



Knockout of crustacean leg patterning genes suggests that insect wings and body walls evolved from ancient leg segments

Heather S. Bruce^{1,2}✉ and Nipam H. Patel^{1,2,3}

The origin of insect wings has long been debated. Central to this debate is whether wings are a novel structure on the body wall resulting from gene co-option, or evolved from an exite (outgrowth; for example, a gill) on the leg of an ancestral crustacean. Here, we report the phenotypes for the knockout of five leg patterning genes in the crustacean *Parhyale hawaiiensis* and compare these with their previously published phenotypes in *Drosophila* and other insects. This leads to an alignment of insect and crustacean legs that suggests that two leg segments that were present in the common ancestor of insects and crustaceans were incorporated into the insect body wall, moving the proximal exite of the leg dorsally, up onto the back, to later form insect wings. Our results suggest that insect wings are not novel structures, but instead evolved from existing, ancestral structures.

The origin of insect wings has fascinated researchers for over 130 years. Two competing theories have developed to explain their emergence. Given that insects evolved from crustaceans¹, one theory (the exite theory) proposes that insect wings evolved from a crustacean exite (a lobe-shaped lateral outgrowth, such as a gill or plate, on the proximal leg)². In this model, the proximal leg became incorporated into the body wall, which moved the leg exite dorsally, up onto the back, to later form insect wings³. The second theory (the paranotal theory) proposes that wings evolved from body wall plates called paranotal lobes^{4,5}, such as those found in wingless insects like silverfish, and may represent a novel structure not found in crustaceans⁵—a result of the body wall co-opting genes that pattern leg exites⁶.

As the evidence for either the leg exite or body wall theories has mounted, they both appeared equally compelling. Recently, researchers proposed a dual-origin model, wherein insect wings may represent a combination of both the body wall and leg^{7,8}. To determine whether wings can be traced to an ancestral crustacean exite or are a novel structure not found in crustaceans, the legs and body walls of insects must be compared with those of a crustacean where the hypothesized proximal leg region is still free and not incorporated into the body wall.

Results

For our studies, we used the amphipod crustacean *Parhyale hawaiiensis* (Fig. 1), which is a genetically tractable crustacean with uniramous walking legs that have clearly identifiable segments. In *Parhyale*, as with many other crustaceans, each leg has seven segments, while in insects, each leg has six segments. We examined five leg patterning genes in *Parhyale*: *Distalless* (*Dll*), *Sp6–9*, *dachshund* (*dac*), *extradenticle* (*exd*) and *homothorax* (*hth*) (Fig. 2). These five genes form the core of our understanding of patterning along the proximal–distal axis in *Drosophila*⁹. Furthermore, they are the only genes that pattern specific leg segments (as opposed to joints) for which functional data are available in at least two insects, facilitating comparisons with *Parhyale*. While we have documented their expression at several developmental stages (Extended Data Fig. 1),

we note that previous comparative expression studies have not yielded answers to how insect and crustacean leg segments correspond to each other¹⁰, owing in part to the dynamic properties of their expression. We have therefore systematically knocked out these five genes in *Parhyale* using CRISPR–Cas9 mutagenesis (Fig. 2, Extended Data Figs. 2–4 and Supplementary Tables 1 and 2) and compared the results with previously published functional data in *Drosophila* and other insects. *Parhyale* thoracic legs 2–8 (T2–8) are uniramous, like insect legs, and comparable for the features discussed here. We used *Parhyale* (T3) for this study because, in addition, each leg segment is easily distinguished by its shape or additional structures such as gills, plates or combs.

The six distal leg segments of *Parhyale* and insects are in alignment. Our comparison of leg gene function revealed that the six distal leg segments of *Parhyale* and insects correspond to each other in a one-to-one fashion (Fig. 2). In both insects and *Parhyale*, *Dll* is required for the development of leg segments 1–5, counting from the distal end of the leg^{11–17} (Fig. 2b). In *Drosophila*, *dac* is required in the trochanter through the proximal tarsus (leg segments 3–5 and proximal leg segment 2; Extended Data Fig. 5)^{18,19}. *Parhyale* has two *dac* paralogues. *dac1* does not seem to have a morphologically visible knockout phenotype. However, *dac2* is required to pattern leg segments 3–5 (Fig. 2c). In both insects and *Parhyale*, *Sp6–9* (ref. 20) is required for the development of leg segments 1–6 (Fig. 2d)^{21–23}. *exd* and *hth* form heterodimers and therefore have similar phenotypes^{24–27}. In both insects and *Parhyale*, *exd* or *hth* knockout results in deletions/fusions of all proximal leg segments up to and including the proximal part of segment 3 (the joint between segments 2 and 3 appears intact)^{24–27} (Fig. 2e–g). In both insects and *Parhyale*, the normal distal segments of several legs may emerge from a single stump representing their fused proximal segments^{24–27}. In making the comparisons described above, we compared the knockout phenotypes we obtained in *Parhyale* with what appear to be null phenotypes in insects, noting that a number of RNA interference (RNAi) experiments have, unsurprisingly, yielded a range of phenotypes in other insects that probably represent a range of partial knockdown

¹University of California, Berkeley, Berkeley, CA, USA. ²Marine Biological Laboratory, Woods Hole, MA, USA. ³Department of Organismal Biology and Anatomy, University of Chicago, Chicago, IL, USA. ✉e-mail: hbruce@mbl.edu

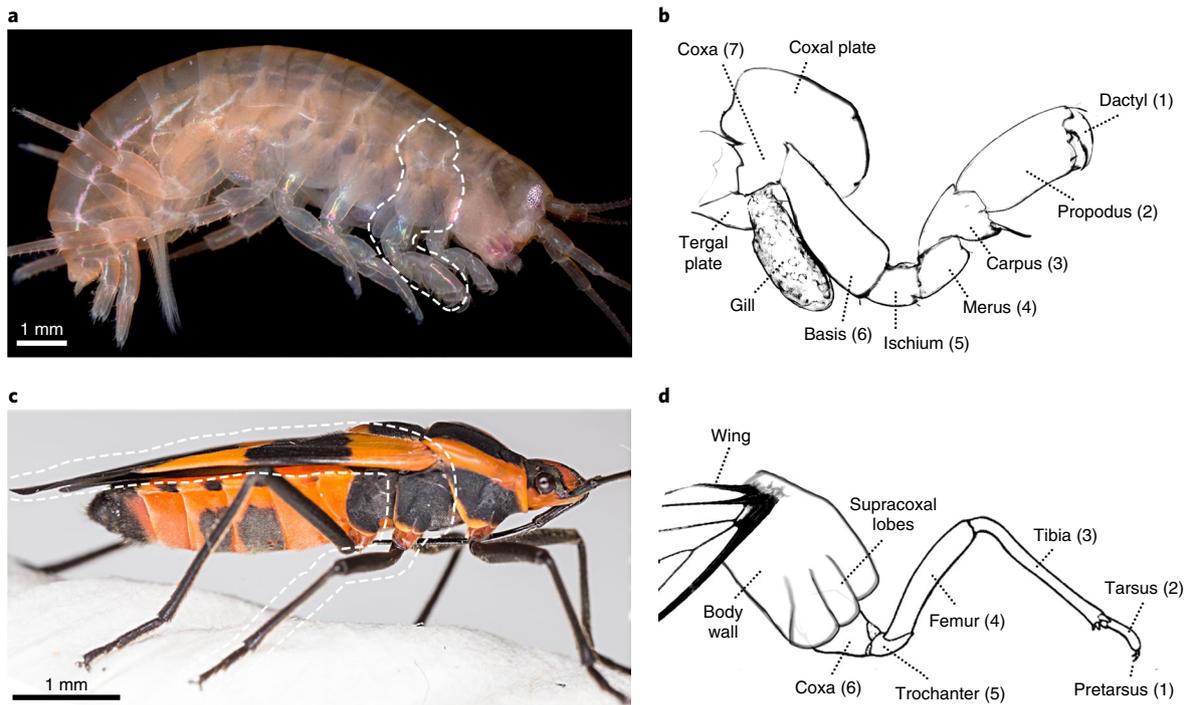


Fig. 1 | Crustacean and insect legs. **a**, An adult *Parhyale* with the third thoracic leg (T3) outlined. **b**, Schematic of the *Parhyale* T3 leg. **c**, An adult *Oncopeltus* (milkweed bug) with T2 outlined. **d**, Schematic of the *Oncopeltus* T2 leg. Images adapted with permission from Aaron Pomerantz (**a,c**); and from ref. ⁴³, Smithsonian Miscellaneous Collection (**d**).

phenotypes. For example, *Dll* genomic deletions in *Drosophila* and *Tribolium* result in complete deletion of leg segments 1–5 (refs. ^{11–14}). In contrast, *Dll* RNAi in *Oncopeltus* (milkweed bug) affects only a subset of leg segments 1–5 (leg segments 1–3 are deleted and leg segment 4 is misshapen), suggesting that this is a partial knockdown¹⁵.

In summary, the functions of *Dll*, *Sp6–9*, *dac*, *exd* and *hth* in the six distal leg segments of *Parhyale* and insects align to each other in a one-to-one fashion (Fig. 2g), such that the insect coxa corresponds to the crustacean basis, the insect trochanter corresponds to the crustacean ischium, the insect femur corresponds to the crustacean merus, and so on for all leg segments (Fig. 1b,d). Interestingly, *Parhyale* has one additional proximal leg segment that is not accounted for in the insect leg: the crustacean coxa. This ancestral leg segment may have simply been deleted in insects. However, following from the previous theory that the wing is derived from an exite on proximal leg segments that were incorporated into the insect body wall, this observation suggests that part of the insect body wall might correspond to the crustacean coxa. Thus, what appears to be the body wall of insects may be divided into two regions: first, a single, dorsal–medial region that does not appear to be homologous to the leg, which we call the true body wall; and second, the incorporated proximal leg, which lies between the true body wall and the free leg segments.

Insects incorporated the ancestral proximal leg region into the body wall. To test the hypothesis that insects incorporated the ancestral crustacean coxa into the body wall, we wanted to find genes that distinguished the true body wall from the hypothesized incorporated leg segment that now functions as body wall. We predicted that genes expressed in a single, dorsal stripe in *Drosophila* should represent the true body wall, and should be expressed in the dorsal-most tissue of *Parhyale*. Conversely, genes expressed in the left and right lateral body wall of *Drosophila*, but not the wing or

leg, might be expressed in the coxa of *Parhyale*. We identified two ideal candidates: *pannier* (*pnr*), which in *Drosophila* is expressed in a single, dorsal stripe²⁸; and the *Iroquois* complex genes, which in *Drosophila* are expressed in the lateral body wall (forming a C shape around the wing²⁹) and are required for specifying the body wall around the wing³⁰.

Drosophila does not form legs as larvae, so we instead used the red flour beetle *Tribolium castaneum* for comparisons with *Parhyale*. The *Iroquois* complex genes we identified in *Parhyale* and *Tribolium* had the best reciprocal blast hit to *Drosophila araucan* (*ara*). Interestingly, the expression of *pnr* and *ara* in *Tribolium* and *Parhyale* was so similar that they could essentially be superimposed (Fig. 3). In both *Tribolium* and *Parhyale*, *pnr* is expressed in the dorsal-most tissue, consistent with a true body wall identity. In both *Tribolium* and *Parhyale*, *ara* is expressed in three domains: two armbands and a circular patch. All three expression domains correspond precisely with our model (Fig. 2g). The circular patch is on leg segment 6. In *Tribolium* this is the insect coxa, whereas in *Parhyale* this is the basis. The *ara* ventral armband is on the proximal side of leg segment 7. In *Parhyale* this is the crustacean coxa, whereas in *Tribolium* this is a leg segment-like structure proximal to the insect coxa (Fig. 3k,l) that will later flatten to form the lateral body wall^{31–33}. Thus, it appears that the ancestral crustacean coxa was indeed incorporated into the insect body wall. These experiments demonstrate the predictive power of our model.

The insect wing is homologous to the *Parhyale* tergal plate. We were intrigued by the two armbands of *ara* expression because they bracket the presumptive wing region in *Tribolium* and the tergal plate in *Parhyale* that express wing genes^{8,34–36}. This suggested that the insect wing may be homologous to the crustacean tergal plate. The *Parhyale* tergal plate is a flange or plate-like expansion of the body wall that looks similar to the plate-like expansions on the *Parhyale* coxa and basis leg segments (coxa plate and basal plate,

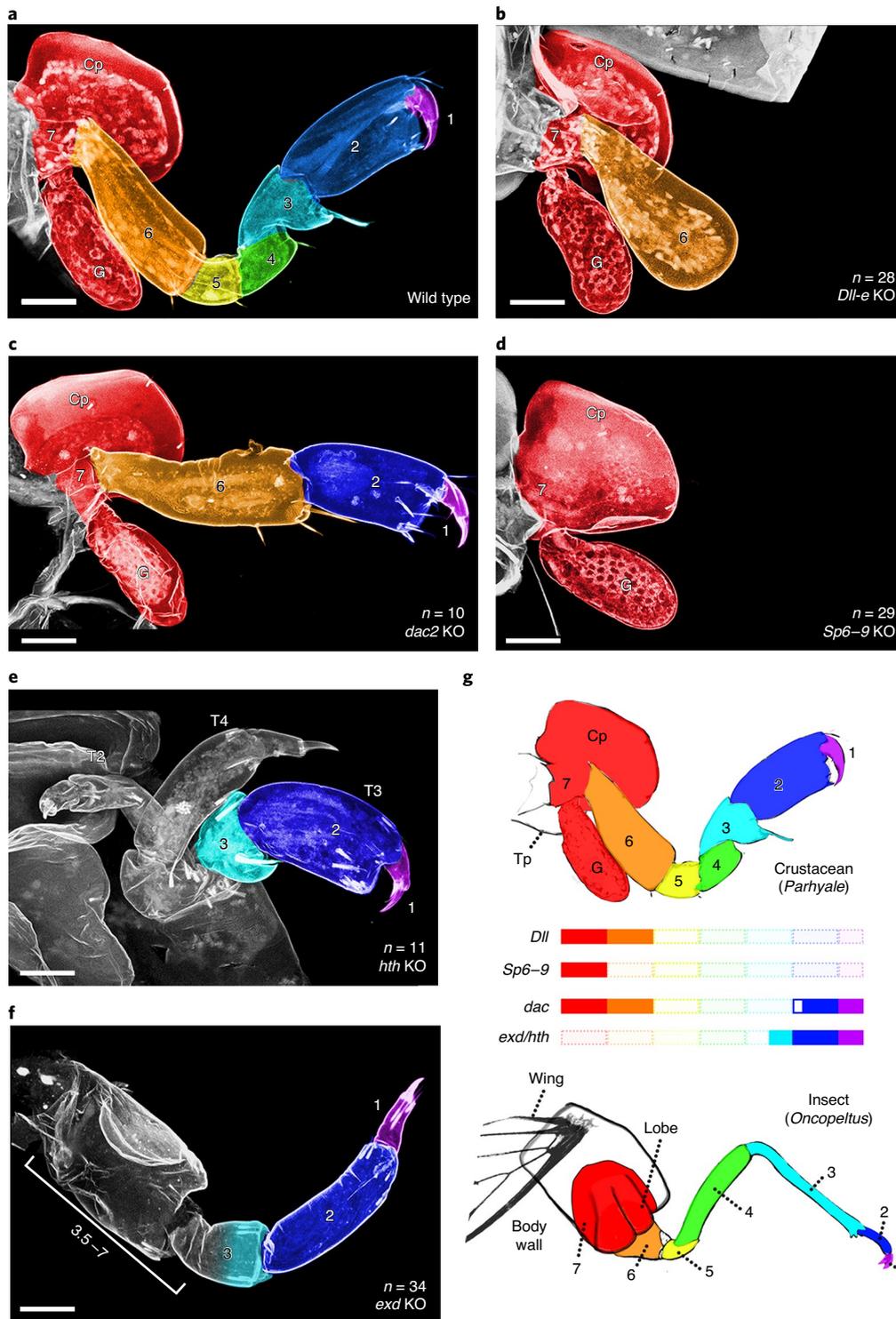


Fig. 2 | Knockout phenotypes of leg patterning genes. a–f, *Parhyale* CRISPR-Cas9 phenotypes in dissected third thoracic legs (T3) of the wild type (a), *Dll-e* knockout (b) *dac2* knockout (c), *Sp6-9* knockout (d), *hth* knockout (e) and *exd* knockout (f). In b, leg segment 6 is rounded, but we do not believe that this is part of the *Dll* phenotype. We interpret this as resulting from the lack of a joint to anchor the tissue, without which the tissue retracts and becomes slightly rounded. Graded cyan in f indicates deletion/fusion of proximal leg segment 3. g, Summary of the leg patterning gene functional phenotypes and leg segment alignment between the crustacean (top) and insect (bottom). The coloured bars correspond to the remaining leg segments following knockout. The semi-transparent bars indicate deleted leg segments. The open bar in *dac* indicates extension of *dac* function into tarsus of insects. Cp, coxal plate; G, gill; Tp, tergal plate. Scale bars, 50 μm. Each *Parhyale* has many pairs of legs, so *n* values show the number of animals for which the majority of the animal's legs displayed the same severe phenotype shown in this figure. For a complete account of all phenotypes, from mild to severe, see Supplementary Table 2. Panel g (bottom) adapted with permission from ref. 43, Smithsonian Miscellaneous Collection.

respectively; Fig. 4a). Notably, Shiga et al.³⁷ and Clark-Hachtel and Tomoyasu³⁶ showed that the *Parhyale* tergal plate, coxal plate and basal plate express and require insect wing genes. Given that exites are defined as a lobe that emerges laterally from the proximal part of the leg and that expresses wing genes³⁸, the morphological and molecular data suggested that all three plates are exites. The tergal plate on the body wall is curious because, traditionally, exites

are considered to be characteristic of leg segments³⁹. We therefore hypothesized that part of the *Parhyale* body wall might have been derived from an additional leg segment, as in insects.

