

Hox gene duplications correlate with posterior heteronomy in scorpions

Prashant P. Sharma, Evelyn E. Schwager, Cassandra G. Extavour and Ward C. Wheeler

Proc. R. Soc. B 2014 **281**, 20140661, published 13 August 2014

Supplementary data

"Data Supplement"

<http://rsb.royalsocietypublishing.org/content/suppl/2014/08/12/rsb.2014.0661.DC1.html>

References

[This article cites 44 articles, 11 of which can be accessed free](#)

<http://rsb.royalsocietypublishing.org/content/281/1792/20140661.full.html#ref-list-1>

Subject collections

Articles on similar topics can be found in the following collections

[evolution](#) (1856 articles)

Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click [here](#)



Research

Cite this article: Sharma PP, Schwager EE, Extavour CG, Wheeler WC. 2014 Hox gene duplications correlate with posterior heteronomy in scorpions. *Proc. R. Soc. B* **281**: 20140661.
<http://dx.doi.org/10.1098/rspb.2014.0661>

Received: 19 March 2014

Accepted: 22 July 2014

Subject Areas:
evolution

Keywords:

Scorpiones, opisthosoma, Arthropoda, bauplan, tagmosis

Author for correspondence:

Prashant P. Sharma
e-mail: psharma@amnh.org

Electronic supplementary material is available at <http://dx.doi.org/10.1098/rspb.2014.0661> or via <http://rsb.royalsocietypublishing.org>.



Royal Society Publishing

Hox gene duplications correlate with posterior heteronomy in scorpions

Prashant P. Sharma¹, Evelyn E. Schwager², Cassandra G. Extavour²
and Ward C. Wheeler¹

¹Division of Invertebrate Zoology, American Museum of Natural History, Central Park West at 79th Street, New York, NY 10024, USA

²Department of Organismic and Evolutionary Biology, Harvard University, 16 Divinity Avenue, Cambridge, MA 02138, USA

The evolutionary success of the largest animal phylum, Arthropoda, has been attributed to tagmatization, the coordinated evolution of adjacent metameres to form morphologically and functionally distinct segmental regions called tagmata. Specification of regional identity is regulated by the Hox genes, of which 10 are inferred to be present in the ancestor of arthropods. With six different posterior segmental identities divided into two tagmata, the bauplan of scorpions is the most heteronomous within Chelicerata. Expression domains of the anterior eight Hox genes are conserved in previously surveyed chelicerates, but it is unknown how Hox genes regionalize the three tagmata of scorpions. Here, we show that the scorpion *Centruroides sculpturatus* has two paralogues of all Hox genes except *Hox3*, suggesting cluster and/or whole genome duplication in this arachnid order. Embryonic anterior expression domain boundaries of each of the last four pairs of Hox genes (two paralogues each of *Antp*, *Ubx*, *abd-A* and *Abd-B*) are unique and distinguish segmental groups, such as pectines, book lungs and the characteristic tail, while maintaining spatial collinearity. These distinct expression domains suggest neofunctionalization of Hox gene paralogues subsequent to duplication. Our data reconcile previous understanding of Hox gene function across arthropods with the extreme heteronomy of scorpions.

1. Introduction

The evolutionary success of Arthropoda is attributed to their segmented bauplan and its modularization through tagmosis, whereby groups of adjacent segments evolve in concert to achieve morphological and functional distinction from other such groups along the anteroposterior (AP) axis [1]. A conserved cluster of transcription factors called the Hox genes play key roles in conferring segmental identity and have been implicated as the driving force in the evolution of tagmata [2–6]. Loss or gain of Hox gene function has been shown to cause homeotic transformations in multiple arthropod species [6–10]. Ten Hox genes are inferred to have been present in the common ancestor of arthropods and their sister group, Onychophora, and are typically closely genetically linked on a chromosome [4,11–14]. They are expressed collinearly along the AP axis, i.e. in the same order as they occur on the chromosome in the cluster [15,16], and spatial shifts in Hox gene expression are associated with morphological change in body plans [9,17–21].

Scorpions are an unusual and iconic lineage of arthropods, due to an ancient origin and marked heteronomy. This lineage is unique among arthropods in possessing a tagma dedicated exclusively to prey capture and defence: the flexible and photosensitive metasoma (tail) of scorpions harbours the venom glands and expresses one of the opsins shared with the eyes, conferring light signal transduction to this eyeless body region [22–25]. Like most Euchelicerata, the prosoma (anterior tagma) of scorpions consists of seven segments: the protocerebral, cheliceral, pedipalpal and four leg-bearing segments (but see [26] for a review of the segmental nature of the protocerebrum). Hox gene expression surveys of multiple chelicerate orders (spiders, mites and harvestmen) have demonstrated the

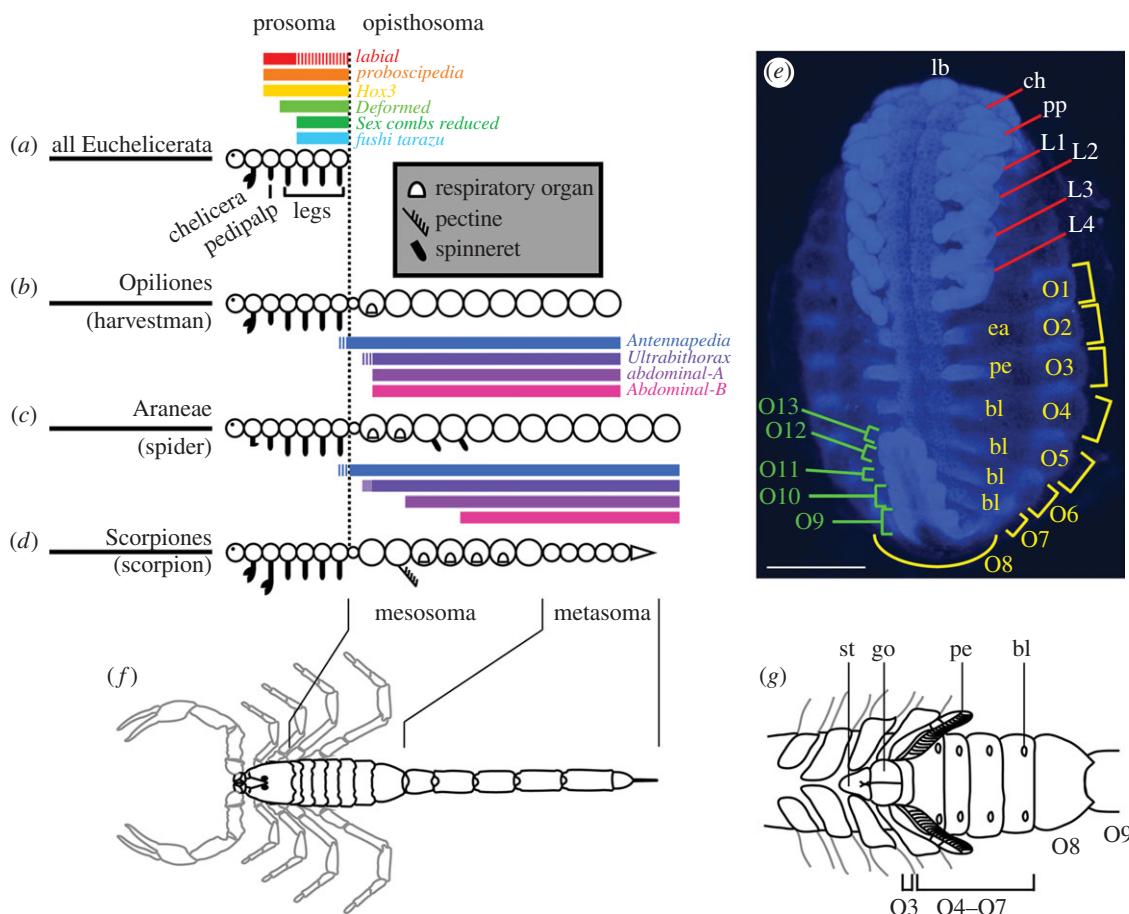


Figure 1. Four Hox genes are expressed in the chelicerate opisthosoma. (a) Summary of Hox gene expression domains in the chelicerate prosoma, based on data from spiders, mites and harvestmen. (b) Opisthosomal Hox expression in Opiliones is correlated with the homonomy of the harvestman posterior tagma. Note coincident anterior expression boundaries of *Ubx*, *abd-A* and *Abd-B*. (c) Differentiation of four distinct anterior opisthosomal segment types in spiders is correlated with staggered expression domains of the opisthosomal Hox group. (d) The bauplan of scorpions has six segmental identities divided into two tagmata, which cannot be reconciled with only four Hox genes in this region. (e) Limb bud stage embryo of *C. sculpturatus*, with segments of the prosoma, mesosoma and metasoma completely formed. (f) Generalized scorpion anatomy in dorsal view. (g) Generalized scorpion mesosomal anatomy in ventral view. bl, book lung; ch, chelicera; ea, embryonic appendage; go, genital operculum; L1, first walking leg; lb, labrum; O1, first opisthosomal segment; pe, pectine; pp, pedipalp; st, sternum. Scale bar in (e): 200 μ m. (Online version in colour.)

evolutionary conservation of Hox patterning in the prosoma. As in the anterior tagma of all arthropods surveyed, the patterning of the prosoma is associated with staggered anterior expression boundaries of the first six Hox genes, engendering a unique combination of Hox transcripts for each segmental identity (figure 1a) [4,27,28]. In chelicerates, few of these anterior six Hox genes are expressed posterior to the last prosomal segment, with transient expression of *proboscipedia* (*pb*), *Hox3*, *Deformed* (*Dfd*) and *Sex combs reduced* (*Scr*) observed in posterior segments of some, but not all, species. This leaves four Hox genes, *Antennapedia* (*Antp*), *Ultrabithorax* (*Ubx*), *abdominal-A* (*abd-A*) and *Abdominal-B* (*Abd-B*), to pattern the posterior tagma, the opisthosoma [4,28].

Functional studies of *Antp* in the spider *Parasteatoda tepidariorum* have demonstrated the role of this gene in enforcing the prosomal–opisthosomal boundary [10]. The earliest expression of *Antp* consistently occurs in the first opisthosomal (O1) segment of multiple chelicerates [10,12,28]. Additionally, a previous study illustrated the lability of *abd-A* and *Abd-B* anterior expression boundaries in some chelicerate orders. In harvestmen (Opiliones), the homonomous opisthosoma is correlated with coincident anterior boundaries of *Ubx*, *abd-A* and *Abd-B* in the second opisthosomal segment (figure 1b) [28]. In spiders, the differentiation of the opisthosoma into segments

forming the pedicel (O1), the respiratory organs (book lungs and/or tracheal tubules; O2–O3), spinnerets (silk-spinning organs; O4–O5) or no external organs at all (O6–O12), is in turn correlated with the staggered expression domains of *Ubx*, *abd-A* and *Abd-B*, which result in unique combinations of Hox transcripts in each segment type (figure 1c) [27,29].

Unlike all other Chelicerata, scorpions have a highly differentiated opisthosoma that contains six segmental identities conventionally grouped into two tagmata: the mesosoma ('abdomen' segments O1 to O8) and the metasoma (segments O9 to O13) (figure 1d). The mesosoma and the metasoma are functionally distinct tagmata, with the former harbouring major organ systems, and the latter dedicated to prey immobilization and defence through envenomation (figure 1f,g). In the mesosoma, O1 develops a transient pair of embryonic appendage rudiments that are resorbed before hatching, and the remainder of the segment becomes part of the genital chamber in adults. O2 bears the genital pores, O3 the pectines (sensory appendages unique to scorpions) and O4–O7 the book lungs. No organs form on O8, the tapering last segment of the mesosoma. O9–O13 narrow to form the characteristic metasoma, which bears the venom glands and the aculeus (stinger).

Among the arthropods, only insects, some crustaceans and scorpions bear three tagmata (reviewed by [4]). Yet it is

unknown how Hox genes regionalize the three tagmata in scorpions. Only the collective protein expression of *Ubx* and *abd-A* (termed ‘*UbdA*’) has been surveyed in the scorpion *Smeringurus mesaensis*, which shows an anterior expression boundary in part of the O2 segment, comparable to harvestmen [28,30]. Scorpion life-history characteristics, including live birth in all extant species and gestation periods lasting several months have heretofore precluded extensive developmental genetic study in this lineage [30–33]. Here, we investigated the genetic basis of posterior tagmosis in the bark scorpion *Centruroides sculpturatus*. We show that *C. sculpturatus* has two paralogues of every Hox gene class excepting *Hox3*, and thus possesses a total of 19 Hox genes. We show that the paralogues of the posterior class Hox genes each have distinct expression patterns that result in unique combinations of Hox transcripts for each segmental identity. The paralogues retain archetypal Hox gene spatial collinearity, reconciling the heteronomy of scorpions with established mechanisms of Hox gene regulation in Arthropoda.

2. Material and methods

Adult females of the scorpion *C. sculpturatus* were purchased from Hatari Invertebrates (AZ, USA). Females were anaesthetized with CO₂ and embryos dissected from the ovary following a published protocol [31]. RNA was extracted from one clutch of limb bud stage embryos using Trizol (Invitrogen) and first strand cDNA synthesis was performed using SuperScriptIII (Invitrogen). A developmental transcriptome of *C. sculpturatus* was generated by sequencing this cDNA in a single flowcell on an Illumina HiSeq 2500 platform subsequent to strand-specific library preparation with the Apollo 324TM system (IntegenX), using paired-end 150 bp-long reads at the Harvard University Bauer Core Facility. Terminal bases with a Phred quality score under 33 were trimmed and over-represented rRNA sequences were discarded. Assembly was conducted using the software TRINITY with a path reinforcement distance of 75 and enforcing strand-specific assembly, with a resulting estimated coverage of 23.9× [34].

Fragments of Hox genes were identified by BLAST. Phylogenetic analysis of Hox gene amino acid sequences was conducted as follows: amino acid sequences of chelicerate and mandibulate Hox orthologues with known expression patterns were aligned using MUSCLE v. 3.6 [35] and culled to 68 conserved, adjacent positions using GBLOCKS v. 0.91b [36]. Maximum-likelihood analysis was performed using RAxML v. 7.7.5 [37] and Bayesian inference analysis using MRBAYES v. 3.1.2 [38]. Details of heuristics and all multiple sequence alignments are provided in the electronic supplementary material.

Templates for riboprobe synthesis were generated by PCR-amplified gene-specific primers, and cloning amplicons using the TOPO TA Cloning Kit with One Shot Top10 chemically competent *Escherichia coli* (Invitrogen, Carlsbad, CA, USA), following the manufacturer’s protocol. Amplicon identities were verified by Sanger sequencing. Primer sequences for templates are provided in the electronic supplementary material, table S1. The developmental transcriptome is deposited in the NCBI Sequence Read Archive, accession no. SRR1515193.

Centruroides sculpturatus embryos were fixed and whole mount *in situ* hybridization performed following a published protocol [39]. Approximately 15–20 embryos were assayed for every parologue. For each pair of paralogues, clutches were divided and expression examined in equivalent stages. Embryos were mounted in glycerol and images were captured using an HrC AxioCam and an Axio Zoom V.16 fluorescence stereomicroscope driven by Zen (Zeiss).

3. Results

(a) Identification of scorpion Hox genes

Phylogenetic analysis of a 68 site conserved region of Hox gene amino acid sequences obtained from a *C. sculpturatus* transcriptome was conducted using maximum-likelihood and Bayesian inference analyses. Tree topologies obtained indicated exactly two copies of all Hox genes in *C. sculpturatus*, except for *Hox3*, for which a single copy was recovered (figure 2). Among chelicerates, duplicates of Hox genes have been previously reported for the spiders *Cupiennius salei* and *P. tepidariorum*, in the genes *labial* (*lab*), *pb*, *Dfd*, *Scr* and *Ubx*; and the horseshoe crab *Limulus polyphemus*, in which up to four putative copies occur of some Hox genes [11,40] (electronic supplementary material, table S2). To assess whether the scorpion and spider paralogues represent lineage-specific duplications, we constructed larger nucleotide sequence alignments for *lab*, *Dfd*, *Scr* and *Ubx* using spider, scorpion and harvestman sequences only. Consideration of only closely related chelicerate Hox sequences enabled retention of more sites for analysis upon masking of ambiguously aligned sites, as non-chelicerate sequences diverge greatly outside of the conserved region. We included paralogues of some Hox genes found in the genome of the scorpion *Mesobuthus martensii* [25] (electronic supplementary material, table S2). In no gene tree do scorpion and spider orthologues form parologue-specific clusters (figure 3). Instead, gene trees generated from these alignments showed that the spider and scorpions sequences frequently form clusters corresponding to their respective orders when rooted with orthologous harvestman sequences, supporting lineage-specific duplications in both spiders and scorpions (figure 3).

