A Unified Framework to Homologize Appendage Segments across Arthropoda

Authors: Heather S. Bruce[1,2]* and Nipam H. Patel[2,3]

Affiliations: 1. University of California Berkeley, Berkeley, CA. 2. Marine Biological Laboratory, Woods Hole, MA. 3. Dept of Organismal Biology and Anatomy, University of Chicago, Chicago, IL
*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 spiders, we show that all arthropod legs can be aligned in a one-to-one fashion. We propose a model wherein insects incorporated two proximal leg segments into the body wall, which moved the ancestral leg lobe (exite) up onto the back to later form wings. For myriapods and 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, and convergent evolution.

Keywords: arthropod; appendage; novelty

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, breathing,

sensing and chemoreception, osmoregulation, copulation, silk-production, venom delivery, and more(*1*). For over a century, researchers have sought to understand the evolutionary trajectories and relationships of all of these structures and their functions. However, this requires knowledge of how to align leg segments between insects, crustaceans, myriapods, and chelicerates, so that the proximal-distal position is known. Such an alignment would answer long-standing questions about the origins and homologies of many fascinating arthropod structures, such as the origin of insect wings, whether the exopods (the lateral branch when a leg is split) of chelicerates and crustaceans are homologous to each other, as well as histories and relationships of leg outgrowths such as the *Limulus* flabellum. The answers to these questions in turn have larger evolutionary implications about novelty, homology, the frequency of gene co-option, as well as the convergent evolution and plasticity of developmental fields across large phylogenetic distances. A framework to homologize arthropod appendages (i.e. the one-by-one alignment of arthropod leg segments of all major arthropod groups) opens up a powerful system for studying these questions.

Arthropod leg terms

An understanding of arthropod appendage diversity requires a few arthropod terms (Fig. 1A). Using crustacean legs as an example, the proximal part ("protopod") of an arthropod leg often bears lobes of various shapes and functions (*2*). When emerging laterally from the proximal part of the leg, 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 split ("biramous"), meaning that two distal leg branches 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 arise when the distal end of the developing limb bud splits in two, while exites emerge later by budding off of the existing proximal leg (*3-5*).

Arthropod legs are divided into segments (Fig. 1E – H). Leg segments are bounded to either side by joints where muscles insert (Fig. 1B) (*6, 7*). Leg segments sometimes have subdivisions within them where no muscle inserts, i.e. muscles pass through the subdivision without inserting. These serve as points of flexion, such as the subdivisions in the tarsus of insects, arachnids, and myriapods (Fig. 1F), but do not represent true segments.

The leg segments of chelicerates, myriapods, crustaceans, and insects have different numbers, shapes, and names. Chelicerates have either 7 or 8 leg segments, myriapods have either 6 or 7, insects have 6, and the crustacean ground plan has 7 or 8 leg segments (Fig. 1E-H) (*2, 6, 8-10*). For over a century, researchers have proposed many different theories to account for this variation (*2, 6, 11, 12*). Using morphology, authors have proposed leg segment deletions, duplications, and fusions to account for the different numbers of leg segments between arthropod taxa. Other authors concluded that arthropod legs cannot be homologized and aligned at all(*13*). More recently, the expression of several leg patterning genes was compared in a chelicerate, a crustacean, and an insect, but this too provided no clear way of aligning and homologizing leg segments, likely due to dynamic changes in gene expression(*14*).

Alignment of insect and crustacean legs reveals the origin of insect wings

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 (*15*), one theory is that insect wings evolved from crustacean exites (outgrowths such as plates or gills on the proximal leg) (*16*). The second theory proposes that crustacean exite genes were co-opted and expressed by an unrelated tissue, the dorsal body wall, in order to form insect wings on the back, meaning wings are a novel structure not present in crustaceans. To test these two hypotheses, Bruce & Patel 2020 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) (*17*).They then wanted to understand the proximal leg segments. To do so, they compared the expression of *pannier* (*pnr*), the Iroquois complex gene *aurucan* (*ara*), and wing genes in *Parhyale* and insects (*17*, *18*). They found that, in both *Parhyale* and insects, the expression of *ara* distinguishes two proximal leg segments (leg segments 7 and 8; Fig. 2), each of which bears an exite that expresses wing genes; and that *pnr* expression marks the true body wall. These data suggested that insects had incorporated two ancestral proximal leg segments, 7 and 8, into the body wall, which carried their exites dorsally. An alignment of the function of five leg patterning genes in conjunction with the expression of *ara*, *pnr*(*17*), and wing genes(*18*) places the insect wing in register with the *Parhyale* tergal plate: both are exites on the proximal-most (8th) leg segment. Therefore

insect wings are not novel, but instead evolved 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, this is the same as the chelicerate ground plan(*6*) as well as the ancestral arthropod ground plan, because the ancestor of all living arthropods also had 8 leg segments (*19*). If insect and crustacean legs can be homologized, this model may extend to myriapods and chelicerates as well, in a grand unified theory of appendages across all four groups of arthropods.

Homologizing chelicerate and pancrustacean legs

To align *Parhyale* and chelicerate legs, we compared our leg segment deletion phenotypes in *Parhyale* to previously published results in chelicerates, as we had done for insects. Functional experiments in chelicerates have been performed for *hth*, *Dll*, *Sp6-9*, and *dac*. While functional data is not available for *exd* in a chelicerate, given that *exd* and *hth* are cofactors, we presume that the leg segment deletion phenotypes of *exd* and *hth* are similar in chelicerates as they are in other arthropods (*20-23*). Based on the leg segment deletion phenotypes of *hth*, *Dll*, *Sp6-9*, and *dac*, 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.