***Parhyale* has an additional proximal leg segment.** In fact, many groups of crustaceans have an additional proximal leg segment: the precoxa^{39–41}. To determine whether *Parhyale* has a precoxa remnant

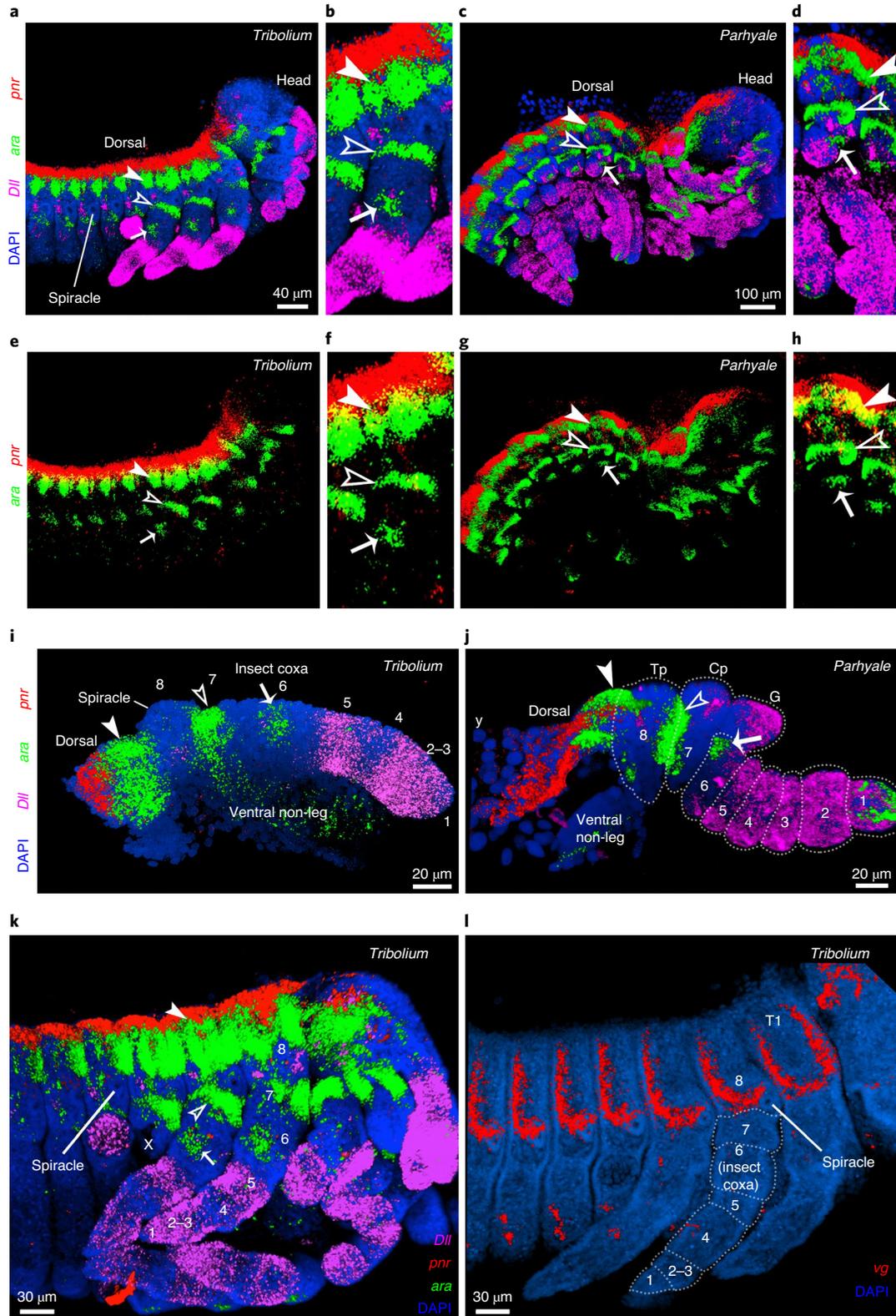


Fig. 3 | Comparison of the expression of *ara* and *pnr* in *Tribolium* and *Parhyale*. **a,b,e,f**, Dissected right half of a *Tribolium* embryo with *pnr* in red, *ara* in green, *Dll* in magenta, DAPI in blue, as indicated; **b** and **f** are zoomed in images of **a** and **e**, respectively. **c,d,g,h**, Dissected right half of a *Parhyale* embryo with *pnr* in red, *ara* in green, *Dll* in magenta, DAPI in blue, as indicated; **d** and **h** are zoomed in images of **c** and **g**, respectively. **i**, Dissected leg of a *Tribolium* embryo. **j**, Dissected T3 leg of a *Parhyale* embryo. **k**, A *Tribolium* embryo near hatching. The X indicates a T3 leg that was accidentally removed during dissection. **l**, *Tribolium* embryo. The spiracle abuts the wing marker *vestigial* (*vg*) and can therefore be used to locate the presumptive wing. *pnr* (red) is expressed in the most dorsal region of the embryo. The large cells dorsal to *pnr* expression in **c** and **j** are extra-embryonic yolk (*y*) that exist before dorsal closure. *ara* is expressed in three domains: a dorsal armband on leg segment 8 (closed arrowhead), an armband straddling leg segments 7 and 8 (open arrowhead) and a circular patch on leg segment 6 (arrow). The armbands bracket the wing region in *Tribolium* and the tergal plate in *Parhyale*. In **i** and **j**, *pnr* and *ara* are expressed in a smattering of ventral non-leg cells. *Tribolium* distal leg numbering is based on *Dll* expression, which begins in the trochanter (leg segment 5) and is reduced in the femur and tibia (leg segments 4 and 3)³¹. *Tribolium* larvae have a fused tibia and tarsus (the tibiotarsus), which is labelled 2–3 (ref. ³¹). In both *Tribolium* (**i**) and *Parhyale* legs (**j**), a muscle expressing *pnr* and *ara* that extends the length of leg segments 7 and 8 was masked to clearly show the ectodermal domains. The *Tribolium* in situ hybridizations are representative of 21 embryos, of which eight were at the stage shown in **a** and **e**, five were at the stage shown in **i** and four were at the stage shown in **j**. The *Parhyale* in situ hybridizations are representative of 12 embryos, of which nine were at the stage shown in **c**, **g** and **j**. DAPI, 4',6'-diamidino-2-phenylindole.

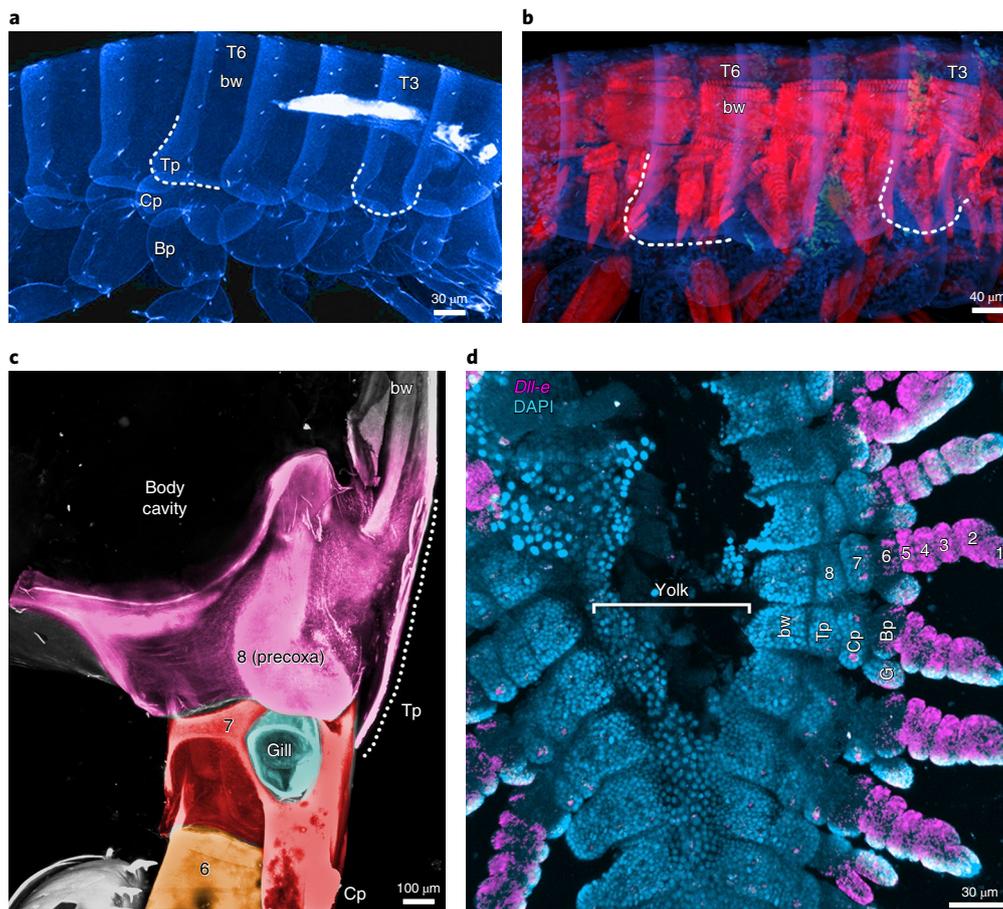


Fig. 4 | Evidence for a precoxal leg segment in *Parhyale*. **a,b**, Confocal images of a *Parhyale* hatchling, with the auto-fluorescent cuticle shown in blue (**a**) and the muscle phalloidin stain in red (**b**). The T3 and T6 tergal plates (Tp) are outlined. Cp, coxal plate; Bp, basal plate. The blocks of simple, anterior-posterior muscles of the body are in contrast with the orthogonal, complexly arranged muscles of the leg segments. Note that the tergal plates (dotted lines) overlay the orthogonal leg muscle of the precoxal protrusion, not the simple body wall (bw) muscles. The image in **b** is representative of 12 phalloidin-stained animals. **c**, Posterior-lateral view of the right T6 leg, looking edge-on at the tergal plate. The tergal plate (dotted outline) emerges from the precoxa (pink). The coxa is shaded red, the gill is teal (partially cut for visibility) and the basis is orange. Note that both the tergal and coxal plates form a contiguous cuticle with their leg segment (that is, there is no distinguishing suture between the leg and exite). **d**, Dorsal view of a *Parhyale* embryo before dorsal closure, with eight leg partitions visible. The tergal plate emerges from the eighth leg partition. The image in **d** is representative of four dorsally mounted and stained embryos. Panels **a,b** adapted with permission from Erin Alberstat.

that was incorporated into the body wall, we examined *Parhyale* embryos, hatchlings and adults using confocal microscopy. We identified a precoxal structure that meets the criteria for a true leg segment^{39,42}: it protrudes conspicuously from the body wall, being

nearly twice the length of the coxa (Fig. 4c; see also the *vestigial* knockout in Fig. 1j of Clark-Hachtel and Tomoyasu³⁶, where the precoxal structure lacks the tergal plate and is easier to see); it forms a true, muscled joint; and it extends musculature to another leg

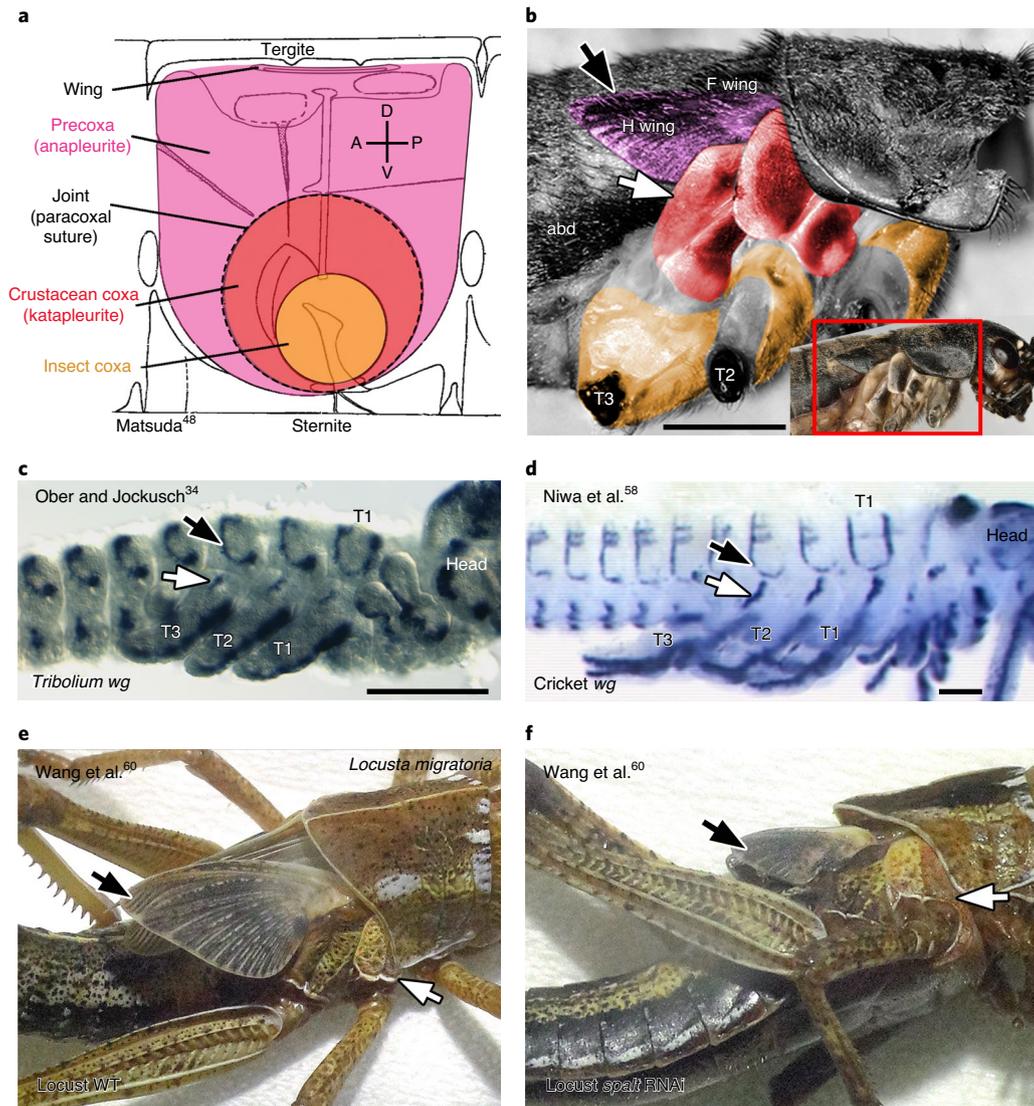


Fig. 5 | The lateral body wall of insects may be composed of two incorporated leg segments and two exites, the wing and the lobe. a, The lateral body wall in winged insects is composed of what appears to be two flattened leg segments (the anapleurite and katapleurite) separated by a vestigial joint (the paracoxal suture). A, anterior; P, posterior; D, dorsal; V, ventral. **b**, Two exites in the body wall of a cricket (*G. assimilis*): the wing (pink) and the lobe (red). abd, abdomen; F wing, forewing; H wing, hindwing. **c,d**, *wingless (wg)* is expressed in two domains, consistent with the location of the wing (black arrow) and the lobe (white arrow), dorsal to the insect coxa of *Tribolium* (**c**) and a cricket (**d**). **e**, Wild-type locust nymph (early fifth instar) injected with double-stranded RNA against green fluorescent protein (GFP). **f**, Locust nymph (early fifth instar) injected with double-stranded RNA against the wing gene *spalt*. Knocking down *spalt* affects both the wing and the lobe. Scale bars, 4 mm (**b**, **e** and **f**) and 100 μ m (**c** and **d**). Images adapted with permission from ref. ⁴⁸, Memoirs of the Entomological Society of Canada (**a**); from ref. ³⁴, Elsevier (**c**); from ref. ⁵⁸, The Company of Biologists (**d**); and from ref. ⁶⁰, Springer Nature Ltd, under a Creative Commons license CC BY 4.0 (**e,f**).

segment (Extended Data Fig. 6). Furthermore, in *Parhyale* embryos, eight subdivisions can be observed in the developing legs (Fig. 4d). Also, *ara* is expressed in a similar armband in both leg segment 7 and this precoxal structure (leg segment 8), suggesting that both are leg segments (Fig. 3). Importantly, the body wall exite (tergal plate) in *Parhyale* emerges from the protrusion that we hypothesize to be the incorporated precoxa, and not from the true body wall (Figs. 3 and 4 and Extended Data Fig. 6). Thus, much of the lateral body wall in *Parhyale* appears to be derived from an ancestral proximal leg segment, as in insects.

Discussion

If the crustacean leg ground plan has two additional proximal leg segments relative to insects—the precoxa (segment 8) and

crustacean coxa (segment 7)—what happened to these leg segments in insects? Our *pnr* and *ara* expression data suggest that these two ancestral leg segments were both incorporated into the insect body wall. These expression data also suggest that the insect wing is homologous to the tergal plate of *Parhyale*, and therefore support the hypothesis that insect wings evolved from a structure that already existed in the crustacean ancestor (Fig. 5).

We note that many previous observations support a model where insects incorporated two leg segments into the body wall (Fig. 6), including from embryology, morphology, molecular work, phylogenetics and precedent in other arthropods, as discussed in brief below. In fact, the idea that insects incorporated the proximal leg into the body wall—now known as the insect subcoxa theory—is over 125 years old^{3,31–33,43–52}. It was first proposed by Hansen in

1893 (ref. ⁵³), and many well-known morphologists since then have further developed the idea, including Heymons in 1899 (ref. ⁵²), Hansen in 1925 (ref. ⁴⁰), Snodgrass in 1927 (ref. ⁴³), Ewing in 1928 (ref. ⁴⁴), Imms in 1937 (ref. ⁴⁵), Weber in 1952 (ref. ⁵⁴), Sharov in 1966 (ref. ⁵⁵), Matsuda in 1970 (ref. ⁴⁸) and Kukalová-Peck in 1983 (ref. ³).

