EvoDevo

CrossMark

Composition and genomic organization of arthropod Hox clusters

Ryan M. Pace[1,4] ⓘ, Miodrag Grbić[2,3] and Lisa M. Nagy[1*]

Abstract

Background: The ancestral arthropod is believed to have had a clustered arrangement of ten Hox genes. Within arthropods, Hox gene mutations result in transformation of segment identities. Despite the fact that variation in segment number/character was common in the diversification of arthropods, few examples of Hox gene gains/losses have been correlated with morphological evolution. Furthermore, a full appreciation of the variation in the genomic arrangement of Hox genes in extant arthropods has not been recognized, as genome sequences from each major arthropod clade have not been reported until recently. Initial genomic analysis of the chelicerate *Tetranychus urticae* suggested that loss of Hox genes and Hox gene clustering might be more common than previously assumed. To further characterize the genomic evolution of arthropod Hox genes, we compared the genomic arrangement and general characteristics of Hox genes from representative taxa from each arthropod subphylum.

Results: In agreement with others, we find arthropods generally contain ten Hox genes arranged in a common orientation in the genome, with an increasing number of sampled species missing either *Hox3* or *abdominal-A* orthologs. The genomic clustering of Hox genes in species we surveyed varies significantly, ranging from 0.3 to 13.6 Mb. In all species sampled, arthropod Hox genes are dispersed in the genome relative to the vertebrate *Mus musculus*. Differences in Hox cluster size arise from variation in the number of intervening genes, intergenic spacing, and the size of introns and UTRs. In the arthropods surveyed, Hox gene duplications are rare and four microRNAs are, in general, conserved in similar genomic positions relative to the Hox genes.

Conclusions: The tightly clustered Hox complexes found in the vertebrates are not evident within arthropods, and differential patterns of Hox gene dispersion are found throughout the arthropods. The comparative genomic data continue to support an ancestral arthropod Hox cluster of ten genes with a shared orientation, with four Hox gene-associated miRNAs, although the degree of dispersion between genes in an ancestral cluster remains uncertain. *Hox3* and *abdominal-A* orthologs have been lost in multiple, independent lineages, and current data support a model in which inversions of the *Abdominal-B* locus that result in the loss of *abdominal-A* correlate with reduced trunk segmentation.

Keywords: Hox, Evolution, Development, Arthropod, Chelicerate, *Tetranychus*, *Ixodes*, *Daphnia*, Segmentation

Background

The Hox genes are a highly conserved set of homeodomain transcription factors that function in fundamental developmental processes in metazoans. This conservation also extends to their genomic arrangement [1–3]. They are typically found clustered in the genome, in the same transcriptional orientation, with their anterior–posterior (A/P) domains of expression and function mirroring their genomic position, commonly referred to as spatial collinearity [4–6], although in rare occurrences Hox genes are not clustered and are spread throughout the genome [7–9]. Among metazoans, Hox genes have been intensely studied in arthropods and it is hypothesized that a genomic cluster of ten genes (*labial, proboscipedia, Hox3, Deformed, Sex combs reduced, fushi tarazu, Antennapedia, Ultrabithorax, abdominal-A,* and *Abdominal-B*)

*Correspondence: lnagy@email.arizona.edu
[1] Department of Molecular and Cellular Biology, University of Arizona, Tucson, AZ 85721, USA
Full list of author information is available at the end of the article

was present ancestrally within the clade [1], a hypothesis further supported by data from a member of the Onychophora, the sister clade to arthropods, that has a full complement of ten Hox genes [10].

While gene-based surveys in phylogenetically diverse arthropods support a general conservation of Hox genes in extant arthropods [11, 12], comparatively little is known about the conservation of the genomic arrangement of Hox genes (i.e., genomic clustering) throughout the phylum as the majority of data comes from insects. Hox genes in insects are generally positioned relatively close to each other on the same chromosome in the same transcriptional orientation, but have different amounts of intervening genes and intergenic space [13–16] (Fig. 1). However, there are several exceptions to this. For example, the *Drosophila melanogaster* Hox genes are split

between the *Antennapedia* and *Ultrabithorax* complexes, separated by a large gap (~9.7 Mb) [17, 18] (Fig. 1). Only one other non-Drosophilid, the silk moth *Bombyx mori*, is known to contain a split genomic arrangement of Hox genes, where a large genic gap (~12 Mb) exists between *labial* and the rest of the Hox cluster [19, 20] (Fig. 1). Additionally, inversions that disrupt the transcriptional orientation of Hox genes are found in several taxa, including *Drosophila* and *Anopheles gambiae* [13, 14, 17, 18, 21, 22] (Fig. 1). These fragmented and inverted organizations appear unusual among insects.