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)(*14*, *17*, *24-31*). In spiders, *Parhyale,* and insects, *Sp6-9* is required for the development of leg segments 1 – 6 (Fig. 3G – L) (*17*, *32-37*). 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) (*17, 38-41*). In spiders, harvestman, and *Parhyale*, a weak dac2 phenotype causes leg segment 4 to be truncated and fused onto leg segment 3 (Fig. 4A - F). In harvestman, *Parhyale*, and insects, a strong dac2 phenotype affects leg segments 3 – 5 (Fig. 4G - L). For the above comparisons, we note that RNAi gives a range of partial knockdowns, but we focus on what appear to be the null phenotypes.

If the six distal leg segments of *Parhyale*, chelicerates, and insects are in alignment, this suggests a model of how to homologize all of their leg segments (Fig. 5). Chelicerates with eight leg segments, such as sea spiders, align one-to-one with the eight leg segments of *Parhyale* and insects. However, in chelicerates with seven leg segments, 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. Another possibility is that the proximal-most leg segment was incorporated into the body wall, similar to how insects incorporated the proximal leg segments into their body wall. We discuss observations from morphology, phylogeny, paleontology, embryology, and molecular studies that argues that the proximal-most leg segment was incorporated into the body wall.

Discussion

Early-branching chelicerates often have eight leg segments like the euarthropod 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 (*6, 42*). An interesting exception is the horseshoe crab, *Limulus* (*6*). *Limulus* is an early-branching chelicerate (*42*)(although see (*43*)) 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. Proximal to the coxa of each walking leg is a small, spiny endite ("free endite", or "epicoxite") that does not belong to any obvious leg segment, but is nevertheless moveable by muscles(*44*). Dorsal to the coxa of each walking leg is a conspicuous Y-shaped pleurite (body wall exoskeleton plate), which articulates with the coxa. As moveable endites and pleurites are characteristic of leg segments, these may represent the remnant of a proximal 8th leg segment in *Limulus* (*45*, *46*).

Support for this hypothesis can be found by examining the fossil xiphosurids *Offacolus* (*47*) and *Dibasterium* (*48*). In stark contrast to *Limulus*, these fossil xiphosurids have a large, segmented exopod on the most proximal leg segment. In *Offacolus*, the limb base is dorsoventrally elongated such that it occupies the entire lateral region of the body. The proximal leg region is not well preserved in these fossils, and thus the authors are uncertain how many proximal leg segments were present. However, a comparison to *Limulus* suggests that the limb base of *Offacolus* may actually be two leg segments (Fig. 6A): the elongate coxa like that seen in *Limulus*(*8*, *44*), and a smaller proximal leg segment from which the fossil exopod emerged. If this smaller proximal leg segment degenerated into the body wall of living horseshoe crabs like *Limulus*, this would explain how extant horseshoe crabs lost both the ancestral eighth leg segment and the ancestral exopod.

Embryological evidence of this proximal 8th leg segment can be observed in modern chelicerates with seven leg segments. In embryos of the 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; (*29*)). 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. 6C; (*30*)), and in *Limulus*, a dot of *Dll* is expressed above each walking leg (*49*). This is reminiscent of insects, which have incorporated two leg segments into the body wall (Bruce 2020): *wg* is expressed in two dots above the leg, one dot on each incorporated leg segment (Fig. 6D (*50*)). In insects, these *wg* dots pattern exites, but in chelicerates, the *wg* dot may be patterning the remnant of the exopod.

Thus it appears that chelicerate, crustacean, and insect legs can be aligned in a one-to-one fashion. We note that no functional data is available to myriapods, but given that they share a common ancestor with chelicerates and pancrustaceans, and the legs of the other three clades of arthropods align, we believe it's reasonable to assume that this leg model applies to myriapods as well.

A grand unified theory of dorsoventral patterning of arthropod appendages

This grand unified theory of arthropod legs allows for a reinterpretation of previous molecular work 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 (*51*), wg is initially expressed in a solid stripe in each body segment, just as it is in insects (*52-54*), myriapods (*55, 56*), and spiders (*30*)(Fig. 7) . The Triops leg grows out like a shelf that wraps around dorsoventrally, rather than a cylinder. 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.

   If insect wing and lateral body wall are interpreted as proximal leg, wg expression in insects mirrors that of crustaceans. In *Tribolium* (*50*, *54*) and cricket (*53*) 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 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 also 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 (Prpic 2004; Coulcher 2013; Jockush 2000).

   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. 3C, (*30*, *57*)). Thus, arachnids have one wg domain in the lateral body wall corresponding to one fused leg segment where the exopod once emerged (Fig. 6C), and insects have two wg domains in the lateral body wall corresponding to two fused leg segments, each with an exite (Fig. 6A). The expression domains of dpp are also consistent with this reinterpretation.

Independent origins of exopods

On the proximal-most leg segment of many fossil arthropods including trilobites, *Leanchoilia*, and *Offacolus*, there is a structure that many authors believe is an exopod (although see (*8*, *12*). The exopods of these fossil arthropods are often thought to be homologous to the exopod of crustaceans (Boxshall 2004; Wallosek 1997; Schram 1986), having a single origin inherited by the common ancestor of all arthropods. This is a reasonable hypothesis when only morphology is considered. However, when the molecular evidence is considered together with morphology, exopod homology becomes less plausible: the crustacean exopod emerges from leg segment 6, while the chelicerate and early arthropod exopod emerges from leg segment 8 (Fig. 8A). 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 splitting of the leg axis (Hejnol 200; Wolff 2008), 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. Thus, if the chelicerate leg becomes split 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. 8). This $1 + 7$ configuration indeed appears to be the case for fossil euarthropods like trilobites and fossil chelicerates (*19*, *47*). In contrast, crustacean legs have a $3 + 5$ configuration, where the leg splits at the third most proximal leg segment, and therefore both endopod and exopod each have up to 5 segments(*9*), mirroring each other, for a total of 8 segments. In fact, this is not the first time that this has been noted. Stormer 1944 also noted that

the crustacean exopod arises from leg segment 6, and concluded it was be homologous to the trilobite lateral appendage(*12*).

