Insect wings and body wall evolved from ancient leg segments

Heather S. Bruce* and Nipam H. Patel
Department of Molecular and Cell Biology, University of California, Berkeley, CA
*Correspondence: hbruce@berkeley.edu

Researchers have long debated the origin of insect wings. One theory proposes that the proximal portion of the ancestral crustacean leg became incorporated into the body[1-3], which moved the leg's epipod (multi-functional lobe, e.g. gill) dorsally, up onto the back to form insect wings[4]. Another theory proposes that the dorsal insect body wall co-opted crustacean epipod genes to form wings[5]. Alternatively, wings may be derived from both leg and body wall (dual origin)[6]. To determine whether wings can be traced to ancestral, pre-insect structures, or arose by co-option, comparisons are necessary between insects and arthropods more representative of the ancestral state, where the hypothesized proximal leg region is not fused to the body wall. To do so, we examined the function of five leg gap genes in the crustacean *Parhyale hawaiensis* and compared this to previous functional data from insects. Here we show, using CRISPR-Cas9 mutagenesis, that leg segment deletion phenotypes of all five leg gap genes in *Parhyale* align to those of insects only by including the hypothesized fused ancestral proximal leg region. We also argue that possession of eight leg segments is the ancestral state for crustaceans. Thus, *Parhyale* incorporated one leg segment into the body, which now bears the tergal plate, while insects incorporated two leg segments into the body, the most proximal one bearing the wing. We propose a model wherein much of the body wall of insects, including the entire wing, is derived from these two ancestral proximal leg segments, giving the appearance of a "dual origin" [6-10]. This model explains many observations in favor of either the body wall, epipod, or dual origin of insect wings.

Arthropod appendages are key to their spectacular success, but their incredible diversity has complicated comparisons between distantly related species. The origin of the most debated appendage, insect wings, pivots on the alignment of leg segments, because wings may be derived from an epipod (e.g. gill or plate, Fig. 1b)[11] of ancestral leg segments that fused to the body[4,12], or alternatively, may represent a co-option of the epipod-patterning pathway by the insect body wall[5], or a combination of both (Clark-Hachtel, accompanying manuscript)[6-10]. To answer this, functional comparisons are necessary between insects and arthropods more representative of the ancestral state, where the hypothesized proximal leg region is not fused to the body wall.

Towards this aim, we examined five leg gap genes, *Distalless* (*Dll*), *Sp6-9*, *dachshund* (*dac*), *extradenticle* (*exd*), and *homothorax* (*hth*), in an amphipod crustacean, *Parhyale hawaiensis*. While we have documented their expression at several developmental stages (Fig. S1), our comparative analysis does not rely solely on these expression patterns, given that expression is not always a reliable indication of function, and expression is often temporally dynamic[13]. Instead, we have systematically knocked out these genes in *Parhyale* using CRISPR-Cas9 mutagenesis and compared this to our understanding of their function in *Drosophila* and other insects (Figs. 2, S2).

Insects have six leg segments, while *Parhyale* has seven (Fig. 1). In insects, *Dll* is required for the development of leg segments 2 – 6[14-17] . In *Parhyale*, the canonical *Dll* gene,

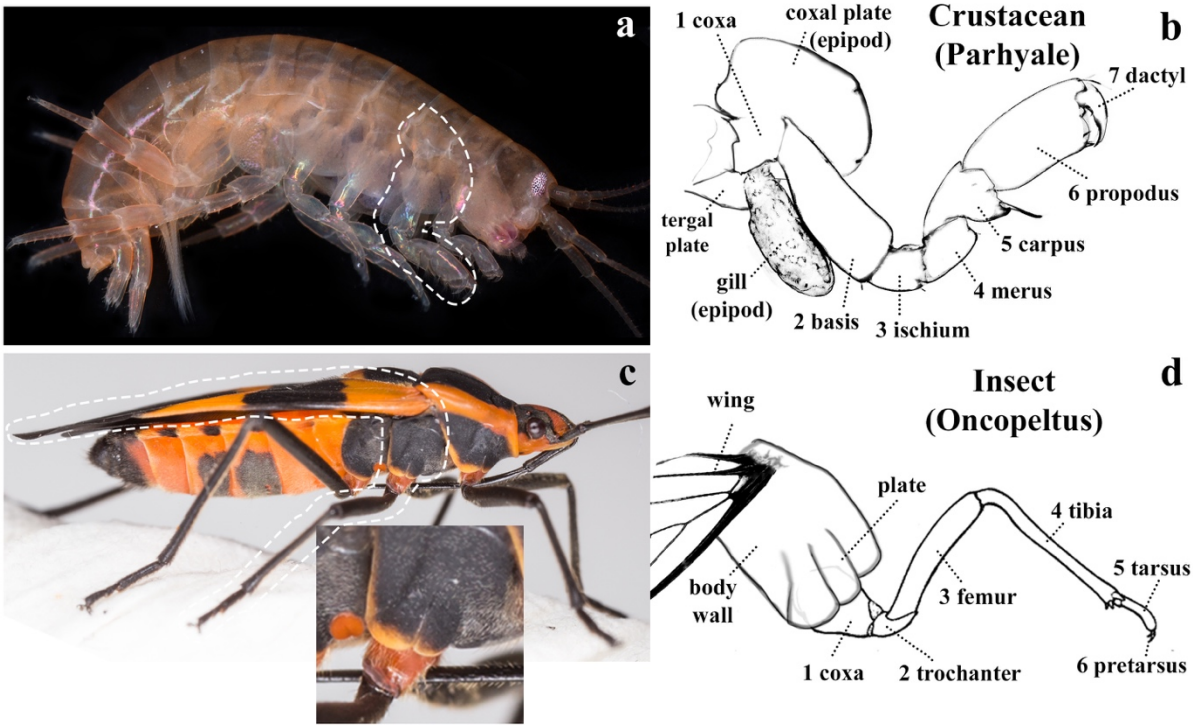


Fig. 1. Crustacean and insect legs. (a) Adult *Parhyale*, with third thoracic leg (T3) outlined. (b) Cartoon of *Parhyale* T3. The coxal plate extends over the leg. (c) Adult *Oncopeltus*, with T2 outlined. Inset shows magnified proximal leg, with body wall plate extending over the leg. (d) Cartoon of *Oncopeltus* T2 leg.

Dll-e[18-20], is required for the development of leg segments 3 – 7 (Fig. 2b). In insects, *Sp6-9* is required for the development of leg segments 1 – 6 [14,21-23], and in addition in *Drosophila*, loss of *Sp6-9* (i.e. D-Sp1[22]) occasionally transforms the leg towards wing and lateral body wall identity[23]. In *Parhyale*, *Sp6-9*[22] is required for the development of leg segments 2 – 7 (Fig. 2c), and in some legs, segment 2 is occasionally homeotically transformed towards a leg segment 1 identity (Fig S3). In *Drosophila*, dac is required in the trochanter through proximal tarsus (leg segments 2 – 4, and first tarsus)[24,25]. *Parhyale* has 2 dac paralogs. Dac1 does not seem to be expressed in the legs or have a knockout phenotype. Dac2 is required to pattern leg segments 3 – 5 (Fig. 2d). *Exd* and *hth* are expressed in the body wall and proximal leg segments of insects[26-29] and *Parhyale*[30] (Fig S1). They form heterodimers[31] and therefore have similar phenotypes[26-29]. In insects, *exd* or *hth* knockout results in deletions/fusions of the coxa through proximal tibia (leg segments 1 – 3, and proximal tibia) [26-29]. In *Parhyale*, *exd* or *hth* knockout results in deletions/fusions of the coxa through proximal carpus (leg segments 1 – 4, and proximal carpus; Figs. 2e, f). In both insects[26,27,32] and *Parhyale*, the remaining distal leg segments are sometimes transformed towards a generalized thoracic leg identity (compare Fig. 2 e, f and Fig S4). In both insects[26-29] and *Parhyale* (Fig. S4), *exd* or *hth* knockout results in deletions/fusions of body segments.