(b) Expression of *Cscu-Antp* paralogues

The anterior expression boundary of *Cscu-Antp-1* occurs in the L4 (fourth walking leg) segment, including in the limb buds (figure 4c). By contrast, the anterior expression boundary of *Cscu-Antp-2* occurs in the posterior part of the L4 segment, and it is absent from the L4 limb buds (figure 4d; electronic supplementary material, figure S1). In other chelicerates, the expression domain boundaries of the opisthosomal Hox group genes (*Antp*, *Ubx*, *abd-A* and *Abd-B*) extend to the posterior terminus of the embryo [10,28,29,40,41]. Consistent with this pattern, expression of *Cscu-Antp-1* and *Cscu-Antp-2* extends to the posterior (figure 4a,b). *Cscu-Antp-1* is expressed at higher levels in O1 and O2 than in O3 through to O13, with the strongest expression occurring in the O2 neural lobes. Expression of *Cscu-Antp-2* in posterior segments is more uniform than that of *Cscu-Antp-1*. Both genes are expressed in the neuroectoderm along the ventral midline.

(c) Expression of *Cscu-Ubx* paralogues

Distinct anterior expression boundaries of *Cscu-Ubx-1* and *Cscu-Ubx-2* occur in ventral portions of O1 and O2, respectively (figure 4e,f). In *Cscu-Ubx-1*, the anterior boundary comprises a single pair of small domains in the posterior of O1, whereas the anterior expression boundary of *Cscu-Ubx-2* consists of a pair of hemispherical domains in O2 that abut the O1–O2 boundary. *Cscu-Ubx-2* expression is stronger in the opisthosomal organs and ventral ectoderm than in the pleura, and both orthologues show similar levels of expression in posterior segments.

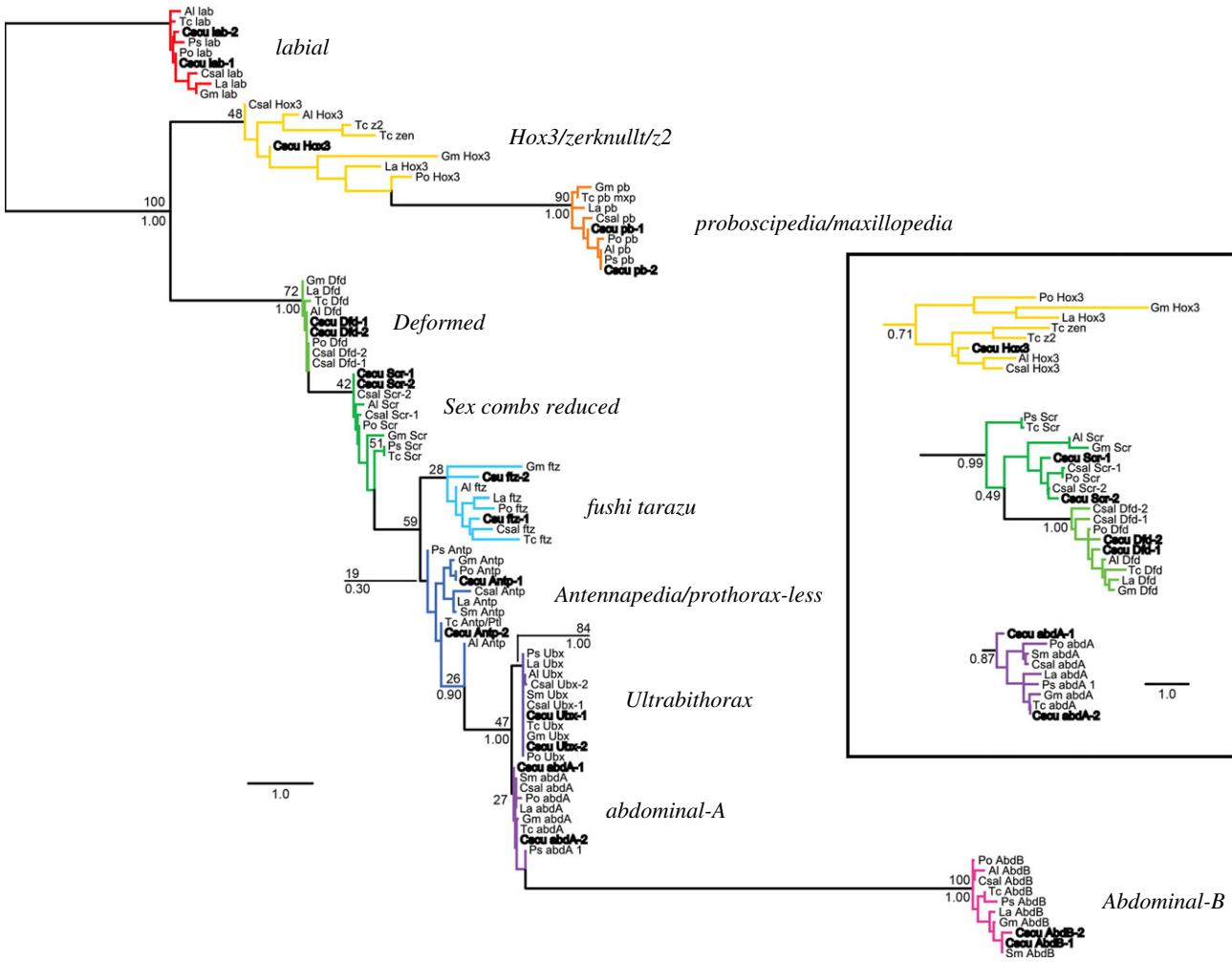


Figure 2. Tree topology inferred from maximum-likelihood analysis of conserved regions (68 amino acid characters), using selected arthropod taxa for which gene expression has been reported ($\ln L = -2074.37$). *Centruroides sculpturatus* sequences are indicated in bold. Colours correspond to Hox gene identities, as in figure 1. Inset: alternative topologies of *Hox3*, *Scr*, *Dfd* and *abd-A* subtrees based on Bayesian inference analyses. Values on nodes indicate bootstrap resampling frequencies (integers) and posterior probabilities (values ≤ 1.00). (Online version in colour.)

(d) Expression of *Cscu-abd-A* paralogues

Anterior expression boundaries of *Cscu-abd-A-1* and *Cscu-abd-A-2* occur in O3 and O4, respectively (figure 4g,h). Clear expression of *Cscu-abd-A-1* is observed in the developing pectines, the largest of the embryonic opisthosomal organs in this species. *Cscu-abd-A-2* is expressed in the first book lung primordia in O4. In posterior segments, expression level of *Cscu-abd-A-1* is higher in the metasoma than in the mesosoma. Both *Cscu-abd-A-1* and *Cscu-abd-A-2* are strongly expressed in the metasoma.

(e) Expression of *Cscu-Abd-B* paralogues

The strong anterior expression boundary of *Cscu-Abd-B-1* in the metasomal segments occurs in O8, the recurved mesosomal segment that occurs in the posterior of the scorpion embryo (but not at the terminus of the AP axis). By contrast, the strong anterior expression boundary of *Cscu-Abd-B-2* occurs in the posterior part of O9 (figure 4i,j). Restricted and unique expression domains of each *Abd-B* parologue are also observed in the ventral ectoderm of mesosomal segments, as follows. *Cscu-Abd-B-1* is additionally expressed in the ventral neuroectoderm of O3–O7, with expression restricted to the medial part of the neuroectoderm. By contrast, additional expression of *Cscu-Abd-B-2* occurs in the ventral ectoderm of O4–O7, but in domains that comprise most of the

neuroectoderm. Expression in the book lungs is observed in later stages (electronic supplementary material, figure S2). Both paralogues are uniformly expressed in the metasoma posterior to O9 (figure 4i,j). Expression of neither parologue was observed in the genital segment (only stages prior to formation of genital pores were examined in this study).

4. Discussion

The degree of tagmatization occurring in the scorpion opisthosoma is exceptional among chelicerates. Our gene expression data demonstrate that both paralogues of each of the last four Hox genes are expressed in a canonical Hox-like pattern in scorpions. The staggering of the anterior expression boundaries, resulting from the offset of the anterior expression boundaries by one or more segments, establishes a unique combination of Hox transcripts for each segmental identity (figure 5a).

Expression data in the spider and the harvestman implicate roles for *abd-A* and/or *Abd-B* in opisthosomal differentiation [28]. Apropos, offsets in anterior expression domains among *Cscu-abd-A* and *Cscu-Abd-B* paralogues appear to delimit the segmental identities of the scorpion pectine- and book lung-bearing segments (O3 and O4–O7, respectively). The weaker expression domains of the *Cscu-Abd-B* paralogues are similar to their counterparts in the spider *C. salei*, wherein *Abd-B* is

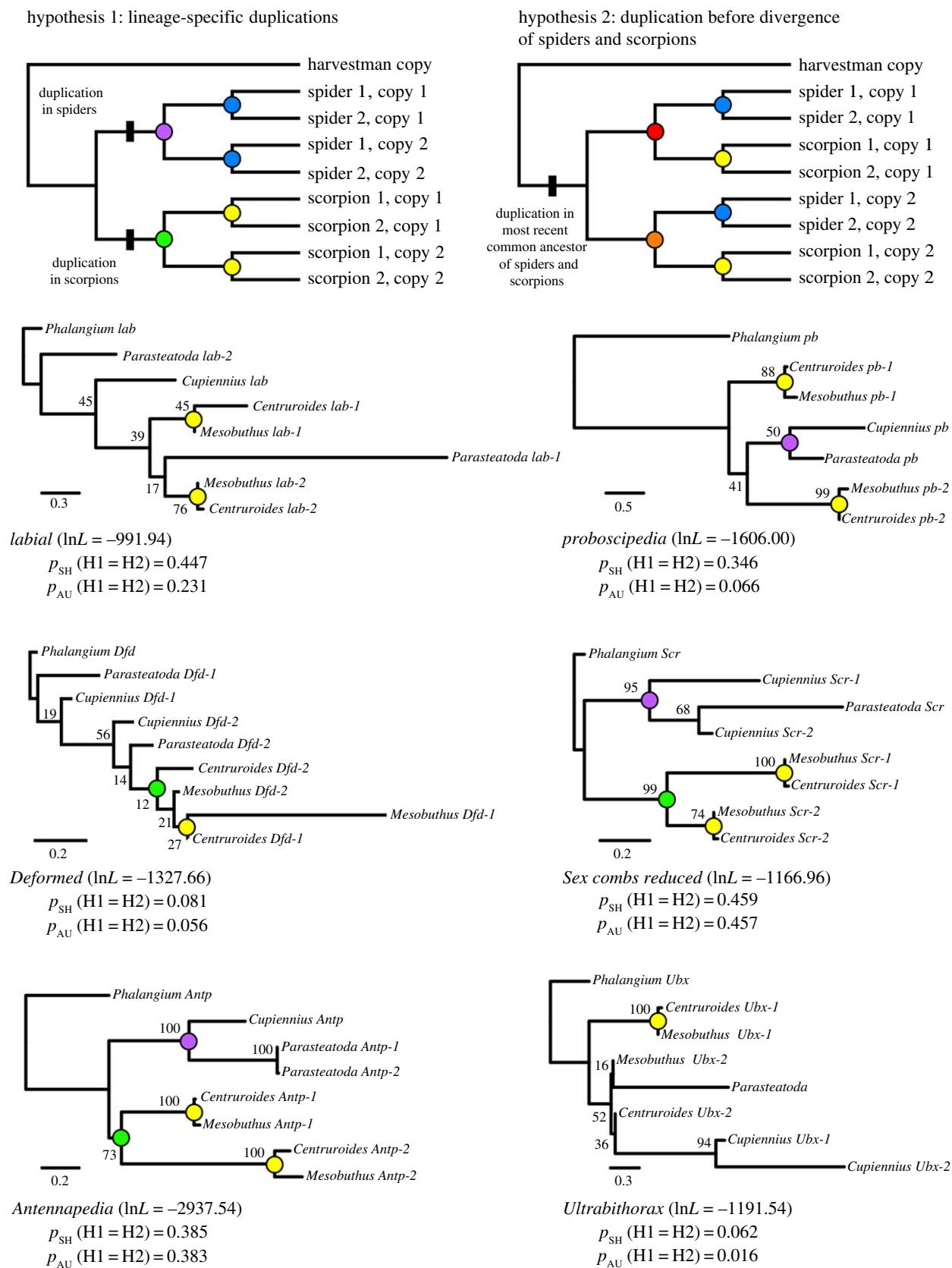


Figure 3. Chelicerate Hox gene trees inferred from maximum-likelihood analysis of nucleotide sequences, wherein multiple paralogues are present in scorpions, spiders or both. Hypothetical trees in the top row indicate alternative scenarios of lineage-specific duplications (left) and a single duplication in the common ancestor of spiders and scorpions (right). Coloured circles indicate nodes expected under a particular hypothesis. Pairs of p -values under each gene tree indicate results of Shimodaira-Hasegawa (SH) and approximately unbiased (AU) tests. Harvestman: *P. opilio*; spiders: *C. salei* and *P. tepidariorum*; scorpions: *C. sculpturatus* and *M. martensii*. (Online version in colour.)

strongly expressed throughout posterior segments (O5–O10), and weakly in O3–O4 [28]. The functional significance of the weaker *Abd-B* domains is not known in chelicerates [10,29].

Paralogues of all four scorpion opisthosomal Hox genes follow spatial collinearity. However, protein expression of

UbdA in another scorpion species (*S. mestaensis*) shows an anterior expression boundary similar to the one observed for *Cscu-Ubx-2* in early stages (i.e. in the ventral part of O2, with expression throughout O3 and posterior segments; compare figure 4f to fig. 2a of [30]). In late stages of *S. mestaensis*,

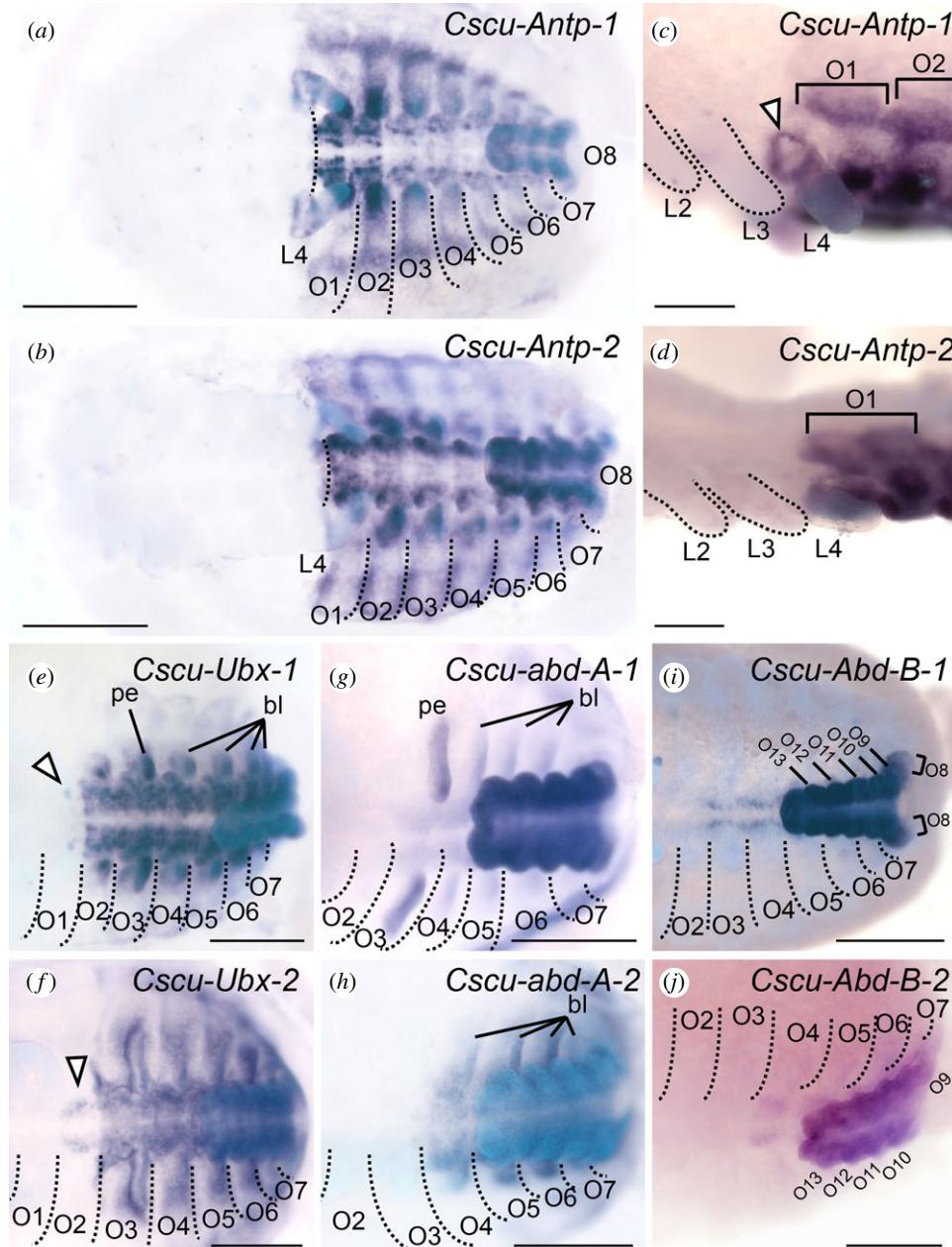


Figure 4. Staggered expression of opisthosomal Hox group paralogues yields unique Hox transcript combinations for every segmental identity. Embryos are shown in ventral view unless otherwise specified. (a,b) Expression of *Cscu-Antp-1* and *Cscu-Antp-2* spans L4 to posterior terminus. (c,d) Higher magnification lateral views of embryos in (a,b). Arrowheads indicate unique anterior expression domains of *Cscu-Antp-1* in the L4 limb bud, and *Cscu-Antp-2* in the posterior part of the L4 neural lobe only. (e,f) Expression of *Cscu-Ubx-1* and *Cscu-Ubx-2* shows anterior expression boundaries offset by one segment, i.e. posterior and ventral part of O1, and ventral part of O2, respectively. Arrowheads show expression in a more restricted area as compared with posterior segments. (g,h) Like the *Ubx* paralogues, the anterior expression boundaries of *Cscu-abd-A-1* and *Cscu-abd-A-2* are offset by a single segment, in O3 and O4, respectively. (i,j) *Cscu-Abd-B-1* and *Cscu-Abd-B-2* are strongly expressed starting from O8 (brackets in (i) show upturning O8 segment) and from O9, respectively. Weaker expression domains occur along the ventral midline in both paralogues, but in distinct tissues. pe, pectine; bl, book lung. Scale bars: 200 µm (c,d); 400 µm (e–j); and 500 µm (a,b). (Online version in colour.)

expression of UbdA is similar to that of *Cscu-Ubx-1* (i.e. with strong expression in O2 and posterior segments; compare figure 4e to fig. 2c of [30], but includes more tissue in the O1 segment. It is unclear whether these differences in expression correspond to differences in embryonic stage or differences between the two lineages.

