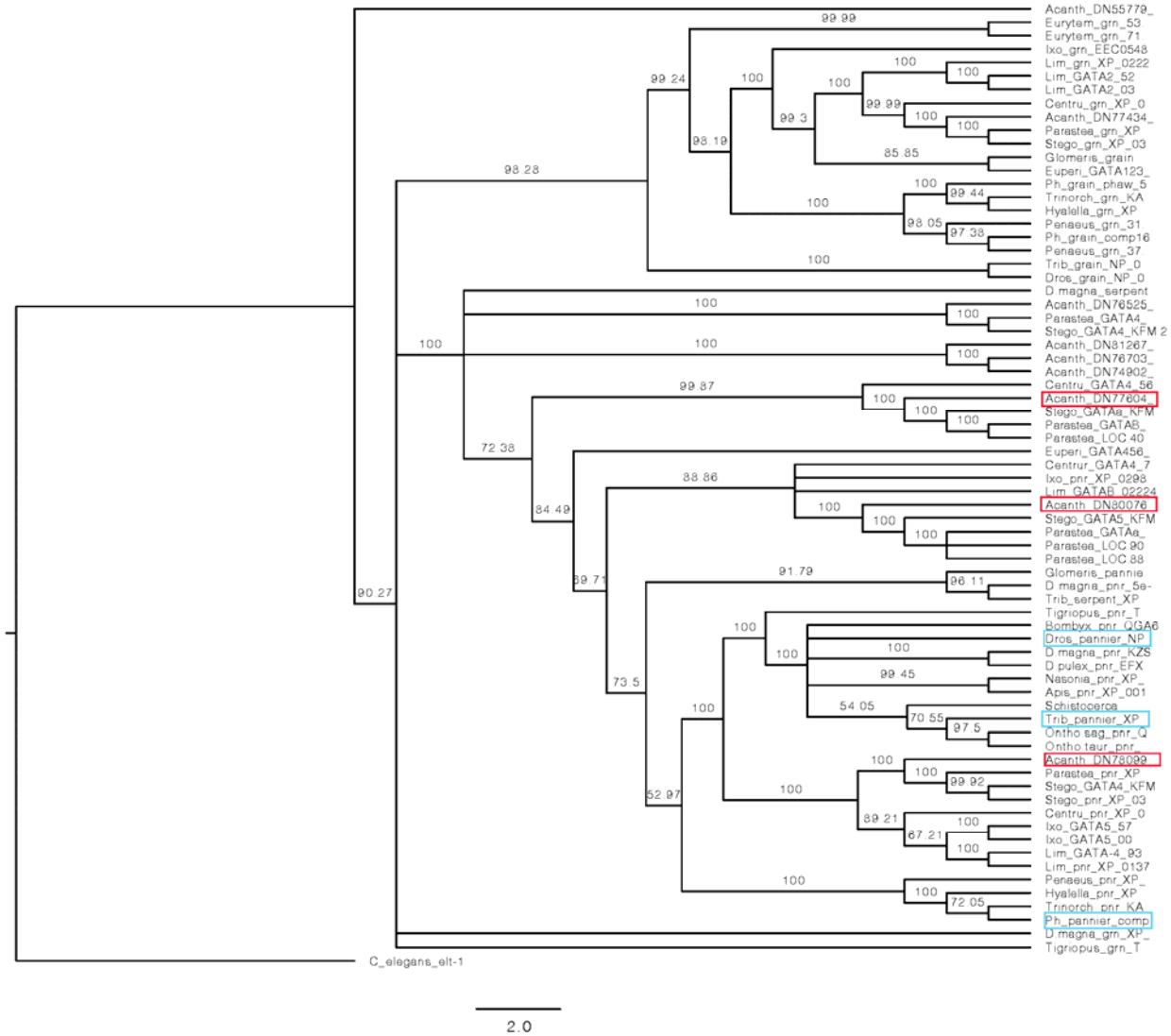


Fig. S1. Three orthologs of *pnr* were identified in *Acanthoscurria* with closest homology to *Drosophila*, *Tribolium*, and *Parhyale pnr* (Fig. S1-2). However, only one of these was expressed at the stages examined, Acanth_DN78099, and was presumed to be *pnr*. Consensus tree generated using Mafft, which gave similar topology to Clustal consensus tree.