In summary, the expression and function of *Dll*, *Sp6-9*, *dac*, *exd*, and *hth* in *Parhyale* are shifted distally by one segment relative to insects. This shift is accounted for if insects fused an ancestral proximal leg segment to the body wall (Fig. 2g). Thus, there is a one-to-one homology between insect and *Parhyale* legs, displaced by one segment, such that the insect coxa is homologous to the crustacean basis, the insect femur is the crustacean ischium, and so on for all

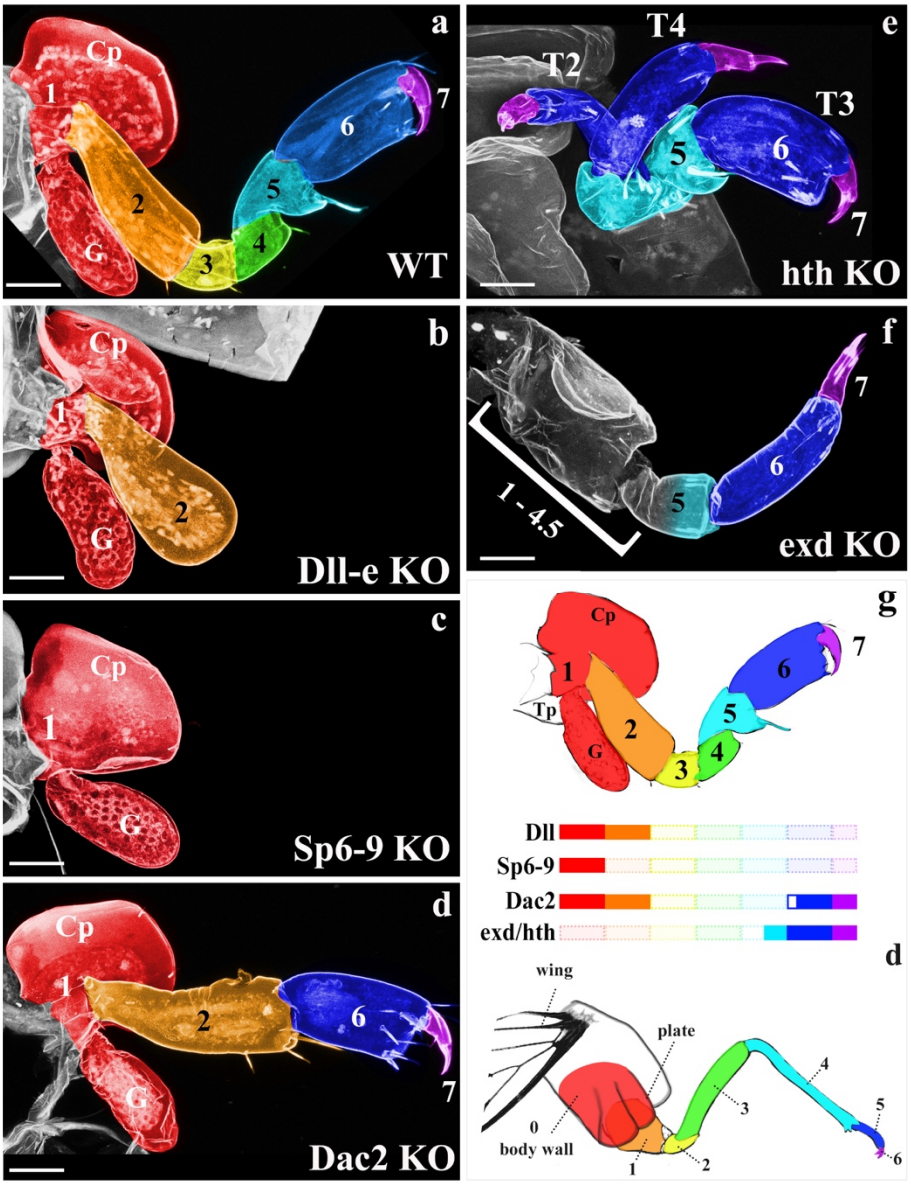
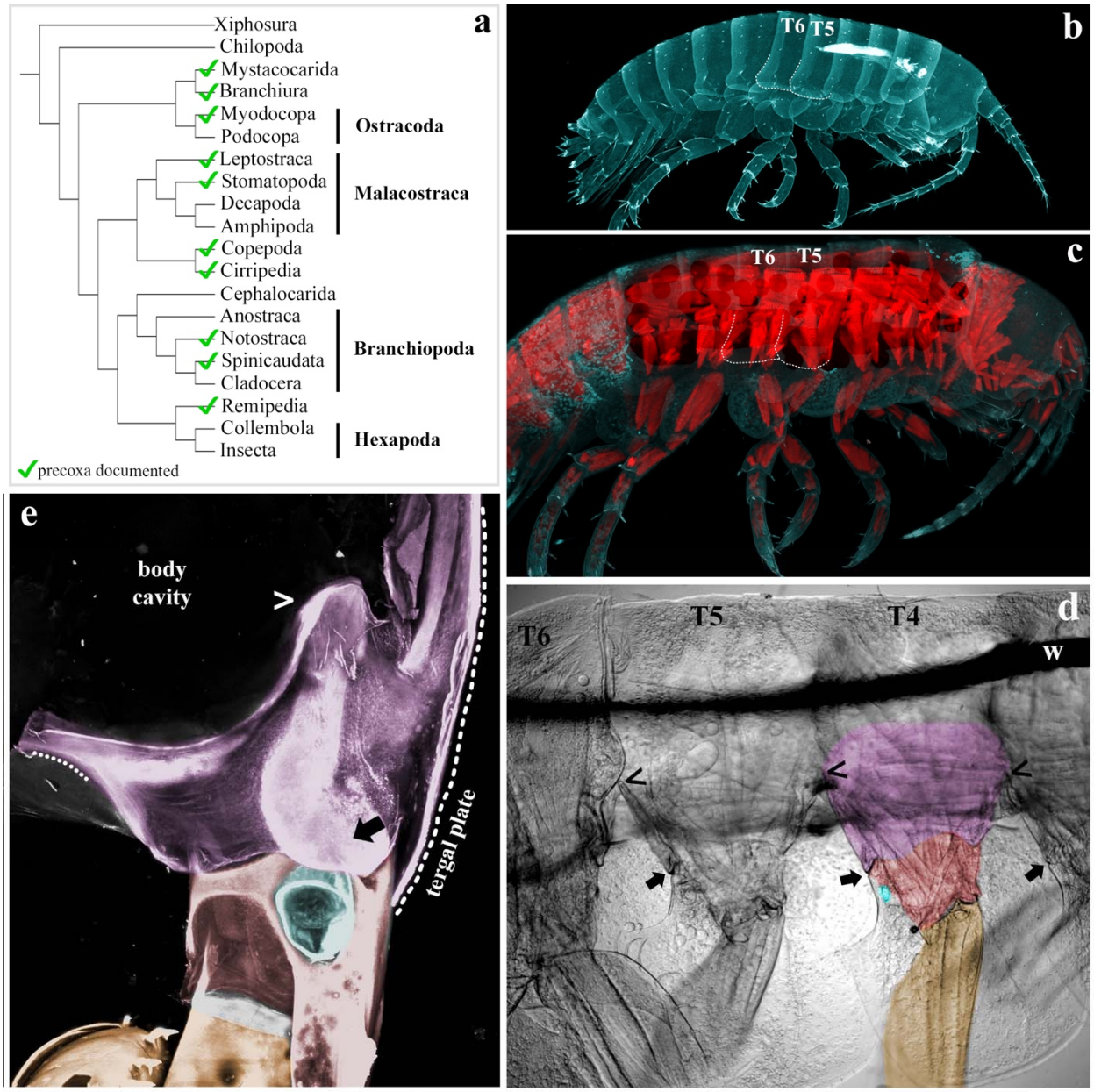


Fig. 2. Knockout phenotypes of leg gap genes. (a-f) *Parhyale* CRISPR-Cas9 phenotypes in dissected third thoracic legs (T3). Graded cyan in f indicates deletion/fusion of proximal leg segment 5. (g) Leg gap gene function in *Parhyale* and insects aligns only if insects incorporated the red leg segment into the body wall (0). Color bars correspond to remaining leg segments following knockout, transparent bars indicate deleted leg segments. Open bar in dac indicates slight extension of dac function into tarsus 1 of insects. Coxal plate (Cp), gill (G), tergal plate (Tp). Scale bar 50um.

leg segments. This also means that at least part of the insect body wall is homologous to the crustacean coxa.

The data thus far is agnostic regarding the origin of the insect wing. However, we noted that *Parhyale* has what appears to be an epipod, the tergal plate, emerging proximal to the coxa. Clark-Hachtel (accompanying manuscript) show that the tergal plate, coxal plate, and basal plate all require the same “wing” genes, indicating that all three are epipods. They also show that nubbin, a marker of arthropod leg joints, is expressed in a distinct stripe above the *Parhyale* tergal plate, suggesting there is a leg segment here. An examination of the crustacean appendage morphology literature in the context of recent phylogenies shows that most crustaceans in fact have an additional proximal leg segment, the precoxa (Fig. 3a), and that the presence of a precoxa is the ancestral state. Although a precoxa has not been previously documented in amphipods, a careful examination using confocal and bright field microscopy reveals that *Parhyale* has a structure between the coxa and body wall that meets the criteria for a leg



segment: it protrudes from the body wall; it forms a true, muscled joint; and it extends musculature to another leg segment (Figs. 3 and S5)[12,33,34]. Furthermore, the tergal plate emerges not from the body wall, but from this precoxa (Fig. 3e). Thus, much of what appears to be lateral body wall in *Parhyale* is in fact proximal leg.