In the harvestman *Phalangium opilio*, a chelicerate with putatively single orthologues of *Antp* and *Ubx*, the anterior expression boundary of both Hox genes shifts during development [28]. Early expression of *Antp* commences in O1, but expands into the L4 segment (including the limb bud) in later stages [28]; the same phenomenon occurs with spiders' *Antp* [10,27]. Curiously, in *C. sculpturatus*, the *Antp* paralogues'

expression domains resemble the two temporally distinct expression profiles of the single harvestman and spider *Antp* orthologues. Similarly, *Ubx* paralogues of both spiders and scorpions have offset anterior expression boundaries (although neither spider *Ubx* parologue is expressed in O1 [40]), whereas *Ubx* expression in the harvestman commences in the posterior part of O2 but shifts forward in later stages [28].

These data are suggestive of subfunctionalization of *Antp* in scorpions and of *Ubx* in both scorpions and spiders. A similar pattern is observed in spider and harvestman *Scr*; the harvestman *Scr* orthologue's expression domain resembles the union of expression domains of the two *Scr* paralogues of *C. salei* [28,40]. However, we add the caveat that gene silencing of

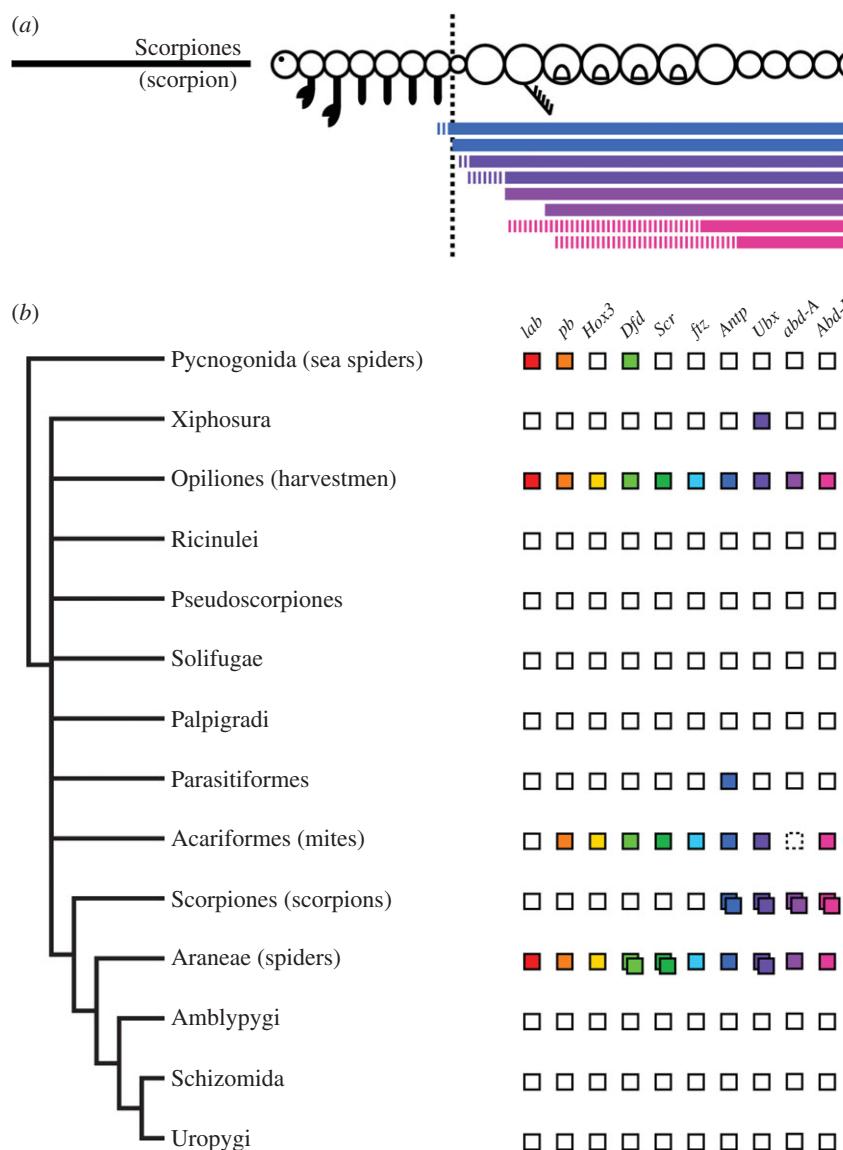


Figure 5. (a) Summary of opisthosomal Hox expression patterns in *C. sculpturatus*. (b) Summary of known Hox expression patterns in Chelicerata. Coloured squares indicate known expression domains. Dashed square for acariform *abd-A* indicates gene loss. Duplicated genes without verified expression data not indicated. Tree topology based on [42]. (Online version in colour.)

spider *Antp*, the only functional data available to date for chelicerate Hox genes, affects only the O1 segment, de-repressing appendage formation [10]. No *Antp* phenotype has been reported in L4, and it is therefore not clear what, if anything, is achieved by expansion of *Antp* expression into the L4 segment, which has now been reported in two spiders (*C. salei* and *P. tepidariorum*), a mite and a harvestman [12,27,28,40,43]. Furthermore, phylogenetic placements of scorpion and spider Hox paralogues, when these are retained and topologically informative (specifically, *Dfd*, *Scr* and *Antp*), do not support with statistical significance a scenario of lineage-specific duplications (figure 3). Likelihood-based tests of tree topology demonstrate that neither lineage-specific duplications nor a common duplication event in the most recent common ancestor of scorpions and spiders is the topologically superior scenario (at $\alpha = 0.01$) in any of the six test cases (figure 3 and table 1; electronic supplementary material). A lack of statistical significance in this case may be attributable to: (i) the paucity of sampled lineages, (ii) rate heterogeneity in one or more arachnid species, (iii) the degree of evolutionary conservation of the homeobox domain in multiple sequence alignments, or (iv) some combination of these systematic biases. In spite of

accruing genomic resources for arachnids, we were unable to discover duplicated scorpion Hox genes in published databases, as well as our own unpublished scorpion libraries (P.P. Sharma and W.C. Wheeler 2014, personal communication), particularly because no other embryonic transcriptomes have been generated for this group of chelicerates.

It therefore remains to be investigated how many other Hox paralogues have been retained in spider genomes, and what function the spider paralogues serve during embryogenesis. Similarly, the degree of retention of Hox gene paralogues within Scorpiones cannot presently be assessed because both *C. sculpturatus* and *M. martensii* are members of the same family of scorpions (Buthidae), and basal relationships of the extant scorpions are unresolved [44,45]. Given that both spiders and scorpions have a conserved prosoma, but a more differentiated opisthosoma, than other euhelicerates, it is possible that some prosomal Hox paralogues in spiders and scorpions have been lost or undergone subfunctionalization, whereas some opisthosomal paralogues have neofunctionalized. The recently sequenced genome of the scorpion *M. martensii* is reported to bear over 32 000 genes, more than any other arthropod genome, and demonstrably accelerated gene family turnover,

Table 1. Likelihood scores for Hox gene tree alignments with topologies unconstrained or constrained to a specific scenario (as in figure 3). (Difference in log likelihoods and standard deviation of this value are indicated for the comparison of hypothesis 1 (lineage-specific duplication) and hypothesis 2 (duplication in most recent common ancestor of spiders and scorpions).)

	InL (unconstrained)	InL (Hypothesis 1)	InL (Hypothesis 2)	Δ(lnL _{H1, H2})	σ (lnL)
<i>labial</i>	−991.94	−1090.34	−1094.68	4.332493	6.291582
<i>proboscipedia</i>	−1606.00	−1715.38	−1717.49	2.112443	2.312595
<i>Deformed</i>	−1327.66	−1447.39	−1463.84	16.450287	11.739177
<i>Sex combs reduced</i>	−1166.96	−1284.29	−1292.70	8.406873	4.858968
<i>Antennapedia</i>	−2937.54	−3157.80	−3158.49	0.694105	1.626799
<i>Ultrabithorax</i>	−1191.54	−1616.86	−1613.57	3.292049	4.592267

but a putative whole genome duplication in scorpions has not been assessed [25].

An alternative route to testing a common mechanism of Hox gene duplications in spiders and scorpions may be concerted expansion of available genomic resources for other arachnid orders. Specifically, the discovery of Hox gene paralogues in non-spider tetrapulmonates (Amblypygi, Schizomida and Thelyphonida) would empower statistical assessments of the likelihood of a common duplication mechanism. Moreover, such a discovery could lend credence to the recently proposed sister relationship between scorpions and Tetrapulmonata [42].

5. Conclusion

We propose that the duplication of scorpion Hox genes and putative neofunctionalization of paralogues of the opisthosomal

group facilitated the extreme heteronomy of scorpions, the only group of arthropods with a tagma dedicated exclusively to prey capture and defence. Our results suggest that gene duplications may have been a plausible mechanism for achieving morphological diversification in this ancient group of arthropods.

Acknowledgements. We are indebted to Philip Donoghue and two anonymous reviewers whose comments greatly improved an earlier draft of this manuscript. Douglas Richardson facilitated imaging at the Harvard Center for Biological Imaging. P.P.S. designed and performed the research, analysed data and wrote the paper. E.E.S. performed research and analysed data. C.G.E. contributed reagents. W.W. contributed analytical tools. All authors edited and approved the manuscript.

Funding statement. This work was partially supported by NSF grant no. IOS-1257217 to C.G.E. and AMNH funds to W.W. P.P.S. was supported by the National Science Foundation Postdoctoral Research Fellowship in Biology under grant no. DBI-1202751.

References

- Cisne JL. 1974 Evolution of the world fauna of aquatic free-living arthropods. *Evolution* **28**, 337–366. (doi:10.2307/2407157)
- Carroll SB. 1995 Homeotic genes and the evolution of arthropods and chordates. *Nature* **376**, 479–485. (doi:10.1038/376479a0)
- Popadić A, Abzhanov A, Rusch D, Kaufman TC. 1998 Understanding the genetic basis of morphological evolution: the role of homeotic genes in the diversification of the arthropod bauplan. *Int. J. Dev. Biol.* **42**, 453–461.
- Hughes CL, Kaufman TC. 2002 Hox genes and the evolution of the arthropod body plan. *Evol. Dev.* **4**, 459–499. (doi:10.1046/j.1525-142X.2002.02034.x)
- Lewis EB. 1978 A gene complex controlling segmentation in *Drosophila*. *Nature* **276**, 565–570. (doi:10.1038/276565a0)
- Denell RE, Hummels KR, Wakimoto BT, Kaufman TC. 1981 Developmental studies of lethality associated with the *Antennapedia* gene complex in *Drosophila melanogaster*. *Dev. Biol.* **81**, 43–50. (doi:10.1016/0012-1606(81)90346-8)
- Pultz MA, Diederich RJ, Cribbs DL, Kaufman TC. 1988 The *proboscipedia* locus of the *Antennapedia* complex: a molecular and genetic analysis. *Genes Dev.* **2**, 901–920. (doi:10.1101/gad.2.7.901)
- Angelini DR, Liu PZ, Hughes CL, Kaufman TC. 2005 Hox gene function and interaction in the milkweed bug *Oncopeltus fasciatus* (Hemiptera). *Dev. Biol.* **287**, 440–455. (doi:10.1016/j.ydbio.2005.08.010)
- Liubicich DM et al. 2009 Knockdown of *Parhyale Ultrabithorax* recapitulates evolutionary changes in crustacean appendage morphology. *Proc. Natl Acad. Sci. USA* **106**, 13 892–13 896. (doi:10.1073/pnas.0903105106)
- Khadjeh S et al. 2012 Divergent role of the Hox gene *Antennapedia* in spiders is responsible for the convergent evolution of abdominal limb repression. *Proc. Natl Acad. Sci. USA* **109**, 4921–4926. (doi:10.1073/pnas.1116421109)
- Cook CE et al. 2001 Hox genes and the phylogeny of the arthropods. *Curr. Biol.* **11**, 759–763. (doi:10.1016/S0960-9822(01)00222-6)
- Telford MJ, Thomas RH. 1998 Expression of homeobox genes shows chelicerate arthropods retain their deutocerebral segment. *Proc. Natl Acad. Sci. USA* **95**, 10 671–10 675. (doi:10.1073/pnas.95.18.10671)
- Telford MJ, Thomas RH. 1998 Of mites and *zen*: expression studies in a chelicerate arthropod confirm *zen* is a divergent Hox gene. *Dev. Genes Evol.* **208**, 591–594. (doi:10.1007/s004270050219)
- Janssen R, Eriksson BJ, Tait NN, Budd GE. 2014 Onychophoran Hox genes and the evolution of arthropod Hox gene expression. *Front. Zool.* **11**, 22. (doi:10.1186/1742-9994-11-22)
- Sánchez-Herrero E, Vernos I, Marco R, Morata G. 1985 Genetic organization of *Drosophila* bithorax complex. *Nature* **313**, 108–113. (doi:10.1038/313108a0)
- Kaufman TC, Seeger MA, Olsen G. 1990 Molecular and genetic organization of the *Antennapedia* gene complex of *Drosophila melanogaster*. *Adv. Gen.* **27**, 309–362. (doi:10.1016/S0065-2660(08)60029-2)
- González-Reyes A, Morata G. 1990 The developmental effect of overexpressing a *Ubx* product in *Drosophila* embryos is dependent on its interactions with other homeotic products. *Cell* **61**, 515–522. (doi:10.1016/0092-8674(90)90533-K)
- Averof M, Patel NH. 1997 Crustacean appendage evolution associated with changes in Hox gene expression. *Nature* **388**, 682–686. (doi:10.1038/41786)
- Abzhanov A, Kaufman TC. 2000 Crustacean (malacostracan) Hox genes and the evolution of the arthropod trunk. *Development* **127**, 2239–2249.
- Tomoyasu Y, Wheeler SR, Denell RE. 2005 *Ultrabithorax* is required for membranous wing