From a molecular standpoint, the alignment of the function of chelicerate and crustacean leg gap genes supports independent origins of their exopods. If chelicerate and crustacean exopods are homologous, then the exopod must have moved from the 8th to the 6th leg segment position. However, such a shift would require four deletions and two additions of leg segments in order to keep the number of leg segments constant between chelicerates and crustaceans (Fig. 8B), as well as the accompanying rearrangements of leg gap gene expression patterns. Given that leg segment additions are rare (Boxshall 2004), 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 split 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 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 the ground states of different arthropod taxa. For example, the problematic fossil arthropod *Agnostus* 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.

If chelicerate and crustacean exopods are not homologous, when did the crustacean exopod evolve? If it evolved in Mandibulata, then we might expect the as-yet-unknown stem myriapod 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.

A third interesting outcome of this model is that the *Limulus* flabellum cannot be the remnant of the chelicerate exopod. The flabellum is an unsegmented lobe without muscle insertions, has a sensory function (*58*), and develops by budding off of the proximal leg(*45*). These are also features of crustacean exites, which has led several authors to interpret the flabellum as an exite (*8, 45*). Other authors have interpreted the flabellum as the remnant of the chelicerate exopod (*2*), because it emerges from what appears to be the most proximal leg segment, and because the flabellum expresses *Dll*(*59*). However, while *Dll* indeed patterns leg segments, it is also expressed in exites (*17*), where it patterns sensory structures (*60*). Furthermore, the data we present here suggests that the proximal-most leg segment, which carried the exopod, was reduced and incorporated into the body wall in *Limulus*. Thus, 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. We therefore support the interpretation of the flabellum as an exite. This is unexpected, given that exites are believed to have evolved in crustaceans(*2, 61*). However, this belief 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*, *Dll* knockout deletes the entire exopod and endopod, but leaves the exites

unaffected (Fig. 9) (*17*). If the *Limulus* flabellum is an exite, then *Dll* knock out will truncate the

distal leg, but leave the flabellum unaffected, except for subtle sensory structures on the

flabellum. 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

reinterpretations of the lateral lobes in fossil arthropods, which may be exites(*62*).

References

1.   R. C. A. Brusca, W. E. A. Moore, S. M. 1. A. Shuster, *Invertebrates* (ed. 3, 2016).

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

3.   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).

4.   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).

5.   G. Boxshall, D. A. JAume, Exopodites, epipodites and gills in crustaceans. *Arthropod Systematics & Phylogeny*. **67 (2) 229 – 254** (2009).

6.   J. W. Shultz, Morphology of locomotor appendages in Arachnida: evolutionary trends and phylogenetic implications. *Zool J Linn Soc*. **97:1-56** (1989).

7.   G. Boxshall, *Arthropod Limbs and their Development* (2013).

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

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

10.   D. Grimaldi, M. S. Engel, *Evolution of the Insects* (Cambridge University Press, 2005).

11.   R. E. Snodgrass, Morphology and mechanism of the insect thorax. *Smithsonian Miscellaneous Collections*. **80** (1927).

12.   L. Størmer, *On the relationships and phylogeny of fossil and recent Arachnomorpha* (Oslo, 1944).

13.   S. M. Manton, Habits, functional morphology and the evolution of pycnogonids. *Zool J Linn Soc*. **63: 1-21** (1978).

14.   A. Abzhanov, T. C. Kaufman, Homologs of Drosophila appendage genes in the patterning of arthropod limbs. *Developmental Biology*. **227**, 673–689 (2000).

15.   J. Lozano-Fernandez *et al.*, Pancrustacean Evolution Illuminated by Taxon-Rich Genomic-Scale Data Sets with an Expanded Remipede Sampling. *Genome Biol Evol*. **11**, 2055–2070 (2019).

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

17.   H. S. Bruce, N. H. Patel, Insect wings and body wall evolved from ancient leg segments. Manuscript in review. *Nature Ecology & Evolution* (2020).

18.   C. M. Clark-Hachtel, Y. Tomoyasu, Two sets of wing homologs in the crustacean, Parhyale hawaiensis. *bioRxivorg* (2017), doi:10.1101/236281.

19.   J. Yang *et al.*, Early Cambrian fuxianhuiids from China reveal origin of the gnathobasic protopodite in euarthropods. *Nat Commun*. **9**, 1–9 (2018).

20.   T. Mito *et al.*, Divergent and conserved roles of extradenticle in body segmentation and appendage formation, respectively, in the cricket Gryllus bimaculatus. *Developmental Biology*. **313**, 67–79 (2008).

21.   M. Ronco *et al.*, Antenna and all gnathal appendages are similarly transformed by homothorax knock-down in the cricket Gryllus bimaculatus. *Developmental Biology*. **313**, 80–92 (2008).

22.   C. Rauskolb, K. M. Smith, M. Peifer, E. Wieschaus, extradenticle determines segmental identities throughout Drosophila development. *Development*. **121**, 3663–3673 (1995).

23.   J. Wu, S. M. Cohen, Proximodistal axis formation in the Drosophila leg: subdivision into proximal and distal domains by Homothorax and Distal-less. *Development*. **126**, 109–117 (1999).

24.   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).

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

26.   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).

27.   A. Beermann *et al.*, 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).

28.   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).

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

30.   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).

31.   M. Pechmann *et al.*, Novel Function of Distal-less as a Gap Gene during Spider Segmentation. *PLoS Genet*. **7**, e1002342 (2011).

32.   N. D. Schaeper, N.-M. Prpic, E. A. Wimmer, A clustered set of three Sp-family genes is ancestral in the Metazoa: evidence from sequence analysis, protein domain structure, developmental expression patterns and chromosomal location. *BMC Evol. Biol.* **10**, 88 (2010).

33.   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).

34.   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).

35.   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).

36.   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).

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

38.   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).

39.   B. C. Tavsanli *et al.*, Structure–function analysis of the Drosophila retinal determination protein Dachshund. *Developmental Biology*. **272**, 231–247 (2004).