Since insects evolved from crustaceans, if the insect coxa is homologous to the crustacean basis, then one would expect to find two leg segments incorporated into the insect body wall, each equipped with an epipod (Fig. 4). As predicted, two leg-like segments can be observed proximal to the coxa in basal hexapods[2] including collembolans[35], as well as in the embryos of many insects[9,36,37]. In insect embryos, these two leg-like segments flatten out before hatching to form the lateral body wall[2,3,9,35-38](Fig 1c). Furthermore, insects indeed have two epipods proximal to the insect coxa. When "wing" genes are depleted in insects via RNAi, two distinct regions are affected: the wing, but also the protruding plate adjacent to the leg

< Fig. 3. *Parhyale* has a precoxa. (a) Phylogeny based on Oakley 2012, precoxa references in supplements. (b) Confocal of Parhyale hatchling. Round T5 tergal plate and pointy T6 tergal plate (dashed outlines). (c) Confocal of *Parhyale* hatchling, cuticle in cyan, muscle in red. Note the blocks of simple, anterior-posterior muscles of the body vs the orthogonal, complexly arranged muscles of the leg segments. Outline of tergal plates (dashed line) relative to orthogonal muscle. (d) BF image of right half of adult *Parhyale*, sagittal dissection, innards removed, lateral view. Wire used to position sample (w). The same orthogonal muscles in b are visible as striations that continue above the wire. The precoxa forms a joint with the coxa, including a gliding articulation (arrow). The dorsal limit of the precoxa is unclear, but the most conservative estimate is to begin at the gliding joint (arrow) and follow the leg up to where it meets the adjacent leg, denoted by (<). By comparing (<) and (→), it can be seen that the precoxa protrudes quite a bit from the body wall. However, the precoxa appears to continue farther up the body wall (compare orthogonal muscle striations). (e) Posterior-lateral view of right T6, looking edge-on at tergal plate. The tergal plate (dotted outline) emerges from the precoxa (contiguous pink between ←, >, and ---). In c, d, coxa is red (coxal plate not shaded, to focus on joints), gills (teal) partially cut for visibility, basis is orange, precoxa is pink. Note that all three plates (tergal, coxal, and basal) form contiguous cuticle with their leg segment, i.e. there is no distinguishing suture.

(Fig. 1c) [39-42]. These data are explained if insects incorporated the ancestral precoxa and crustacean coxa into the body wall, with the precoxa epipod later forming the wing and the crustacean coxa epipod later forming the plate.

The results presented here may settle a long-standing debate concerning the origin of insect wings as derived from (a) the epipod of the leg, (b) the body wall, or, more recently, (c) from both (dual-origin hypothesis; see Clark-Hachtel, accompanying manuscript)[6]. Our model accounts for all observations in favor of either the body wall or epipod origin of insect wing evolution, including the dorsal position of insect wings relative to their legs, the loss of ancestral leg segments in insects, the two-segmented morphology of the insect subcoxa in both embryos and adults, the complex musculature for flight, and the shared gene expression between wings and epipods. The realization that crustaceans have a precoxa accounts for the apparent “dual origin” of insect wings: much of what appears to be insect body wall is in fact the crustacean precoxa.

In fact, a number leg-associated outgrowths in arthropods are explained by this model, in addition to insect wings. The Daphnia carapace[43] is the epipod of the precoxa{Hansen:1925tba}; the Oncopeltus small plate outgrowth (Fig. 1c) is the epipod of the crustacean coxa; and the thoracic stylus of jumping bristletails (Fig. 4, st) is the epipod of the crustacean basis[10,44]. This also explains many insect abdominal appendages, like gills[45], gin traps[40], prolegs[46], and sepsid fly appendages[47], which are often proposed as de novo structures[48-50]. However, most insects form abdominal appendages as embryos[45,51], some even with an epipod nub, but these fuse to the body wall before hatching to form the sternites[37]. This is supported by a re-analysis of the expression of *Sp6-9* and its paralog, *buttonhead*, in insect embryos[22]. According to the leg segment homology model presented here (Fig. 4), the paired dots of btd expression in each abdominal segment of insect embryos demonstrates that these appendages are comprised of a minimum of three leg segments: the precoxa (pink), crustacean coxa (red), and insect coxa (orange). Thus, rather than de novo co-options, abdominal appendages were always there,

persisting in a truncated, highly modified state, and de-repressed in various lineages to form apparently novel structures. This provides a model for how insect wings can be both homologous to the epipod of the crustacean precoxa, and yet not be continuously present in the fossil record: epipod fields may persist in a truncated state, perhaps only visible as a nub in the embryo. We propose this as a general mechanism for the origin of novel structures that appear to be derived from serial homologs, rather than co-option.

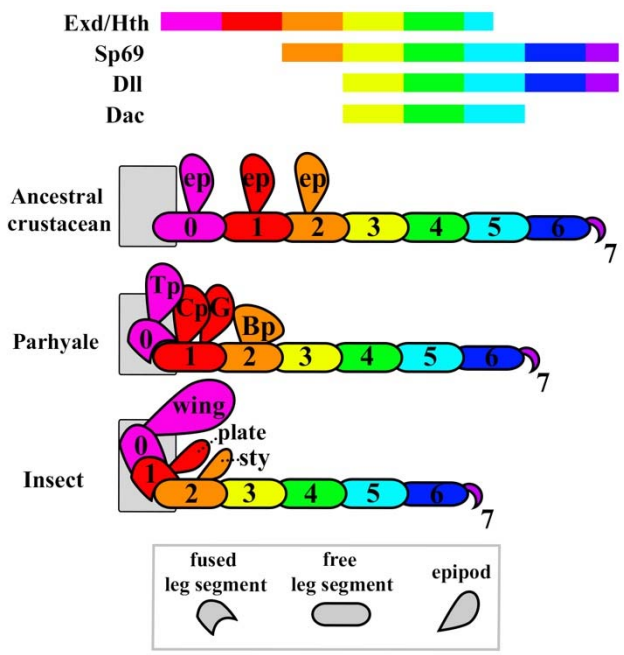
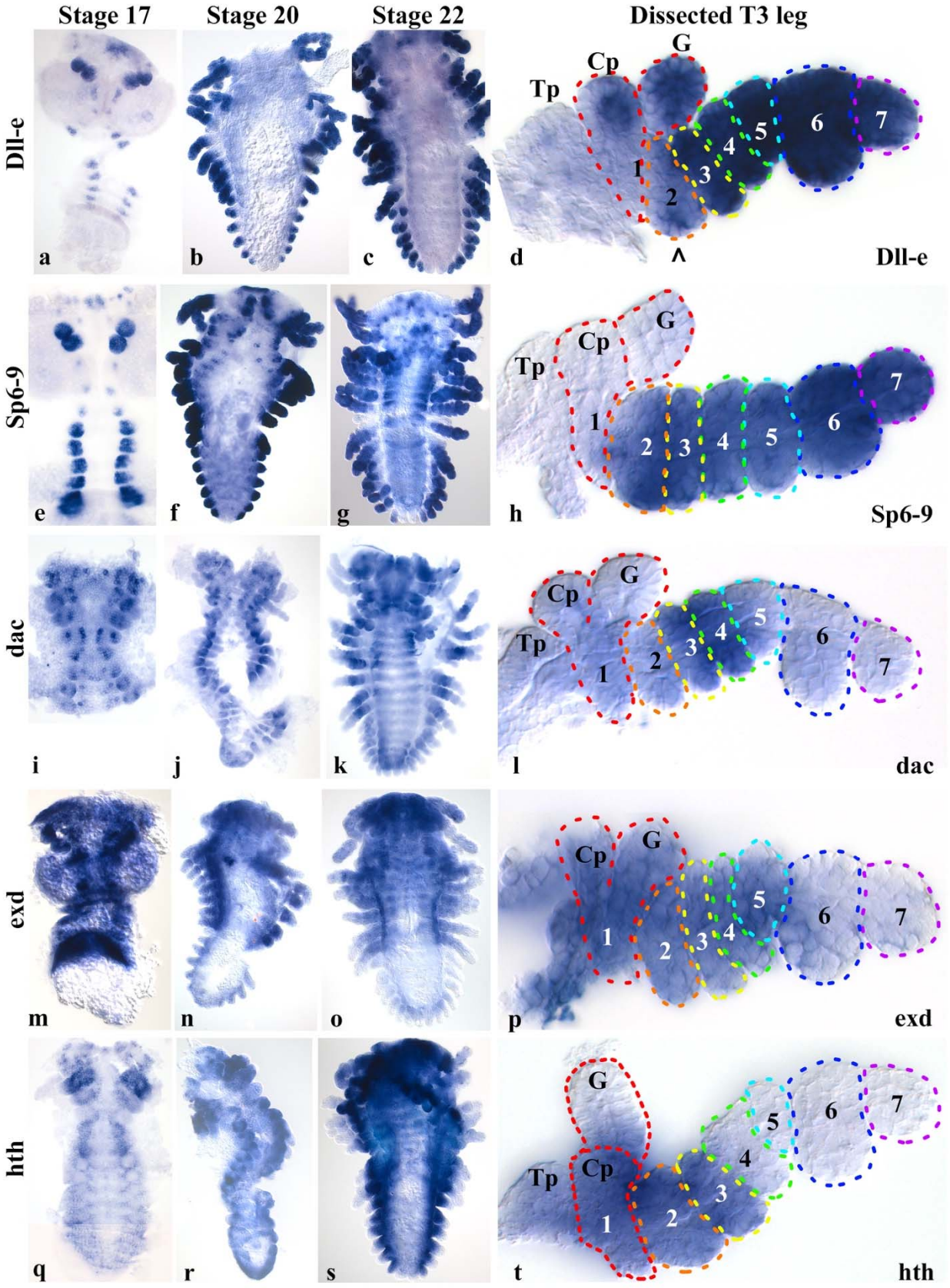