There have been few descriptions of the genomic organization of arthropod Hox genes outside of insects. Two of the better characterized sets of Hox genes from genomic assemblies in non-insects include the myriapod *Strigamia maritima* [23] and the chelicerate *Tetranychus*
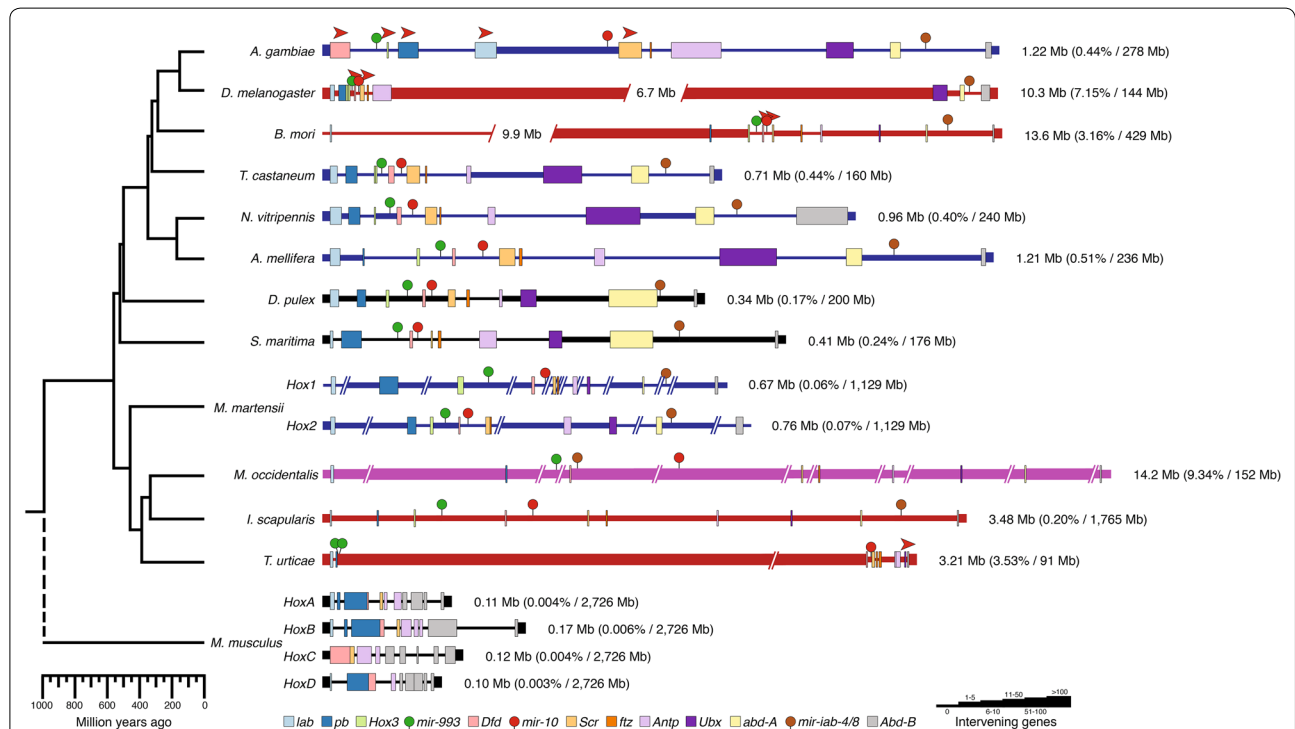