- identity in the beetle *Tribolium castaneum*. *Nature* **433**, 643–647. (doi:10.1038/nature03272)
21. Pavlopoulos A *et al.* 2009 Probing the evolution of appendage specialization by Hox gene misexpression in an emerging model crustacean. *Proc. Natl Acad. Sci. USA* **106**, 13 897–13 902. (doi:10.1073/pnas.0902804106)
 22. Zwicky KT. 1968 A light response in the tail of *Urodacus*, a scorpion. *Life Sci.* **7**, 257–262. (doi:10.1016/0024-3205(68)90020-9)
 23. Zwicky KT. 1970 The spectral sensitivity of the tail of *Urodacus*, a scorpion. *Experientia* **26**, 317. (doi:10.1007/BF01900120)
 24. Rao G, Rao KP. 1973 A metasomatic neuronal photoreceptor in the scorpion. *J. Exp. Biol.* **58**, 189–196.
 25. Cao Z *et al.* 2013 The genome of *Mesobuthus martensii* reveals a unique adaptation model of arthropods. *Nat. Commun.* **4**, 2602. (doi:10.1038/ncomms3602)
 26. Scholtz G, Edgecombe GD. 2006 The evolution of arthropod heads: reconciling morphological, developmental and palaeontological evidence. *Dev. Genes Evol.* **216**, 395–415. (doi:10.1007/s00427-006-0085-4)
 27. Damen WGM, Hausdorf M, Seyfarth E-A, Tautz D. 1998 A conserved mode of head segmentation in arthropods revealed by the expression pattern of Hox genes in a spider. *Proc. Natl Acad. Sci. USA* **95**, 10 665–10 670. (doi:10.1073/pnas.95.18.10665)
 28. Sharma PP, Schwager EE, Extavour CG, Giribet G. 2012 Hox gene expression in the harvestman *Phalangium opilio* reveals divergent patterning of the chelicerate opisthosoma. *Evol. Dev.* **14**, 450–463. (doi:10.1111/j.1525-142X.2012.00565.x)
 29. Damen WGM, Tautz D. 1999 *Abdominal-B* expression in a spider suggests a general role for *Abdominal-B* in specifying the genital structure. *J. Exp. Zool.* **285**, 85–91. (doi:10.1002/(SICI)1097-010X(19990415)285:1<85::AID-JEZ10>3.0.CO;2-N)
 30. Popadić A, Nagy L. 2001 Conservation and variation in *Ubx* expression among chelicerates. *Evol. Dev.* **3**, 391–396. (doi:10.1046/j.1525-142X.2001.01049.x)
 31. Farley RD. 2001 Development of segments and appendages in embryos of the desert scorpion *Paruroctonus mesaensis* (Scorpiones: Vaejovidae). *J. Morphol.* **250**, 70–88. (doi:10.1002/jmor.1060)
 32. Simonnet F, Deutsch J, Quéinnec E. 2004 *hedgehog* is a segment polarity gene in a crustacean and a chelicerate. *Dev. Genes Evol.* **214**, 537–545. (doi:10.1007/s00427-004-0435-z)
 33. Simonnet F, Célérier M-L, Quéinnec E. 2006 *Orthodenticle* and *empty spiracles* genes are expressed in a segmental pattern in chelicerates. *Dev. Genes Evol.* **216**, 467–480. (doi:10.1007/s00427-006-0093-4)
 34. Grabherr MG *et al.* 2011 Full-length transcriptome assembly from RNA-seq data without a reference genome. *Nat. Biotechnol.* **29**, 644–652. (doi:10.1038/nbt.1883)
 35. Edgar RC. 2004 MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797. (doi:10.1093/nar/gkh340)
 36. Castresana J. 2000 Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* **17**, 540–552. (doi:10.1093/oxfordjournals.molbev.a026334)
 37. Stamatakis A. 2006 RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**, 2688–2690. (doi:10.1093/bioinformatics/btl446)
 38. Huelsenbeck JP, Ronquist F. 2005 Bayesian analyses of molecular evolution using MrBayes. In *Statistical methods in molecular evolution* (ed. R Nielsen), pp. 1–52. New York, NY: Springer.
 39. Sharma PP, Gupta T, Schwager EE, Wheeler WC, Extavour CG. 2014 Subdivision of *cap-n-collar* domains is restricted to Mandibulata. *EvoDevo* **5**, 3. (doi:10.1186/2041-9139-5-3)
 40. Schwager EE, Schoppmeier M, Pechmann M, Damen WGM. 2007 Duplicated Hox genes in the spider *Cupiennius salei*. *Front. Zool.* **4**, 10. (doi:10.1186/1742-9994-4-10)
 41. Barnett AA, Thomas RH. 2013 Posterior Hox gene reduction in an arthropod: *Ultrabithorax* and *Abdominal-B* are expressed in a single segment in the mite *Archegozetes longisetosus*. *EvoDevo* **4**, 23. (doi:10.1186/2041-9139-4-23)
 42. Regier JC, Shultz JW, Zwicky A, Hussey A, Ball B, Wetzer R, Martin JW, Cunningham CW. 2010 Arthropod relationships revealed by phylogenomic analysis of nuclear protein-coding sequences. *Nature* **463**, 1079–1083. (doi:10.1038/nature08742)
 43. Schwager EE, Pechmann M, Feitosa NM, McGregor AP, Damen WGM. 2009 *hunchback* functions as a segmentation gene in the spider *Achaearanea tepidariorum*. *Curr. Biol.* **19**, 1333–1340. (doi:10.1016/j.cub.2009.06.061)
 44. Prendini L, Volschenk ES, Maaliki S, Gromov AV. 2006 A ‘living fossil’ from Central Asia: the morphology of *Pseudochactas ovchinnikovi* Gromov, 1998 (Scorpiones: Pseudochactidae), with comments on its phylogenetic position. *Zool. Anz.* **245**, 211–248. (doi:10.1016/j.jcz.2006.07.001)
 45. Soleglad ME, Fet V. 2003 High-level systematics and phylogeny of the extant scorpions (Scorpiones: Orthosterni). *Euscorpius* **11**, 1–57.

Electronic Supplementary Material

Methods

(a) Phylogenetic analysis of 68-site dataset

For the maximum likelihood search with rapid bootstrapping in RAxML v. 7.7.5 [1, 2], a WAG model of sequence evolution [3] with corrections for a discrete gamma distribution (WAG + Γ) [4] was specified, as selected by ProtTest v. 3 [5]. 500 independent searches were conducted and nodal support estimated with 500 bootstrap replicates. The alignment is provided as Dataset S1.

Bayesian inference analyses were conducted with MrBayes v.3.1.2 [7] using a mixed model of protein sequence evolution with corrections for a discrete gamma distribution. Two runs, each with four chains of default temperature distribution, were performed for 5×10^6 generations. Stationarity was assessed using Tracer v. 1.5 [8] and 1.25×10^6 generations were discarded as burnin.

(b) Phylogenetic analysis of individual gene trees

Chelicerate sequences of *labial*, *proboscipedia*, *Deformed*, *Sex combs reduced*, *Antennapedia*, and *Ultrabithorax* were aligned using MUSCLE v. 3.6 [9], and alignments were confirmed using amino acid sequence translations. Ambiguously aligned sites were masked with GBlocks v. 0.91b [10]. Sequences of *Limulus polyphemus* were not included in this analysis, as they are too short to yield reliable phylogenetic signal. Inclusion of

pycnogonid and mite sequences was attempted, but these sequences are highly prone to long-branch attraction artifacts, and were thus also excluded.

For nucleotide sequence alignments, a general time reversible (GTR) [6] model with corrections for a discrete gamma distribution (GTR + Γ) was specified. Tree topology and nodal support were inferred using RAxML v. 7.7.5, with heuristic strategy as above [1, 2]. Alignments are provided as Datasets S2-S7.

(c) Likelihood ratio tests

For each of the six nucleotide alignments (above), we assessed the strength of phylogenetic evidence for evolutionary scenarios using Shimodaira–Hasegawa (SH; [11]) and Approximately Unbiased (AU; [12]) tests in RAxML v. 7.7.5 [1]. We enforced topological constraints consistent with two scenarios: hypothesis 1 constrained the tree topologies to reflect lineage-specific duplications, and hypothesis 2 constrained the tree topologies to reflect a common mechanism of duplication in the MRCA of scorpions and spiders (illustrated in figure 3). To avoid incorrect assignment of orthology under hypothesis 2, all possible topologies conforming to hypothesis 2 were used as inputs. To obtain *p*-values for the test, per-site log likelihood values were computed using the –f g command in RAxML v. 7.7.5 [1]. The resulting likelihoods were analyzed using CONSEL [13].

References

- 1 Stamatakis A. 2006 RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**, 2688–2690. (doi: 10.1093/bioinformatics/btl446)
- 2 Stamatakis A, Hoover P, Rougemont J. 2008 A rapid bootstrap algorithm for the RAxML Web servers. *Syst. Biol.* **57**, 758–771. (doi: 10.1080/10635150802429642)
- 3 Whelan S, Goldman N. 2001 A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol. Biol. Evol.* **18**, 691–699.
- 4 Yang Z. 1996 Among-site rate variation and its impact on phylogenetic analyses. *Trends Ecol. Evol.* **11**, 367–372.
- 5 Abascal F, Zardoya R, Posada D. 2005 ProtTest: Selection of best-fit models of protein evolution. *Bioinformatics* **21**, 2104–2105.
- 6 Tavaré S. 1986 Some probabilistic and statistical problems in the analysis of DNA sequences. *Lect. Math Life Sci.* **17**, 57–86.
- 7 Huelsenbeck JP, Ronquist F. 2005 Bayesian analyses of molecular evolution using MrBayes. In: Nielsen R (ed.) Statistical Methods in Molecular Evolution. Springer, New York.
- 8 Rambaut A, Drummond AJ. 2009 Tracer v. 1.5. Program. <<http://beast.bio.ed.ac.uk/Tracer/>>.
- 9 Edgar RC. 2004 MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nuc. Acids Res.* **32**, 1792–1797. (doi: 10.1093/nar/gkh340)

- 10 Castresana J. 2000 Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* **17**, 540–552.
- 11 Shimodaira, H, Hasegawa, M. 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol. Biol. Evol.* **16**, 1114–1116.
- 12 Shimodaira H. 2002. An approximately unbiased test of phylogenetic tree selection. *Syst. Biol.* **51**, 492–508.
- 13 Shimodaira H, Hasegawa M. 2001. CONSEL: for assessing the confidence of phylogenetic tree selection. *Bioinformatics* **17**, 1246–1247.

Multiple sequence alignments (Data sets S1-S7)

#NEXUS

BEGIN DATA;

DIMENSIONS NTAX=92 NCHAR=68;

FORMAT DATATYPE=PROTEIN INTERLEAVE MISSING=-;

[DATASET S1: Peptide alignment of Hox genes with known expression patterns,
in nexus format.]

MATRIX

Porcellio_scaber_AbdB RKKRKPYSKFQTLELEKEFL

YNAYVSKQKRWELARNLNLT ERQVKIWFQNRRMKKKNSQ

RQAAQEGR

Phalangium_opilio_AbdB RKKRKPYSKFQTLELEKEFL

YNAYVSKQKRWELARNLNLT ERQVKIWFQNRRMKSKKNSQ

RNAENNQN

Archegozetes_longisetosus_AbdB RKKRKPYSKFQTLELEKEFL

FNAYVSKQKRWELARNLNLT ERQVKIWFQNRRMKSKKTSQ RNSDRNK-

Cupiennius_salei_AbdB RKKRKPYSKFQTLELEKEFL
FNAYVSKQKRWELARNLNLT ERQVKIWFQNRRMKSKKTSQ
RNAENNQN

Strigamia_maritima_AbdB RKKRKPYSKFQTLELEKEFL
FNAYVSKQKRWELARNLNLT ERQVKIWFQNRRMKNKKNSQ
RNQTDSAK

Tribolium_castaneum_AbdB RKKRKPYSKFQTLELEKEFL
FNAYVSKQKRWELARNLNLT ERQVKIWFQNRRMKNKKNSQ
RQAAQQQN

Centruroides_sculpturatus_AbdB2 RKKRKPYSKFQTLELEKEFL
FNAYVSKQKRWELARNLNLT ERQVKIWFQGR-----

Lithobius_atkinsoni_AbdB -----L FNAYVSKQKRWELARNLNLT
ERQVKIWFQNRRMKNKKNSQ RNTADTTA

Glomeris_marginata_AbdB -----L FNAYVSKQKRWELARNLNLT
ERQVKIWFQNRRMKNKKNSQ RNQQDTSS

Centruroides_sculpturatus_AbdB1 RKKRKPYSKFQTLELEKEFL
FNAYVSKQKRW-----

Glomeris_marginata_Hox3 -----H FNPYICRPRRVEMALVLGLQ
ERQIKIWFQNRRMFKKKDNR HRDNHNCL

Lithobius_atkinsoni_Hox3 -----SAQLVELEKEFH

FNRYLCRPRRVELAKQLGLT DRQIKIWFQNRRMKYKKEKL HREGCQTA

Tribolium_castaneum_zen KRARTAYTSAQLVELEREFH
HGKYLSRPRRIQIAENLNLS ERQIKIWFQNRRMKHKKEQM NKVSTPRS

Tribolium_castaneum_z2 KRARTAYTSSQLVELEREFH
RSKYLCRPRRIQMAQNLNLT ERQIKIWFQNRRMKFKKEEK NKVVTPKT
Centruroides_sculpturatus_Hox3 KRARTAYTSAQLVELEKEFH
FNRYLCRPRRIEMAALLNLT ERQIKIWFQNRRMKFKKEQK SKGLSVDD
Archegozetes_longisetosus_Hox3 KRARTAYTSSQLVELEKEFH
TSRYLCRPRRIEMASLLKLS ERQIKIWFQNRRMKFKKEQ- -----
Cupiennius_salei_Hox3 KRARTAYTSAQLVELEKEFH
FNRYLCRPRRIEMANLLNLS ERQIKIWFQNRRMKYKKEQK SKGLYIQQ
Tribolium_castaneum_lab NTGRTNFTNKQLTELEKEFH
FNKYLTRARRIEIASALQLN ETQVKIWFQNRRMKQKKRMK EGLIPPEP
Archegozetes_longisetosus_lab NTGRTNFTTNQLTELEKEFH
FNKYLTRARRIEIATALQLN ETQVK-----
Lithobius_atkinsoni_lab -----H FNKYLTRARRIEIATALQLN
ETQVKIWFQNRRMKQKKRLK EGSLVSE
Porcellio_scaber_lab GTGRTNFTTKQLTELEKEFH
FNKYLTRARRIEIASALQLN ETQVKIWFQNRR-----
Cupiennius_salei_lab GSGRTNFTTKQLTELEKEFH
YNKYLTRARRIEIATALQLN ETQVKIWFQNRRMKQKKRMK EPLLVT--
Centruroides_sculpturatus_lab2 GTGRTNFTTKQLTELEKEFH
FNKYLTRARRIEIATALQLN ETQVKIWFQNRRMKQKKRMK EGLVPPEN
Phalangium_opilio_lab GSGRTNFTTKQLTELEKEFH
FNKYLTRARRIEIATALQLN ETQVKIWFQNRRMKQKKRMK EGLIPPEP

Glomeris marginata _lab -----TKQLTELEKEFH
FNKYLTRARRIEATALQLN ETQVKIWFQNRRMKQKKRMK EGLVIVKE
Centruroides sculpturatus _lab1 GSGRTNFTTKQLTELEKEFH
FNKYLTRARRIEATALQLN ETQVKIWFQNRRMKQKKRMK EGL-----
Glomeris marginata _ftz KRTRQTYTRYQTLELEKEFH
SNRYLNRRRIEATSLTLT ERQVKIWFQNRRMKAKREPK MVVHAGGN
Cupiennius salei _Ubx1 RRGRQTYTRYQTLELEKEFH
TNHYLTRRRRIEMAHALCLT ERQIKIWFQNRRMKLKKEIQ AIKELNEQ
Glomeris marginata _Ubx RRGRQTYTRYQTLELEKEFH
TNHYLTRRRRIEMAHALCLT ERQIKIWFQNRRMKLKKEIQ AIKELNEQ
Lithobius atkinsoni _Ubx -----H TNHYLTRRRRIEMAHALCLT
ERQIKIWFQNRRMKLKKEIQ AIKELNEQ
Phalangium opilio _Ubx RRGRQTYTRYQTLELEKEFH
TNHYLTRRRRIEMAHALCLT ERQIKIWFQNRRMKLKKEIQ AIKELNEQ
Porcellio scaber _Ubx -----H TNHYLTRRRRIEMAHALCLT
ERQIKIWFQNRRMKLKKEIQ AIKELNEQ
Strigamia maritima _Ubx RRGRQTYTRYQTLELEKEFH
TNHYLTRRRRIEMAHALCLT ERQIKIWFQNRRMKLKKEIQ AIKELNEQ
Tribolium castaneum _Ubx RRGRQTYTRYQTLELEKEFH
TNHYLTRRRRIEMAHALCLT ERQIKIWFQNRRMKLKKEIQ AIKELNEQ
Centruroides sculpturatus _Ubx1 RRGRQTYTRYQTLELEKEFH
TNHYLTRRRRIEMAHALCLT ERQIKIWFQNRRMKLKKEIQ AIKELNEQ

Centruroides_sculpturatus_Ubx2 RRGRQTYTRYQTLELEKEFH
TNHYLTRRRRIEMAHALCLT ERQIKIWFQNRRMKLKKEIQ AIKELNEQ

Cupiennius_salei_Ubx2 RRGRQTYTRYQTLELEKEFH
TNHYLTRRRRIEMAHSLCLT ERQIKIWFQNRRMKLKKEAQ AIKELNEQ

Archegozetes_longisetosus_Ubx RRGRQTYTRYQTLELEKEFH
TNHYLTRRRRIEMAHSLCLT ERQIKIWFQNRRMKLKKEIQ AIKELNEQ

Phalangium_opilio_AbdA RRGRQTYTRFQTLELEKEFH
FKQYLTRRRRIEIAHALCLT ERQIKIWFQNRRMKHKKEMR AVKEINEQ

Glomeris_marginata_AbdA -----H FNHYLTRRRRIEIAHAFCLT
ERQIKIWFQNRRMKLKKELR AVKEINEQ

Porcellio_scaber_AbdA_1 RRGRQTYTRFQTLELEKEFH
FNHYLTRRRRIEIAHALCLT ERQVKIWFQNRRMKLKKELR -----