Fig 4. Gene expression alignment and proposed leg segment homologies (colors) between an ancestral crustacean, *Parhyale*, and insects. Ancestral precoxa epipod (ep), *Parhyale* tergal plate (Tp), and insect wing are homologous (pink). Ancestral coxa epipod, *Parhyale* coxal plate (Cp) and gill (G), and insect plate (see Fig. 1c) are homologous (red). Ancestral basis epipod, *Parhyale* basal plate (Bp), and jumping bristletail stylus (sty) are homologous (orange).



^ Fig S1. 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[18-20], *Sp6-9*[22], and *exd* and *hth*[30] have been previously characterized, but not at the level of individual leg segments. (d) *Dll*-e is expressed in leg segments 3 – 7; in the interior of the tergal plate (Tp), coxal plate (Cp), and gill (G), where it may be playing a sensory role, similar to the expression of Dll that patterns sensory hairs in the Drosophila wing margin[15]; and marks the bristle (^) of leg segment 2. This bristle is deleted in *Dll*-e KO (compare Fig. 2a, b). (h) *Sp6-9* is expressed in leg segments 2 – 7. (l) dac2 is expressed in leg segments 3 – 5. Expression in segment 5 may be stronger at other time points. (p) *exd* is expressed in the body wall through leg segment 5, and perhaps a little in 6. *Exd* is not expressed in the gill (not visible here). (t) *hth* is expressed in the body wall through leg segment 3. *Hth* is not expressed in the gill. Note that both insects and *Parhyale* share a peculiar disparity between *hth* expression and function, wherein *hth* knockout deletes one more leg segment than would be predicted by the *hth* expression domain.

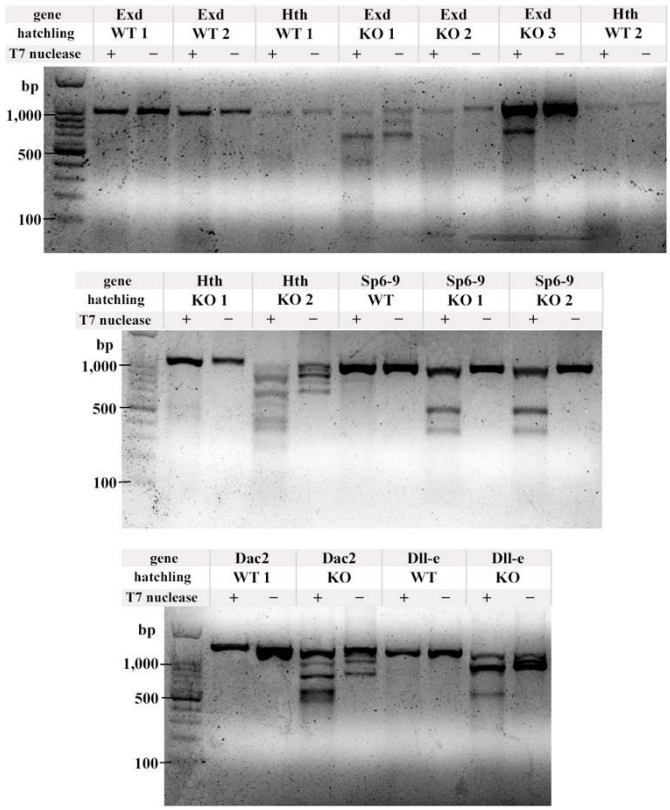


Fig. S2. T7 endonuclease assay to confirm CRISPR-Cas9 mutagenesis. For each gene, one or two wild type (WT) hatchlings were assayed, and one, two, or three KO hatchlings were assayed. T7 endonuclease was either added (+) or not added (–) to the heteroduplex mixture. In brief, a ~1kb region flanking the CRISPR-Cas9 target site by at least 300bp 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 1kb (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).

Gene	sgRNA	total injected	# dead	death %	# hatch w/phenotype	% phenotype of hatched
Dll-e	1+2	151	45	30%	57	54%
exd	1+2	206	90	44%	86	74%
exd	1	204	102	50%	84	82%
exd	2	173	36	21%	85	62%
hth	1+2	124	71	57%	32	60%
hth	1	131	30	23%	36	36%
hth	2	99	62	63%	22	59%
dac2	1+2	80	28	35%	41	79%
dac2	1	84	31	37%	9	17%
dac2	2	88	18	20%	16	23%
Sp6-9	1+2	165	88	53%	51	66%
Sp6-9	1	54	22	41%	9	28%
Sp6-9	2	37	3	8%	15	44%

Table 1. CRISPR-Cas9 injection numbers. Two sgRNAs per gene were made, and either one or both were injected as indicated. Both guides for each gene gave the same phenotype. # dead is the number of embryos that did not survive to hatching. For each gene, sgRNA 1 and 2 produced the same phenotypes.

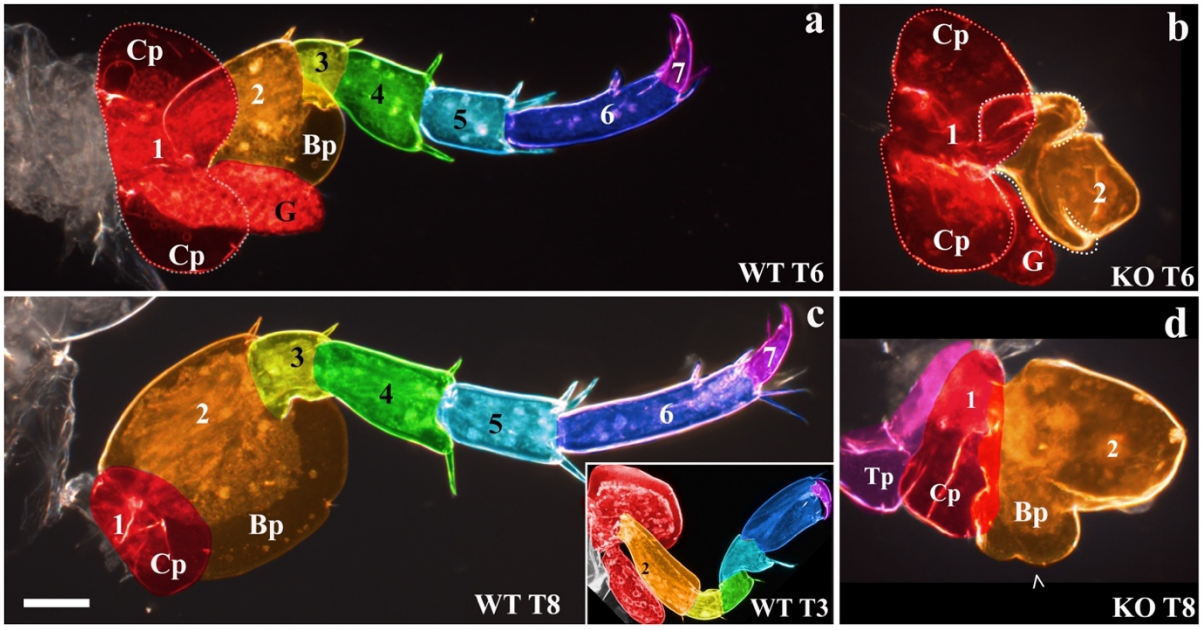


Fig S3. *Sp6-9* knockout sometimes causes a homeotic transformation of orange leg segment 2 towards a red leg segment 1 identity in jumping legs (thoracic legs T6 – 8). In WT jumping legs (a, c), orange leg segment 2 is very large and wide, due to the epipod on this segment (compare to skinny orange leg segment of WT T3 leg, inset in c). In WT T6 legs (a), the red coxal plate is bilobed, while in the WT T8 legs (c), the coxal plate is small and oval. In T6 *Sp6-9* KO (b), the epipod of orange leg segment 2 is bilobed, indicating a transformation towards red leg segment 1. In T8 *Sp6-9* KO (d), the large epipod of orange leg segment 2 has been reduced to the size and shape of the coxal plate, indicating a transformation towards red leg segment 1. Note that the tergal plates are unaffected (d, pink, Tp), which is similar to *Drosophila Sp6-9* knockouts, where the wings are unaffected[23]. The bilobed shape of the transformed T6 basal plate demonstrates that these are transformations towards a coxal plate rather than tergal plate, because the tergal plates are never bilobed. Therefore, these represent a homeotic transformation of one leg segment into another. This argues that the transformation of *Drosophila* leg to wing following loss of *Sp6-9* is also a transformation of one leg segment into another, and thus that insect wings are appendicular. Scale bar 50um.