Fig. 1 Overall size and genomic organization of arthropod Hox genes varies. On the *left* is a representative phylogenetic tree depicting relationships among the arthropod taxa used in the comparative analysis, as depicted in [32, 100–102]. *Mus musculus* is used as an out-group. *Colored boxes* represent Hox genes (and mice homologs according to [103]) and miRNAs, with numbers to the *right* of the *black line* indicating approximate size of the genomic region displayed for individual taxa. All Hox genes are depicted in the same transcriptional orientation, except where indicated with a *red arrowhead*. Data on inversions within the *Anopheles* Hox cluster are conflicting, as published data only show a single microinversion of *ftz* [13, 14, 21]; however, the most recent genome assembly shows a large inversion from *labial* to *Deformed* as depicted here. *Broken lines* indicate large genomic spans in *Bombyx* (12 Mb from *lab* to *pb*, with 9.9 Mb removed here for ease of view), *Drosophila* (9.7 Mb from *Antp* to *Ubx*, with 6.7 Mb removed here for ease of view), and *Tetranychus* (2.9 Mb from *pb* to *Dfd*). The *Anopheles*, *Tribolium*, *Nasonia*, *Apis*, and *Mesobuthus* Hox gene clusters are depicted at 1/2 scale (denoted by the *horizontal blue line*), the *Drosophila*, *Bombyx*, *Ixodes*, and *Tetranychus* Hox gene clusters at 1/6 scale (denoted by the *horizontal red line*), and the *Metaseiulus* Hox gene cluster at 1/20 scale (denoted by the *horizontal purple line*). The number of intervening protein coding genes between Hox genes is indicated by *horizontal line thickness*. *Numbers* to the *right* indicate the respective length of the Hox clusters in the genome in megabase pairs (Mb), as calculated from the transcriptional start of the most 3′ Hox gene to the transcriptional stop of the most 5′ Hox gene, and the proportion of the genome that contains the Hox cluster is indicated as a percentage along with the genome size in parentheses, respectively
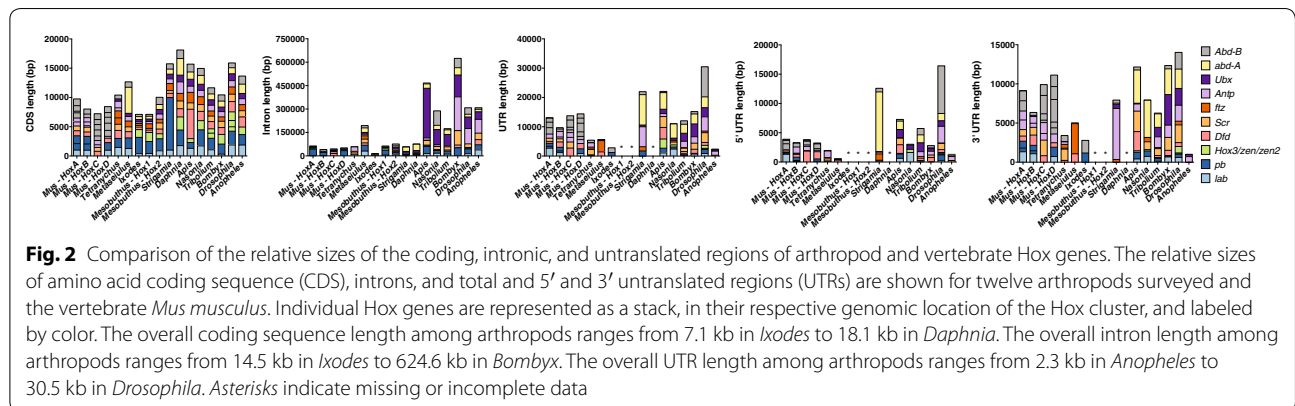
urticae [24]. The *Strigamia* genome contains nine of the ten canonical arthropod Hox genes grouped together (missing *Hox3*), in the same transcriptional orientation [23] (Fig. 1). In *Tetranychus*, there is a large gap between *proboscipedia* and *Deformed*, duplications of *fushi tarazu* and *Antennapedia*, losses of *Hox3* and *abdominal-A*, and an inversion of *Abdominal-B* [24] (Fig. 1). These data from non-insect species suggest there is more variation in the genomic organization of arthropod Hox genes than previously appreciated. It is perhaps unsurprising then that an incomplete set of the canonical ten arthropod Hox genes is often found when performing degenerate PCR surveys [11, 25].