Cupiennius_salei_AbdA RRGRQTYTRFQTLELEKEFH
FNHYLTRRRRIEIAHALCLT ERQIKIWFQNRRMKLKEMR AVKEINEQ

Lithobius_atkinsoni_AbdA -----
LKKEMR AVKEINEQ

Strigamia_maritima_AbdA RRGRQTYTRFQTLELEKEFH
FNHYLTRRRRIEIAHALCLT ERQIKIWFQNRRMKLKEMR AVKEINEQ

Tribolium_castaneum_AbdA RRGRQTYTRFQTLELEKEFH
FNHYLTRRRRIEIAHALCLT ERQIKIWFQNRRMKLKKELR AVKEINEQ

Centruroides_sculpturatus_abdA2 RRGRQTYTRFQTLELEKEFH
FNHYLTRRRRIEIAHALCLT ERQIKIWFQNRRMKLKKELR AVKEINEQ

Centruroides_sculpturatus_abdA1 RRGRQTYTRFQTLELEKEFH
FNHYLTRRRRIEIAHALCLT ERQIKIWFQNRRMKLKKEIR AVKEINEQ

Glomeris_marginata_Scr ----SYTRYQTLELEKEFH
FNRYLTRRRRIEIAHSLCLT ERQIKIWFQNRRMKWKKEHK LAHHLPPP

Phalangium_opilio_Scr KRQRTSYTRYQTLELEKEFH
FNRYLTRRRRIEIAHALCLS ERQIKIWFQNRRMKWKKEHK MASTMPPQ

Cupiennius_salei_Scr1 KRQRTSYTRYQTLELEKEFH
FNRYLTRRRRIEIAHALCLS ERQIKIWFQNRRMKWKKEHK MASPIPPQ

Centruroides_sculpturatus_Scr2 KRQRTSYTRYQTLELEKEFH
FNRYLTRRRRIEIAHALCLS ERQIKIWFQNRRMKWKKEHK MANTVPPQ

Centruroides_sculpturatus_Scr1 KRQRTSYTRYQTLELEKEFH
FNRYLTRRRRIEIAHALCLS ERQIKIWFQNRRMKWKKEHK MAN----

Cupiennius_salei_Scr2 KRQRTSYTRYQTLELEKEFH
FNRYLTRRRRIEIAHALCLS ERQIKIWF-----

Lithobius_atkinsoni_Antp -----H FNRYLTRRRRIEIAHALCLT
ERQIKIWFQNRRMKWKKENK AKLEGAGG

Strigamia_maritima_Antp -----ELEKEFH FNRYLTRRRRIEIAHALCLT
ERQIKIWFQNRRMKWKKENK AKLERAGG

Tribolium_castaneum_Scr KRQRTSYTRYQTLELEKEFH
FNRYLTRRRRIEIAHALCLT ERQIKIWFQNRRMKWKKEHK MASMNIVP

Porcellio_scaber_Scr KRQRTSYTRYQTLELEKEFH
FNRYLTRRRRIEIAHALCLT ERQIKI-----

Phalangium_opilio_Antp KRGRQTYTRYQTLELEKEFH
FNRYLTRRRRIEIAHALCLT ERQIKIWFQNRRMKWKKENK AKLEAGLA
Centruroides_sculpturatus_Antp1 KRGRQTYTRYQTLELEKEFH
FNRYLTRRRRIEIAHALCLT ERQIKIWFQNRRMKWKKENK AKLEAGL-
Archegozetes_longisetosus_Antp KRGRQTYTRYQTLELEKEFH
FNRYLTRRRRIEIAHALCLT ERQIKIWFQNRRMKWKKEIR TR-----
Cupiennius_salei_Antp KRGRQTYTRYQTLELEKEFH
FNRYLTRRRRIEIAHALCLT ERQIKIWFQNRRMKWKKENK AKEPAAGF
Porcellio_scaber_Antp KRGRQTYTRYQTLELEKEFH
FNRYLTRRRRIEIAHALCLT ERQIKIWFQNRRMKWKKENK TKVENGNP
Tribolium_castaneum_Antp_Pt1 KRGRQTYTRYQTLELEKEFH
FNRYLTRRRRIEIAHALCLT ERQIKIWFQNRRMKWKKENK TKGEGGSE
Glomeris_marginata_Antp KRGRQTYTRYQTLELEKEFH
FNRYLTRRRRIEIAHALCLT ERQIKIWFQNRRMKWKKENK AKIEAGVD
Lithobius_atkinsoni_ftz -----H FNRYLTRRRRIEIAHALGLT
ERQIKIWFQNRRMKAKKENK LQNGQTSP
Centruroides_sculpturatus_Antp2 KRGRQTYTRYQTLELEKEFH
FNRYLTRRRRIEIAHAL--- -----
Archegozetes_longisetosus_Scr KRQRTSYTRYQTLXLEKEFH
FNRYLTRRRRIEIAHSLCLS ERQIKIWFQNRRMKWKKEH- -----
Tribolium_castaneum_Dfd KRQRTAYTRHQILELEKEFH
YNRYLTRRRRIEIAHTLVLS ERQIKIWFQNRRMKWKKDNK LPNTKNVR

Cupiennius_salei_Dfd KRQRTAYTRHQILELEKEFH
FNRYLTRRRRIEIAHALCLS ERQIKIWFQNRRMKWKKDNK LPNTKNVK

Cupiennius_salei_Dfd2 KRQRTAYTRHQILELEKEFH
FNRYLTRRRRIEIAHALCLS ERQIKIWFQNRRMKWKKDNK LPNTKNVK

Phalangium_opilio_Dfd KRQRTAYTRHQILELEKEFH
FNRYLTRRRRIEIAHALCLS ERQIKIWFQNRRMKWKKDNK LPNTKNVK

Centruroides_sculpturatus_Dfd2 KRQRTAYTRHQILELEKEFH
FNRYLTRRRRIEIAHSLCLS ERQIKIWFQNRRMKWKKDNK LPNTKNVK

Centruroides_sculpturatus_Dfd1 KRQRTAYTRHQILELEKEFH
FNRYLTRRRRIEIAHSLCLS ERQIKIWFQNRRMKWKKDNK LPNTKNVK

Archegozetes_longisetosus_Dfd KRQRTAYTRHQILELEKEFH
FNRYLTRRRRIEIAHSLCLS ERQIKIWFQNRRMKWKKDN- -----

Lithobius_atkinsoni_Dfd -----RQQILELEKEFH FNRYLTRRRRIEIAHSLCLS
ERQIKIWFQNRRMKWKKDNK LPNTKNVR

Glomeris_marginata_Dfd -----H FNRYLTRRRRIEIAHSLCLS
ERQIKIWFQNRRMKWKKDNK LPNTKNVR

Tribolium_castaneum_ftz KRTRQTYTRYQTLELEKEFH
FNKYLTRRRRIEIAESLRLT ERQIKIWFQNRRMKAKKDTK FTEQSVTS

Cupiennius_salei_ftz KRSRQTYTRYQTLELEKEFH
FNQYLTRRRRIEIAHTLGLT ERQIKIWFQNRRMKAKKENK FPISSSNS

Centruroides_sculpturatus_ftz2 KRTRQTYTRYQTLELEKEFH
FNRYLTRRRRVEIARMLSLT ERQIKIWFQNRRMKAKKE-- -----

Phalangium_opilio_ftz KRTRQTYTRIQTLELEKEFH
FNRYLTRRRRIEIAHSLGLT ERQIKIWFQNRRMKAKKETN LQPTASST
Archegozetes_longisetosus_ftz KRTRQTYTRYQTLELEKEFH
FNRYLTRRRRIEIAHSLGLT ERQIKIWFQNRRMKAKKENK IKVDPNSA
Centruroides_sculpturatus_ftz1 KRTRQTYTRYQTLELEKEFH
FNHYLTRRRRIEIAHSLGLT E-----
Phalangium_opilio_Hox3 KRPRTAYTNNQLVELEKEFH
FNRYLCRPRRVELAAQLSLS ERQIKVWFQNRRMKYKKDSK SRGGGSIY
Porcellio_scaber_pb RRLRTAYTNTQLLELEKEFH
FNKXLCRPRRIXXAASLDLT ERQVKVWF-----
Lithobius_atkinsoni_pb -----LLELEKEFH FNKYLCPRIIEIAASLDLA
ERQVKVWFQNRRMKHKRQTM GKGSDDDGG
Glomeris_marginata_pb -----H FNKYLCPRIIEIAASLDLT
ERQVKVWFQNRRMKHKRQTV GKGEDGA
Cupiennius_salei_pb RRLRTAYTNTQLLELEKEFH
FNKYLCPRIIEIAASLDLT ERQVKVWFQNRRMKHKRQTS VMKDDDKD
Tribolium_castaneum_pb_mxp RRLRTAYTNTQLLELEKEFH
FNKYLCPRIIEIAASLDLT ERQVKVWFQNRRMKHKRQTL GKQGDDGD
Centruroides_sculpturatus_pb1 RRLRTAYTNTQLLELEKEFH
FNKYLCPRIIEIAASLDLT ERQVKVWFQNRRMKHKRQTM
MNKNDDKN
Phalangium_opilio_pb RRLRTAYTNTQLLELEKEFH
FNKYLCPRIIEIAASLDLT ERQVKVWFQNRRMKHKRQSL MTKNGDEK

Centruroides_sculpturatus_pb2 RRLRTAYTNTQLLELEKEFH
FNKYLCRPRRIEIAASLDLT ERQVKVWFQNRRMKHKRQSM
MNKNGEDK

Archegozetes_longisetosus_pb RRLRTAYTNTQLLELEKEFH
FNKYLCRPRRIEIAASLDLT ERQVKVWFQNRRMKHKRQS- -----
;
END;

#NEXUS

BEGIN DATA;

DIMENSIONS NTAX=8 NCHAR=228;

FORMAT DATATYPE=DNA INTERLEAVE MISSING=-;

[DATASET S2: Nucleotide alignment of scorpion, spider, and harvestman labial sequences.]

MATRIX

Phalangium_opilio_lab CCAACTTATAAATGGATGCA
AGTGAAAAGAAATGTTCCAA AACCGAGGTTGGGTGGACA
AACTTTACGACGAAACAAC AACCGAAACTCGAAAAAGAAT

Cupiennius_salei_lab CCCACTTACAAATGGATGCA
AGTGAAAAGGAGTCTGCCGA AGCCAGGCTCCGGCCGGACA
AACTTCACAACCAAACAGCT TACAGAACTGGAAAAGGAGT
Centruroides_sculpturatus_lab1 CCTACTTACAAATGGATGCA
AGTCAAGAGGAGCGCACCTA AACCAGGATCGGGCGTACG
AATTTCACGACCAAACAGTT AACAGAACTTGAGAAAGAGT
Centruroides_sculpturatus_lab2 CCTACCTACAAGTGGATGCA
GGTGAAAAGGAGTGTGCCCA AAACAGGAACAGGTAGGACC
AACTTTACAACAAAACAAC AACAGAACTCGAAAAAGAAT
Mesobuthus_martensii_lab1 CCTACTTATAAGTGGATGCA
AGTCAAGAGGAGCGCACCTA AACCAGGTTGGTCGAACC
AATTTCACAAACAAACATT AACAGAACTGGAAAAAGAAT
Mesobuthus_martensii_lab2 CCTACCTACAAGTGGATGCA
GGTGAAAAGGAGTGTGCCCA AAACAGGAACAGGTAGGACC
AACTTTACTACAAAACAAC AACAGAACTAGAAAAAGAAT
Parasteatoda_tepidariorum_lab1 CCCACATTCAAATGGATGCA
AATTAAAAGGAGTGTACCTA AACCTGCTTCAGGAAGGACG
AATTTCAGCACAAAGCAACT GACTGAATTAGAAAAAGAAT
Parasteatoda_tepidariorum_lab2 CCCACATATAAGTGGATGCA
AGTAAAACGAAGTGTCCCA AGCCTGGAACGGGACGAACG
AATTTCACGACGAAACAGTT GACAGAACTCGAAAAAGAAT

Phalangium_opilio_lab TTCACTTTAACAAAGTACCTG
ACCAGGGCCAGGCGAATCGA AATGCCACCGCCTGCAAT
TGAACGAGACGCAGGTCAAA ATATGGTTCAAAATCGTCG
Cupiennius_salei_lab TCCACTATAACAAATACCTG
ACACGAGCCAGGAGAATAGA GATTGCAACGGCTTGCAGC
TCAACGAAACTCAAGTGAAA ATATGGTTCAGAATCGAAG
Centruroides_sculpturatus_lab1 TTCACTTCAACAAATATTG
ACGAGGGCAAGAAGGATCGA AATTGCCACCGCCCTCAAC
TTAACGAGACTCAAGTGAAA ATCTGGTTCCAGAACAGGAG
Centruroides_sculpturatus_lab2 TTCATTCAACAAATATTG
ACAAGAGCAAGGAGGATCGA AATAGCCACAGCTCTCAAC
TAAATGAGACCCAAGTCAAG ATATGGTTCCAAAATAGAAG
Mesobuthus_martensii_lab1 TTCATTCAACAAATATTG
ACGAGGGCAAGAAGGATCGA AATAGCCACAGCGTTGCAAC
TCAACGAAACTCAAGTGAAA ATTTGGTTCCAAAACAGGAG
Mesobuthus_martensii_lab2 TTCATTCAACAAATATCTA
ACAAGAGCAAGAAGGATCGA AATAGCTACAGCACTCAAC
TAAACGAGACCCAAGTCAAG ATATGGTTCCAAAATAGAAG
Parasteatoda_tepidariorum_lab1 TTCATTCCACAAATATCTA
AACCGGGCTCGTAGAATAGA AATAGCCTCTCTTGGCC
TCAATGAAACGCAAGTGAAA ATTTGGTTCCAAAACAGGCG

Parasteatoda_tepidariorum_lab2 TTCATTTAACAAATACCTA
ACAAGAGCTAGGCGAATTGA AATTGCTACTGCTCTTCATT
TAAATGAGACACAAGTTAA- -----

Phalangium_opilio_lab AATGAAACAAAAGAAACGTA TGAAAGAG

Cupiennius_salei_lab GATGAAGCAGAAGAAGCGTA TGAAAGAG

Centruroides_sculpturatus_lab1 AATGAAGCAGAAGAAGAGGA

TGAAAGAG

Centruroides_sculpturatus_lab2 GATGAAGCAAAAGAAAAGAA

TGAAAGAG

Mesobuthus_martensii_lab1 AATGAAGCAGAAGAAGAGAA TGAAAGAA

Mesobuthus_martensii_lab2 GATGAAGCAAAAGAAAAGAA TGAAAGAG

Parasteatoda_tepidariorum_lab1 AATGAAAGCCAAGAAAAGGC

TAAAAGAC

Parasteatoda_tepidariorum_lab2 ----- -----

;

END;

#NEXUS

BEGIN DATA;

DIMENSIONS NTAX=7 NCHAR=375;

FORMAT DATATYPE=DNA INTERLEAVE MISSING=-;

[DATASET S3: Nucleotide alignment of scorpion, spider, and harvestman proboscipedia sequences.]