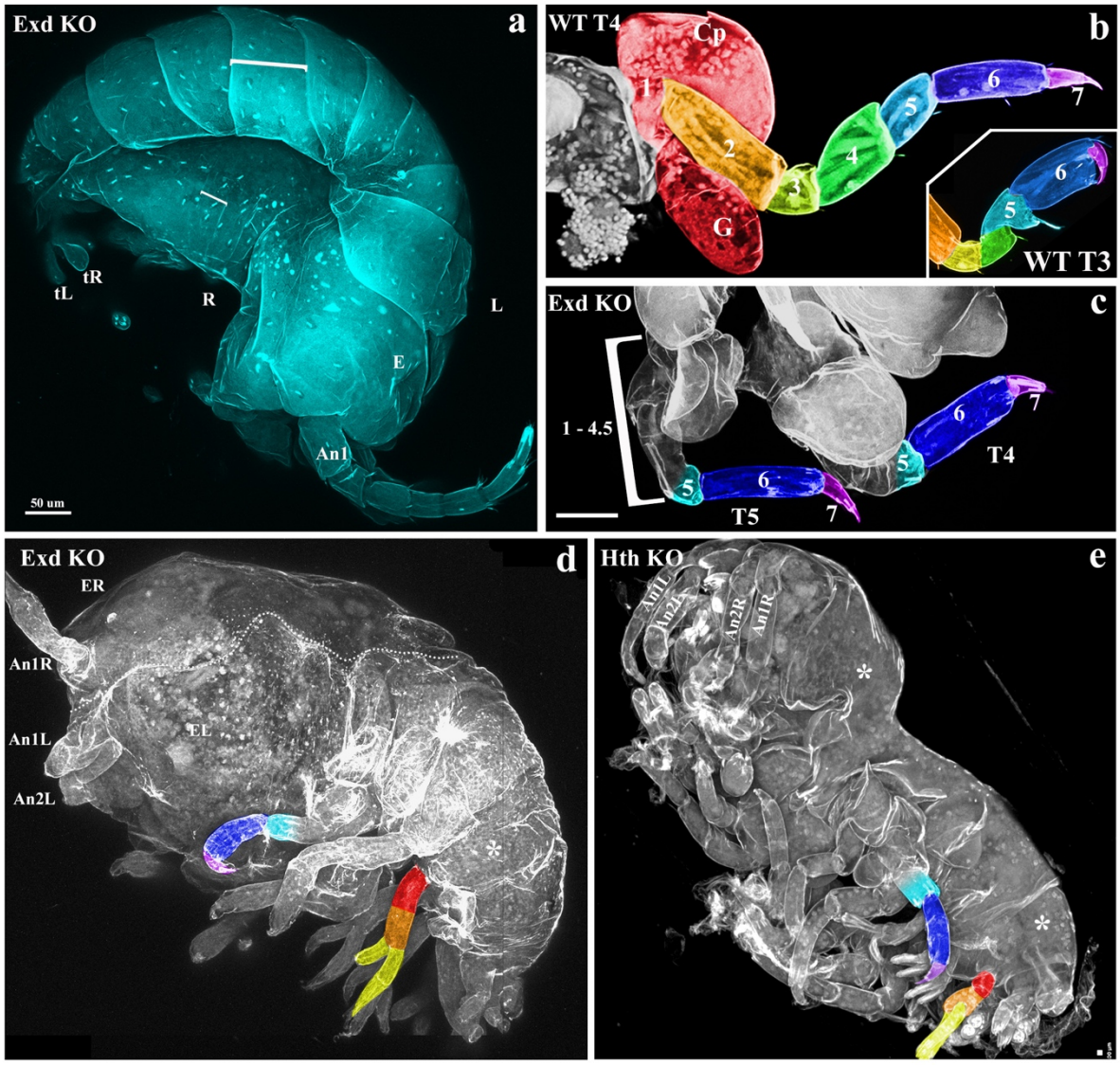


Fig. S4. *Exd* and *hth* phenotypes continued. (a) Body segment fusions/deletions in *exd* knockout whole hatchling. Confocal of unilaterally affected hatchling, dorsal view, anterior at bottom, posterior at left. Left side of animal (L) appears WT. The foreshortening of only the right (R) half of the body results in hatchlings with bodies twisted 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. Inset, WT T3 leg. Note broad shape of WT T3 blue leg segment 6 to skinny shape in WT T4/5. Also note triangle shape of WT T3 cyan leg segment 5 vs cylinder shape in WT T4/5, and presence of bristle in T3. (c) *exd* KO T4 and T5 legs. Loss of *exd* deletes/fuses leg segments 1 – 4 and proximal 5, leaving the distal half of leg segment 5 (indicated by fading cyan), and all of leg segments 6 and 7. Note that the joint between leg segments 5 and 6 is normal, but there is no apparent joint on the proximal side of leg segment 5. *Exd* KO also transforms the remaining T3 leg segments towards a T4/5 identity: *exd* KO T3 blue leg segment 6 is skinny, and cyan leg segment 5 is cylindrical and lacks the bristle (see Fig 2f). (d) Lateral view of *exd* KO hatchling. Hatchling died before cuticle growth. Dorsal midline indicated

with dashed white line. Left and right positions of eye in WT animals (EL, ER). (e) Lateral view of *hth* KO hatchling. *Exd* and *hth* KO produce the same body segment deletions/fusions, indicated with (*), compare to WT body segments in a, Left side, and in Figs. 1A and 3B. Neither *exd* nor *hth* KO appears to affect abdominal legs, because all abdominal proximal leg segments (red and orange) are intact in the same severely affected hatchlings where all thoracic proximal leg segments are deleted/fused, leaving only the distal thoracic leg segments (cyan, blue, purple). Lack of phenotype in abdominal legs is not due to knockout mosaicism: *exd* and *hth* are indeed knocked out in the abdomens of these hatchlings, because the body segments of the abdomen are fused together (*). Antenna (An).