40.   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).

41.   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).

42.   J. Lozano-Fernandez *et al.*, Increasing species sampling in chelicerate genomic-scale datasets provides support for monophyly of Acari and Arachnida. *Nat Commun*, 1–8 (2019).

43.   P. P. Sharma *et al.*, Phylogenomic interrogation of arachnida reveals systemic conflicts in phylogenetic signal. *Mol. Biol. Evol.* **31**, 2963–2984 (2014).

44.   S. M. Manton, J. P. Harding, Mandibular mechanisms and the evolution of arthropods. *… of the Royal …*. **247** (1964).

45.   L. Størmer, *Studies on Trilobite Morphology. Part I: The Thoracic Appendages and Their Phylogenetic Significance* (1939).

46.   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).

47.   M. D. Sutton, D. E. G. Briggs, 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).

48.   D. E. G. Briggs *et al.*, Silurian horseshoe crab illuminates the evolution of arthropod limbs. **109**, 15702–15705 (2012).

49.   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).

50.   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).

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

52. S. M. Cohen, Specification of limb development in the Drosophila embryo by positional cues from segmentation genes. *Nature*. **343**, 173–177 (1990).

53. N. Niwa *et al.*, Correlation of diversity of leg morphology in Gryllus bimaculatus (cricket) with divergence in dpp expression pattern during leg development. *Development*. **127**, 4373–4381 (2000).

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

55. 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).

56. 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).

57. 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).

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

59. 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).

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

61. 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).

62. D. Zhai *et al.*, Three-Dimensionally Preserved Appendages in an Early Cambrian Stem-Group Pancrustacean. *Curr. Biol.* **29**, 171–177.e1 (2019).