MATRIX

Centruroides_sculpturatus_pb1 GTACCGGAGTACCCATGGAT
GAAGGAGAAGAACGACACGC GTAAGCAACACCAAGAGAAAT

GGGATGCCACGACGTCTCG AACTGCTTACACAAACACAC

Centruroides_sculpturatus_pb2 GTCCCGGAATATCCTTGGAT
GAAAGAGAAGAACGACCA GAAAACCACACCAAGATAAT

GGGATGCCTCGTCGACTCCG TACAGCCTACACTAATACTC

Mesobuthus_martensi_pb1 GTGCCGGAGTACCCATGGAT
GAAAGAGAAGAACGACACGC GCAAGCAGCACCAAGAGAAAT

GGAATGCCTCGTCGACTCCG AACTGCTTACGAACACAC

Mesobuthus_martensi_pb2 GTCCCGGAATATCCTTGGAT

GAAGGAGAAGAACGACCA GAAAACCACACCAAGATAAT

GGTATGCCTCGTCGACTCCG CACAGCATACTAATACTC

Cupiennius_salei_pb -----TGGAT GAAGGAGAAGAACGACCA

GAAAGCAACAGCAAGACAAT GGTATGCCTCGACGACTACG

AACCGCATATACCAATACTC

Parasteatoda_tepidariorum_pb -----TGGAT
GAAGGAGAAGAAAACATCTA NGAACAGCATCAAGAAAAC
GGGATGCCTCGTCGCCTGAG GACTGCCTACACTAACACACGC
Phalangium_opilio_pb -----

Centruroides_sculpturatus_pb1 AGTTGTTGGAAGTGGAGAAG
GAGTTTCATTTAATAAAATA TCTGTGCAGACCCCCGGAGGA
TAGAAATTGCAGCATCTCTG GATCTAACAGAGAGACAAGT
Centruroides_sculpturatus_pb2 AGTTATTGGAAGTGGAGAAA
GAATTCCATTCAACAAATA TCTGTGCCGTCCCCGTGCA
TTGAAATTGCTGCATCTTG GATTGACAGAACGCCAAGT
Mesobuthus_martensii_pb1 AGTTATTGGAATTGGAGAAA
GAGTTTCATTTAATAAAATA TCTGTGCAGACCACGGAGGA
TAGAAATTGCAGCATCTCTG GACTAACAGAAAGGCAAGT
Mesobuthus_martensii_pb2 AGTTATTGGAGTTGGAGAAA
GAATTCCACTCAACAAATA CCTGTGCCGTCCCCGTGTA
TTGAAATAGCTGCATCTTA GATTGACAGAACGACAAGT
Cupiennius_salei_pb AGCTTTAGAGCTGGAGAAA
GAATTTCATTCACAAAGTA TCTATGTCGACCGAGGAGGA
TAGAAATTGCTGCAAGCCTC GATCTGACAGAGAGACAAGT

Parasteatoda_tepidariorum_pb AGCTGCTAGAGCTGGAGAAG
GAATTCCACTTCAACAAATA TTTGTGTCGGCCGANGCGCA
TTGAAATTGCTGCTTCTCTT GACCTTACAGANCAGTCAGGT
Phalangium_opilio_pb -----

Centruroides_sculpturatus_pb1 GAAGGTGTGGTTCCAGAATC
GTAGAACATGAAACACAAGAGG CAAACTATGATGAACAAAAAA
TGACGATAAGAACATGGGGATG CGGGAAGTAGAACGCCGCAC
Centruroides_sculpturatus_pb2 AAAGGTGTGGTTTCAGAATA
GAAGAACATGAAACATAAACGA CAATCTATGATGAATAAAAAA
CGGTGAAGATAAAAGTGACA ATGAGAGTATCAAAGACGTA
Mesobuthus_martensii_pb1 GAAGGTGTGGTTCCAGAATC
GAAGAACATGAAACACAAGAGA CAAACTATGATGAATAAGAA
TGATGATAAAATGGTGATG CAGGAAGCAGAACGCCAGAT
Mesobuthus_martensii_pb2 GAAGGTGTGGTTTCAGAACA
GAAGAACATGAAACATAAACGA CAATCTATGATGAATAAAAAA
CGGCGAAGATAAAAGTGACA ATGAGAGCATCAAAGATGTA
Cupiennius_salei_pb GAAAGTATGGTTCCAGAACCC
GTCGTATGAAGCACAAAGAGA CAGACTTCTGTTATGAAGGA
TGACGATAAAAGATGGGGTTG ACAGAAGTCTAGATAAGTGAT
Parasteatoda_tepidariorum_pb CAAAGTCTGGTTCCAAAACC GCCGA-----

Phalangium_opilio_pb -----CCAAAATC
GGAGGATGAAGCACAAAGAGA CAGTCTTGATGACTAAGAA
CGGCGATGAAAAGGGTAGCG GACAGTCGATCGGGAGTCCC

Centruroides_sculpturatus_pb1 GGGGACCAGAGTAGCAACAG
CGGGACCAAAGAACATGTACCT CGCTGCAGTCACCGTGACG
GTGAGCTCTACACCG
Centruroides_sculpturatus_pb2 GCTGATGAACGTTCAAGGAC
TCCAACAATCCAAACCGATA GTTTAAGTTCTCGGATACA
GAAACACATGCCCT

Mesobuthus_martensii_pb1 GGGGACCAGAGTAGCAATAG
TGGGACCAAAGAACATGTACCT CGCTTCAGTCACCGTGACG
GTGAGTTCAACACCA

Mesobuthus_martensii_pb2 GTCGATGAACGATCAAGGAC
TCCAACAATTCAAACAGATA GCTTGAGTTCTCGGATACA
GAAACACATCGACCA

Cupiennius_salei_pb CACGGAATGCCCTCCCGAAC
GCCAGAACATTGAAACACAT CG-----

Parasteatoda_tepidariorum_pb -----

Phalangium_opilio_pb TGTTCCGATCAGACTTCTAC
CAGCCAGGATGGATCGACGC CTCTCATTGCCATCGACG
CACGTACCTAGCTCA

;
END;

#NEXUS

BEGIN DATA;
DIMENSIONS NTAX=9 NCHAR=279;
FORMAT DATATYPE=DNA INTERLEAVE MISSING=-;

[DATASET S4: Nucleotide alignment of scorpion, spider, and harvestman
Deformed sequences.]

MATRIX

Phalangium_opilio_Dfd -----
----- GCGTACACGAGGCATC
Centruroides_sculpturatus_Dfd1 TACCCGTGGATGAAGAAAGT
ACACGTCAATTCACTGCT---- --AATGGGAACTTTCGGGT
GTGGAACCAAAAAGACAGAG GACCGCTTATACTAGGCATC
Cupiennius_salei_Dfd2 TACCCGTGGATGAAGAAAGT
CCACGTCCGGCTCAGTTGCTG CCAATGGGAATTCCCCCGGA
GTTGAGGCCAAAAGGCAAAG GACGGCATACTCGTCATC

Cupiennius_salei_Dfd1 TATCCCTGGATGAAGAAAGT
CCACAGCAATCCAGTG---- --AATGGAAGCTTCCCAGGT
ATCGAACCGAAACGGCAGCG GACAGCCTACACGAGGCACC
Centruroides_sculpturatus_Dfd2 TATCCCTGGATGAAGAAAGC
CCACGTGGGTACAGCT---G CAAATGGAAACTTACTGGA
ATGGAACCTAAAAGACAGAG GACAGCGTACACTAGACATC
Parasteatoda_tepidariorum_Dfd1 TATCCATGGATGAGAAAAGT
TCATACAAATAATCCCGGTG TGAATGGAGTATATCCAGGC
CTTGAACCAAAACGACAACG AACTGCCTACACAAGACACC
Parasteatoda_tepidariorum_Dfd2 TATCCCTGGATGAAAAAAAGT
ACATGTTGGGACAGTTGCTG CCAATGGAAATTCCCTGGA
GTAGAACCTAAAAGGCAAAG GACAGCCTACTCGACACC
Mesobuthus_martensii_Dfd2 -----G
CTAACGGGAACCTCTCCGGT GTGGAGCCAAAAGACAAAG
GACCGCTTACACTAGGCACC
Mesobuthus_martensii_Dfd1 TACCCGTGGATGAAGAAGGT
GCACGTCAATTCAAGGT---- --AAGGCGCGACCACGTGGC GTC-----

Phalangium_opilio_Dfd AGATTCTCGAGCTAGAAAAA
GAGTTCCATTTAACCGGTA CCTGACTAGAAGACGACGGA
TTGAAATCGCTCACGCGCTC TGTTATCCGAGAGGCAAAT

Centruroides_sculpturatus_Dfd1 AGATACTCGAACTGGAGAAA
GAATTCCATTCAATCGCTA CCTGACGCGTAGGAGAAGGA
TCGAAATTGCTCACTCCCTT TGCCTGAGCGAACGACAGAT

Cupiennius_salei_Dfd2 AGATTCTCGAGTTAGAGAAA
GAGTTCACTCAACCGTA CCTTACACGGCGGCGACGCA
TCGAGATAGCGCACGCCCTC TGTCTCTCAGAAAGGCAGAT

Cupiennius_salei_Dfd1 AGATCCTCGAGCTAGAGAAG
GAGTTCACTCAATCGCTA CTTGACGAGACGGAGGCGGA
TCGAGATCGCTCACGCATT TGTTGTCCGAGAGACAGAT

Centruroides_sculpturatus_Dfd2 AAATTTAGAACTTGAAAAG
GAATTCCACTTCAACCGTTA CCTGACGCGGAGGAGAAGAA
TCGAAATAGCTCATTCCCTG TGTCTCAGTCAAAGGCAGAT

Parasteatoda_tepidariorum_Dfd1 AAATTCTAGAACTCGAAAAAA
GAGTTCACTTCAACCGCTA CCTGACGCGGAGGAGACGAA
TCGAAATCGCCCATTCCCTC TGCCTTCAGAAAGACAAAT

Parasteatoda_tepidariorum_Dfd2 AGATCCTAGAATTGGAAAAAA
GAATTCCATTCAACAGATA CCTAACTCGCGAAGAAGAA
TAGAGATTGCTCACGCGTTA TGCCTCTCTGAAAGGCAGAT

Mesobuthus_martensii_Dfd2 AGATCCTAGAGCTGGAGAAA
GAATTCACTTCAACCGATA TCTGACGCGCAGGAGAAGGA
TCGAAATCGCTCACTCGCTG TGTTGAGCGAGAGACAGAT

Mesobuthus_martensii_Dfd1 -----

Phalangium_opilio_Dfd TAAGATTGGTCCAAAATA
GGCGAATGAAATGGAAAAAG GACAATAAATTGCCTAATAC
CAAAAACGTCAAGAAAAAC

Centruroides_sculpturatus_Dfd1 TAAGATCTGGTCCAAAACA
GGAGGATGAAATGGAAGAAG GACAACAAATTACCCAATAC
TAAGAACGTTAAGAAGAAG

Cupiennius_salei_Dfd2 CAAAATATGGTCCAGAAC
GACGAATGAAGTGGAGAAG GACAATAAACTTCCAAACAC
AAAGAACGTCAAGAAGAAA

Cupiennius_salei_Dfd1 CAAAATCTGGTCCAGAAC
GACGGATGAAGTGGAGAAA GACAATAAACTTCCAAATAC
TAAGAACGTTAAGAAGAAA

Centruroides_sculpturatus_Dfd2 AAAAATATGGTTCAAAATA
GGAGAATGAAATGGAAAAAAA GACAACAAATTACCGAAC
GAAAAATGTTAAAAG---

Parasteatoda_tepidariorum_Dfd1 CAAGATTGGTTCAAAATA
GGCGGATGAAGTGGAGAAA GACAATAAGCTACCAAATAC
AAAGAACGTGAAGAAGAAA

Parasteatoda_tepidariorum_Dfd2 CAAGATTGGTTCAGAAC
GAAGGATGAAATGGAAAAAAA GATAATAAGTTACCAAATAC
GAAGAACGTCAAGAAGAAA

Mesobuthus_martensii_Dfd2 CAAGATATGGTTCCAAAACA
GGAGGATGAAGTGGAAAAAG GACAACAAGCTACCCAACAC
CAAGAACGTTAAGAAGAAA

Mesobuthus_martensii_Dfd1 -----

;
END;

#NEXUS

BEGIN DATA;
DIMENSIONS NTAX=8 NCHAR=330;
FORMAT DATATYPE=DNA INTERLEAVE MISSING=-;

[DATASET S5: Nucleotide alignment of scorpion, spider, and harvestman Sex
combs reduced sequences.]

MATRIX

Phalangium_opilio_Scr AACCCGCCGCAAATATATCC
TTGGATGAGGAAGGTTCACG TCGGACAAAATGGAGTCAAC
GCAATGGCGAAACTAACG CCAGAGAACGTACACTC

Parasteatoda_tepidariorum_Scr AATCCTCCTCAGATTATCC
CTGGATGAGAAAAGTCATG TCGGGCAAAATGGTGTGAAC
TCCATGGGAGAACAAAGCG CCAAAGAACATCCTACACCC

Cupiennius_salei_Scr1 -----TCCTCAGATTATCC
TTGGATGAGAAAAGTACACG TTGGGCAAAATGGTGTCAAT
TCAATGGGTGAGACGAAACG CCAAAGAACCTCGTATACTC

Cupiennius_salei_Scr2 AACCCCTCCCCAAATCTACCC
CTGGATGAGGAAGGTTCACG TCGGGCAAAATGGTGTGAAT
GCGATGGCGAACGAAACG CCAAAGAACGTCCTATACTC

Centruroides_sculpturatus_Scr2 AATCCACCGCAGATTATCC
ATGGATGAGGAAAGTTCACG TCGGTCAAGACGGCGTCAAC
GCTATGGGTGAGACAAAACG CCAGCGGACGTCATACACTC

Centruroides_sculpturatus_Scr1 ACAACACCACAGATTATCC
ATGGATGAGAAAAGTCCACG TCGGTCAAGACGGTGTCAAC
GCCATGGGTGAGACCAAGAG GCAGCGAACATCCTACACAC

Mesobuthus_martensii_Scr1 -----
ACGGTGTCAAC GCCATGGGTGAGACCAAGAG
GCAGCGAACATCCTATACAC

Mesobuthus_martensii_Scr2 AATCCACCGCAGATTATCC
ATGGATGAGGAAAGTTCACG TCGGTCAAGACGGCGTCAAC
GCTATGGGTGAGACAAAACG CCAGCGGACGTCATACACTC

Phalangium_opilio_Scr GCTACCAAACTCTGGAATTG
GAAAAAGAGTTTCATTCAA CCGTTACCTCACCAAGGCGGA
GGAGGATCGAAATAGCCC-- -----

Parasteatoda_tepidariorum_Scr GGTATCAAACATTGGAGCTC
GAGAAAGAGTTGGTTTAA CCGGTATTAACCAGACGAC
GGCGGATTGAGATAGCTCAT GCCCTGTGTCTGACAGAAAG

Cupiennius_salei_Scr1 GATACCAGACCTGGAGTTG
GAGAAAGAATTCCATTCAA CCGGTACCTCACCCGACGAA
GACGAATTGAAATTGCGCAC GCCCTCTGCCTCAGCGAAAG

Cupiennius_salei_Scr2 GATACCAGACGCTGGAGCTG
GAAAAAGAATTCACTTCAA CCGGTATCTTACGAGGCGGC
GGAGGATCGAAATAGCACAC GCTCTATGCCTGTCCGAAAG

Centruroides_sculpturatus_Scr2 GTTACCAGACCCTGGAGTTA
GAGAAGGAGTTCCACTTCAA TCGCTACCTGACTCGAAGGA
GGAGGATCGAGATAGCCCAC GCTCTTGCTAAGTGAAAG

Centruroides_sculpturatus_Scr1 GCTACCAGACTTTAGAATTG
GAGAAAGAGTTTCATTCAA CCGTTATCTGACGCGGAGAA
GAAGGATAGAAATAGCTCAC GCTCTATGTCTATCTGAACG

Mesobuthus_martensii_Scr1 GCTACCAGACTTAGAATTG
GAGAAAGAGTTCACTTCAA CCGTTATCTGACGCGGAGAA
GAAGGATAGAAATAGCTCAC GCTCTTGCTATCCGAACG

Mesobuthus_martensii_Scr2 GTTACCAGACCCTGGAGTTA
GAGAAGGAGTTCCACTCAA TCGCTACCTGACTCGAAGGA
GGAGGATCGAGATAGCCCAC GCTCTTGCTAAGTGAAAG

Phalangium_opilio_Scr -----

Parasteatoda_tepidariorum_Scr ACAAATTAAAATC-----

Cupiennius_salei_Scr1 ACAGATCAAAATCTGGTTCC
AGAACCGGCGAATGAAGTGG AAAAAGGAGCACAAAGATGGC
TTCACCCATACCCCCTCAA TACCTCAGGTCATACCGGAT
Cupiennius_salei_Scr2 ACAGATCAAGATCTGGTTC-----

Centruroides_sculpturatus_Scr2 GCAGATCAAGATATGGTTTC
AAAACAGGAGAATGAAATGG AAGAAAGAACACAAGATGGC
TAACACGGTGCCACCTCAGA TACCTCAGGTGATGGCAGAT

Centruroides_sculpturatus_Scr1 ACAGATCAAGATATGGTTTC
AGAATCGACGGATGAAATGG AAGAAAGAGCATAAGATGGC
CAACGCTGTGCCACCACAGA TACCTCAGATGATGCCAGAT

Mesobuthus_martensii_Scr1 ACAGATCAAGATATGGTTTC
AGAATCGACGGATGAAATGG AAGAAAGAGCATAAGATGGC
CAACGCTGTGCCACCACAGA TACCTCAGATGATGCCAGAT

Mesobuthus_martensii_Scr2 GCAGATCAAGATATGGTTCC
AAAACAGGAGAATGAAATGG AAGAAGGAGCACAAAGATGGC
TAACACGGTGCCACCTCAGA TACCTCAGGTGATGGCAGAT

Phalangium_opilio_Scr -----
Parasteatoda_tepidariorum_Scr -----
Cupiennius_salei_Scr1 CACCTTCCTCACCATGCA TGGCGAGACT
Cupiennius_salei_Scr2 -----
Centruroides_sculpturatus_Scr2 CATCATCATCCTCACCATCT
ACACAGCGAG
Centruroides_sculpturatus_Scr1 CATCATCACCATCTTCACAA
TGAAACCAAA
Mesobuthus_martensii_Scr1 CATCATCACCATCTTCACAA TGAAACCAAA
Mesobuthus_martensii_Scr2 CACCACCATCCTCACCATCT ACACAGCGAG
;
END;

#NEXUS

BEGIN DATA;
DIMENSIONS NTAX=8 NCHAR=699;
FORMAT DATATYPE=DNA INTERLEAVE MISSING=-;

[DATASET S6: Nucleotide alignment of scorpion, spider, and harvestman
Antennapedia sequences.]