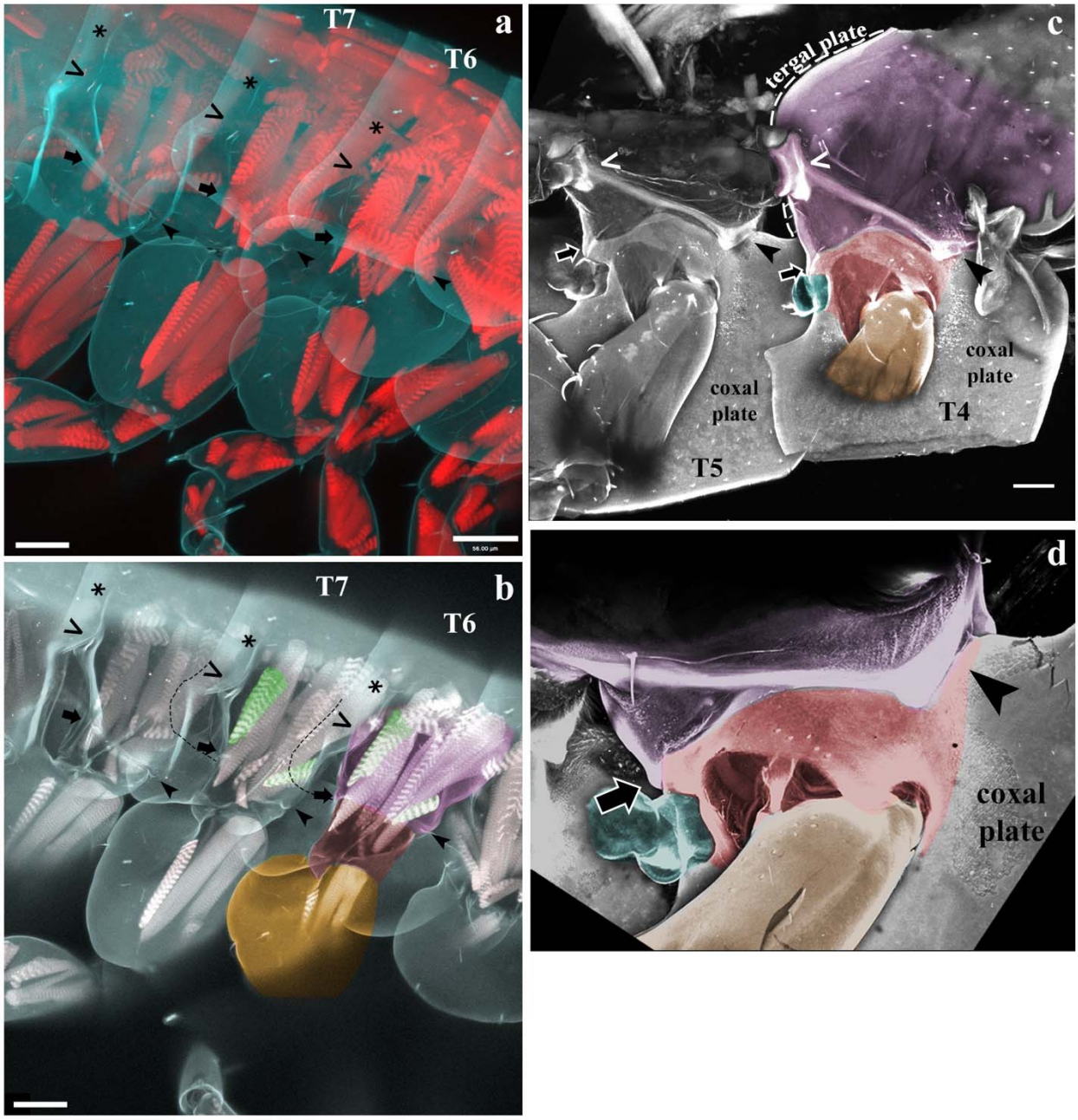


Fig. S5. *Parhyale* precoxa forms a true, muscled joint and extends musculature to another leg segment. Confocal images. (a) Phalloidin stain of muscle in right half of *Parhyale* hatchling. Contrast simple, anterior-posterior body muscles to orthogonal, complexly arranged leg muscles. No muscles cross the coxa-basis joint, as noted by Boxshall 1998. Note that all three plates (tergal, coxal, and basal) form contiguous cuticle with their leg segment, i.e. there is no distinguishing suture. (b) Optical section showing superficial muscles of right half. Confocal colors are partially desaturated: cuticle in grey-blue, muscle in grey-pink. The precoxa forms two articulations with the coxa: an anterior, bifurcated, load-bearing hinge articulation (arrowhead), and a posterior gliding articulation (→) (see also Fig. 3e). Coxa is red (coxal plate not shaded, to focus on joints), basis is orange, precoxa is magenta pink. Adjacent legs meet on their ventral sides at (<) and on their dorsal sides at (*). Outline of tergal plate (dashed line) relative to muscle

and joints shows that tergal plate emerges from precoxa. 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)[12,33,34]. 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[12,33,34]. (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.

METHODS

BIOINFORMATICS

Partial or complete sequences for *Parhyale Dll*, *Sp6-9*, *Exd*, and *Hth* have been previously identified. These were >99% identical at the nucleotide level to sequences in the *Parhyale* assembled transcriptome. In order to confirm their orthology, identify potential *Parhyale* paralogs and identify *Parhyale dac*, we ran reciprocal best Blast hit searches. For each gene, orthologs from several arthropods and vertebrates were downloaded from NCBI and EMBL and aligned against the *Parhyale* transcriptome[52] 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 orthologs from several arthropods were the best hits to our *Parhyale* candidates (i.e. were each other's reciprocal best Blast hits). These reciprocal best Blast hits are listed in the tables below, and were deposited in Genbank under Accession Numbers MG457799 - MG457804.

No *Parhyale* buttonhead/Sp5 was recovered in the assembled transcriptome. Buttonhead/Sp5 was also not found in the genome of the related amphipod Hyalella azteca. The assembled transcriptome only recovered fragments of *Parhyale* Sp1-4, so the previously sequenced *Parhyale* Sp1-4 (CBH30980.1) was used for the table below (asterisk).

Parhyale has three Dll paralogs, which appear to be an amphipod-specific duplication, because a related amphipod, Hyalella azteca, also has these same three Dll paralogs. The three *Parhyale* Dll paralogs had the lowest E-values to all Dll orthologs examined, but which of the three *Parhyale* Dll paralogs had the lowest E-value was variable, as expected for a clade-specific duplication.

The coding region for *Parhyale* exd and hth in the assembled transcriptome are longer than those previously identified. 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.

Extradenticle		
Query id	Subject id	E-value
Daphnia_pulex exd EFX62563.1	*Parhyale* exd MG457802	8.00E-177
Drosophila exd AAF48555.1	*Parhyale* exd MG457802	7.00E-173
Hyalella exd XP_018011298.1	*Parhyale* exd MG457802	2.00E-166
Parhyale exd CAO98909.1	*Parhyale* exd MG457802	6.00E-126
Tribolium exd NP_001034501.1	*Parhyale* exd MG457802	1.00E-173
Homo Pbx1 NP_002576.1	*Parhyale* exd MG457802	3.00E-166

Homothorax		
Query id	Subject id	E-value
Daphnia hth EFX75948.1	*Parhyale* hth MG457803	0
Drosophila hth NP_476578.3	*Parhyale* hth MG457803	6.00E-179
Homo Meis2 AAH07202.1	*Parhyale* hth MG457803	1.00E-148

Hyalella hth XP_018016731.1	*Parhyale* hth MG457803	0
Parhyale hth CAO98908.1	*Parhyale* hth MG457803	0
Tribolium hth NP_001034489.1	*Parhyale* hth MG457803	0

276
277

Sp6-9, Sp1-4, buttonhead/Sp5		
Query id	Subject id	E-value
Drosophila btd NP_511100.1	*Parhyale* Sp6-9 MG457804	4.00E-47
Drosophila Sp1-4 NM_142975.3	* *Parhyale* Sp1-4 CBH30980.1	5.00E-62
Drosophila Sp6-9 NP_727360.1	*Parhyale* Sp6-9 MG457804	6.00E-109
Homo Sp4 NP_003103.2	* *Parhyale* Sp1-4 CBH30980.1	2.00E-66
Homo Sp5 NP_001003845.1	*Parhyale* Sp6-9 MG457804	7.00E-62
Homo Sp8 NP_874359.2	*Parhyale* Sp6-9 MG457804	3.00E-105
Hyalella Sp1-4 XP_018012207.1	* *Parhyale* Sp1-4 CBH30980.1	0
Hyalella Sp6-9 XP_018014881.1	*Parhyale* Sp6-9 MG457804	0
Parhyale Sp1-4 CBH30980.1	* *Parhyale* Sp1-4 CBH30980.1	0
Parhyale Sp6-9 CBH30981.1	*Parhyale* Sp6-9 MG457804	0
Tribolium btd NP_001107792.1	*Parhyale* Sp6-9 MG457804	7.00E-59
Tribolium Sp1-4 XP_015833716.1	*Parhyale* Sp6-9 MG457804	3.00E-62
Tribolium Sp6-9 XP_008198341.1	*Parhyale* Sp6-9 MG457804	6.00E-159

278
279

Distalless		
Query id	Subject id	E-value
Drosophila Dll ACL83212.1	PhDllL2	2.00E-54
Drosophila Dll ACL83212.1	PhDllL1	2.00E-48
Drosophila Dll ACL83212.1	PhDlle MG457801	4.00E-42
Homo DLX-2 AAB40902.1	PhDlle MG457801	3.00E-35
Homo DLX-2 AAB40902.1	PhDllL2	6.00E-35
Homo DLX-2 AAB40902.1	PhDllL1	3.00E-34
Hyalella DLX-2 XP_018023955.1	PhDlle MG457801	0
Hyalella DLX-2 XP_018023955.1	PhDllL1	1.00E-49
Hyalella DLX-2 XP_018023955.1	PhDllL2	3.00E-45
Hyalella DLX-6 XP_018023956.1	PhDllL2	4.00E-102
Hyalella DLX-6 XP_018023956.1	PhDllL1	1.00E-51
Hyalella DLX-6 XP_018023956.1	PhDlle MG457801	1.00E-40
Hyalella unchar. protein XP_018023484.1	PhDllL1	8.00E-83
Hyalella unchar. protein XP_018023484.1	PhDllL2	0.89
Parhyale Dll-e ACT78885.1	PhDlle MG457801	0

Parhyale Dll-e ACT78885.1	PhDllL1	7.00E-48
Parhyale Dll-e ACT78885.1	PhDllL2	1.00E-44
Tribolium Dll AAG39634.1	PhDllL1	7.00E-48
Tribolium Dll AAG39634.1	PhDllL2	1.00E-46
Tribolium Dll AAG39634.1	PhDlle MG457801	5.00E-39