While the general paradigm is that most arthropods contain a similar set of ten Hox genes, there are examples where Hox gene loss or duplication (and divergence) has occurred [11, 25]. The best studied of the examples of gene loss are the *fushi tarazu* (*ftz*) and *Hox3* genes [26–31]. In the case of the *ftz* gene, loss of homeotic function, without a loss of the gene itself, appears to have been relatively common and there is currently no evidence that its loss of homeotic function had any phenotypic consequence [32]. However, here we report that loss of *abd-A* in *Tetranychus* [24], the oribatid mite *Archegozetes longisetosus* [11], a pycnogonid [33], and three species of barnacle [34–36]—all of which have very reduced trunk segmentation—suggest some Hox gene losses correlate with discrete morphological change along the A/P body axis.

In contrast to the uncommon association of a loss of a Hox gene with morphological variation, there are numerous examples where morphological diversification along the A/P body axis is achieved through changes in the regulation of the Hox genes, leading to variation in the A/P expression boundaries of the Hox genes [37, 38] (see [12] for review), as well as to changes in their downstream targets [39, 40] (see [41] for review). The intergenic regions between Hox genes are thus important

sites of regulation [42] and also include microRNAs (miRNAs), small non-coding RNAs, known to play essential roles in Hox gene regulation [43]. While there is no consensus on the number of conserved arthropod Hox gene-associated miRNAs, more than twenty are functionally annotated in *Drosophila melanogaster* between the Hox genes *labial* and *Abdominal-B* [44], with four of these miRNAs—*miR-993*, *miR-10*, *miR-iab-4*, and *miR-iab-8*—found in conserved positions within arthropod genomes [45–51] (Fig. 2). The last two, *miR-iab-4* and *miR-iab-8*, reside at the same locus and produces sense (*miR-iab-4*) and antisense (*miR-iab-8*) transcripts [52, 53]. However, these views of arthropod Hox gene regulation are mainly derived from insects, leaving relatively little known about the extent of miRNA conservation and divergence throughout the Arthropoda.

With representative sequenced genomes from each major arthropod clade now available, we can examine the variation in the genomic organization and evolution of arthropod Hox gene clusters within a greater phylogenetic context, including the degree to which chromosomal arrangement, transcriptional orientation, and regulatory elements such as miRNAs are conserved. Using a comparative genomic approach, we examined how the spatial organization of Hox genes has changed during arthropod evolution. To this end, we performed a comparative analysis of the genomic structure of Hox genes that have been previously characterized in six insects, the centipede *Strigamia maritima* (Myriapoda) [23], several chelicerates including the scorpion *Mesobuthus martensii*, the spider mite *Tetranychus*, the predatory mite *Metaseiulus occidentalis*, and the vertebrate *Mus musculus*. In addition, we included in our comparative analysis the water flea *Daphnia pulex* (Crustacea) [54] and the deer tick *Ixodes scapularis* (Chelicerata) [55], arthropods that have yet to have the genomic organization of their Hox genes well characterized.

Fig. 2 Comparison of the relative sizes of the coding, intronic, and untranslated regions of arthropod and vertebrate Hox genes. The relative sizes of amino acid coding sequence (CDS), introns, and total and 5′ and 3′ untranslated regions (UTRs) are shown for twelve arthropods surveyed and the vertebrate *Mus musculus*. Individual Hox genes are represented as a stack, in their respective genomic location of the Hox cluster, and labeled by color. The overall coding sequence length among arthropods ranges from 7.1 kb in *Ixodes* to 18.1 kb in *Daphnia*. The overall intron length among arthropods ranges from 14.5 kb in *Ixodes* to 624.6 kb in *Bombyx*. The overall UTR length among arthropods ranges from 2.3 kb in *Anopheles* to 30.5 kb in *Drosophila*. *Asterisks* indicate missing or incomplete data

Methods

Hox genes and genomic sequence collection

Species choice

We focused our analysis on non-insect Hox clusters for which less information on Hox cluster organization has been published. Reliable analysis of genomic organization of a Hox cluster requires sufficiently long contigs and confidence in the accuracy of the assembly. Consequently, we limited our analysis to those genomes with sufficient quality coverage and assembly (at least 7X coverage with Sanger sequencing, or 15X coverage with 454 pyrosequencing) and for which the data were publically available. At present, six of the published genomes of non-insect arthropods adhere to that standard and are included in our analysis. We note, however, that for none of these species is there a chromosome linkage map; so validating contig linkage relationships is not possible. We also included representatives of the major clades of insects for which a completed sequence was available and focused on the best-quality genome assemblies (6 species).