In the embryos of many insects, the proximal part of the developing leg is observed to fuse into the body wall^{31–33,46,52,56}. In adult insects, the morphology of the body wall surrounding the leg looks like two flattened leg segments separated by a joint with muscle insertions (Fig. 5a)^{4,32,48,57}. This joint remnant can be observed molecularly: *Serrate*, which marks leg segment joints, is also expressed at the boundary between the hypothesized two incorporated leg segments in the body wall of *Tribolium*^{21,31}.

If insects incorporated two leg segments into the body wall, and the wing is the exite of the eighth leg segment, then some insects might also express an exite on the seventh leg segment (Fig. 5). One example of this may be the supracoxal lobe of orthopterans (Fig. 5b), hemipterans and coleopterans⁴⁸. This is a lateral plate-like lobe emerging from leg segment 7 of our model (Fig. 1c,d and 5b and Supplementary Video 1). Mashimo and Machida⁵¹ described the embryonic development of the cricket subcoxa (our leg segment 7), from which the lobe can be observed to develop. Notably, the wing/exite genes *wingless* and *vestigial* are expressed not only in the presumptive wing tissue, but also where the lobe would form in the flour beetle (*T. castaneum*; Fig. 5c)^{8,34,35}, cricket (*Gryllus assimilis*; Fig. 5d)⁵⁸, fruit fly (*D. melanogaster*)⁵⁹ and jumping bristletail (*Pedetontus unimaculatus*)⁷. Importantly, when wing/exite genes are depleted via RNAi, both the wing and the lobe are reduced (Fig. 5e,f) in the locust (*L. migratoria*)⁶⁰, milkweed bug (*Oncopeltus fasciatus*)⁶¹, flour beetle (*T. castaneum*)⁸, mealworm beetle (*Tenebrio molitor*)⁶² and scarab beetle (*Onthophagus sagittarius*)⁶³. This suggests that the lobe, like the wing, is an exite, consistent with our hypothesis that two leg segments were incorporated into the insect body wall.

Phylogenetically, recent rearrangements¹ within crustacean groups allow the exite on leg segment 8 to be traced from crustaceans to winged insects. For example, the *Parhyale* tergal plate (Fig. 4a) looks similar to the flange-shaped pleural lobes of cephalocarid crustaceans^{41,64} and paranotal lobes of wingless hexapods such as silverfish, the latter of which many authors have hypothesized to be the precursors of insect wings (paranotal theory^{5,65}). Molecularly, wing genes are expressed in a wingless insect, the jumping bristletail (*P. unimaculatus*), in each thoracic segment in a dorsal region, consistent with later development of the paranotal lobes⁷. It would

be useful to test the cephalocarid link in this series by examining the expression and function of wing genes in cephalocarid pleural lobes.

Finally, incorporation of proximal leg segments into the body wall is not uncommon in arthropods. Centipedes and symphylans (myriapods) also appear to have incorporated a proximal leg segment into the body wall^{39,43,44,50}. Terrestrial isopod crustaceans have fused not one but two proximal leg segments into the body (the precoxa and coxa), such that the basis (leg segment 6) now forms the first free leg segment^{40,66–68}. Notably, these are the same two proximal leg segments that are hypothesized to have been incorporated

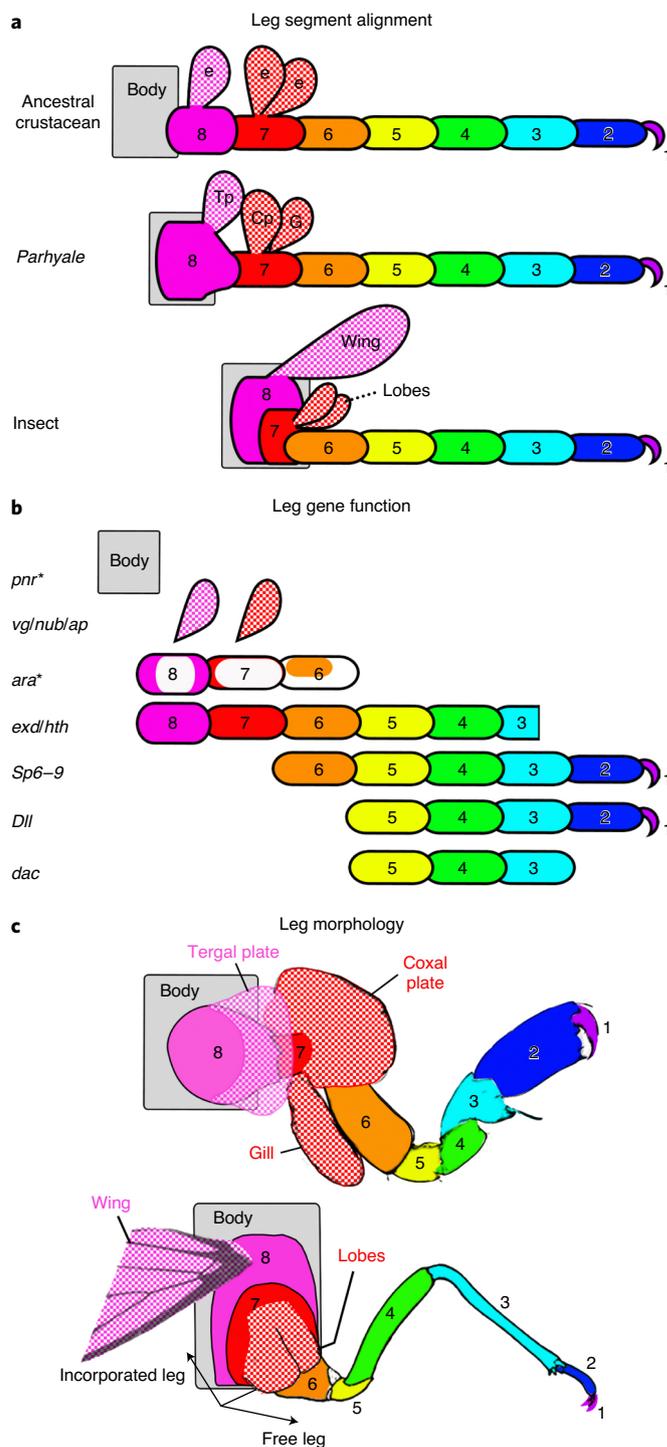


Fig. 6 | Proposed leg segment homologies. **a**, Proposed leg segment homologies (colours) between insects, *Parhyale* and an ancestral crustacean, based on leg gene function alignment. e, exite. **b**, Schematic showing which structures require each gene in crustaceans and insects, providing a model for how to align crustacean and insect legs. 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 to each other in a one-to-one fashion. The expression of *pnr* and *ara* (where an asterisk indicates comparisons based on expression rather than data from genetic knockout), as well as the expression and function of wing genes such as *nubbin* (*nub*) and *apterous* (*ap*), suggest that insect wings are derived from crustacean exites. In this model, the checkered pink exites of leg segment 8 are homologous (that is, the ancestral crustacean precoxa exite, the *Parhyale* tergal plate and the insect wing) and the checkered red exites of leg segment 7 are homologous (that is, the ancestral crustacean coxa exites, the *Parhyale* coxal plate and gill, and the insect supracoxal lobes; see Figs. 1c,d and 5b,e). **c**, Leg segment morphologies with proposed homologies (colours) in *Parhyale* (top) and insects (bottom). Panel **c** (bottom) adapted with permission from ref. ⁴³, Smithsonian Miscellaneous Collection.

into the insect body wall. It is interesting to note that, for all of these groups, incorporation of proximal leg segments appears to be associated with supporting more of the weight of the body on the legs. Perhaps incorporating proximal leg segments served to reinforce and brace the leg to be more load bearing.

Wings did not arise by fusion of leg segments and exites and are not novel. Clark-Hachtel and Tomoyasu's interpretation³⁶ of the dual-origin theory proposes that two ancestral exites and leg segments fused to form the insect wing. While previous research supports the existence of two ancestral exite fields (that is, two regions in the body wall of insects that express wing genes⁷), it has not been demonstrated that these two tissues fuse to form the wing (see the Supplementary Discussion and Extended Data Fig. 7 for an alternative interpretation of ectopic wing fusion in ref. ⁸). If wings were a fusion of two exites, there should be only one exite remaining (the wing) that emerges from the body wall. However, as discussed above, loss of function of wing genes in insects disrupts two separate structures (the wing and the lobe), which argues that the two ancestral exite fields have remained separate in insects^{8,60–62}. Additionally, if leg segments 7 and 8 and their exites had fused to form the insect wing, the armband of *ara* expression between leg segments 7 and 8 would be reduced or lost, resulting in only two domains of *ara* expression: a single, dorsal armband and the circular patch on leg segment 6 (Fig. 3). However, in both *Tribolium* and *Parhyale*, *ara* is expressed in a nearly identical configuration and spacing, even at late stages (Fig. 3 and Extended Data Fig. 8).

We note that the leg segment position of an exite (that is, whether it emerges from leg segment 6, 7 or 8) should be considered of prime importance when determining its homology across insects and crustaceans. Many exites will share the same genes, by virtue of being exites. Thus, an examination of wing/exite genes alone will not be informative about whether any particular exite-like structure arose by co-option or is homologous to an ancestral structure, nor to which exite it may correspond in other arthropods. Therefore, in this study, rather than focusing on wing/exite genes, we have instead determined a set of molecular markers for each leg segment. We thus go beyond merely demonstrating that a structure is an exite, and instead identify which of the three proximal leg segments each exite emerges from. By comparing the expression patterns of *pnr* and *Iroquois* genes, and joint patterning genes such as *Serrate*, a more precise statement about the homologies of any exite can be made.

In summary, our results provide a compelling solution to the long-standing debate regarding the origin of insect wings as derived from either an exite of the leg or paranotal lobes of the body wall. Our model accounts for all observations in favour of either of these hypotheses, including the apparent dorsal position of insect wings relative to their legs, the apparent loss of ancestral leg segments in insects, the two-segmented morphology of the insect body wall in both embryos and adults, the origin of the complex musculature necessary for flight, and the shared gene expression between insect wings and crustacean exites. Our model also explains the apparent dual origin of insect wings from both the body wall and leg exite, not because of a blending of the body wall and leg, but because much of the insect body wall is derived from leg segments. We propose that both the leg exite and body wall theories are correct, but each is relevant to different phylogenetic time points: crustacean leg exites evolved into body wall lobes, then subsequently into wings. While wings are an outgrowth of what is now the insect body wall, they owe their origin to the leg segment of an ancestral arthropod.

Methods

Sequence identification and analysis. Partial or complete sequences for *Parhyale* *Dll*, *Sp6–9*, *Exd* and *Hth* have been identified previously. These were >99% identical at the nucleotide level to sequences in the *Parhyale* assembled transcriptome.

To confirm their orthology and identify potential *Parhyale* paralogues, we performed reciprocal best Blast hit (RBBH) searches. For each gene, orthologues from several arthropods and vertebrates were downloaded from the National Center for Biotechnology Information (NCBI) and the European Bioinformatics Institute of the European Molecular Biology Laboratory and aligned against the *Parhyale* transcriptome¹⁷ using standalone NCBI blastp. The *Parhyale* hits with the lowest e-values were used to run a blastp against the NCBI database, restricted to Arthropoda. We confirmed that the original set of orthologues from several arthropods were the best hits to our *Parhyale* candidates (that is, were each other's RBBHs). These RBBHs are listed in the Supplementary Information, and were deposited in GenBank under accession numbers MG457799–MG457804, MT103930 and MT103931.

No *Parhyale* buttonhead/Sp5 was recovered in either of two independently assembled transcriptomes, nor in the *Parhyale* genome. Buttonhead/Sp5 was not found in NCBI for the related amphipod *Hyalella azteca*, nor in the insect groups Hemiptera or Orthoptera. The assembled transcriptome only recovered fragments of *Parhyale* Sp1–4, so the previously sequenced *Parhyale* Sp1–4 (CBH30980.1) was used.

Parhyale has three *Dll* paralogues, but the one required for leg patterning is *Dll-e*^{16,17}. The related amphipod, *H. azteca*, also appears to have these same three *Dll* paralogues. The three *Parhyale* *Dll* paralogues had the lowest e-values to all *Dll* orthologues examined, but which of the three *Parhyale* *Dll* paralogues had the lowest e-value was variable, as expected for a clade-specific duplication.

The coding regions for *Parhyale* *Exd* and *Hth* in the assembled transcriptome are longer than those previously identified⁶⁹. *Exd* is 204 amino acids longer and *Hth* is 166 amino acids longer. This explains the higher-than-expected e-values between the *Parhyale* *Exd* and *Hth* sequences identified previously and the *Parhyale* *Exd* and *Hth* sequences used in this study. An e-value of zero means less than 2×10^{-308} , but is reported as zero by BLAST.

The *Parhyale* and *Tribolium* sequences for *pnr* and *aurucan* were confirmed by RBBH to the *Drosophila* RefSeq sequences, since these genes have not been adequately characterized in other organisms. *Parhyale ara* was also confirmed by RBBH to the Kerner⁷⁰ database of 67 *Iroquois* sequences.

Cloning and RNA probe synthesis. Total RNA was extracted from a large pool of *Parhyale* embryos at multiple stages of embryogenesis, from stages 12–26 using TRIzol. Complimentary DNA was generated using SuperScript III. Primers were generated with Primer3 (<http://bioinfo.ut.ee/primer3-0.4.0>), with a preferred product size of 700 base pairs, and we avoided evolutionarily conserved domains to avoid possible probe cross-reactivity. Inserts were amplified with Platinum Taq (Thermo Fisher Scientific; 10966026), ligated into pGEM-T Easy vectors (Promega; A1360) and transformed into *Escherichia coli*. The resulting plasmids were cleaned with a QIAprep Miniprep kit (Qiagen; A1360) and sequenced to verify the correct insert and determine sense and anti-sense orientation. In situ templates were generated by PCR from these plasmids using M13F/R primers and purified with a Qiagen PCR Purification kit (Qiagen; 28104). The resulting PCR products were used to generate digoxigenin-labelled RNA probes (Roche; 11175025910) using either T7 or Sp6 RNA polymerase. RNA probes were precipitated with LiCl, resuspended in water and run on an agarose gel to check that the probes were the correct size, and the concentration was determined using a NanoDrop 10000. The probes were used at a concentration of 1–5 ng μl^{-1} .

In situ protocol. Embryo collection, fixation and dissection were performed as previously described⁷¹. In situ hybridization was performed as previously described⁷². In brief, embryos were fixed in 4% paraformaldehyde in artificial seawater for 45 min, dehydrated to methanol and stored overnight at -20°C to discourage the embryos from floating during the later hybridization step. The embryos were rehydrated using 1× phosphate-buffered saline (PBS) with 0.1% Tween 20 (PBST), post-fixed for 30 min in 9:1 PBST:paraformaldehyde and washed in PBST. The embryos were incubated in hybridization solution at 55°C for at least 36 h. The embryos were blocked with 5% normal goat serum and 1× Roche blocking reagent (Roche; 11096176001) in PBST for 30 min. Sheep anti-DIG-AP antibody (Roche; 11093274910) was added at 1:2,000 and incubated for 2 h at room temperature. The embryos were developed in BM Purple (Roche; 11442074001) for a few hours or overnight. After the embryos were sufficiently developed, they were dehydrated using methanol to remove any pink background, then rehydrated using PBST. The embryos were then moved to 1:1 PBS:glycerol with 0.1 mg ml^{-1} 4',6'-diamidino-2-phenylindole, then 70% glycerol in PBS.

In situ HCR version 3.0. *Parhyale* and *Tribolium* complimentary DNA sequences were submitted to Molecular Instruments⁷³ and the probe sets are available from the company. In situ hybridizations were performed using their whole-mount *Drosophila* embryo protocol, with the following exceptions: (1) the embryos were left on the bench, not rocked back and forth; (2) the embryos were permeabilized in sodium dodecyl sulfate detergent solution (1% sodium dodecyl sulfate, 0.5% Tween 20, 50 mM Tris-HCl (pH 7.5), 1 mM EDTA (pH 8.0) and 150 mM NaCl) for 30 min instead of proteinase K, which improved the morphology in our hands; and (3) the embryos were in pre-hybridization solution and pre-amplification solution

for 1–2 h, to allow complete settling. The probe set numbers from Molecular Instruments are provided in Supplementary Table 11.

CRISPR–Cas9 guide RNA generation, injection and imaging. Guide RNAs were generated using ZiFit^{4,73}, as previously described⁷⁶. Single guide RNAs (sgRNAs) were ordered from Synthego. The injection mixes had a final concentration of 333 ng μl⁻¹ Cas9 protein, 150 ng μl⁻¹ sgRNA (for both single- and double-guide injection mixes) and 0.05% phenol red for visualization during injection, all suspended in water. One- or two-cell embryos were injected with approximately 40–60 pl sgRNA mixture, as previously described⁷⁶. The resulting knockout hatchlings were fixed in 4% paraformaldehyde in artificial seawater at 4 °C for 1–2 d, then moved to 70% glycerol in 1× PBS. Dissected hatchling limbs were visualized with Zeiss LSM 700 and 780 confocal microscopes using ultraviolet autofluorescence to visualize the cuticle. Z-stacks were assembled with Velocity version 6.0 (PerkinElmer). False-coloured images were created using the Overlay command in Adobe Photoshop CS6. Staining for muscle was performed by incubating fixed 1-day-old hatchlings with Alexa Fluor 555 Phalloidin at 1 U ml⁻¹ in PBST.

T7 endonuclease I assay. Genomic primers were designed using Primer3 and flanked the target site by at least 400 base pairs to either side. DNA isolation and subsequent PCR amplification of the region of interest was modified from previously described protocols⁷⁷. Genomic DNA was amplified directly from fixed hatchlings in 70% glycerol using Ex Taq (Takara; RR001A). The resulting PCR products were purified with the QIAquick PCR Purification kit (Qiagen; 28104). Heteroduplexes were annealed and digested by T7 endonuclease I according to New England Biolab protocols (NEB; M0302L). The digested products were run out on a 1.5% agarose gel. The genomic primers used for the T7 endonuclease I assay are listed in Supplementary Table 12.