Dachshund		
Query id	Subject id	E-value
Daphnia pulex dac EFX90187.1	*Parhyale* Dac1 MG457799	3.00E-67
Drosophila dac AAF53538.3	*Parhyale* Dac2 MG457800	2.00E-64
Homo dach2 Q96NX9	*Parhyale* Dac1 MG457799	4.00E-52
Hyalella Dac1 XP_018011787.1	*Parhyale* Dac1 MG457799	7.00E-109
Hyalella Dac1 XP_018011787.1	*Parhyale* Dac2 MG457800	2.00E-55
Hyalella Dac2 XP_018011801.1	*Parhyale* Dac2 MG457800	0
Hyalella Dac2 XP_018011801.1	*Parhyale* Dac1 MG457799	1.00E-59
Tribolium dac1 XP_015834662.1	*Parhyale* Dac2 MG457800	6.00E-72

IN SITU PRIMER SEQUENCES

Primer name	product size	seq
hth FORWARD	941	GTTATGGGCTCCGTACCTGA
hth REVERSE	941	GCCAGCTGTTTCTTCTGGTC
exd FORWARD	734	AGCGAGTCCTCAACAAAGGA
exd REVERSE	734	AGGAGGCGTGTGCTATTCTG
Dll FORWARD	725	TGGGTCCAGTTCAACCTCTC
Dll REVERSE	725	GACATCGTCCTCCAAAGCAT
dac 1 FORWARD	638	GGAGAGCAGAGGGGACTTTT
dac 1 REVERSE	638	CCACTTCACGACCTCCTCAT
dac 2 FORWARD	699	CTTCAACCCCCTCCAGTACA
dac 2 REVERSE	699	TGTCTGTCGTCGTCTTCCTG
Sp6-9 FORWARD	789	CAAATGGCTCGCATGTATTG
Sp6-9 REVERSE	789	CAGTGCGTTCAAACTTCCAA

CLONING AND RNA PROBE SYNTHESIS

Total RNA was extracted from a large pool of *Parhyale* embryos at multiple stages of embryogenesis, from Stages 12 to 26 using Trizol. cDNA was generated using Superscript III. Primers were generated with Primer3 (http://bioinfo.ut.ee/primer3-0.4.0), with a preferred product size of 700bp, and did not include the DNA binding domain. Inserts were amplified with Platinum Taq (ThermoFisher 10966026), ligated into pGem T-Easy vectors (ProMega A1360), and transformed into E coli. The resulting plasmids were cleaned with a QiaPrep mini-prep kit (Qiagen A1360), and sequenced to verify the correct insert and determine sense and anti-sense promoters. In situ templates were generated by PCR from these plasmids using M13F/R primers and purified with Qiagen PCR Purification kit (Qiagen 28104). The resulting PCR products were used to make DIG-labeled 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 probes were the correct size, and concentration was determined using a Nanodrop 10000. Probes were used at 1-5ng/uL concentration.

IN SITU PROTOCOL

Embryo collection, fixation, and dissection as previously described[53]. In situ performed as previously described[54]. In brief, embryos were fixed in 4% paraformaldehyde (PFA) in artificial seawater for 45 minutes, dehydrated to methanol, and stored overnight at -20C to discourage embryos from floating in later hybridization solution (Hyb) step. Embryos were rehydrated to 1xPBS with 0.1% Tween 20 (PTw), post-fixed for 30 minutes in 9:1 PTw:PFA, and washed in PTw. Embryos were incubated in Hyb at 55C for at least 36 hours. Embryos were blocked with 5% normal goat serum and 1x Roche blocking reagent (Roche 11096176001) in PTw for 30 minutes. Sheep anti-DIG-AP antibody (Roche 11093274910) was added at 1:2000 and incubated for 2 hours at room temperature. Embryos were developed in BM Purple (Roche 11442074001) for a few hours to overnight. After embryos were sufficiently developed, they were dehydrated to methanol to remove any pink background, then rehydrated to PTw. Embryos were then moved to 1:1 PBS:glycerol with 0.1mg/mL DAPI, then 70% glycerol in PBS.

CRISPR-CAS9 GUIDE RNA GENERATION, INJECTION, AND IMAGING

Guide RNAs were generated using ZiFit[55,56] as previously described[57]. sgRNAs were ordered from Synthego. Injection mixes had a final concentration of 333ng/uL Cas9 protein, 150ng/uL 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 picoliters of sgRNA mixture as previously described[57]. Resulting knockout hatchlings were fixed in 4% paraformaldehyde in artificial seawater at 4C for 1 – 2 days, then moved to 70% glycerol in 1xPBS. Dissected hatchling limbs were visualized with Zeiss 700 and 780 confocal microscopes using the autofluorescence in the DAPI channel. Z-stacks were assembled with Volocity. Hatchling images were desaturated, levels adjusted, and false-colored using Overlay with Adobe Photoshop CS6.

T7 ENDONUCLEASE I ASSAY

Genomic primers were designed using Primer3, and flanked the target site by at least 400bp to either side. DNA isolation and subsequent PCR amplification of the region of interest was modified from previously described protocols[58]. Genomic DNA was amplified directly from

fixed hatchlings in 70% glycerol using ExTaq (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 NEB protocols (NEB M0302L). The digested products were run out on a 1.5% agarose gel. Genomic primers used for the T7 endonuclease I assay are listed below.

GENOMIC DNA PRIMERS

Primer name	product size	seq
exd left	907	CTTGAGATTCGTTCAGGTGCA
exd right	907	TTCTCCCCAGTTCCTTGCAA
hth left	943	TGTTCGTGTACCCGCAGAT
hth right	943	TCGGGCATACTAGAAGGCAG
Sp6-9 left	935	GCCCAGCTACTAACGATTTTCA
Sp6-9 right	935	GATCCGCTTCCTGACAGTTG
Dll-e left	922	GGAATGGTGAAGGAAGAGCG
Dll-e right	922	TCAGCAGTGCAGACTCATGT
dac2 left	983	CACGCGACACTCATACACAG
dac2 right	983	GATGCTCCTCCCACCGAATA

PRECOXA PHYLOGENY REFERENCES

Branchiura[59-61]. Mystacocarida[62,63]. Ostracoda[62,64-66]. Copepoda[62,67,68]. Cirripedia[66]. Decapoda[62,69,70]. Leptostraca[66]. Stomatopod[59,71]. Amphipoda[72,73]. Cephalocarida[59,74]. Notostraca[66]. Spinicaudata[75]. Remipedia[35,62]. Collembola[35]. Insecta[2,3,12,36,37,76].

AUTHOR CONTRIBUTIONS

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

References

1. Matsuda, R. Morphology and Evolution of the Insect Thorax. *https://doi.org/10.4039/entm10276fv* **102,** 1–427 (1970).
2. Snodgrass, R. E. *Morphology and mechanism of the insect thorax*. **80(1),** (1927).
3. DEUVE, T. The epipleural field in hexapods. *Annales de la Société entomologique de France* (2001).
4. Kukalová-Peck, J. Origin of the insect wing and wing articulation from the arthropodan

leg. *Can. J. Zool.* **61,** 1618–1669 (1983).

5. Snodgrass, R. E. *Principles of insect morphology.* (McGRAW-HILL BOOK COMPANY, INC., 1935).

6. Clark-Hachtel, C. M. & Tomoyasu, Y. Exploring the origin of insect wings from an evo-devo perspective. *Curr Opin Insect Sci* **13,** 77–85 (2016).

7. Requena, D. *et al.* Origins and Specification of the Drosophila Wing. *Current Biology* 1–16 (2017). doi:10.1016/j.cub.2017.11.023

8. Prokop, J. *et al.* Paleozoic Nymphal Wing Pads Support Dual Model of Insect Wing Origins. *Curr. Biol.* **27,** 263–269 (2017).

9. Mashimo, Y. & Machida, R. Embryological evidence substantiates the subcoxal theory on the origin of pleuron in insects. *Sci Rep* **7,** 1–9 (2017).

10. Niwa, N. *et al.* Evolutionary origin of the insect wing via integration of two developmental modules. *Evol. Dev.* **12,** 168–176 (2010).

11. Averof, M. & Cohen, S. M. Evolutionary origin of insect wings from ancestral gills. *Nature* **385,** 627–630 (1997).

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

13. Estella, C. A dynamic network of morphogens and transcription factors patterns the fly leg. *Curr. Top. Dev. Biol.* **98,** 173–198 (2012).

14. Estella, C., Rieckhof, G., Calleja, M. & Morata, G. The role of buttonhead and Sp1 in the development of the ventral imaginal discs of Drosophila. *Development* **130,** 5929–5941 (2003).

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

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

17. 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).

18. Serano, J. M. *et al.* Comprehensive analysis of Hox gene expression in the amphipod crustacean Parhyale hawaiensis. *Developmental Biology* **409,** 297–309 (2015).

19. Browne, W. E., Price, A. L., Gerberding, M. & Patel, N. H. Stages of embryonic development in the amphipod crustacean,Parhyale hawaiensis. *genesis* **42,** 124–149 (2005).

20. Liubicich, D. M. *et al.* Knockdown of Parhyale Ultrabithorax recapitulates evolutionary changes in crustacean appendage morphology. *Proc. Natl. Acad. Sci. U.S.A.* **106,** 13892–13896 (2009).