MATRIX

Parasteatoda_tepidariorum_Antp1 TACCTTACAGACATCCGCCA

ACGTCAGTACATGCAGCCTC AGTACGCTTCTCGCCAGTA

CAAGGGCGACTTACCCCCG GTTTCCACCTTACGACAGAC

Phalangium_opilio_Antp -----

Cupiennius_salei_Antp -----

Parasteatoda_tepidariorum_Antp2 TACCTTACAGACATCCGCCA

ACGTCAGTACATGCAGCCTC AGTACGCTTCTCGCCAGTA

CAAGGGCGACTTACCCCCG GTTTCCACCTTACGACAGAC

Mesobuthus_martensii_Antp1 -----

Mesobuthus_martensii_Antp2 TATCTTCCGGATATAAGGCA

GAGACAGTACTTGCAACCTC AATATGCATCGTGCCTCG

CAAGGTGCAACCTATCCTCG ATTTCCACCGTACGATCGAC

Centruroides_sculpturatus_Antp1 TATCTGACGGACCTGAGACA

GAGACAGTATCTTCAACCTC AATATGCATCATCACCAGCC

CAAGGTGCCACTTATCCACG CTTTCCACCATATGACCGAT

Centruroides_sculpturatus_Antp2 TACCTCCGGATATAcgaca
AAGACAGTACCTCAACCCCC AATACGCATCATCTCCAGCT
CAAGGTGCTACATACCCACG TTTCCCACCTTACGATCGAT

Parasteatoda_tepidariorum_Antp1 TGGAAATACGACCTATCTCC
CAAAGCCCACCAGGACACTA TTACCAATCTCAGGGCAACA
TGCCCCAACCGGCACATCAG ACACCACAACACACACACCTTT
Phalangium_opilio_Antp -----

Cupiennius_salei_Antp -----

Parasteatoda_tepidariorum_Antp2 TGGAAATACGACCTATCTCG
CAAAGCCCACCAGGACACTA TTACCAATCTCAGGGCAACA
TGCCCCAACCTGCACATCAG ACACCGCAACACACACACCTTT

Mesobuthus_martensii_Antp1 -----

Mesobuthus_martensii_Antp2 TAGAAATTAGGCCATCTCA
CCGAGCCCTCCAGGGCAGTA TTATCAACAACCACCTTCGC
ATCCTCAG-----CAGCAG CAACAACAACAGGCACCACT

Centruroides_sculpturatus_Antp1 TGGAAATTCTGTCCCATATCT
CCAAGTCCCCCGGGCATT A TTATCAGCAGCCA-----
CCAGGTCCCATGCCACCA CAACCTCAGCATGCACCACT

Centruroides_sculpturatus_Antp2 TAGAAATTAGGCCCATCTCC
CCGAGCCCACCTGGTCAATA TTATCACAGGCCACCTTCTC
ATCCTCAGCAGCAGCCACAA CAACAACAACACGCACCACT

Parasteatoda_tepidariorum_Antp1 AACACATCCTAACGCCTATG
TACCACAGGATGGTCAGAAC TGTAGGGGCTCCCCGACGGA
AAGTTGCAAACCTAACAGC AACCGGTACATCGGCCTGTG
Phalangium_opilio_Antp -----AACAAATTACG

TTCCTCAGGACGCGCAGAAC TGTCGGGGTTCCCCAACGA
TAGTTGCAAATGCAACAGG GACCGGTCAACCGGGGAGGG
Cupiennius_salei_Antp -----
---ATGCAGCAGC AGCAGGTTCACAGGCCCGTG

Parasteatoda_tepidariorum_Antp2 AACACATCCTAACGCCTATG
TACCACAGGATGGTCAGAAC TGTAGGGGTTCCCCGACGGA
AAGTTGCAAACCTAACAGC AACCGGTACATCGGCCTGTG

Mesobuthus_martensii_Antp1 -----

Mesobuthus_martensii_Antp2 CAATCATCCAAATGCATACG
TGCCACAGGATGGTCAGAAC TGCGTGGATCACCTAACGG
CAGTTGCAAATTACAGGGGC AACAA----CAGCCACAG

Centruroides_sculpturatus_Antp1 CACTCACCCGAACGCTTACG
TACCTCAGGATGGACAAAAT TGCGTGGTTCTCCTAACGA
TAGTTGCAAGATGCAACAAAC AACCTGTCCAGCGACCTTG

Centruroides_sculpturatus_Antp2 CAATCATCCCAATGCCTATG
TGCGCAGGATGGTCAAAAT TGCGTGGATCACCCAACGG
TAGTTGCAAATTACAGGGGC AACAA-----CAACCACAG

Parasteatoda_tepidariorum_Antp1 AATCCGGATTGTGCTGCCAA
CAATATGCATCCGCAACACT GTCAAAGTCCTGTACATTCC
CCGAATCACGTACAACAACA GCCTCCAATAACCAAGTTC
Phalangium_opilio_Antp GTACCAGTCGATTGTCCGAA
CAACATTCAACCAGCAG---T GTCAGAGCCC GGTTCACTCC
AACCCCGCCGTGGCTCAAGG AAACCCCGGCAACCCAGCTC
Cupiennius_salei_Antp AACCCAGACTGTGCGGGCAA
TAACATGCACCCCCAGCATT GTCAAAGCCCTGTGCATTG
CCAAATCATGTGCCACAGCA ACCACCGAATAACCAAGTTC
Parasteatoda_tepidariorum_Antp2 AATCCGGATTGTGCTGCCAA
CAATATGCATCCGCAACACT GTCAAAGTCCTGTACATTCC
CCGAATCACGTACAACAACA GCCTCCAATAACCAAGTTC
Mesobuthus_martensii_Antp1 -----

Mesobuthus_martensii_Antp2 ATGCAAGACTGTGGCCCTCC
TGCACCTCCACCGCCTCATT GTCAGAACCCATCGGCG---
CCAAATCATATGCACCAATC TAATCAAGCAAATCAATTAC

Centruroides_sculpturatus_Antp1 ACCTCTGACTGCCCACCTAA
CAACATGCACCAGGCCCTT GCCAGAGTCCTGTACATTCT
CCCAACCACGTGACTCCCTC CAATCCCAACCAGCAGGTCC
Centruroides_sculpturatus_Antp2 ATGCAAGACTGTGGTCCCC TGCA---
CCTCCTCCTCATT GTCAGAATCCCACGGGG---
CCTAATCATATGCACCAATC CAATCAACCAAACCAACTAC

Parasteatoda_tepidariorum_Antp1 AACCAGCTAGTGGTGGA--- ---
CCTAGTCCTTGTATCC ATGGATGCGGAGTCAGTTG
AGAGGAAACGTGGTCGGCAG ACCTACACTCGCTACCAGAC
Phalangium_opilio_Antp AAACCCCCAACTCCGGATCT
TTACCCAGTCGCTATAACCC CTGGATGAGAAGTCAGTTG
AAAGGAAAAGGGTCTGCAA ACGTACACTCGTTACCAAAC
Cupiennius_salei_Antp AGTCTCCGAGCGGAGGA--- ---
CCGAGTCCATTATATCC ATGGATGAGGAGTCAGTCG
AAAGGAAAAGAGGGCGGCAA ACTTACACTCGTTACCAGAC
Parasteatoda_tepidariorum_Antp2 AACCAGCTAGTGGTGGA--- ---
CCTAGTCCTTGTATCC ATGGATGCGGAGTCAGTTG
AGAGGAAACGTGGTCGGCAG ACCTACACTCGCTACCAGAC
Mesobuthus_martensii_Antp1 -----
G AACGAAAGCGCGGACGCCAG ACCTATAACCGCTACCAGAC

Mesobuthus_martensii_Antp2 ACACGCAGACTTCCAGCACA
ATGCCTAGTCCTCTATATCC GTGGATGAGAAGTCAATTG
AGAGGAAAAGAGGGAGGCAG ACTTATAACCGCTACCAAAC
Centruroides_sculpturatus_Antp1 AGCCTCCGAGCTCGGGCAAC
TTGCCAAGTCCACTGTATCC ATGGATGAGGAGTCAGTTG
AACGAAAGCGCGGACGCCAG ACCTATAACCGCTACCAGAC
Centruroides_sculpturatus_Antp2 ACACGCAGACTTCGAGTACA
ATGCCTAGTCCTTGATCC ATGGATGAGAAGTCAATTG
AAAGGAAAAGAGGGAGGCAG ACATACACCCGTTACCAAAC

Parasteatoda_tepidariorum_Antp1 CCTTGAATTAGAGAAGGAAT
TTCATTCAATCGCTACTTG ACGCGTCGGCGTCGGATCGA
GATCGCTCACACTCTTGCT TGACGGAGCGACAGATCAAG
Phalangium_opilio_Antp TCTGGAACTGGAGAAAGAAT
TTCACCAACCGATATTAA ACAAGACGGCGGCGGATTGA
AATAGCGCATGCGCTTGCC TGACGGAAAGGCAGATAAAG
Cupiennius_salei_Antp TCTAGAGTTAGAAAAAGAAT
TCCACCAACCGTTATCTT ACACGTAGACGCCGAATCGA
AATAGCACATGCCCTTGCC TCACCGAGCGCCAGATCAAG
Parasteatoda_tepidariorum_Antp2 CCTTGAATTAGAGAAGGAAT
TTCATTCAATCGCTACTTG ACGCGTCGGCGTCGGATCGA
GATCGCTCACACTCTTGCT TGACGGAGCGACAGACCAAG

Mesobuthus_martensii_Antp1 ATTGGAATTGGAGAAAGAGT
TCCACTTCAACCGCTATCTG ACGCGGAGAAGAAGGATAGA
GATAGCTCACGCTTATGTC TAACGGAACGTCAGATCAAG

Mesobuthus_martensii_Antp2 ACTGGAATTAGAGAAAGAGT
TCCACTTAATCGATATTAA ACCCGGAGGAGAAGAATCGA
AATAGCCCACGCCCTTGCT TAACCGAACGACAAATCAAG

Centruroides_sculpturatus_Antp1 ATTGGAATTGGAGAAAGAGT
TCCACTTCAACCGCTATCTG ACGCGGAGAAGAAGGATAGA
AATAGCTCACGCTCTATGTC TAACGGAACGTCAGATCAAG

Centruroides_sculpturatus_Antp2 ATTGGAATTAGAGAAAGAAT
TCCACTTAATCGATACTTA ACCCGGAGAAGAAGAATCGA
AATAGCTCACGCCCTTGCT TAACCGAACGACAAATCAAG

Parasteatoda_tepidariorum_Antp1 ATCTGGTTCCAGAACCGACG
CATGAAGTGGAAAAAAGAAA ACAAGTCAAAGATGGAAGCA
GGCCTAGTCATGGGACCTGG TGGCCCGAATTGGTCCAC

Phalangium_opilio_Antp ATCTGGTTCCAAAATCGGCG
GATGAAATGGAAGAAGGAAA ACAAAAGCGAAACTAGAAGCC
GGATTGCCATGGGTGCCAC GCCTCCGAACTAACGCAT

Cupiennius_salei_Antp ATCTGGTTCCAGAACAGAAG
GATGAAGTGGAAAAAGGAAA ATAAAGCTAACACCAGCAGCT
GGATTCAATTGGAAACA-- -----GAACTGGTTCAC

Parasteatoda_tepidariorum_Antp2 ATCTGGTCCAGAACCGACG
CATGAAGTGGAAAAAAGAAA ACAAGTCAAAGATGGAAGCA
GGCCTAGTCATGGGACCTGG TGGCCCGAATTGGTCCAC

Mesobuthus_martensii_Antp1 ATATGGTTTCAGAATCGGAG
GATGAAATGGAAAAAAGAAA ATAAAGCTAAATTGGAAGCT
GGTCTGGCATTGGGACCAGC TCCTCCGAACTAACACAC

Mesobuthus_martensii_Antp2 ATATGGTCCAAAACAGAAG
AATGAAATGGAAAAAAGAAA ACAAAAGCTAAACTGGAAGCA
GGACTTCTATGGGAGGCC TACACCAGAATTAACACTC

Centruroides_sculpturatus_Antp1 ATATGGTTTCAGAATCGGAG
GATGAAATGGAAAAAAGAAA ATAAAGCTAAACTGGAAGCT
GGCCTGGCATTAGGCCAGC TCCACCGAACTAACACAC

Centruroides_sculpturatus_Antp2 ATCTGGTCCAAAACAGAAG
AATGAAGTGGAAAAGGAGA ACAAAAGCCAACTGGAAGCG
GGACTTCAATGGGAGGACC TACACCTGAATTAACACTG
;
END;

#NEXUS

BEGIN DATA;

DIMENSIONS NTAX=8 NCHAR=315;
FORMAT DATATYPE=DNA INTERLEAVE MISSING=-;

[DATASET S7: Nucleotide alignment of scorpion, spider, and harvestman
Ultrabithorax sequences.]

MATRIX

Centruroides_sculpturatus_Ubx2 CAAGGCCAGCCAACGCTTT
CTATCCGTGGATGGCCATTG CA-----GGTCCAAC
GGTTTGCAGAGACGAGGTCG CCAAACCTACACACGTTACC
Centruroides_sculpturatus_Ubx1 CAAGTACCCACCAACGCATT
CTATCCCTGGATGGCTATCG CAGGTCTAAAGGGGCAAAC
GGTTTAAGAAGGAGAGGAAG GCAGACATAACCCGATATC
Phalangium_opilio_Ubx CAAACGCCAACCAACGCTTT
CTACCCCTGGATGGCCATCG CCGGTTATTCCCGCGCTAAC
GGTTTACGGCGTCGAGGACG ACAAACGTACACTCGGTACC
Parasteatoda_tepidariorum_Ubx CAAACTCAGAGCAACACATT
TTATCCATGGATGGCCGTAC AAGGTCCA-----AAT
GGCATGAGGAGAAGAGGTCG CCAAACCTACACACGTTACC
Cupiennius_salei_Ubx1 -----GCCAAC
GGACTGCGTCGCAGGGGCCG CCAGACCTACACCGCTACC

Cupiennius_salei_Ubx2 GCGGATAATTACAAGAAGCT
GTACGAAATATTGTGGCTTG AGAGTTATTATGGAGCCAAT
GGAGTCCGACGACGAGGTCG ACAAACGTACACCCGCTACC

Mesobuthus_martensii_Ubx1 CAAGTGCCCACCAACGCTTT
CTATCCCTGGATGGCTATCG CA-----GGAGCAAAT
GGTTTAAGAAGAAGAGGAAG ACAGACATACACCCGATATC

Mesobuthus_martensii_Ubx2 -----
GCTAAC GGTTTGCAGACGAGGTCG CCAAACCTACACACGTTATC

Centruroides_sculpturatus_Ubx2 AAACCTGGAGCTAGAGAAA
GAGTTCATACCAACCATTAA TCTGACACGGAGAAGAAGAA
TCGAAATGGCACACGCCCTA TGTCTAACAGAAAGACAGAT

Centruroides_sculpturatus_Ubx1 AAACCCTGGAATTAGAAAAAA
GAGTTCACACGAACCATTAA TCTGACACGGCGAAGAAGGA
TCGAAATGGCACATGCCCTG TGTCTGACAGAACGACAGAT

Phalangium_opilio_Ubx AAACGCTGGAATTGGAGAAG
GAATTCCACACGAACCATTAA CCTCACTCGGAGACGAAGGA
TCGAGATGGCGCACCGCGCTA TGCCTTACCGAGAGACAGAT

Parasteatoda_tepidariorum_Ubx AGACTCTGGAACTAGAAAAAA
GAATTCCACACCAATCATTAA CCTAACACGAAGAAGACGGA
TTGAAATGGCGCATTCCCTT TGTTAACAGAGCGCAAAT

Cupiennius_salei_Ubx1 AAACCCTGGAGCTGGAGAAG
GAGTTCCATACCAATCACTA CCTGACCCGACGGAGACGCA
TTGAGATGGCCCACGCCCTT TGCCTCACCGAACGCCAAAT

Cupiennius_salei_Ubx2 AAACTTGGAACTCGAGAAA
GAGTTCCACACCAACCAACTA CCTAACACGCCGTCGGAGAA
TCGAAATGGCCCACTCCCTG TGTCTCACCGAACGCCAAAT

Mesobuthus_martensii_Ubx1 AAACCTTGGATTAGAAAAAA
GAGTTTCACACGAACCATT A TCTGACACGGCGAAGAAGGA
TCGAAATGGCTCATGCCCTG TGTCTGACAGAACGACAGAT

Mesobuthus_martensii_Ubx2 AAACCTTGGAGCTGGAGAAA
GAGTTCCATACCAATCATTA TCTGACACGGAGAAGAAGAA
TCGAAATGGCACATGCTCTT TGCCTAACAGAAAGACAGAT

Centruroides_sculpturatus_Ubx2 CAAGATATGGTTTCAGAATC
GAAGGATGAAACTGAAGAAA GAGATTCAAGCTATCAAAGA
GCTAAATGAACAAGAAAGAC AGGCACAGGCAGCTAAAGTA
Centruroides_sculpturatus_Ubx1 CAAGATCTGGTTCCAAAATA
GGAGGATGAAGTTGAAAAAAA GAGATCCAAGCCATCAAAGA
ACTGAATGAACAAGAGAGAC AGTCCAAAGGATCTTCCCAA
Phalangium_opilio_Ubx TAAAATATGGTTCCAAAATC
GACGGATGAAATTGAAAAAG GAAATTCAAGGCCATCAAGGA
ACTCAACGAGCAGGAAAGAC AGGCGCAGGCCTCGAAGGTT