63. L. B. Holthuis, C. H. J. M. Fransen, *The Recent Genera of the Caridean and Stenopodidean Shrimps (Crustacea, Decapoda)* (Nationaal Natuurhistorisch Museum, 1993).
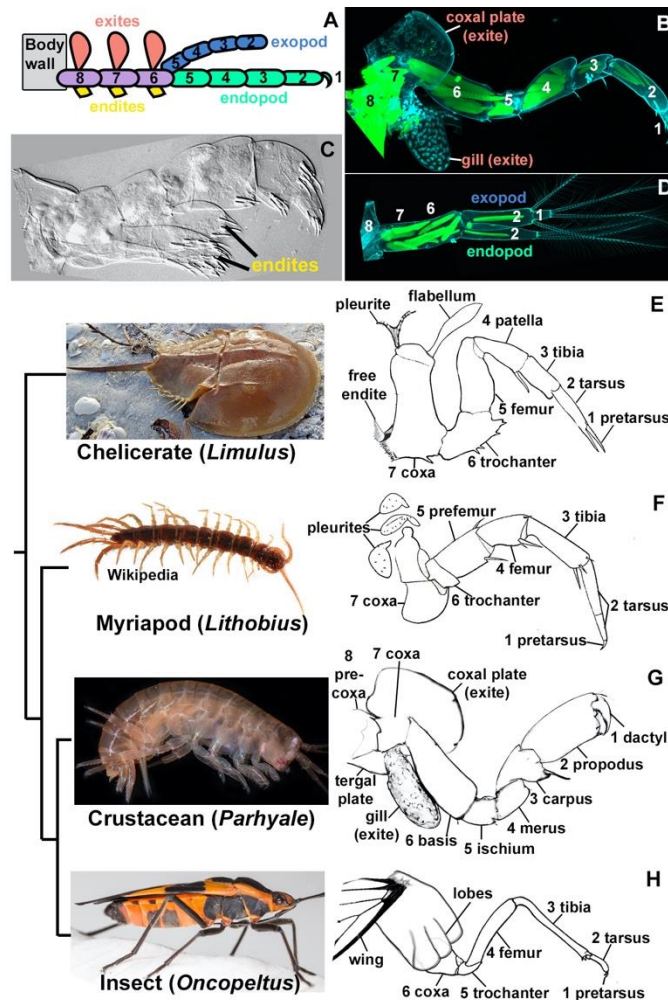
Figure 1. Arthropod legs. (A) Generalized crustacean leg based on (*63*). Proximal leg segments (purple), exopod (blue), endopod (teal), exites (peach), endites (yellow). (B) Confocal image of thoracic walking leg of a crustacean (*Parhyale*). Muscle (green phalloidin), autofluorescent cuticle. Exites (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 (*15*)