21. 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).

22. 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).

23. 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).

24. 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).
25. Tavsanli, B. C. *et al.* Structure–function analysis of the Drosophila retinal determination protein Dachshund. *Developmental Biology* **272,** 231–247 (2004).
26. Mito, T. *et al.* Divergent and conserved roles of extradenticle in body segmentation and appendage formation, respectively, in the cricket Gryllus bimaculatus. *Developmental Biology* **313,** 67–79 (2008).
27. Ronco, M. *et al.* Antenna and all gnathal appendages are similarly transformed by homothorax knock-down in the cricket Gryllus bimaculatus. *Developmental Biology* **313,** 80–92 (2008).
28. Rauskolb, C., Smith, K. M., Peifer, M. & Wieschaus, E. extradenticle determines segmental identities throughout Drosophila development. *Development* **121,** 3663–3673 (1995).
29. 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).
30. Prpic, N.-M. & Telford, M. J. Expression of homothorax and extradenticle mRNA in the legs of the crustacean Parhyale hawaiensis: evidence for a reversal of gene expression regulation in the pancrustacean lineage. *Dev Genes Evol* **218,** 333–339 (2008).
31. Rieckhof, G. E., Casares, F., Ryoo, H. D., Abu-Shaar, M. & Mann, R. S. Nuclear translocation of extradenticle requires homothorax, which encodes an extradenticle-related homeodomain protein. *Cell* **91,** 171–183 (1997).
32. Casares, F. & Mann, R. S. The ground state of the ventral appendage in Drosophila. *Science* **293,** 1477–1480 (2001).
33. Boxshall, G. in 241–267 (Springer Berlin Heidelberg, 2013). doi:10.1007/978-3-642-36160-9_11
34. Shultz, J. W. Morphology of locomotor appendages in Arachnida: evolutionary trends and phylogenetic implications. *Zool J Linn Soc* **97:1-56,** (1989).
35. Bäcker, H., Fanenbruck, M. & Wägele, J. W. A forgotten homology supporting the monophyly of Tracheata: The subcoxa of insects and myriapods re-visited. *Zoologischer Anzeiger - A Journal of Comparative Zoology* **247,** 185–207 (2008).
36. KOBAYASHI, Y. Formation of Subcoxae-1 and 2 in Insect Embryos: The Subcoxal Theory Revisited. *Proc Arthropod Embryol Soc Jpn* (2017).
37. 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).
38. Hansen, H. J. *Studies on Arthropoda*. (1930).
39. 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. U.S.A.* **110,** 16951–16956 (2013).
40. Ohde, T., Yaginuma, T. & Niimi, T. Insect morphological diversification through the modification of wing serial homologs. *Science* **340,** 495–498 (2013).

41. Medved, V. *et al.* Origin and diversification of wings: Insights from a neopteran insect. *Proceedings of the …* (2015).
42. Wang, D. *et al.* spalt is functionally conserved in Locusta and Drosophila to promote wing growth. *Sci Rep* **7,** 1–9 (2017).
43. 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. doi:10.1101/160010
44. Bitsch, J. The controversial origin of the abdominal appendage-like processes in immature insects: Are they true segmental appendages or secondary outgrowths? (Arthropoda hexapoda). *J. Morphol.* **273,** 919–931 (2012).
45. Komatsu, S. & Kobayashi, Y. Embryonic development of a whirligig beetle, Dineutus mellyi, with special reference to external morphology (insecta: Coleoptera, Gyrinidae). *J. Morphol.* **273,** 541–560 (2012).
46. Suzuki, Y., Squires, D. C. & Riddiford, L. M. Developmental Biology. *Developmental Biology* **326,** 60–67 (2009).
47. Bowsher, J. H. & Nijhout, H. F. Partial co-option of the appendage patterning pathway in the development of abdominal appendages in the sepsid fly Themira biloba. *Dev Genes Evol* **219,** 577–587 (2010).
48. Hoch, H. *et al.* Non-sexual abdominal appendages in adult insects challenge a 300 million year old bauplan. *Curr. Biol.* **24,** R16–7 (2014).
49. Angelini, D. R. & Kaufman, T. C. Comparative developmental genetics and the evolution of arthropod body plans. *Annu. Rev. Genet.* **39,** 95–119 (2005).
50. Moczek, A. P. On the origins of novelty in development and evolution. *Bioessays* **30,** 432–447 (2008).
51. Matsuda, R. *Morphology and Evolution of the Insect Abdomen*. (Elsevier, 1976).
52. Kao, D. *et al.* The genome of Parhyale hawaiensis: a model for animal development, regeneration, immunity and ligno-cellulose digestion. *Elife* 1–76 (2016).
53. Rehm, E. J., Hannibal, R. L., Chaw, R. C., Vargas-Vila, M. A. & Patel, N. H. Fixation and Dissection of Parhyale hawaiensis Embryos. *Cold Spring Harbor Protocols* **2009,** pdb.prot5127–pdb.prot5127 (2009).
54. 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 hawaiensis Embryos. *Cold Spring Harbor Protocols* **2009,** pdb.prot5130–pdb.prot5130 (2009).
55. Sander, J. D. *et al.* ZiFiT (Zinc Finger Targeter): an updated zinc finger engineering tool. *Nucleic Acids Res.* **38,** W462–W468 (2010).
56. 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).
57. 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).
58. 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).
59. Schram, F. R. *Crustacea*. (Oxford University Press, USA, 1986).
60. AGUIAR, J. C. *et al.* A new species of Argulus (Crustacea, Branchiura, Argulidae) from the skin of catfish, with new records of branchiurans from wild fish in the Brazilian

Pantanal wetland. 1–23 (1999). doi:10.11646/zootaxa.4320.3.3

61. TANZOLA, R. D., of, M. V.-O. P.-A. J.2017. Argulus ventanensis sp. n.(Crustacea, Branchiura) parasite of Hypsiboas pulchellus tadpoles (Anura, Hylidae). *panamjas.org*

62. Boxshall, G. Comparative limb morphology in major crustacean groups: the coxa-basis joint in postmandibular limbs. *Arthropod Relationships* (1998).

63. Stachowitsch, M. & Proidl, S. *The Invertebrates*. (University of Texas Press, 1992).

64. Horne, D. J. Homology and homoeomorphy in ostracod limbs. *Hydrobiologia* **538,** 55–80 (2005).

65. Cohen, A. C., Martin, J. W. & Kornicker, L. S. Homology of Holocene ostracode biramous appendages with those of other crustaceans: the protopod, epipod, exopod and endopod. *Lethaia* **31:251-265,** (1998).

66. Hansen, H. J. *Studies on Arthropoda. II. At the Expense of the Rask-Ørsted Fund*. (1925).

67. Karanovic, T. Two new genera and three new species of subterranean cyclopoids (Crustacea, Copepoda) from New Zealand, with redescription of Goniocyclops silvestris Harding, 1958. *Contributions to Zoology* **74, 3/4,** 223–254 (2005).

68. FIERS, F. & JOCQUE, M. Leaf litter copepods from a cloud forest mountain top in Honduras (Copepoda: Cyclopidae, Canthocamptidae). *Zootaxa* **3630,** 270–22 (2013).

69. Borradaile, L. A. XXVII.— Notes upon Crustacean limbs. *Journal of Natural History Series 11* **17,** 193–213 (1926).

70. Snodgrass, R. E. *Textbook of arthropod anatomy*. (1952).

71. Watling, Les, Hof, C. H. J. & Schram, F. R. THE PLACE OF THE HOPLOCARIDA IN THE MALACOSTRACAN PANTHEON. *Journal of Crustacean Biology* **20,** 1–11 (2000).

72. Ungerer, P. & Wolff, C. External morphology of limb development in the amphipod Orchestia cavimana (Crustacea, Malacostraca, Peracarida). *Zoomorphology* **124,** 89–99 (2005).

73. Wolff, C. & Scholtz, G. Cell lineage analysis of the mandibular segment of the amphipod *Orchestia cavimana* reveals that the crustacean paragnaths are sternal outgrowths and not limbs. *Front. Zool.* **3,** 19 (2006).

74. Olesen, J., Haug, J. T., Maas, A. & Waloszek, D. External morphology of Lightiella monniotae (Crustacea, Cephalocarida) in the light of Cambrian ‘Orsten’ crustaceans. *Arthropod Structure and Development* **40,** 449–478 (2011).

75. Ferrari, F. D. & GRYGIER, M. J. Comparative morphology among trunk limbs of Caenestheriella gifuensis and Leptestheria kawachiensis (Crustacea: Branchiopoda: Spinicaudata). *Zool J Linn Soc* **139,** 547–564 (2003).

76. Coulcher, J. F., Edgecombe, G. D. & Telford, M. J. Molecular developmental evidence for a subcoxal origin of pleurites in insects and identity of the subcoxain the gnathal appendages. *Sci Rep* **5,** 1–8 (2015).