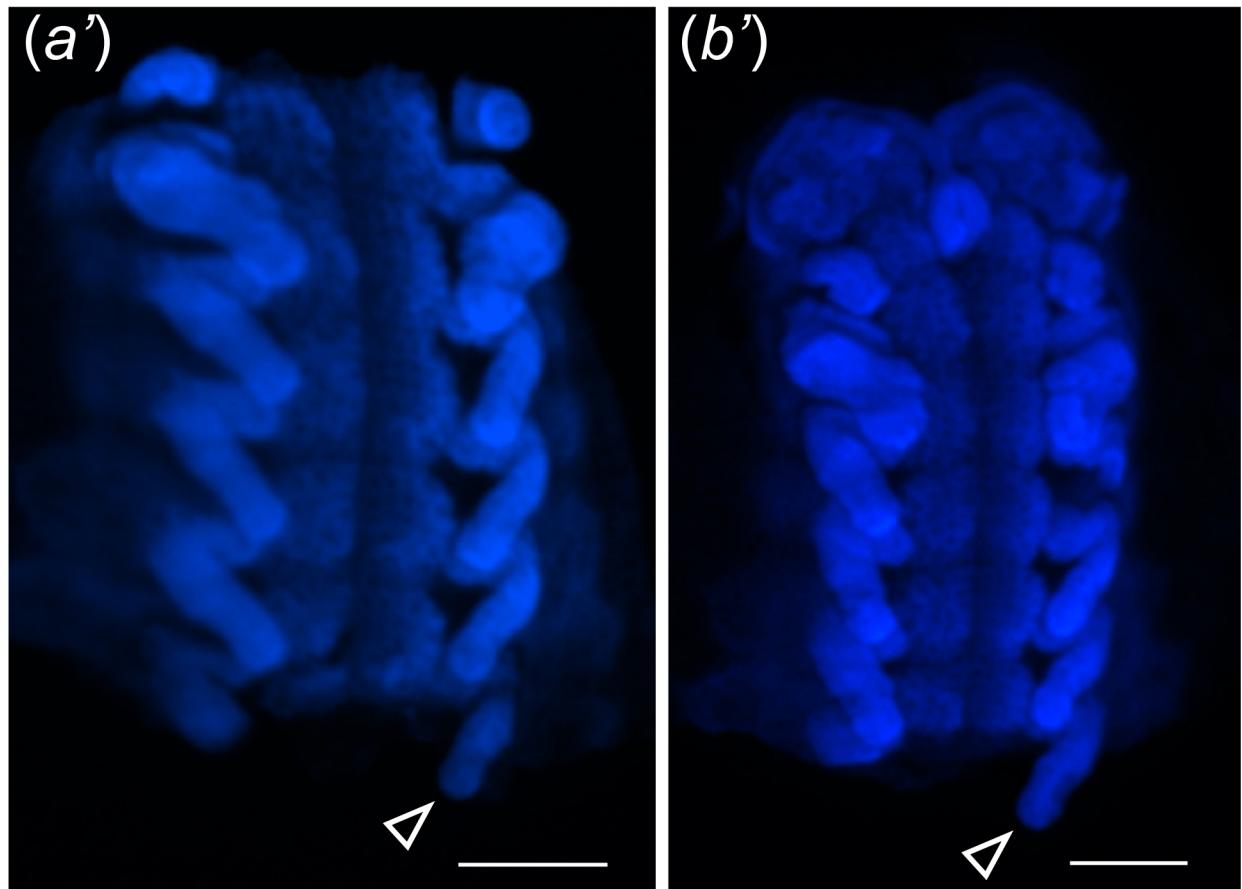
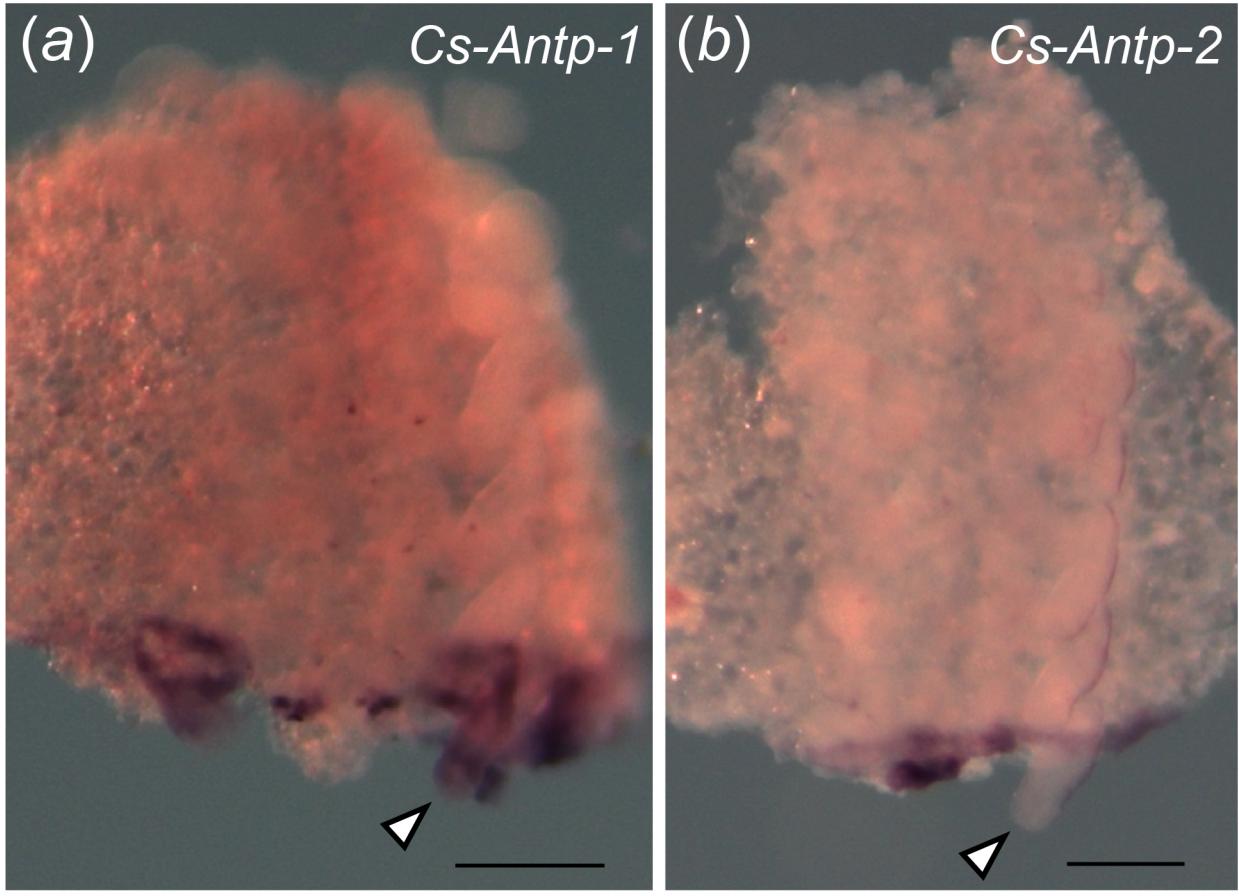
Parasteatoda_tepidariorum_Ubx TAAGATTGGTTTCAGAAC
GGCGGATGAAATTGAAGAAA GAAATTCAAGGCCATCAAAGA
ATTGAGTGAACAAGAACGCC AATCTCAACAAGCCAAGGCC
Cupiennius_salei_Ubx1 TAAGATTGGTTCCAGAAC
GGAGGATGAAGCTCAAGAAA GAGATACAGGCCATCAAGGA
ACTCAACGAACAAGAGAGGC AAGCCCAGGCCGCTAAGCTG
Cupiennius_salei_Ubx2 CAAGATATGGTTTCAGAAC
GGAGGATGAAGCTCAAGAAC AGAGCTCAGGCCATCAAAGA
GCTAAACGAACAAGAACGCC AGGCCAAGCTGAAAAACT
Mesobuthus_martensii_Ubx1 CAAGATCTGGTTCCAAAATA
GGAGGATGAAGTTGAAAAAAA GAGATCCAAGGCCATCAAAGA
ACTGAATGAACAGGAGAGAC AGTCCAAAGGATCTTCCCAA
Mesobuthus_martensii_Ubx2 CAAGATATGGTTCCAAAATC
GAAGGATGAAATTGAAGAAA GAGATTCAAGCTATCAAAGA
GCTAAATGAACAAGAACAGAC AAGCACAGGCAGCTAAAGTA

Centruroides_sculpturatus_Ubx2 GCAGCACACAGCAG
Centruroides_sculpturatus_Ubx1 GGGAACCCAGAAC
Phalangium_opilio_Ubx GTCTCCTCCCATAAT
Parasteatoda_tepidariorum_Ubx ACTACCACTGTATCA
Cupiennius_salei_Ubx1 GCCGCGCACCAAAAA
Cupiennius_salei_Ubx2 GCCTCAACTTCGACT
Mesobuthus_martensii_Ubx1 GGGAACCCAGAAC

Mesobuthus_martensii_Ubx2 GCAGCACAAACAGCAG

;

END;



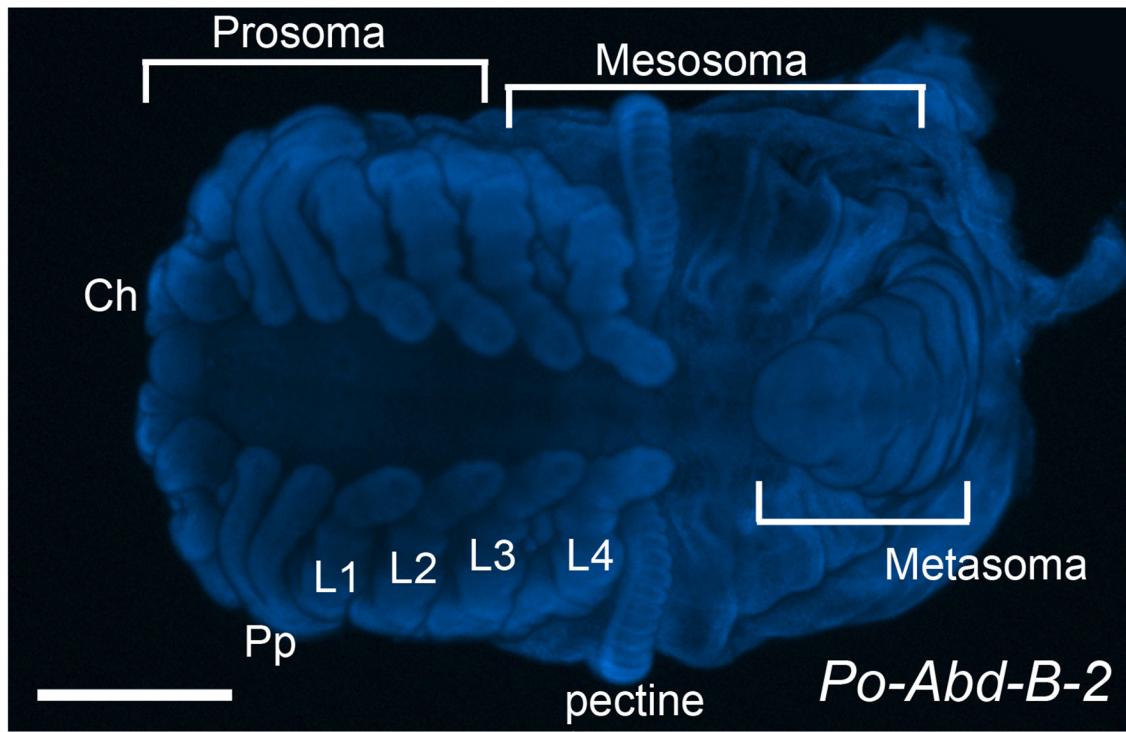
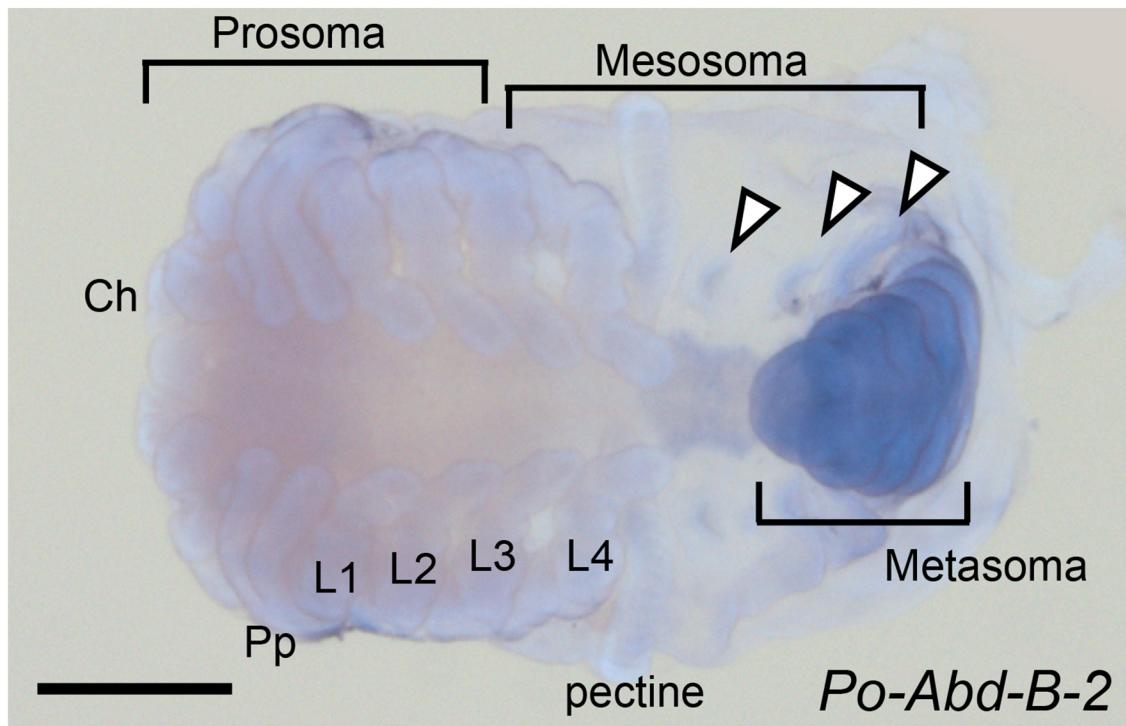


Table S1

Gene	Primer Name	Primer Sequence	Amplicon Length
<i>Cscu-Antp-1</i>	Cscu_Antp1_for Cscu_Antp1_rev	CTCAGCATGCACCACTCACT CAGGCCAGCTTCCAGTTAG	689 bp
<i>Cscu-Antp-2</i>	Cscu_Antp2_for Cscu_Antp2_rev	TACCTTCAACCCCCAATACGC CAAAGGGCGTGAGCTATTTC	635 bp
<i>Cscu-Ubx-1</i>	Cscu_Ubx1_for Cscu_Ubx1_rev	GTCGGCTCACGATAACCACCT CCTTGATTCTGGTTCCCTTG	609 bp
<i>Cscu-Ubx-2</i>	Cscu_Ubx2_for Cscu_Ubx2_rev	ACCGCGTGCAAACCTCTACTC TGCAGATGCCTGATTGCTAC	615 bp
<i>Cscu-abdA-1</i>	Cscu_abdA1_for Cscu_abdA1_rev	GGCTGATATCGGTTCGTGTGTT CAGCTGGGTGTTGAGAAT	658 bp
<i>Cscu-abdA-2</i>	Cscu_abdA2_for Cscu_abdA2_rev	GCCTTGTCTCCTCAGTCGTC TTGCTCTGTTGTCGTCGTC	600 bp
<i>Cscu-AbdB-1</i>	Cscu_AbdB1_for Cscu_AbdB1_rev	GCCAAGAGAGTGACCGCTAC CCCATCGTTCTGTTCGAT	688 bp
<i>Cscu-AbdB-2</i>	Cscu_AbdB2_for Cscu_AbdB2_rev	CARACNTGGTGYAAYTA GRAACCADATYTTNACYT	408 bp

Table S2.

Order	Species	<i>lab</i>	<i>pb</i>	<i>Hox3</i>	<i>Dfd</i>	<i>Scr</i>	<i>ftz</i>	<i>Antp</i>	<i>Ubx</i>	<i>abdA</i>	<i>AbdB</i>
Acariformes	<i>Archegozetes longisetosus</i>	1	1	1	1	1	1	1	1	0	1
Araneae	<i>Cupiennius salei</i>	1	1	1	2	2	1	1	2	1	1
Araneae	<i>Parasteatoda tepidariorum</i>	2	1	1	2	1	1	2	1	1	1
Opiliones	<i>Phalangium opilio</i>	1	1	1	1	1	1	1	1	1	1
Scorpiones	<i>Centruroides sculpturatus</i>	2	2	1	1	1	1	1	1	1	1
Scorpiones	<i>Mesobuthus martensii</i>	2	2	1	2	2	1	2	2	1	1