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

Data availability

All of the data that support the findings of this study are available in the main text or the Supplementary Information.

Received: 14 July 2019; Accepted: 15 October 2020;

Published online: 1 December 2020

References

- Lozano-Fernandez, J. et al. Pancrustacean evolution illuminated by taxon-rich genomic-scale data sets with an expanded remipede sampling. *Genome Biol. Evol.* **11**, 2055–2070 (2019).
- Averof, M. & Cohen, S. M. Evolutionary origin of insect wings from ancestral gills. *Nature* **385**, 627–630 (1997).
- Kukalová-Peck, J. Origin of the insect wing and wing articulation from the arthropodan leg. *Can. J. Zool.* **61**, 1618–1669 (1983).
- Snodgrass, R. E. *Principles of Insect Morphology* (Smithsonian Institution, 1935).
- Crampton, G. C. The phylogenetic origin and the nature of the wings of insects according to the paranotal theory. *J. N. Y. Entomol. Soc.* **23**, 1–39 (1916).
- Jockusch, E. L. & Nagy, L. M. Insect evolution: how did insect wings originate? *Curr. Biol.* **7**, R358–R361 (1997).
- Niwa, N. et al. Evolutionary origin of the insect wing via integration of two developmental modules. *Evol. Dev.* **12**, 168–176 (2010).
- Clark-Hachtel, C. M., Linz, D. M. & Tomoyasu, Y. Insights into insect wing origin provided by functional analysis of *vestigial* in the red flour beetle, *Tribolium castaneum*. *Proc. Natl Acad. Sci. USA* **110**, 16951–16956 (2013).
- Ruiz-Losada, M., Blom-Dahl, D., Córdoba, S. & Estella, C. Specification and patterning of *Drosophila* appendages. *J. Dev. Biol.* **6**, 17 (2018).
- Abzhanov, A. & Kaufman, T. C. Homologs of *Drosophila* appendage genes in the patterning of arthropod limbs. *Dev. Biol.* **227**, 673–689 (2000).
- Cohen, S. M. & Jürgens, G. Proximal–distal pattern formation in *Drosophila*: cell autonomous requirement for *Distal-less* gene activity in limb development. *EMBO J.* **8**, 2045–2055 (1989).
- Cohen, B., Simcox, A. A. & Cohen, S. M. Allocation of the thoracic imaginal primordia in the *Drosophila* embryo. *Development* **117**, 597–608 (1993).
- Campbell, G. & Tomlinson, A. The roles of the homeobox genes *aristalless* and *Distal-less* in patterning the legs and wings of *Drosophila*. *Development* **125**, 4483–4493 (1998).
- Beermann, A. 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).
- Angelini, D. R. & Kaufman, T. C. Functional analyses in the hemipteran *Oncopeltus fasciatus* reveal conserved and derived aspects of appendage patterning in insects. *Dev. Biol.* **271**, 306–321 (2004).
- Liubicich, D. M. et al. Knockdown of *Parhyale Ultrabithorax* recapitulates evolutionary changes in crustacean appendage morphology. *Proc. Natl Acad. Sci. USA* **106**, 13892–13896 (2009).
- Kao, D. et al. The genome of the crustacean *Parhyale hawaiiensis*, a model for animal development, regeneration, immunity and lignocellulose digestion. *eLife* **5**, e20062 (2016).
- Mardon, G., Solomon, N. M. & Rubin, G. M. *dachshund* encodes a nuclear protein required for normal eye and leg development in *Drosophila*. *Development* **120**, 3473–3486 (1994).
- Tavsanli, B. C. et al. Structure–function analysis of the *Drosophila* retinal determination protein *Dachshund*. *Dev. Biol.* **272**, 231–247 (2004).
- Schaeper, N. D., Prpic, N.-M. & Wimmer, E. A. 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).
- Beermann, A., Aranda, M. & Schröder, R. The *Sp8* zinc-finger transcription factor is involved in allometric growth of the limbs in the beetle *Tribolium castaneum*. *Development* **131**, 733–742 (2004).
- Schaeper, N. D., Prpic, N.-M. & Wimmer, E. A. 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).
- Estella, C. & Mann, R. S. Non-redundant selector and growth-promoting functions of two sister genes, *buttonhead* and *Sp1*, in *Drosophila* leg development. *PLoS Genet.* **6**, e1001001 (2010).
- Mito, T. et al. Divergent and conserved roles of *extradenticle* in body segmentation and appendage formation, respectively, in the cricket *Gryllus bimaculatus*. *Dev. Biol.* **313**, 67–79 (2008).
- Ronco, M. et al. Antenna and all gnathal appendages are similarly transformed by *homothorax* knock-down in the cricket *Gryllus bimaculatus*. *Dev. Biol.* **313**, 80–92 (2008).
- Rauskolb, C., Smith, K. M., Peifer, M. & Wieschaus, E. *extradenticle* determines segmental identities throughout *Drosophila* development. *Development* **121**, 3663–3673 (1995).
- Wu, J. & Cohen, S. M. Proximodistal axis formation in the *Drosophila* leg: subdivision into proximal and distal domains by *Homothorax* and *Distal-less*. *Development* **126**, 109–117 (1999).
- Calleja, M. et al. Generation of medial and lateral dorsal body domains by the *pannier* gene of *Drosophila*. *Development* **127**, 3971–3980 (2000).
- Ikmi, A., Netter, S. & Coen, D. Pre-patterning the *Drosophila* notum: the three genes of the *iroquois* complex play intrinsically distinct roles. *Dev. Biol.* **317**, 634–648 (2008).
- Diez del Corral, R., Aroca, P., Gmez-Skarmeta, J. L., Cavodeassi, F. & Modolell, J. The Iroquois homeodomain proteins are required to specify body wall identity in *Drosophila*. *Genes Dev.* **13**, 1754–1761 (1999).
- Coulcher, J. F., Edgecombe, G. D. & Telford, M. J. Molecular developmental evidence for a subcoxal origin of pleurites in insects and identity of the subcoxa in the gnathal appendages. *Sci. Rep.* **5**, 15757 (2015).
- Kobayashi, Y., Niikura, K., Oosawa, Y. & Takami, Y. 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).
- Kobayashi, Y. Formation of Subcoxae-1 and 2 in insect embryos: the subcoxal theory revisited. *Proc. Arthropod Embryol. Soc. Jpn* **48**, 33–38 (2017).
- Ober, K. A. & Jockusch, E. L. The roles of *wingless* and *decapentaplegic* in axis and appendage development in the red flour beetle, *Tribolium castaneum*. *Dev. Biol.* **294**, 391–405 (2006).
- Beermann, A., Prühs, R., Lutz, R. & Schröder, R. A context-dependent combination of Wnt receptors controls axis elongation and leg development in a short germ insect. *Development* **138**, 2793–2805 (2011).
- Clark-Hachtel, C. M. & Tomoyasu, Y. Two sets of candidate crustacean wing homologues and their implication for the origin of insect wings. *Nat. Ecol. Evol.* <https://doi.org/10.1038/s41559-020-1257-8> (2020).
- Shiga, Y. et al. Repeated co-option of a conserved gene regulatory module underpins the evolution of the crustacean carapace, insect wings and other flat outgrowths. Preprint at *bioRxiv* <https://doi.org/10.1101/160010> (2017).
- Boxshall, G. A. & Jaume, D. Exopodites, epipodites and gills in crustaceans. *Arthropod Syst. Phylogeny* **67**, 229–254 (2009).
- Boxshall, G. A. The evolution of arthropod limbs. *Biol. Rev.* **79**, 253–300 (2004).
- Hansen, H. J. *Studies on Arthropoda II* (Gyldendalske Boghandel, 1925).
- Schram, F. R. *Crustacea* (Oxford Univ. Press, 1986).
- Shultz, J. W. Morphology of locomotor appendages in Arachnida: evolutionary trends and phylogenetic implications. *Zool. J. Linn. Soc.* **97**, 1–56 (1989).
- Snodgrass, R. E. Morphology and mechanism of the insect thorax. *Smithson. Misc. Collect.* **80**, 1–108 (1927).
- Ewing, H. E. *The Legs and Leg-Bearing Segments of Some Primitive Arthropod Groups: With Notes on Leg-Segmentation in the Arachnida* (Smithsonian Institution, 1928).

45. Imms, A. D. *Recent Advances in Entomology* (P. Blakiston's Son & Co. Inc., 1937).
46. Roonwal, M. L. Studies on the embryology of the African migratory locust, *Locusta migratoria migratorioides* Reiche and Frm. (Orthoptera, Acrididae). II—Organogeny. *Phil. Trans. R. Soc. Lond. B Biol. Sci.* **227**, 175–244 (1937).
47. Carpentier, F. & Barlet, J. The first leg segments in the Crustacea Malacostraca and the insects. *Smithson. Misc. Collect.* **137**, 99–115 (1959).
48. Matsuda, R. Morphology and evolution of the insect thorax. *Mem. Entomol. Soc. Can.* **102**, 5–431 (1970).
49. Deuve, T. The epipleural field in hexapods. *Ann. Soc. Entomol. Fr.* **37**, 195–231 (2001).
50. Bäcker, H., Fanenbruck, M. & Wägele, J. W. A forgotten homology supporting the monophyly of Tracheata: the subcoxa of insects and myriapods re-visited. *Zool. Anz.* **247**, 185–207 (2008).
51. Mashimo, Y. & Machida, R. Embryological evidence substantiates the subcoxal theory on the origin of pleuron in insects. *Sci. Rep.* **7**, 12597 (2017).
52. Heymons, R. *Beiträge zur Morphologie und Entwicklungsgeschichte der Rhyngchoten* (Druck von Ehrhardt Karras, 1899).
53. Hansen, H. J. Zur Morphologie der Gliedmassen und Mundtheile bei Crustaceen und Insecten. *Zool. Anz.* **16**, 193–198, 201–212 (1893).
54. Weber, H. *Morphologie, Histologie und Entwicklungsgeschichte Der Articulaten* (G. Fischer, 1952).
55. Sharov, A. G. *Basic Arthropodan Stock: With Special Reference to Insects* (Pergamon Press, 1966).
56. Uchifune, T. & Machida, R. Embryonic development of *Galloisiana yuasai* Asahina, with special reference to external morphology (insecta: Grylloblattodea). *J. Morphol.* **266**, 182–207 (2005).
57. Kobayashi, Y. The paracoxal suture in insect embryos: its state and importance for understanding the basalmost podomeres. *Proc. Arthropod. Embryol. Soc. Jpn* **52**, 43–44 (2018).
58. Niwa, N. 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).
59. Requena, D. et al. Origins and specification of the *Drosophila* wing. *Curr. Biol.* **27**, 3826–3836.e5 (2017).
60. Wang, D. et al. *spalt* is functionally conserved in *Locusta* and *Drosophila* to promote wing growth. *Sci. Rep.* **7**, 44393 (2017).
61. Medved, V. et al. Origin and diversification of wings: insights from a neopteran insect. *Proc. Natl Acad. Sci. USA* **112**, 15946–15951 (2015).
62. Ohde, T., Yaginuma, T. & Niimi, T. Insect morphological diversification through the modification of wing serial homologs. *Science* **340**, 495–498 (2013).
63. Hu, Y., Linz, D. M. & Moczek, A. P. Beetle horns evolved from wing serial homologs. *Science* **366**, 1004–1007 (2019).
64. Olesen, J., Haug, J. T., Maas, A. & Waloszek, D. External morphology of *Lightiella monniotae* (Crustacea, Cephalocarida) in the light of Cambrian 'Orsten' crustaceans. *Arthropod Struct. Dev.* **40**, 449–478 (2011).
65. Alexander, D. E. A century and a half of research on the evolution of insect flight. *Arthropod Struct. Dev.* **47**, 322–327 (2018).
66. Gruner, H.-E. Über das Coxalglied der pereopoden der isopoden. *Zool. Anz.* **152**, 312–317 (1954).
67. Dreyer, H. & Wägele, J. W. The Scutocoxifera tax. nov. and the information content of nuclear SSU rDNA sequences for reconstruction of isopod phylogeny (Peracarida: Isopoda). *J. Crustac. Biol.* **22**, 217–234 (2002).
68. Wolff, C. The embryonic development of the malacostracan crustacean *Porcellio scaber* (Isopoda, Oniscidea). *Dev. Genes Evol.* **219**, 545–564 (2009).
69. Prpic, N.-M. & Telford, M. J. Expression of *homothorax* and *extradenticle* mRNA in the legs of the crustacean *Parhyale hawaiiensis*: evidence for a reversal of gene expression regulation in the pancrustacean lineage. *Dev. Genes Evol.* **218**, 333–339 (2008).
70. Kerner, P., Ikmi, A., Coen, D. & Vervoort, M. Evolutionary history of the *iroquois/Irx* genes in metazoans. *BMC Evol. Biol.* **9**, 74 (2009).
71. Rehm, E. J., Hannibal, R. L., Chaw, R. C., Vargas-Vila, M. A. & Patel, N. H. Fixation and dissection of *Parhyale hawaiiensis* embryos. *Cold Spring Harb. Protoc.* **2009**, pdb.prot5127 (2009).
72. Rehm, E. J., Hannibal, R. L., Chaw, R. C., Vargas-Vila, M. A. & Patel, N. H. In situ hybridization of labeled RNA probes to fixed *Parhyale hawaiiensis* embryos. *Cold Spring Harb. Protoc.* **2009**, pdb.prot5130 (2009).
73. Choi, H. M. T. et al. Third-generation in situ hybridization chain reaction: multiplexed, quantitative, sensitive, versatile, robust. *Development* **145**, dev165753 (2018).
74. Sander, J. D., Zaback, P., Joung, J. K., Voytas, D. F. & Dobbs, D. Zinc Finger Targeter (ZiFIT): an engineered zinc finger/target site design tool. *Nucleic Acids Res.* **35**, W599–W605 (2007).
75. Sander, J. D. et al. ZiFIT (Zinc Finger Targeter): an updated zinc finger engineering tool. *Nucleic Acids Res.* **38**, W462–W468 (2010).
76. Martin, A. et al. CRISPR/Cas9 mutagenesis reveals versatile roles of Hox genes in crustacean limb specification and evolution. *Curr. Biol.* **26**, 14–26 (2016).
77. Gloor, G. B., Nassif, N. A., Johnson-Schlitz, D. M., Preston, C. R. & Engels, W. R. Targeted gene replacement in *Drosophila* via P element-induced gap repair. *Science* **253**, 1110–1117 (1991).
78. Linz, D. M. & Tomoyasu, Y. Dual evolutionary origin of insect wings supported by an investigation of the abdominal wing serial homologs in *Tribolium*. *Proc. Natl Acad. Sci. USA* **115**, E658–E667 (2018).

Acknowledgements

We thank C. Clark-Hachtel and Y. Tomoyasu for sharing results before publication. We thank E. Jarvis for the *Parhyale* phalloidin muscle stains and A. Pomerantz for the *Oncopeltus* image and whole adult *Parhyale* image. We thank C. Wolff, T. Deuve, Y. Kobayashi, F. Schram and J. W. Shultz for helpful discussion. We thank J. Shen for confirming that *spalt* RNAi reduces both the wing and lobe in *Locusta*. We thank G. Mardon for the image of the *Drosophila dac* knockout adult leg. This work is supported by the National Science Foundation (NSF; IOS-1257379 to N.H.P.) and NSF Graduate Research Fellowship (to H.S.B.).

Author contributions

H.S.B. and N.H.P. conceived the experiments. H.S.B. performed the experiments, conceived the model and wrote the manuscript. N.H.P. edited and revised the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at <https://doi.org/10.1038/s41559-020-01349-0>.

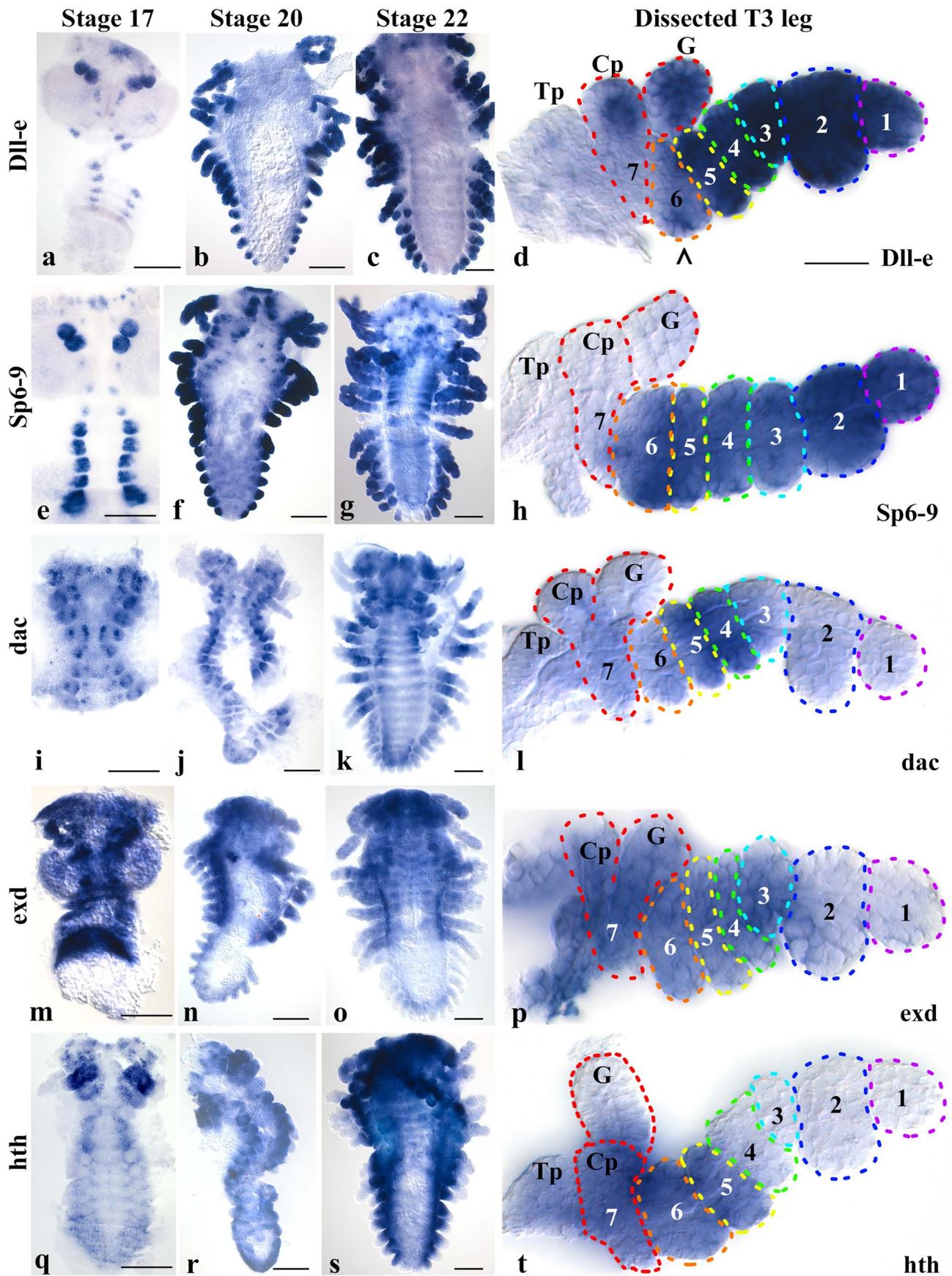
Supplementary information is available for this paper at <https://doi.org/10.1038/s41559-020-01349-0>.