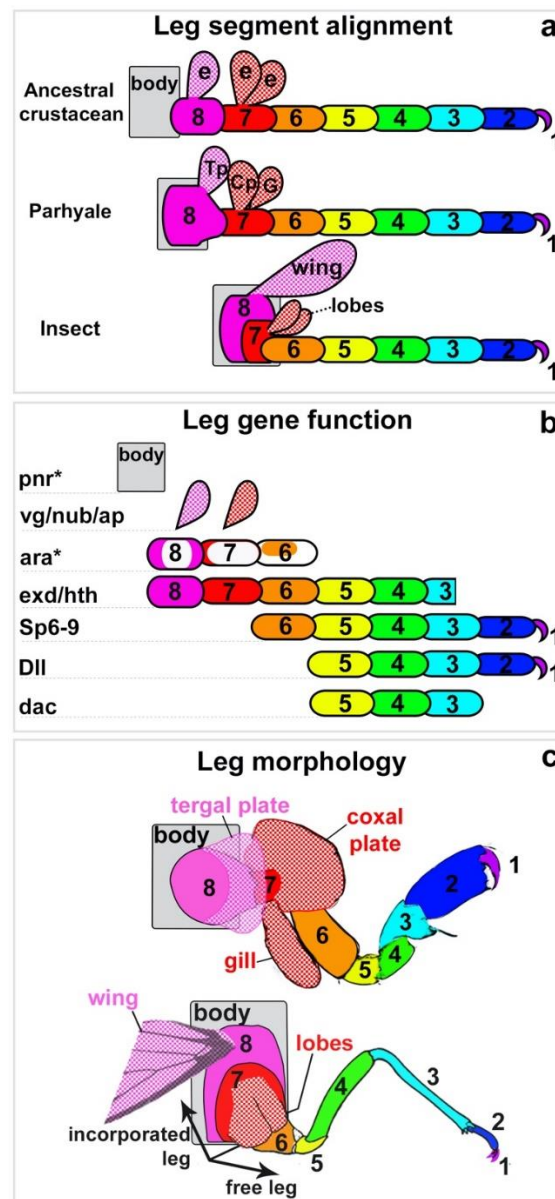
Fig. 2. Leg segment homologies (colours) between insects, *Parhyale,* and a hypothetical ancestral crustacean modified from (*63*)(a) based on leg gene function alignment. (b) A schematic of each leg segment that requires each leg patterning gene in crustaceans and insects. Based on the function of *exd*, *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.
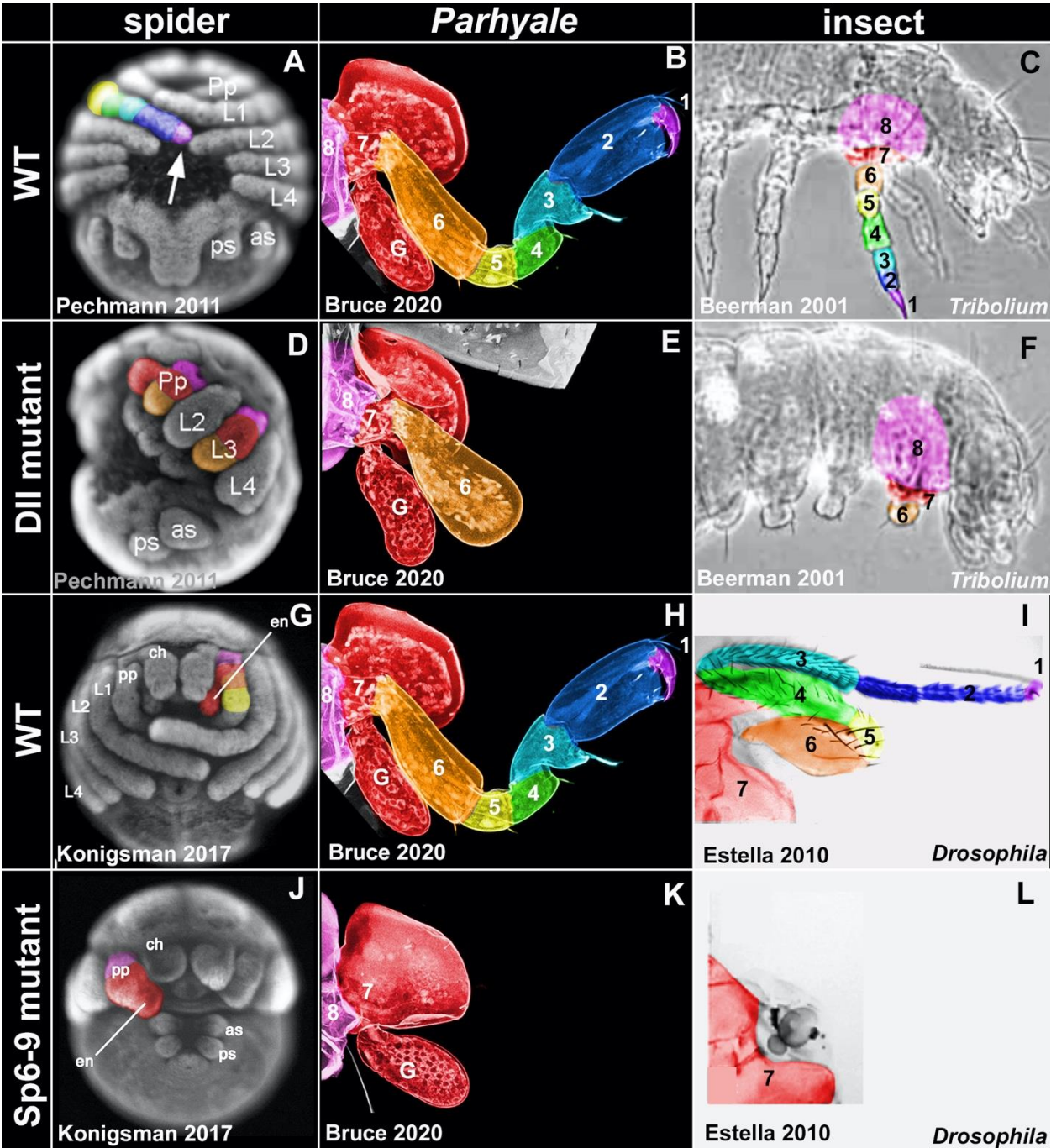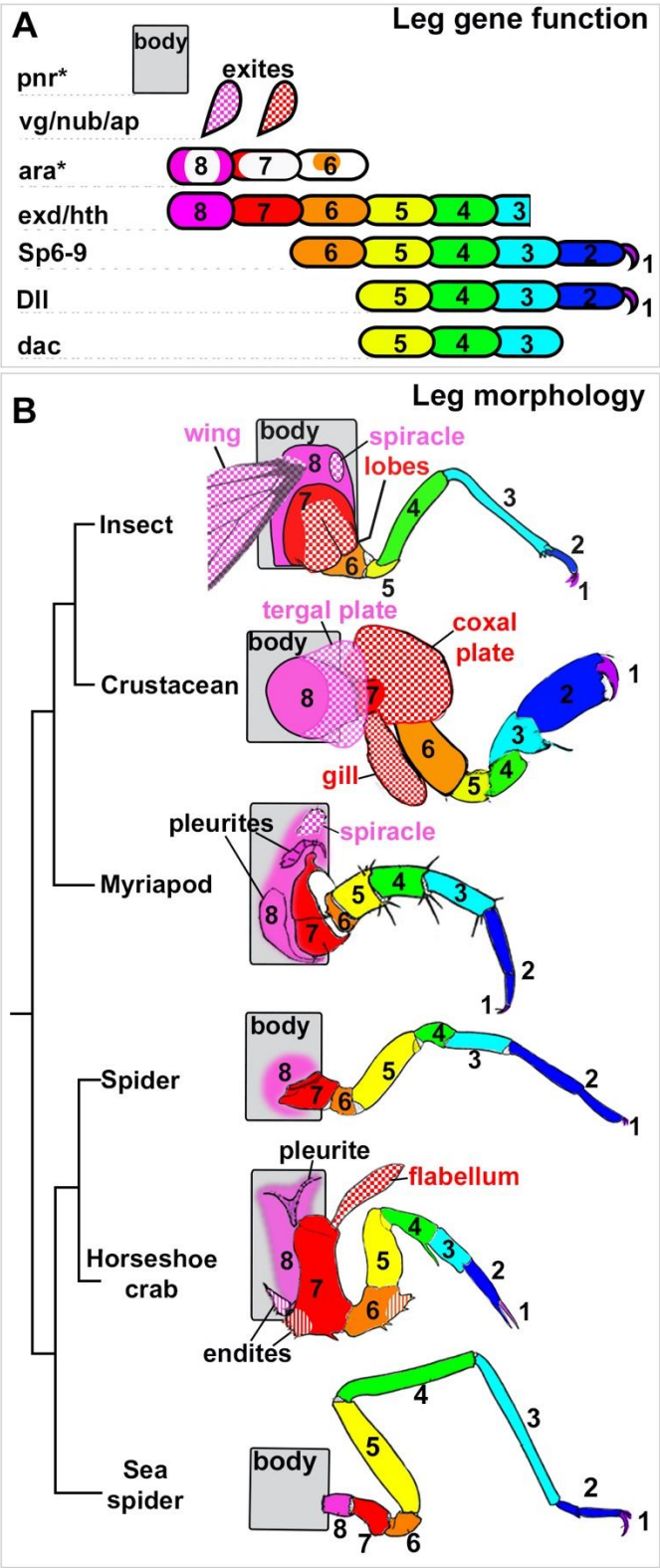