Databases used: *Tetranychus urticae* data were collected from the Online Resource for Community Annotation of Eukaryotes (OrcAE, http://bioinformatics.psb.ugent.be/orcae/) [56]. *Daphnia pulex* data were collected from the JGI Daphnia pulex v1.0 genome database (http://genome.jgi-psf.org) [54]. *Strigamia maritima* data were collected from the EnsemblGenomes database (http://metazoa.ensembl.org/) [23]. *Ixodes scapularis* and *Anopheles gambiae* data were collected from the VectorBase IscaW1.4 and AgamP4 databases, respectively (https://www.vectorbase.org) [55, 57, 58]. *Bombyx mori* data were collected from the Silkworm Genome Database (http://www.silkdb.org/silkdb) [59] and from the work of [19, 20]. *Apis mellifera* data were collected from the Hymenoptera Genome Database (http://hymenopteragenome.org/) [60]. *Drosophila melanogaster*, *Nasonia vitripennis*, *Tribolium castaneum*, *Metaseiulus occidentalis*, *Mesobuthus martensii*, and *Mus musculus* data were collected from NCBI. Gene accession identification numbers are included in Additional file 1: Table S1.

Annotation of conserved arthropod Hox gene-associated miRNAs

In identifying conserved miRNAs, we undertook a sequence homology-based approach that was consistent with previous studies [45, 61–64]. Briefly, Hox gene-associated miRNAs were first curated from miRBase [44] by searching the genomic positions of the Hox genes for *Bombyx*, *Drosophila*, *Anopheles*, *Apis*, *Nasonia*, *Tribolium*, *Daphnia*, and *Ixodes*. *Drosophila* Hox gene-associated miRNA stem-loop sequences were then downloaded from miRBase and used in BLAST analyses to identify conserved Hox gene-associated miRNAs in *Tetranychus*,

Mesobuthus, *Metaseiulus*, and *Strigamia*. Identification of putative precursor miRNA sequences was based on a BLAST hit with an alignment length greater than or equal to 20 nucleotides and greater than or equal to 80 percent identity (Additional file 2: Table S4). BLAST hits were then analyzed for predicted secondary structure using minimum free energy (MFE) with RNAfold [65] and miRAlign [61]. Only sequences with a predicted stem-loop structure with a MFE less than or equal to -20 kcal/mol and that contained the mature sequence on the stem were considered putative miRNAs (Additional file 2: Tables S4, Additional file 3: Table S5). Predicted precursor miRNA sequences were aligned to other arthropod miRNA sequences using MUSCLE [66], trimmed, and subsequently used in generating a maximum likelihood phylogenetic tree [67] to support orthology as previously described [68], using the HKY85 substitution model (Additional file 4: Figure S1; Additional file 5). Accession numbers or genomic locations of miRNAs are included in Additional file 1: Table S1.

Results

Comparison of the genomic arrangement of arthropod Hox genes

In Fig. 1, we compare key features of the genomic organization of the insect, non-insect and mouse Hox genes, including relative position on the chromosome, intergenic spacing, number of intervening genes, transcriptional orientation, and the position of Hox gene-associated miRNAs.

Overall size of the Hox cluster: None of the arthropod Hox clusters analyzed show the tightly linked genomic organization as seen in the vertebrate *Mus musculus,* in which Hox clusters range from 0.10 to 0.17 Mb in size (Fig. 1 and Additional file 6: Table S2). *Daphnia* contains the tightest linked arrangement of Hox genes observed in any arthropod with a sequenced genome to date (0.34 Mb; Fig. 1; Additional file 6: Table S2). The ten *Daphnia* Hox genes are located on a single genomic scaffold (scaffold 7), which measures 2.3 Mb in length. The *Ixodes* Hox genes are also located on a single genomic scaffold (DS891538), which measures 3.9 Mb in length, but the *Ixodes* Hox genes span ten times the genomic distance of the *Daphnia* Hox genes (3.48 Mb; Additional file 6: Table S2). The increase in the spatial arrangement of the *Ixodes* Hox cluster is correlated with its particularly large genome (1.7 Gb) [57] relative to other arthropods with sequenced genomes, which range from 91 Mb in *Tetranychus* [24] to 1.3 Gb in *Aedes aegypti* [69]. However, the *Ixodes* Hox cluster occupies a similar percentage of the genome (0.20 %) as other arthropod Hox clusters that lack large genomic gaps (0.3 ± 0.2 %, mean ± standard deviation).