Correspondence and requests for materials should be addressed to H.S.B.

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

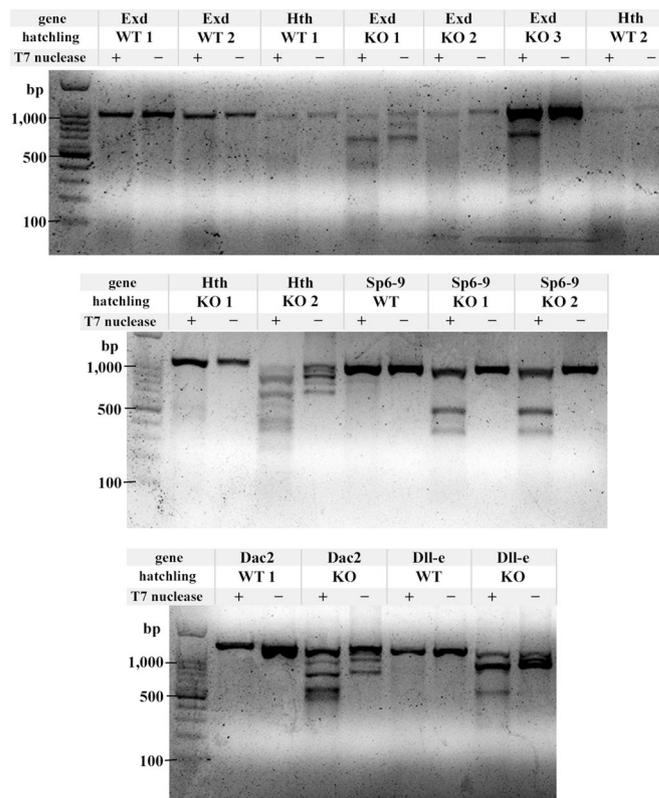
Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2020

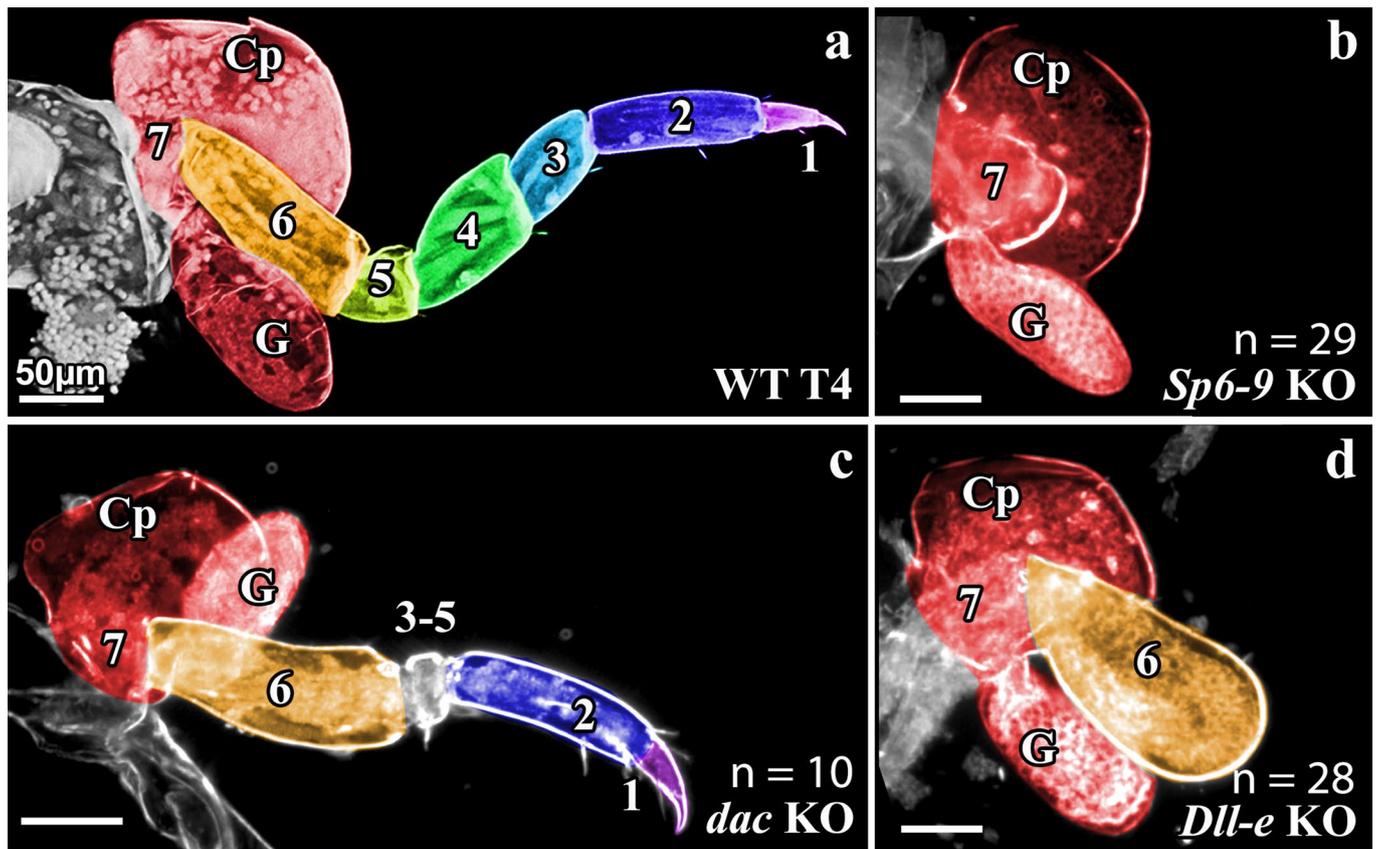


Extended Data Fig. 1 | See next page for caption.

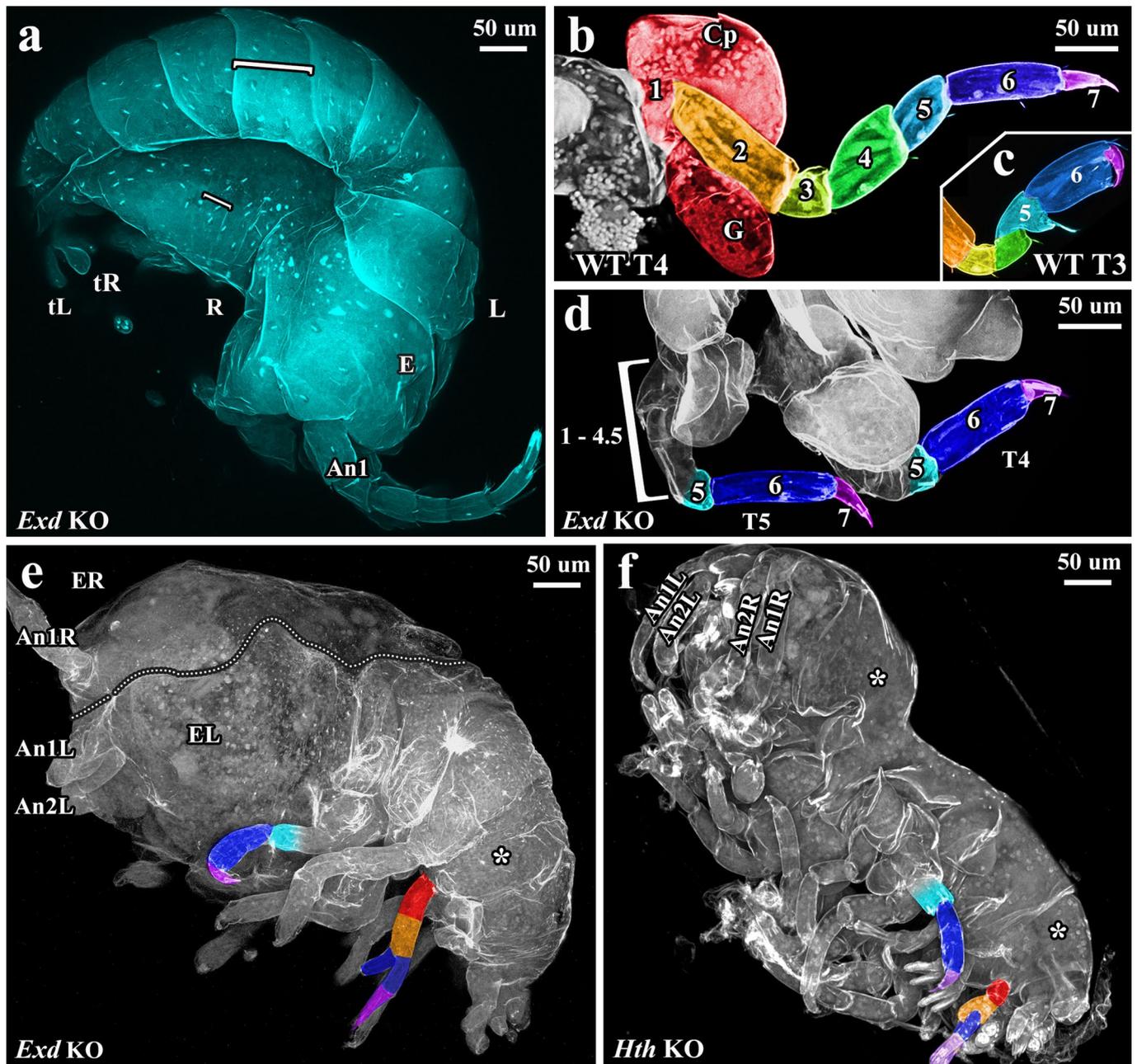
Extended Data Fig. 1 | Expression of leg gap genes in whole embryos and dissected third thoracic legs (T3). **a-d**, *Dll-e*. **e-h**, *Sp6-9*. **i-l**, *dac2*. **m-p**, *exd*. **q-t**, *hth*. Embryonic expression data for *Dll-e* (Browne 2005; Serano 2015), *Sp6-9* (Schaeper 2010), and *exd* and *hth* (Prpic 2008) have been previously characterized, but not at the level of individual leg segments. (d) *Dll-e* is expressed in leg segments 1-5; in the interior of the tergal plate (Tp), coxal plate (Cp), and gill (G). (h) *Sp6-9* is expressed in leg segments 1-6. (l) *dac2* is expressed in leg segments 3-5. Expression in segment 3 may be stronger at other time points. (p) *exd* is expressed in the body wall through leg segment 3. *Exd* is expressed only in the base of the gill (distal gill not visible here). (t) *hth* is expressed in the body wall through leg segment 5. *Hth* is expressed only in the base of the gill. Note that both insects and *Parhyale* share a peculiar disparity between *hth* expression and function, wherein *hth* knockout deletion extends one more leg segment further than the *hth* expression domain. Whole embryo scale bars = 100 μm . Dissected leg scale bar = 25 μm . a, n = 5; b, n = 6; c, n = 7; e, n = 6; f, n = 8; g, n = 8; i, n = 4; j, n = 4; k, n = 5; m, n = 3; n, n = 6; o, n = 4; q, n = 3; r, n = 3; s, n = 8.



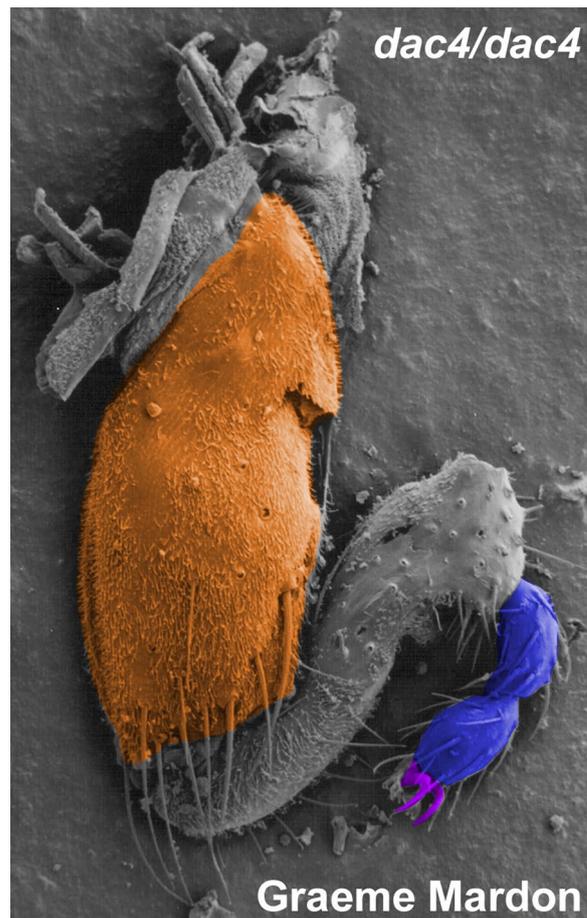
Extended Data Fig. 2 | T7 endonuclease assay to confirm CRISPR-Cas9 mutagenesis. For each gene, one or two wild type (WT) hatchlings were assayed, and one, two, or three KO hatchlings were assayed. T7 endonuclease was either added (+) or not added (-) to the heteroduplex mixture. In brief, a ~1-kb region flanking the CRISPR-Cas9 target site by at least 300 bp to either side was amplified by PCR from either WT or KO hatchlings. The purified PCR products were denatured, then slowly cooled to allow WT DNA and mutant DNA with indels to anneal, resulting in a 'bubble' of unpaired DNA (heteroduplex) at the target site. T7 endonuclease was added to the (+) samples, incubated, and run on a 1.5% agarose gel. KO animals are mosaic, so if the target site was cut, the indels will cause heteroduplexes when annealed with either a WT strand, or a different indel. When a single deletion is present, each half of the cut heteroduplex adds up to approximately 1 kb (see *Sp6-9* KO 1 and 2). Some deletions are large enough to be seen without the T7 endonuclease assay (see *Dll-e* KO), and some hatchlings had multiple deletions which produced multiple bands when cut with T7 (see *exd* KO 1, *hth* KO 2, *dac2* KO). In hatchlings that have a phenotype but only the WT band (*Exd* KO 2 and *Hth* KO 1), the deletion may be so large that one or both of the primer sites may have been deleted.