Fig. 3. In spiders, *Parhyale*, and insects, *Dll* is required for the development of leg segments 1 - 5, counting from the distal end of the leg. A, D from Pechmann 2011. B, E, H, K from Bruce&Patel (?) 2020. C, F From Beerman 2001. G, J from Konigsman 2017. I, L from Estella 2010.
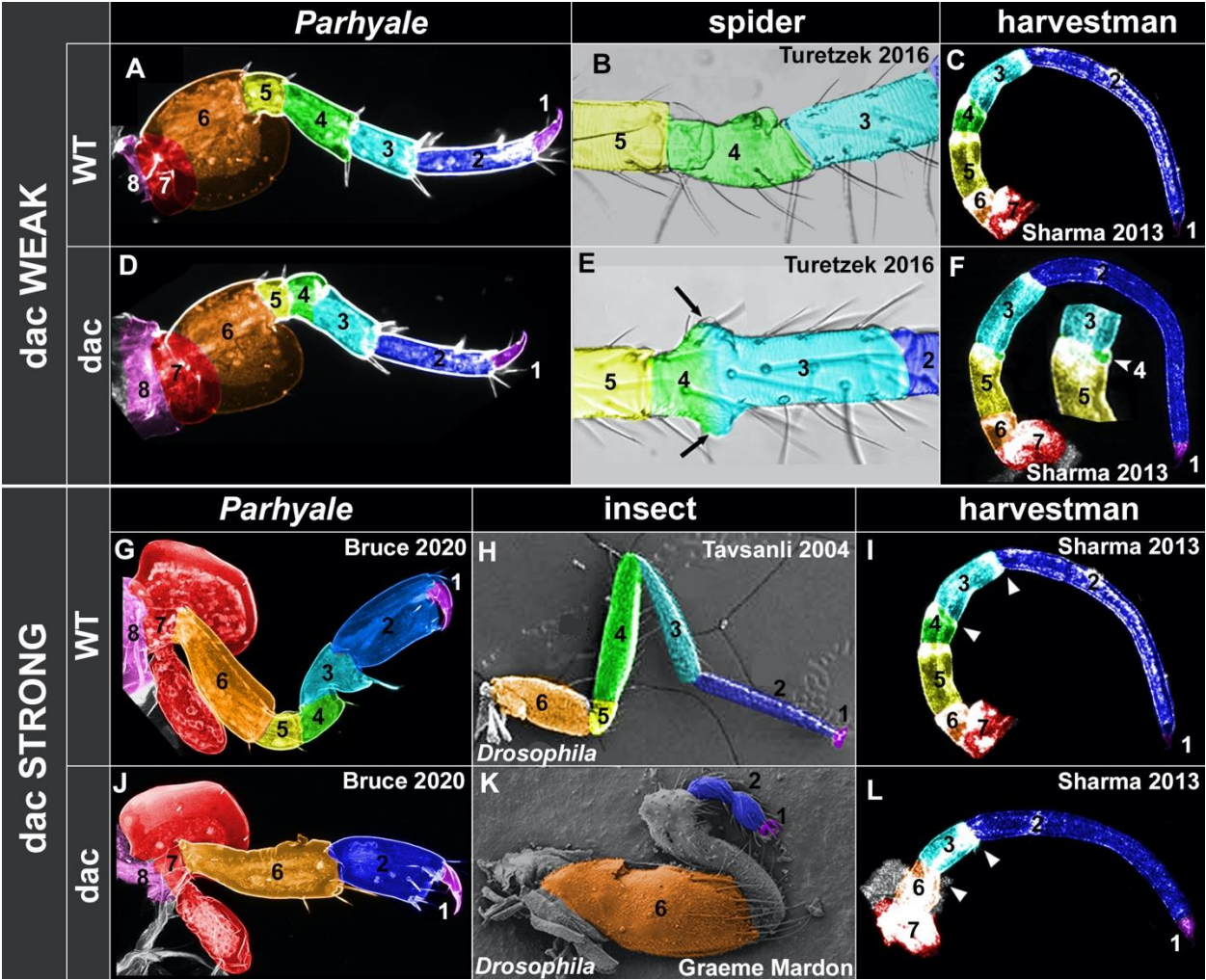
Fig. 4. Dac function in *Parhyale*, spider, harvestman, and insect (*Drosophila*). In spiders, harvestman, and *Parhyale*, a weak dac2 phenotype causes green leg segment 4 to be truncated and fused onto cyan leg segment 3. In harvestman, *Parhyale*, and *Drosophila*, a strong dac2 phenotype affects leg segments 3 – 5.