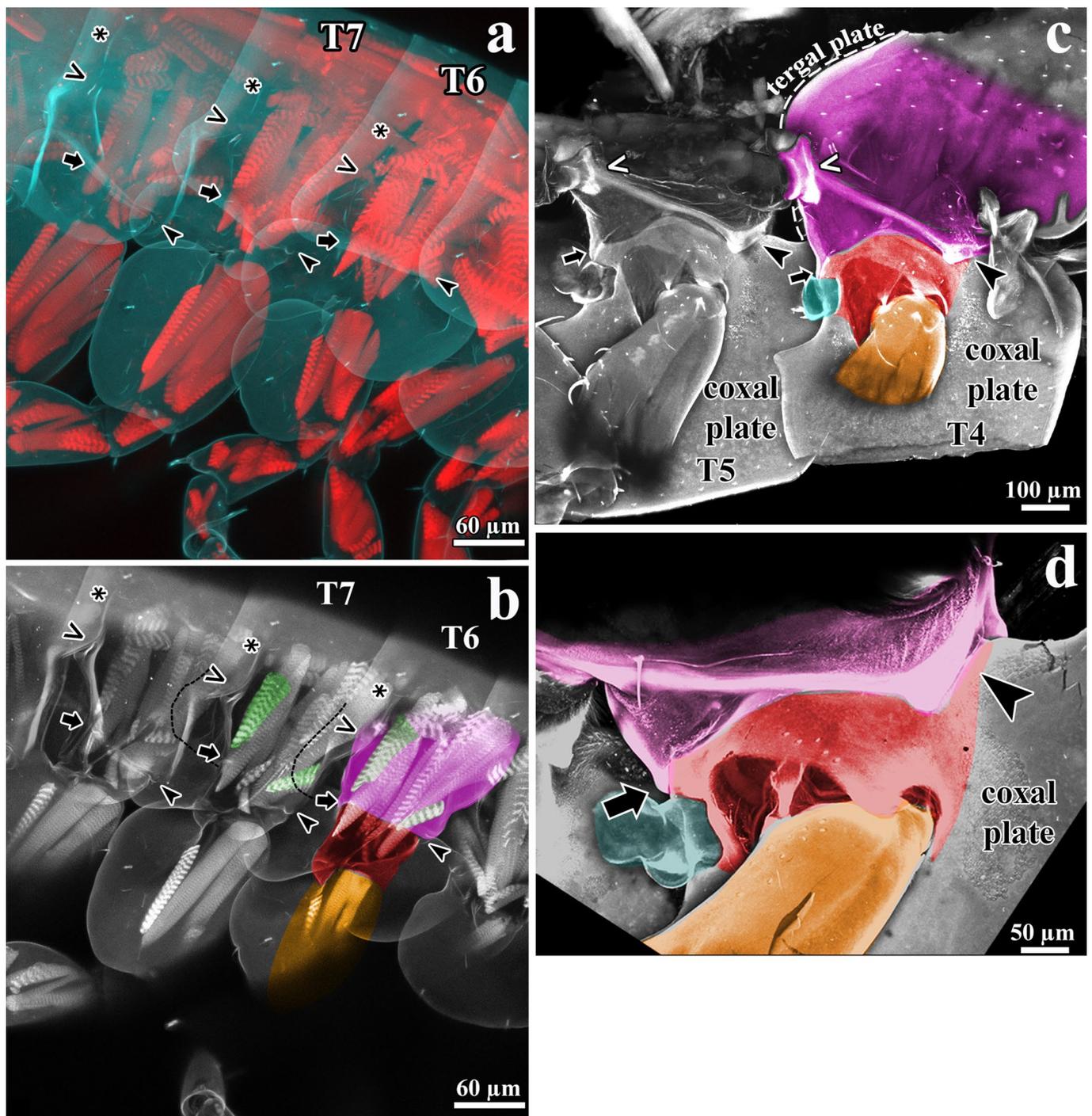
Extended Data Fig. 3 | CRISPR-Cas9 knockout phenotypes in thoracic legs 4 and 5 (T4, T5). Note that T4 and T5 are morphologically and molecularly indistinguishable, and are treated as interchangeable here. **a**, Wild type T4 leg. **b**, *Sp6-9* KO T4/T5 leg. **c**, *dac* KO T4/T5 leg. The fused and nearly deleted remnant of leg segments 3-5 are in gray. Gill is unaffected, but became oriented upward during specimen mounting. **d**, *Dll-e* KO T4/T5 leg. Scale bar = 50 μ m.



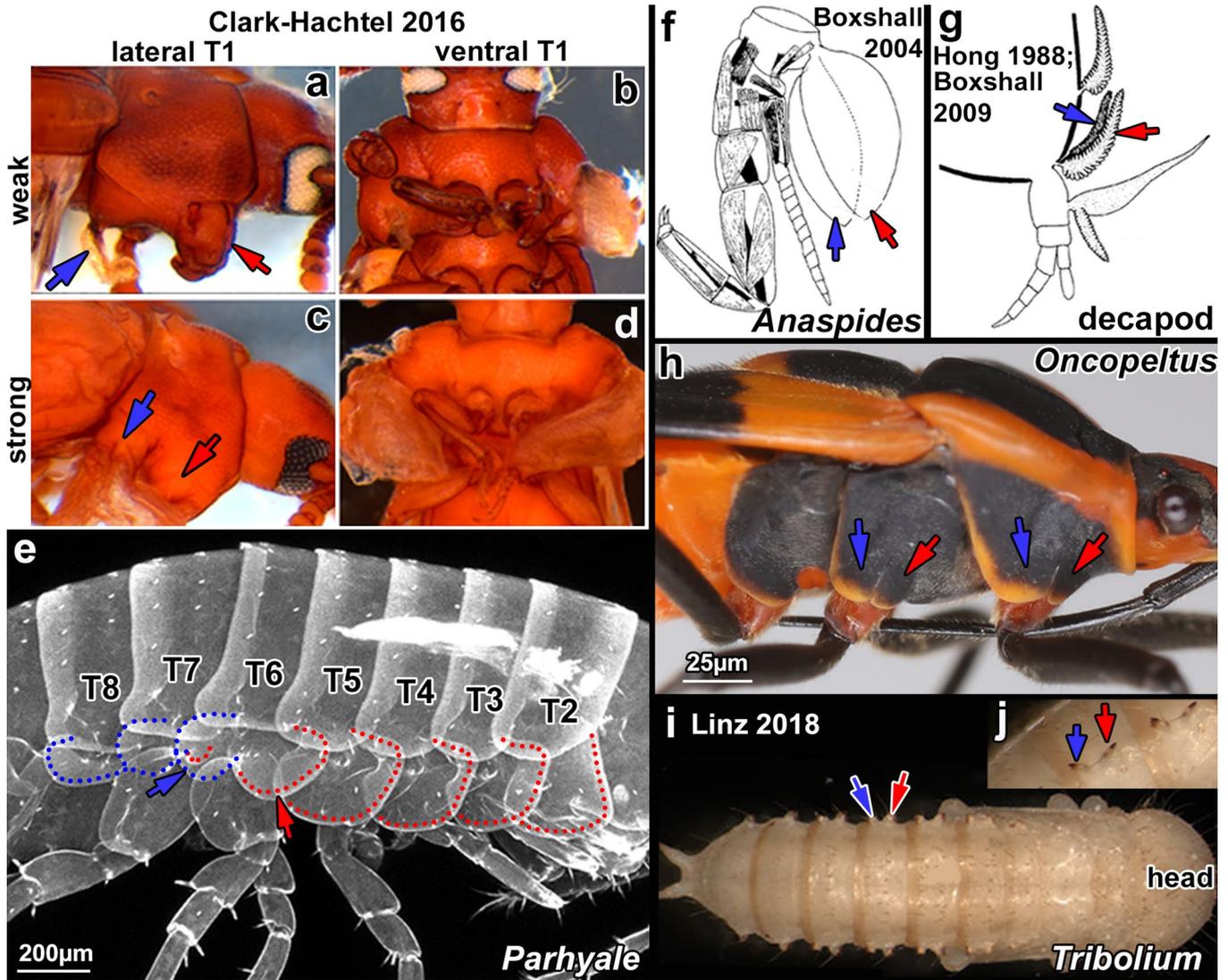
Extended Data Fig. 4 | Exd and Hth phenotypes. Body segment fusions are due to the interaction of *exd* and *hth* with *engrailed* during segmentation (Kobayashi 2003). **a**, Body segment fusions/deletions in *Exd* knockout whole hatchling. Confocal of unilaterally affected hatchling, dorsal view, anterior at bottom, posterior at left. Left side of animal (L) appears WT. The foreshortening of only the right (R) half twists the body laterally into a nearly spiral shape. The tissue where the eye (E) would have been located is deleted, leaving a recess. Left first antenna (An1), left and right telson (tL, tR). White brackets compare the length of the body segments in right fused and left unfused segments. **b**, WT T4 leg. **c**, WT T3 leg, same as Fig. 2a. **d**, *Exd* KO T4 leg. Loss of *exd* deletes/fuses leg segments 1-4 and proximal 5, while the distal half of leg segment 5, and all of leg segments 6 and 7 are WT, because the joint between leg segments 5 and 6 is normal, but there is no joint on the proximal side of leg segment 5, indicated by the proximally faded cyan shading. *exd* KO causes transformation of remaining T3 leg segments towards a T4/5 identity: compare blue leg segment 6 of WT T3 (Fig. 2) and T4 legs to that of *exd* KO T3 (Fig. 2) and T4 legs. **e**, **f**, *Exd* and *Hth* KO produce similar body segment fusions/deletions and proximal thoracic leg segment fusions/deletions. Colours in leg segments are as in Fig. 2.



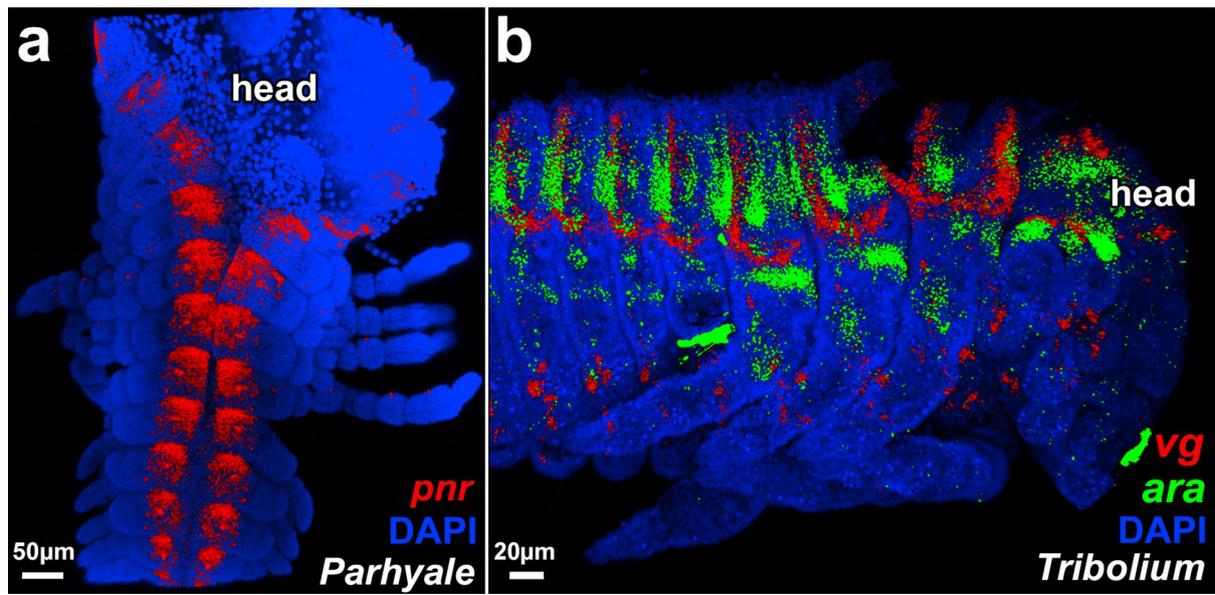
Extended Data Fig. 5 | *Drosophila dachshund* KO leg. *dac4/dac4* homozygote from *dac* mutant lines created by Professor Iain Dawson (Mardon 1994). The trochanter through proximal tarsus (leg segments 3–5, and proximal tarsus) are affected, forming a single, fused tissue. Figure adapted with permission from Graeme Mardon.



Extended Data Fig. 6 | *Parhyale* precoxa forms a true, muscled joint and extends musculature to another leg segment. Confocal images. **a**, Phalloidin stain of muscle in right half of *Parhyale* hatching. Contrast simple, anterior-posterior body muscles to orthogonal, complexly arranged leg muscles. No muscles cross the coxa-basis joint, as noted by Boxshall 1997. Note that all three plates (tergal, coxal, and basal) form contiguous cuticle with their leg segment, that is there is no distinguishing suture between leg segment and exite. **b**, Optical section showing superficial muscles of right half. Cuticle in grey, muscle in pink. **c**, Confocal of dissected left half, medial view. Coxal plate and basis partially cut. The precoxa forms a joint with two articulations with the coxa: an anterior, bifurcated, load-bearing hinge articulation (arrowhead), and a posterior gliding articulation (arrow). Orthogonal muscles visible as striations on T4 precoxa. **d**, Close-up of left T4, medial-anterior view, showing bifurcated hinge articulation. In **a-d**, the precoxa forms two articulations with the coxa: an anterior, bifurcated, load-bearing hinge articulation (arrowhead), and a posterior gliding articulation (arrow). Coxa is red (coxal plate not shaded, to focus on joints), basis is orange, precoxa is magenta pink. The proximal-most region of the legs brace against each other at (<) and (*). The length of the protrusion is twice that of the coxa: compare the beginning (< and *) and end of the protrusion where it forms a joint (arrow and arrowhead) with the coxa (most visible in **b**). Muscles in green insert on the precoxa-coxa joint, indicating that this is a true joint, and not merely a point of flexure in the exoskeleton (annulation; Boxshall 2004, 2013; Shultz 1989). The shorter, anterior muscle originates in the protruding precoxa to insert on the rim of the next leg segment, the coxa. This muscle is therefore an intrinsic muscle, a hallmark of a true leg segment (Boxshall 2004, 2013; Shultz 1989). Panels **a,b** adapted with permission from Erin Alberstat.



Extended Data Fig. 7 | Exites can be split into anterior and posterior lobes in crustaceans and insects. a-d, Ectopic wing formation on T1 following Scr RNAi in *Tribolium*. Mildly affected individuals have wing tissue emerging from unconnected anterior and posterior regions of the body wall. In more severe phenotypes, these anterior and posterior tissues are fused into a more completely transformed wing. **e-j,** Split exites in crustaceans and insects. Anterior lobe (red outline or arrow), posterior lobe (blue outline or arrow). **e,** Split coxal plates in *Parhyale*. **f,** A malacostracan crustacean, *Anaspides*, with split anterior and posterior exites. **g,** Decapod with split anterior and posterior exites (arthrobranches). **h,** *Oncopeltus* with split anterior and posterior supracoxal lobes. **i,j,** Gin traps require the wing genes *vg* and *ap*, consistent with an exite identity. Gin traps have split anterior and posterior jaws. Images adapted with permission from ref. ⁸, PNAS (**a-d**); from Erin Alberstat (**e**); from ref. ³⁹, John Wiley and Sons (**f**); from ref. ³⁸, Arthropod Systematics & Phylogeny (**g**); from Aaron Pomerantz (**h**); and from ref. ⁷⁸, PNAS (**i,j**).



Extended Data Fig. 8 | Proximal gene expression in later *Parhyale* and *Tribolium* embryos. **a**, *pannier* is expressed in the dorsal-most tissue in late *Parhyale* embryos during dorsal closure. **b**, *Tribolium* embryo, *vestigial* marks the future wing region adjacent to the spiracle. The three domains of *araucan* expression remain distinct even at later stages.

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- | | | |
|-------------------------------------|-------------------------------------|--|
| <input type="checkbox"/> | <input checked="" type="checkbox"/> | The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> | A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> | The statistical test(s) used AND whether they are one- or two-sided
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> | A description of all covariates tested |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> | A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> | For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
<i>Give P values as exact values whenever suitable.</i> |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> | Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated |

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

Data analysis

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	A total of 1,596 Parhyale embryos were injected, varying from 37 - 206 embryos per CRISPR sgRNA. We used 2 different guides per gene, and also co-injected the two guides. For each sgRNA or pair of sgRNAs, 66 - 255 of the embryos had a phenotype to score. We confirmed that all guides gave the same qualitative phenotypes.
Data exclusions	No data were excluded.
Replication	A total of 1,596 Parhyale embryos were injected, varying from 37 - 206 embryos per CRISPR sgRNA. We used 2 different guides per gene, and also co-injected the two guides. For each sgRNA or pair of sgRNAs, 66 - 255 of the embryos had a phenotype to score. We confirmed that all guides gave the same qualitative phenotypes.
Randomization	Embryos for injection were collected on different days from different females.
Blinding	Each leg gap gene gives a specific and distinct phenotype (unique sets of leg segments are affected), so blinding wasn't necessary or possible.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involvement in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input type="checkbox"/>	<input checked="" type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern

Methods

n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Antibodies

Antibodies used	Sheep anti-DIG-AP antibody (Roche 11093274910). Alexafluor 555 Phalloidin #A34055
Validation	<p>Analysis Note</p> <ul style="list-style-type: none"> • Cross reactivity to digitoxin and digitoxigenin: <1 % • No cross reactivity with other human estrogen or androgen steroids, e.g. estradiol or testosterone • Cross reactivity with digoxin: not known • Conjugate does not bind to itself at all • Normally one molecule of the conjugate binds to one molecule digoxigenin, although there are two possible binding sites for digoxigenin • Nonspecific binding to RNA is not expected

Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals	Amphipod crustacean Parhyale hawaiiensis Chicago F line. Embryos collected from females. Embryos were male and female, unknown ratio but presumably about 50:50.
Wild animals	study did not involve wild animals.

Field-collected samples

study did not involve field samples

Ethics oversight

no ethical approval required for research on lab populations of crustaceans

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Supplementary information

Knockout of crustacean leg patterning genes suggests that insect wings and body walls evolved from ancient leg segments

In the format provided by the authors and unedited

Supplementary Materials for

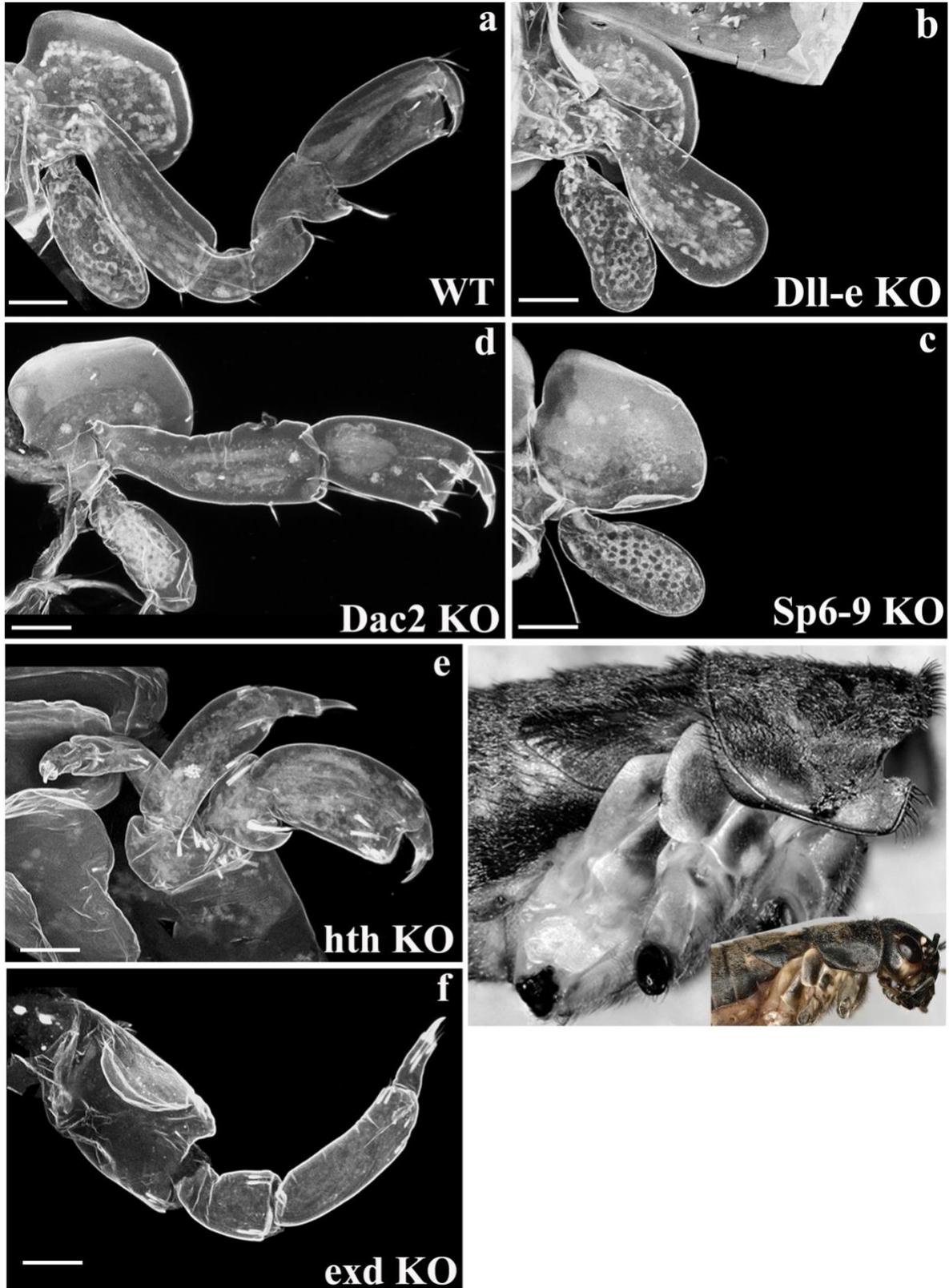
5 Insect wings and body wall evolved from ancient leg segments

Heather S. Bruce and Nipam H. Patel
Correspondence to: hbruce@mbi.edu

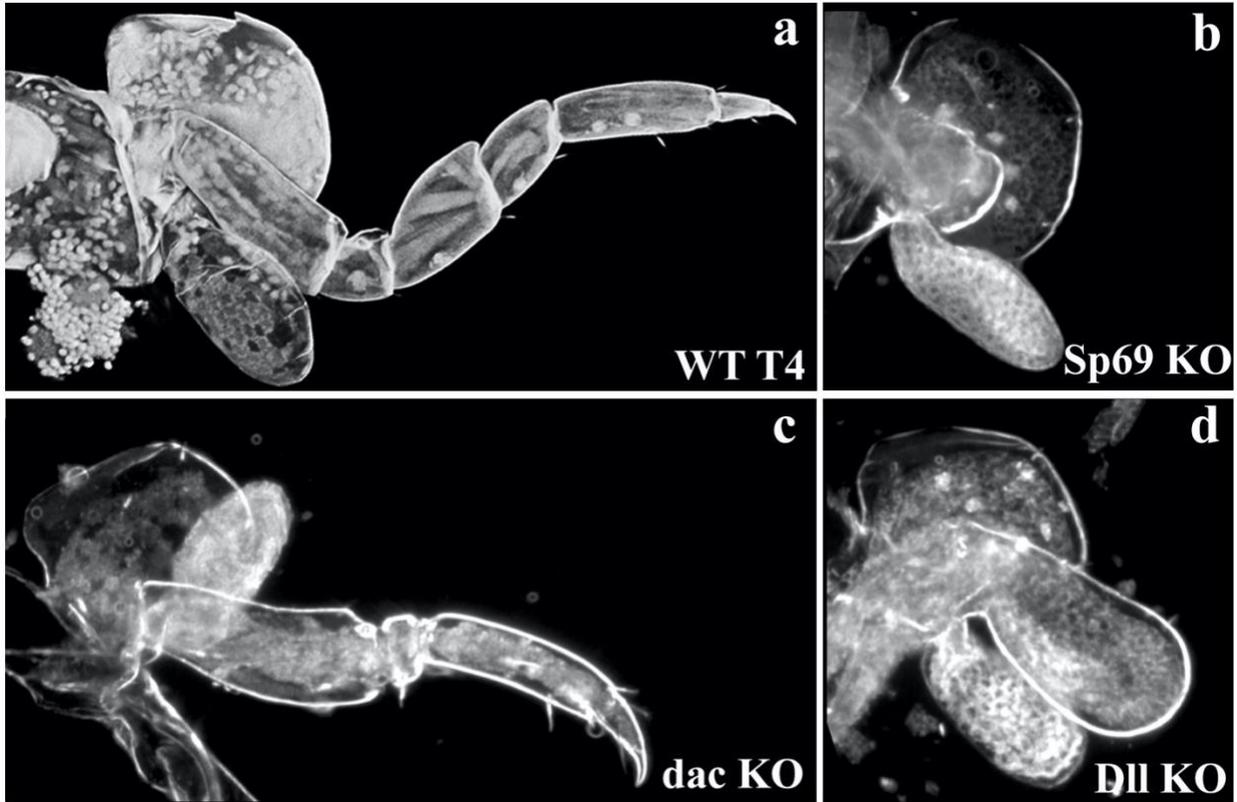
10

This PDF file includes:

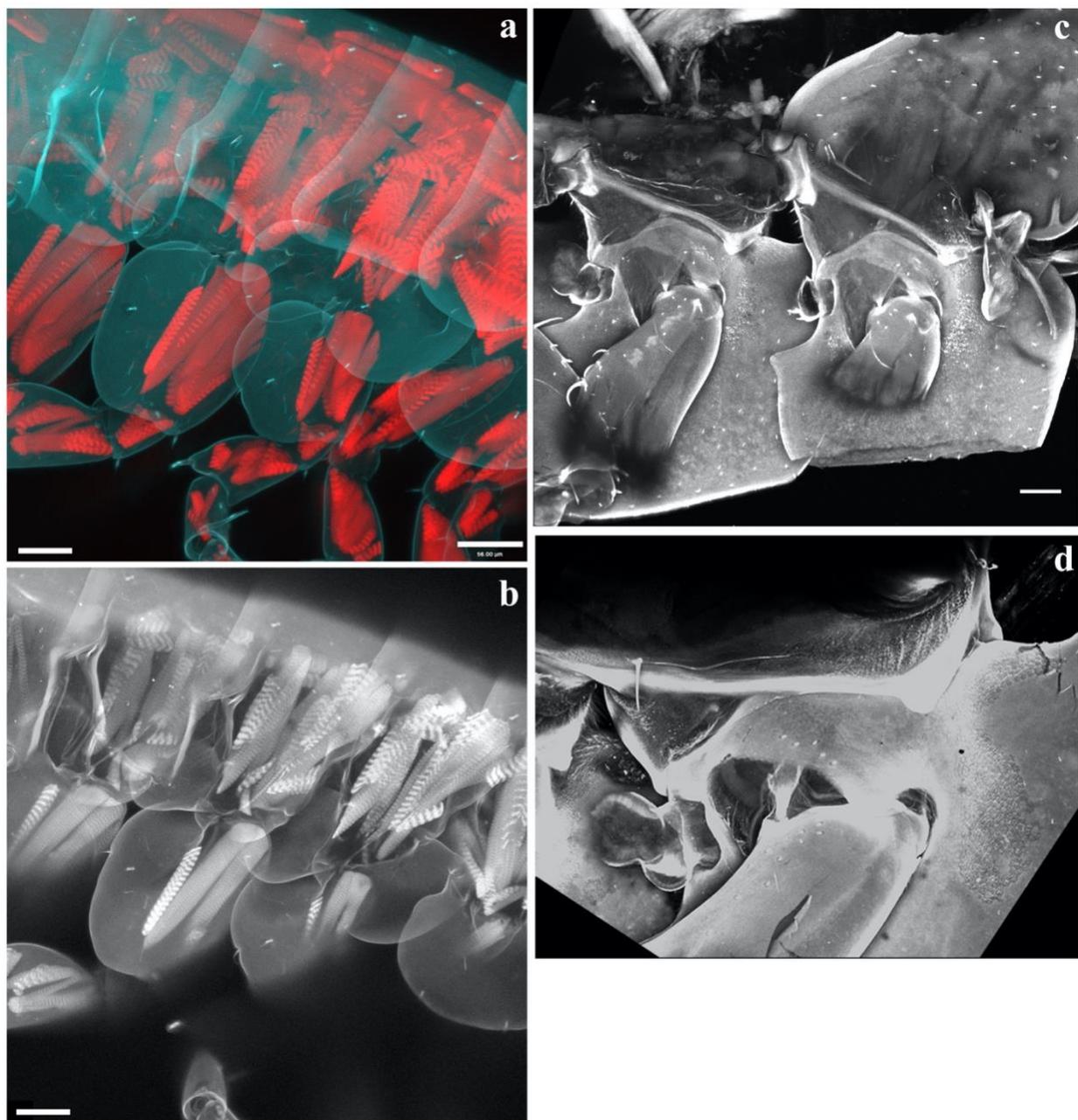
Supplementary Figures 1 to 4
Supplementary Tables 1 to 12
15 Supplementary Discussion



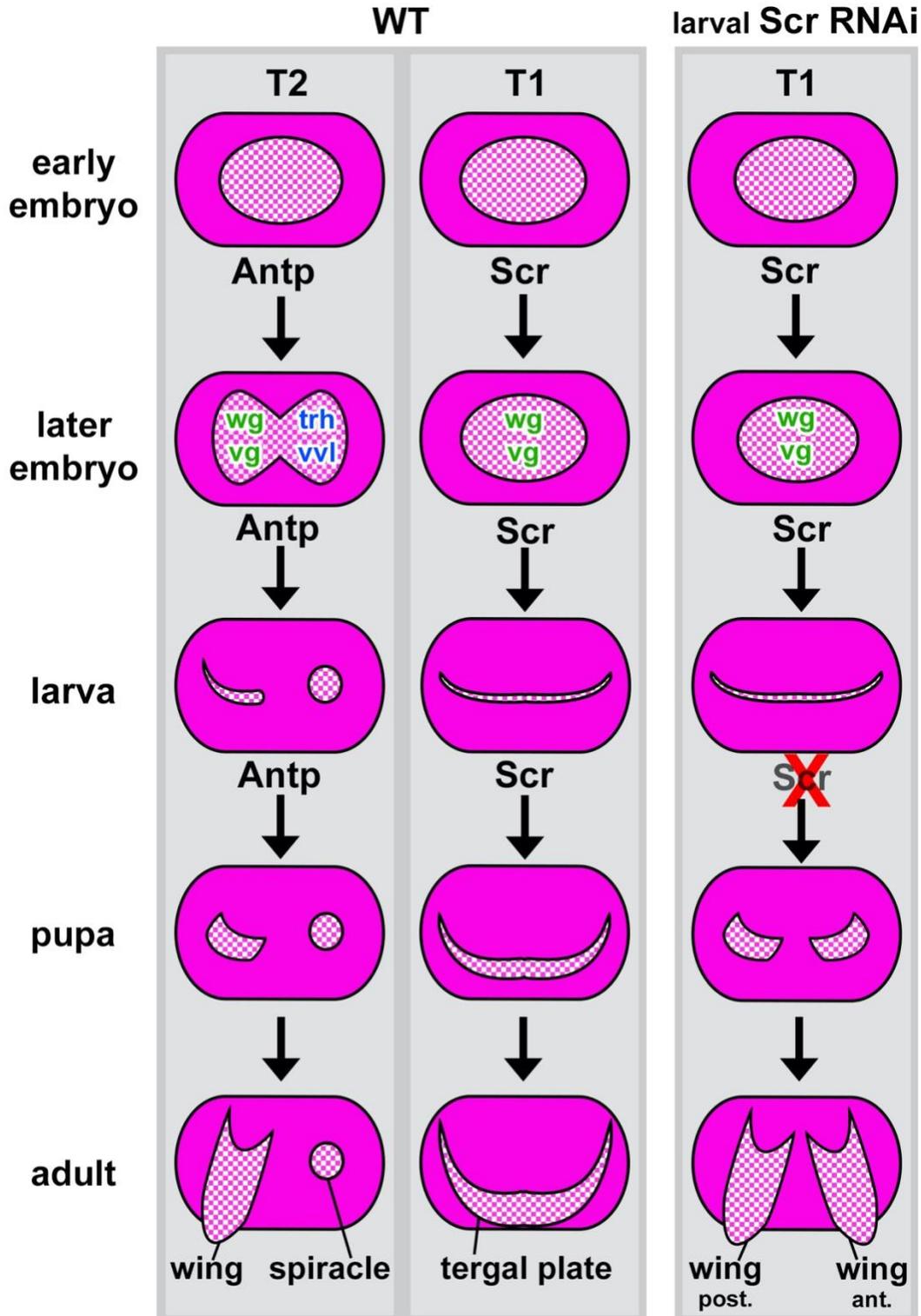
Supplementary Fig. 1. Figs. 2 and 5 with colours and labels removed.



Supplementary Fig. 2. Extended Data 3 with colours and labels removed.



25 **Supplementary Fig. 3.** Extended Data 7 with colours and labels removed.



Supplementary Fig. 4. Proposed exite patterning. *wingless* (*wg*), *vestigial* (*vg*) and *scalloped* appear to pattern plate-like outgrowths, while *trachealess* (*trh*) and *ventral veins lacking* (*vvl*) pattern respiratory structures. In T2 of WT embryos, *Antp* expression causes the exite to split. The anterior exite lobe expresses *trh* and *vvl* and forms the respiratory trachea, while the posterior exite lobe expresses *wg* and *vg* and forms a flat plate, the wing. In contrast, in T1 of WT embryos, *Scr* does not split the exite. The exite lobe expresses both *wg* and *vg* in a continuous U-shape from anterior to posterior, and thus forms a flat plate that encircles T1, the tergal plate. Thus, at the larval stage, the T1 exite is already patterned by *wg*, *vg*, and *sd* to make a flat exite, so the anterior exite lobe is not capable of forming a respiratory trachea to recapitulate the T2 condition (other genes may also be repressing formation of the anterior trachea on T2). However, if *Scr* is knocked down in the larva, the anterior and posterior region of the exite that were pre-patterned to be a single flat exite (the tergal plate) interrupt their program and become split into anterior and posterior wings. This produces two wing structures that, as they get bigger, run into each other and fuse together.

45

Gene sgRNA	Target site
Homothorax sgRNA1 Reverse	GGAAGCCATGTACGACCACA
Homothorax sgRNA2	GGTACGGAGCCCATAACA
Extradenticle sgRNA1	GGCAGTAGGAGTAGCACCGG
Extradenticle sgRNA2 Reverse	GGTGCTGTCGACCCAACCAC
Distalless early sgRNA1	GGGTACAGTAACATGGGGTA
Distalless early sgRNA2	GGCTTCCCCGCCGCCATGTA
Dachshund1 sgRNA1	GGTGGTGGGGAGGCATAATG
Dachshund1 sgRNA2	GGAGGTCGTGAAGTGGGTGG
Dachshund2 sgRNA1	GGCATGCCTCCCAAAGACC
Dachshund2 sgRNA2	GGTGGGGATGAGGTCGGCAA
Sp6-9 sgRNA1	GGCGACCACTTCTGGTGGAC
Sp6-9 sgRNA2 Reverse	GGTGTTAGATAAAAATCTGA

50 **Supplementary Table 1.** Target sites for sgRNAs ordered from Synthego.

gene	sgRNA	total injected	phenotype	#phenotype	#no phenotype	total hatched	% hatched w/phenotype
Dlle	1 + 2	151	truncated or shortened leg	57	56	113	50%
exd	1 + 2	103	fused body segments	30	11	53	80%
			ectopic eye	6			
			mouth transformed but no or few fused body segments	7			
exd	1	102	fused body segments	31	16	59	73%
			ectopic eye	9			
			mouth transformed but no or few fused body segments	3			
exd	2	173	fused body segments	60	16	120	87%
			ectopic eye	19			
			mouth transformed but no or few fused body segments	25			
hth	1 + 2	124	fused body segments	7	8	22	64%
			ectopic eye	2			
			mouth transformed but no or few fused body segments	5			
hth	1	131	fused body segments	12	45	66	32%
			ectopic eye	2			
			mouth transformed but no or few fused body segments	7			
hth	2	99	fused body segments	18	10	36	72%
			ectopic eye	2			
			mouth transformed but no or few fused body segments	6			
dac2	1 + 2	80	animal has more than 3 short legs	11	13	35	63%

			animal has up to three short legs, or twisted legs	11			
dac2	1	84	animal has more than 3 short legs	6	30	39	23%
			animal has up to three short legs, or twisted legs	3			
dac2	2	88	animal has more than 3 short legs	4	24	40	40%
			animal has up to three short legs, or twisted legs	12			
Sp69	1 + 2	165	truncated legs	51	33	84	61%
Sp69	1	54	truncated legs	9	1	10	90%
Sp69	2	37	truncated legs	15	4	19	79%

55 **Supplementary Table 2.** *Parhyale* CRISPR-Cas9 injection numbers. Two sgRNAs per gene were made, and either one or both were injected as indicated. For each gene, sgRNA 1 and 2 produced the same phenotypes. Only the strongest phenotypes (presumably null) were included in the figures, but the individual phenotypes were quite consistent from leg to leg and embryo to embryo. Moderate and weak phenotypes differed from the strongest phenotypes in that the same set of leg segments were deleted, but to greater or lesser degree, or in a certain order, suggesting that some leg segments are more sensitive to loss of each gene than other leg segments. These results are consistent with the range of phenotypes we have documented for *Parhyale* Hox gene knockouts (Martin 2016).