A — Leg gene function

pnr*
vg/nub/ap
ara*
exd/hth
Sp6-9
Dll
dac

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. 5. 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.
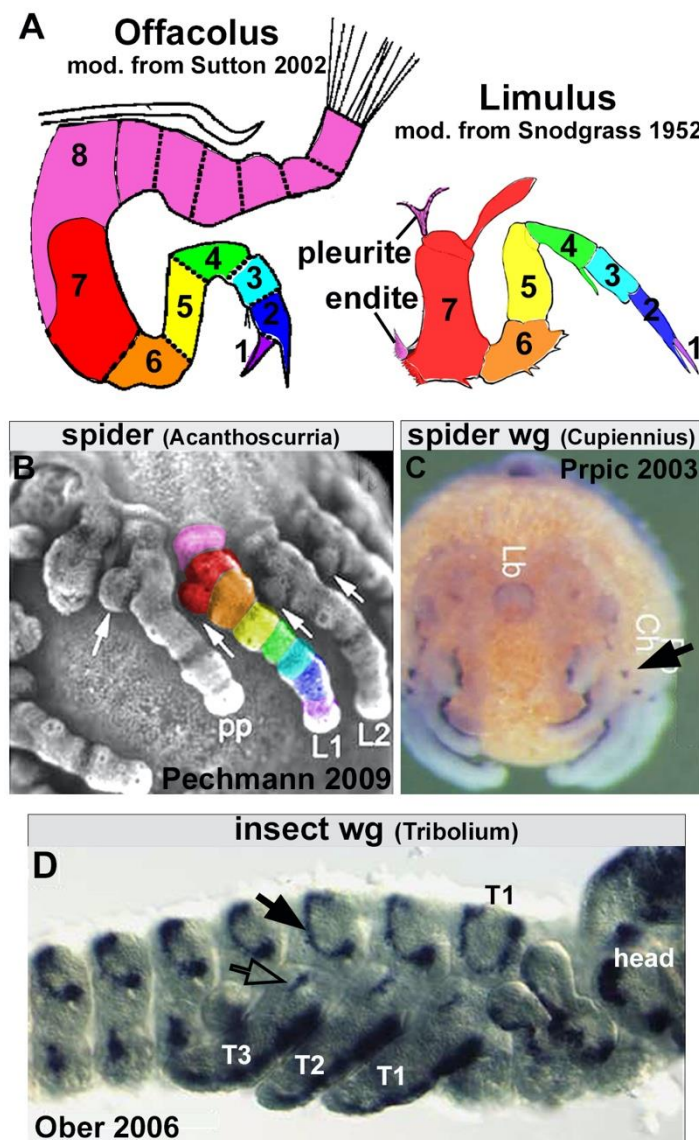
Fig. 6. The ancestral leg segment 8 may have been incorporated into the body wall of chelicerates with seven leg segments. A. Comparison of limbs in *Limulus* and the fossil xiphosurid *Offacolus*. Drawings of legs of *Limulus* (after Snodgrass 1952) and *Offacolus* (after Sutton 2002) were scaled to 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
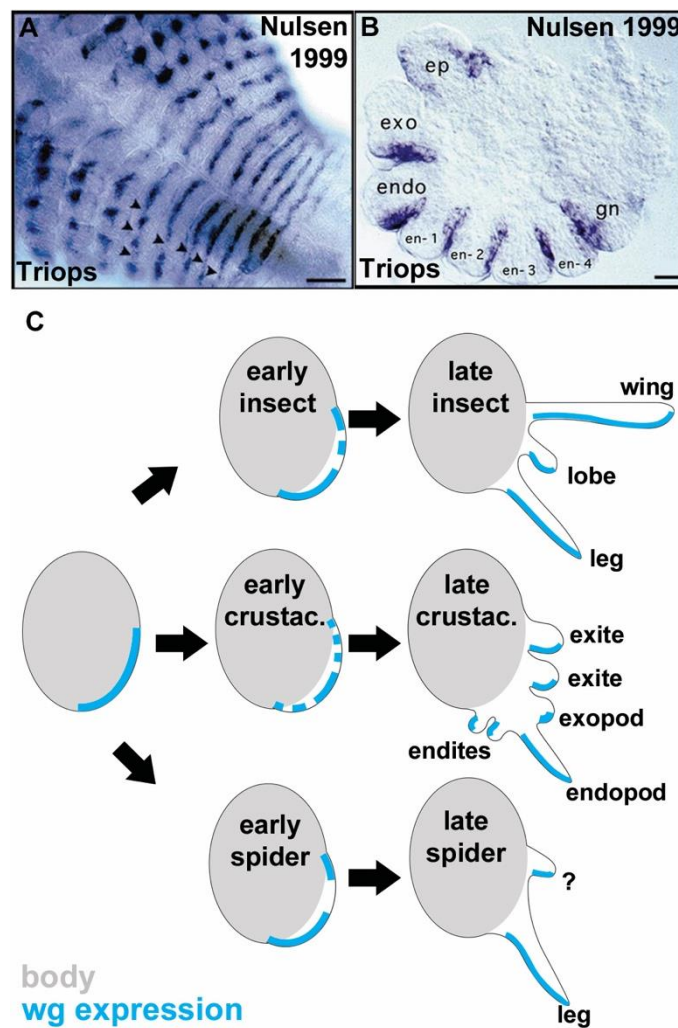
Fig. 7. Wg expression across all arthropods makes sense from the standpoint of our model. A, B. *Triops* crustacean, from Nulsen and Nagy 1999. C. In all arthropods, wg is initially expressed in a solid stripe in each body segment. The crustacean leg grows out like a shelf that wraps around dorsoventrally. 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.
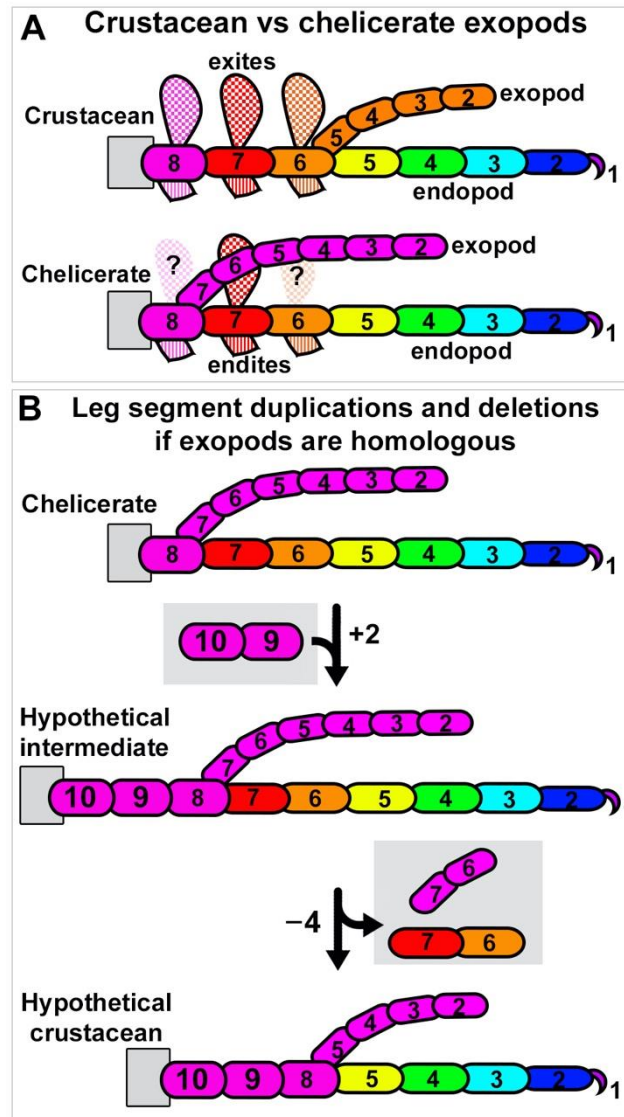
Fig. 8. Exopods of chelicerates and crustaceans emerge from different leg segments. A. Chelicerate and crustacean exopods emerge from different leg segments. B. Illustration of leg segment deletions and duplications that would be necessary in order for chelicerate and crustacean exopods to be homologous but still maintain the observed number of 8 leg segments.
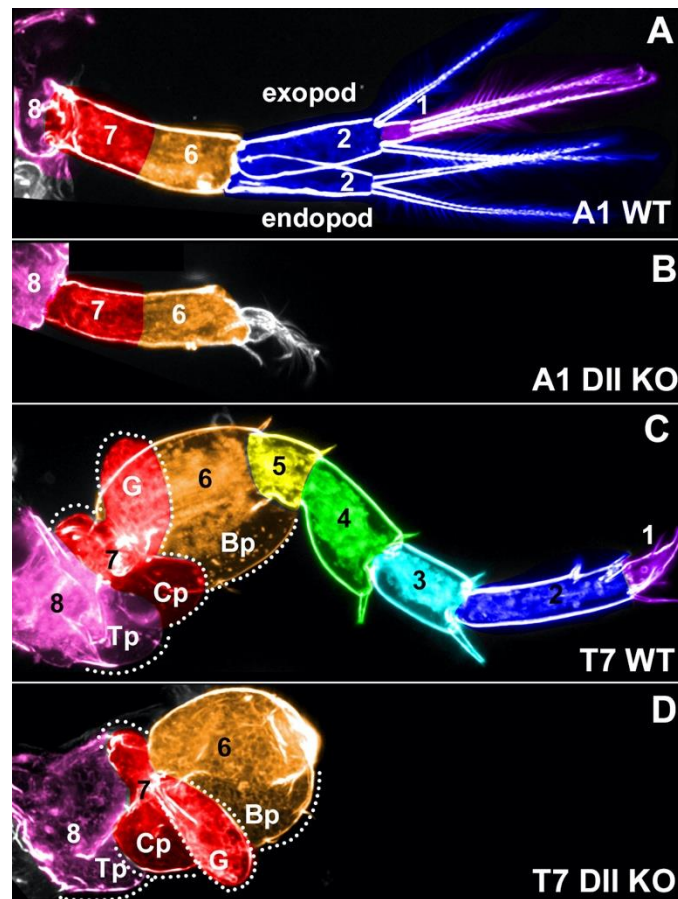
Fig. 9. Dll KO in *Parhyale* deletes the entire exopod and endopod, but leaves the exites unaffected. A, B first abdominal appendage (A1, biramous swimmeret), exopod and endopod are deleted. Patch of bristles on leg segment 6 that velcros left and right appendages together. C, D. seventh thoracic appendage (T7, jumping leg), endopod (telopod) is deleted, exites are unaffected. Gills point in different directions due to sample mounting.