60

Query id	Subject id	E-value
<i>Daphnia_pulex</i> exd EFX62563.1	<i>Parhyale</i> exd MG457802	8e-177
<i>Drosophila</i> exd AAF48555.1	<i>Parhyale</i> exd MG457802	7e-173
<i>Hyalella</i> exd XP_018011298.1	<i>Parhyale</i> exd MG457802	2e-166
<i>Parhyale</i> exd CAO98909.1	<i>Parhyale</i> exd MG457802	6e-126
<i>Tribolium</i> exd NP_001034501.1	<i>Parhyale</i> exd MG457802	1e-173
Homo Pbx1 NP_002576.1	<i>Parhyale</i> exd MG457802	3e-166

65

Supplementary Table 3. Reciprocal best blast hits for *Parhyale* exd.

Query id	Subject id	E-value
<i>Daphnia</i> hth EFX75948.1	<i>Parhyale</i> hth MG457803	0
<i>Drosophila</i> hth NP_476578.3	<i>Parhyale</i> hth MG457803	6e-179
Homo Meis2 AAH07202.1	<i>Parhyale</i> hth MG457803	1e-148
<i>Hyalella</i> hth XP_018016731.1	<i>Parhyale</i> hth MG457803	0
<i>Parhyale</i> hth CAO98908.1	<i>Parhyale</i> hth MG457803	0
<i>Tribolium</i> hth NP_001034489.1	<i>Parhyale</i> hth MG457803	0

70

Supplementary Table 4. Reciprocal best blast hits for *Parhyale* hth.

Query id	Subject id	E-value
<i>Daphnia_pulex</i> dac EFX90187.1	<i>Parhyale</i> Dac1 MG457799	3e-67
<i>Drosophila</i> dac AAF53538.3	<i>Parhyale</i> Dac2 MG457800	2e-64
Homo dach2 Q96NX9	<i>Parhyale</i> Dac1 MG457799	4e-52
<i>Hyalella</i> Dac1 XP_018011787.1	<i>Parhyale</i> Dac1 MG457799	7e-109
<i>Hyalella</i> Dac1 XP_018011787.1	<i>Parhyale</i> Dac2 MG457800	2e-55
<i>Hyalella</i> Dac2 XP_018011801.1	<i>Parhyale</i> Dac2 MG457800	0
<i>Hyalella</i> Dac2 XP_018011801.1	<i>Parhyale</i> Dac1 MG457799	1e-59
<i>Tribolium</i> dac1 XP_015834662.1	<i>Parhyale</i> Dac2 MG457800	6e-72

75

Supplementary Table 5. Reciprocal best blast hits for *Parhyale* dac.

Query id	Subject id	E-value
<i>Drosophila</i> btd NP_511100.1	<i>Parhyale</i> Sp6-9 MG457804	4e-47
<i>Drosophila</i> Sp1-4 NM_142975.3	* <i>Parhyale</i> Sp1-4 CBH30980.1	5e-62
<i>Drosophila</i> Sp6-9 NP_727360.1	<i>Parhyale</i> Sp6-9 MG457804	6e-109
<i>Homo</i> Sp4 NP_003103.2	* <i>Parhyale</i> Sp1-4 CBH30980.1	2e-66
<i>Homo</i> Sp5 NP_001003845.1	<i>Parhyale</i> Sp6-9 MG457804	7e-62
<i>Homo</i> Sp8 NP_874359.2	<i>Parhyale</i> Sp6-9 MG457804	3e-105
<i>Hyaella</i> Sp1-4 XP_018012207.1	* <i>Parhyale</i> Sp1-4 CBH30980.1	0
<i>Hyaella</i> Sp6-9 XP_018014881.1	<i>Parhyale</i> Sp6-9 MG457804	0
<i>Parhyale</i> Sp1-4 CBH30980.1	* <i>Parhyale</i> Sp1-4 CBH30980.1	0
<i>Parhyale</i> Sp6-9 CBH30981.1	<i>Parhyale</i> Sp6-9 MG457804	0
<i>Tribolium</i> btd NP_001107792.1	<i>Parhyale</i> Sp6-9 MG457804	7e-59
<i>Tribolium</i> Sp1-4 XP_015833716.1	<i>Parhyale</i> Sp6-9 MG457804	3e-62
<i>Tribolium</i> Sp6-9 XP_008198341.1	<i>Parhyale</i> Sp6-9 MG457804	6e-159

Supplementary Table 6. Reciprocal best blast hits for *Parhyale* Sp genes.

80

Query id	Subject id	E-value
<i>Drosophila</i> pannier AAN13693.1	<i>Parhyale</i> pannier MT103930	4e-62
<i>Tribolium</i> pannier EFA04531.2	<i>Parhyale</i> pannier MT103930	1e-66

Supplementary Table 7. Reciprocal best blast hits for *Parhyale* pnr.

85

Query id	Subject id	E-value
<i>Drosophila</i> araucan NP001163432.1	<i>Parhyale</i> araucan MT103931	4e-62
<i>Tribolium</i> araucan XP008192408.1	<i>Parhyale</i> araucan MT103931	2e-101
<i>Parhyale</i> araucan MT103931	<i>Drosophila</i> araucan NP_001163432.1	2e-87
<i>Parhyale</i> araucan MT103931	<i>Tribolium</i> araucan XP008192408.1	6e-103
<i>Drosophila</i> araucan NP001163432.1	<i>Tribolium</i> araucan XP008192408.1	4e-105
<i>Tribolium</i> araucan XP008192408.1	<i>Drosophila</i> araucan NP001163432.1	5e-116

Supplementary Table 8. Reciprocal best blast hits for *Parhyale* ara.

90

Query id	Subject id	E-value
<i>Drosophila</i> Dll ACL83212.1	PhDIIIL2	2e-54
<i>Drosophila</i> Dll ACL83212.1	PhDIIIL1	2e-48
<i>Drosophila</i> Dll ACL83212.1	PhDIIe MG457801	4e-42
Homo DLX-2 AAB40902.1	PhDIIe MG457801	3e-35
Homo DLX-2 AAB40902.1	PhDIIIL2	6e-35
Homo DLX-2 AAB40902.1	PhDIIIL1	3e-34
<i>Hyalella</i> DLX-2 XP_018023955.1	PhDIIe MG457801	0
<i>Hyalella</i> DLX-2 XP_018023955.1	PhDIIIL1	1e-49
<i>Hyalella</i> DLX-2 XP_018023955.1	PhDIIIL2	3e-45
<i>Hyalella</i> DLX-6 XP_018023956.1	PhDIIIL2	4e-102
<i>Hyalella</i> DLX-6 XP_018023956.1	PhDIIIL1	1e-51
<i>Hyalella</i> DLX-6 XP_018023956.1	PhDIIe MG457801	1e-40
<i>Hyalella</i> unchar. protein XP_018023484.1	PhDIIIL1	8e-83
<i>Hyalella</i> unchar. protein XP_018023484.1	PhDIIIL2	0.89
<i>Parhyale</i> Dll-e ACT78885.1	PhDIIe MG457801	0
<i>Parhyale</i> Dll-e ACT78885.1	PhDIIIL1	7e-48
<i>Parhyale</i> Dll-e ACT78885.1	PhDIIIL2	1e-44
<i>Tribolium</i> Dll AAG39634.1	PhDIIIL1	7e-48
<i>Tribolium</i> Dll AAG39634.1	PhDIIIL2	1e-46
<i>Tribolium</i> Dll AAG39634.1	PhDIIe MG457801	5e-39

Supplementary Table 9. Reciprocal best blast hits for *Parhyale* Dll.

Primer name	product size	sequence
hth FORWARD	941	GTTATGGGCTCCGTACCTGA
hth REVERSE	941	GCCAGCTGTTTCTTCTGGTC
exd FORWARD	734	AGCGAGTCCTCAACAAAGGA
exd REVERSE	734	AGGAGGCGTGTGCTATTCTG
DII FORWARD	725	TGGGTCCAGTTCAACCTCTC
DII REVERSE	725	GACATCGTCCTCCAAAGCAT
dac 1 FORWARD	638	GGAGAGCAGAGGGGACTTTT
dac 1 REVERSE	638	CCACTTCACGACCTCCTCAT
dac 2 FORWARD	699	CTTCAACCCCCTCCAGTACA
dac 2 REVERSE	699	TGCTGTGTCGTCGTCTTCCTG
Sp6-9 FORWARD	789	CAAATGGCTCGCATGTATTG
Sp6-9 REVERSE	789	CAGTGCGTTCAAACCTCCAA

100 **Supplementary Table 10.** In situ primer sequences

Number	Organism	Target
PRD002	Parhyale hawaiiensis	DII
PRD221	Parhyale hawaiiensis	araucan
PRD223	Parhyale hawaiiensis	pannier
PRD232	Tribolium castaneum	vg
PRD736	Tribolium castaneum	pannier
PRD737	Tribolium castaneum	araucan
PRD738	Tribolium castaneum	DII

105 **Supplementary Table 11.** In situ HCR probe set numbers from Molecular Instruments

Primer name	product size	sequence
exd left	907	CTTGAGATTCG TTCAGGTGCA 110
exd right	907	TTCTCCCCAGTTCCTTGCAA
hth left	943	TGTTTCGTGTACCCGCAGAT
hth right	943	TCGGGCATACTAGAAGGCAG
Sp6-9 left	935	GCCCAGCTACTAACGATTTTCA ₁₁₅
Sp6-9 right	935	GATCCGCTTCCTGACAGTTG
Dll-e left	922	GGAATGGTGAAGGAAGAGCG
Dll-e right	922	TCAGCAGTGCAGACTCATGT
dac2 left	983	CACGCGACACTCATACACAG 120
dac2 right	983	GATGCTCCTCCCACCGAATA

Supplementary Table 12. Genomic DNA primers for T7 endonuclease assay.

Supplementary Discussion

125 Clark-Hachtel and Tomoyasu's interpretation of the dual origin theory proposes that both
ancestral exites and/or both ancestral leg segments fused to form the wing. This interpretation is
based on their observations in Clark-Hachtel 2013¹. They show that, when the Hox gene *Scr* is
knocked down in *Tribolium*, ectopic wings form on T1, producing a phenotypic series where
more mildly affected individuals have wing tissue emerging from unconnected anterior and
130 posterior positions of the body wall (Extended Data 8a – d). As the phenotype gets more severe,
these anterior and posterior tissues appear to fuse into a more completely transformed wing. We
agree with their interpretation that the ectopic wing is derived from two tissues. However, the
two tissues that fuse to form ectopic wings in their study are not consistent with a crustacean
coxa and precoxa identity (leg segments 7 and 8). If the wing were a fusion of exites from leg
135 segments 7 and 8, the two outgrowths of the ectopic wings should be oriented more
dorsoventrally, as the leg segments themselves are, and as the cricket wing and lobe are, rather
than the more anteroposterior orientation observed in the two ectopic wing tissues. More
importantly, as detailed in the main Discussion, it appears that both ancestral leg segments and
their respective exites have remained separate and distinct.

140 What, then, explains the observation that ectopic wings form by the fusion of two tissues?
In both crustaceans and insects, exites can be divided into an anterior and posterior lobe. In
crustaceans at least, these lobes appear to develop from a single bud that later splits^{2,3}. The coxal
plates of *Parhyale* are an example of exites that are split into anterior and posterior lobes. They
form a natural phenotypic series on the thoracic legs (Extended Data 8e): on thoracic legs T2-5,
145 the coxal plate consists of only an anterior lobe; on T6, the coxal plate is divided into equally
sized anterior and posterior lobes; on T7, the coxal plate is divided but the anterior lobe is

smaller than the posterior lobe; and on T8, the coxal plate is divided but the anterior lobe is miniscule compared to the large posterior lobe. Similarly, the exites of many other malacostracan crustaceans are split into anterior and posterior lobes (Extended Data 8f, g)². Importantly, it appears that the different lobes of the split exite can adopt different functional fates: in *Parhyale* embryos, the single coxal exite splits to form the flat, protective coxal plate and also the respiratory gill³.

Split exites can also be seen in insects. The lobes above the legs of cricket (Fig. 5b), locust (Fig. 5e), and *Oncopeltus* (Extended Data 8h) are divided into anterior and posterior lobes. The gin traps of beetles (ectodermal outgrowths that emerge in the same dorsoventral position as wings and require wing genes^{4,5}, consistent with an exite identity) are also split into an anterior jaw and a posterior jaw (Extended Data 8i, j).

The trachea, the insect respiratory organ, also appears to be an exite^{6,7}. *tracheiless* (*trh*), which is expressed in and required for trachea formation, is also expressed in the crustacean gill⁶, an exite. *ventral veins lacking* (*vvl*) is also expressed in both insect trachea and crustacean gills⁶. Based on our *pnr* and *ara* expression studies, the spiracle/trachea form on the remnant of leg segment 8. We therefore propose that insect spiracle/trachea are invaginated exites on the leg segment 8 remnant.

Different types of exites appear to be patterned by different genes (Supplementary Figure 4). *wingless* (*wg*) seems to pattern the ventral side of legs and flat exites. *vestigial* (*vg*) and *scalloped* appear to pattern plate-like outgrowths, such as the insect wing, insect T1 tergal plate (Fig. 3h), *Daphnia* carapace⁸, and *Parhyale* tergal plate, coxal plate, and basal plate⁹. Notably, *vg* is not expressed in the trachea (Fig. 3h) or in the *Parhyale* gill⁹, which makes sense, since the

trachea and gill are not a plate-like outgrowth. *trh* and *vvl* pattern respiratory structures, such as
 170 the insect trachea and crustacean gill, and are not expressed in plate-like structures⁶.

This model is informative for understanding the expression of *wg* (Fig. 5c, d)¹⁰ and *vg* (Fig. 3h) in insect T1 vs T2/3. In *Tribolium* T1, *wg* and *vg* are both expressed in a U-shape (Fig. 3h, 5c, d), presumably patterning the tergal plate that extends anteriorly and posteriorly, completely encircling T1. In contrast, in T2 and T3, *wg* and *vg* are expressed in a posterior C-
 175 shape, where the wing will form, but stops short of the spiracle. Importantly, insects do not form spiracles/trachae on T1^{11,12}. We interpret these observations to mean that the exite on leg segment 8 in insects is split into anterior and posterior portions: in T2/3, the anterior part forms the spiracle, while the posterior part forms the wing; in T1, both anterior and posterior parts form plate-like extensions, the tergal plate (Supplementary Figure 4, WT).

180 This model leads to the following interpretation. Reduction of *Scr* allows for the T1 exite on leg segment 8 to become wing tissue (Supplementary Figure 4, *Scr* RNAi). Normally, on T2/3, the posterior part of the exite becomes wing, while the anterior part becomes trachae. However, since the RNAi experiment in Clark-Hachtel 2013 was done in larvae, the anterior exite of T1 was already patterned to be plate-like, not respiratory. Thus, reduction in *Scr* at the
 185 larval stage caused both the anterior and posterior parts of the exite to become transformed towards wing.

This model predicts that wing can be transformed into trachae, and vice versa. In *Drosophila*, *wg* is expressed in the posterior where the wing will form, while *vvl* and *trh* are expressed in the anterior where trachea form⁶. Notably, *wg* represses tracheal formation, because
 190 when *wg* is overexpressed, *vvl* and *trh* expression are abolished. Conversely, when *wg* is knocked out, *trh* and *vvl* become expressed in both the anterior and posterior tissue, where the

future wing will form, and the posterior tissue invaginates like trachea. This suggests that both trachae and wing are exites. If trachae and wing are anterior and posterior lobes of an exite, and can be transformed into one another, this supports the explanation provided here for the
195 observation that two tissues, anterior and a posterior, on T1 appear to fuse to form ectopic wings.

1. Clark-Hachtel, C. M., Linz, D. M. & Tomoyasu, Y. Insights into insect wing origin provided by functional analysis of vestigial in the red flour beetle, *Tribolium castaneum*. **110**, 16951–
200 16956 (2013).
2. Boxshall, G. A. & Jaume, D. Exopodites, epipodites and gills in Crustaceans. *Arthropod Systematics & Phylogeny* 1–27 (2009).
3. Ungerer, P. & Wolff, C. External morphology of limb development in the amphipod *Orchestia cavimana* (Crustacea, Malacostraca, Peracarida). *Zoomorphology* **124**, 89–99
205 (2005).
4. Ohde, T., Yaginuma, T. & Niimi, T. Insect morphological diversification through the modification of wing serial homologs. *Science* **340**, 495–498 (2013).
5. Linz, D. M. & Tomoyasu, Y. Dual evolutionary origin of insect wings supported by an investigation of the abdominal wing serial homologs in *Tribolium*. **115**, E658–E667 (2018).
- 210 6. Franch-Marro, X., Martín, N., Averof, M. & Casanova, J. Association of tracheal placodes with leg primordia in *Drosophila* and implications for the origin of insect tracheal systems. *Development* **133**, 785–790 (2006).
7. Sánchez-Higueras, C., Sotillos, S. & Castelli-Gair Hombría, J. Common origin of insect trachea and endocrine organs from a segmentally repeated precursor. *Curr. Biol.* **24**, 76–81
215 (2014).

8. Shiga, Y. *et al.* Repeated co-option of a conserved gene regulatory module underpins the evolution of the crustacean carapace, insect wings and other flat outgrowths. **120**, 621–24 (2017).
9. Clark-Hachtel, C. M. & Tomoyasu, Y. Two sets of candidate crustacean wing homologues
220 and their implication for the origin of insect wings. *Nat Ecol Evol* (2020)
doi:10.1038/s41559-020-1257-8.
10. Ober, K. A. & Jockusch, E. L. The roles of wingless and decapentaplegic in axis and
appendage development in the red flour beetle, *Tribolium castaneum*. *Developmental
Biology* **294**, 391–405 (2006).
- 225 11. Gullan, P. J. & Cranston, P. S. *The Insects: An Outline of Entomology*. Willey. (2010).
12. Grimaldi, D. & Engel, M. S. *Evolution of the Insects*. (Cambridge University Press, 2005).