Splits in the Hox cluster

The *Drosophila* Hox cluster is the best-known example of a split Hox cluster, in which the *Antennapedia* complex is separated from the genes in the *Bithorax* complex by 9.6 Mb. A similarly large split is evident in the *Bombyx* Hox cluster (12 Mb), but separates *labial* from the rest of the Hox genes [17–20]. We also observed a split in the Hox gene organization of *Tetranychus* [24]. The *Tetranychus* Hox genes are located on two genomic scaffolds: A 2.7-Mb scaffold (genomic scaffold 11) contains orthologs of *proboscipedia* (*Tu-pb*) and *labial* (*Tu-lab*) in a shared 5′ to 3′ orientation, and a 1.6-Mb scaffold (genomic scaffold 20) contains the remaining Hox genes *Deformed* (*Tu-Dfd*), *Sex combs reduced* (*Tu-Scr*), *fushi tarazu* (paralogs 20g02520 and 20g02530, respectively), *Antennapedia* (paralogs 20g02430 and 20g02440, respectively), *Ultrabithorax* (*Tu-Ubx*), and *Abdominal-B* (*Tu-Abd-B*) (Fig. 1). At this time it is not possible to further align the genomic scaffolds into a contiguous sequence due to the holocentric nature of the three chromosomes of *Tetranychus* [24]. Presuming the scaffolds are contiguous in the genome, the gap between the *Tu-pb* locus and the *Tu-Dfd* locus is a minimum distance of ~2.9 Mb (2.47 Mb from the stop codon in *pb* to the 3′ end of scaffold 11 and 0.47 Mb from start codon in *Tu-Dfd* to the 5′ end of scaffold 20) and contains more than 100 predicted and manually annotated genes (Fig. 1). Thus, the *Tetranychus* Hox cluster split is unique in both where it is located (between *pb* and *Dfd*) and the number of genes interspersed in the region (Fig. 1). However, when the region between *Tu-pb* and *Tu-Dfd* is excluded, the length from the most 5′ to the most 3′ Hox gene in *Tetranychus* is reduced compared to most insects. For example, the *Scr* and *ftz* genes are a mere 1.1 kb apart (Fig. 1 and Additional file 6: Table S2).

We note that for species where Hox genes are found on separate genomic scaffolds (*Tetranychus*, *Metaseiulus*, and *Mesobuthus*), we have arranged the genomic scaffolds to reflect the expected genomic organization of Hox genes (Fig. 1). These data should be interpreted to represent the *minimum* genomic arrangement, and chromosome linkage mapping is needed to confirm this arrangement. It is possible that there are large genomic regions between Hox gene-containing scaffolds in both the scorpion and predatory mite, and an even larger than depicted gap between the *Tetranychus pb* and *Dfd* genes, but until the scaffolds can be further linked it is not possible to ascertain.

Transcriptional orientation

With very few exceptions, the Hox genes share the same transcriptional orientation on their respective chromosomes. In our sample, inversions are limited to the derived insects, *Drosophila*, *Bombyx*, and *Anopheles* [17–20]; and the *Abd-B* gene in *Tetranychus* [24] (Fig. 1). The inversion of *Tu-Abd-B* is also found in two other tetranychid mites (*T. lintearius and T. evansi*, separated by 0.8 and 3 MYA, respectively; data not shown). This inversion is consistent with a model in which the loss of *abdominal-A* resulted from a chromosomal inversion that spanned both the *abd-A* and *Abd-B* loci at the base of this lineage.

Lineage-specific Hox gene duplications

As previously reported [24], several instances of Hox gene duplications were identified in *Tetranychus*. *Tetranychus* harbors two copies of both *fushi tarazu* and *Antennapedia* as tandem duplications, not present in other arthropods (Fig. 1). Needleman–Wunsch global alignment of the coding sequence shows the *Tu-ftz1 and Tu-ftz2* and *Tu-Antp1* and *Tu-Antp2* orthologs share 74 and 55 percent identity, respectively, and at the amino acid level 67 and 33 percent identity, respectively (Additional file 7). Phylogenetic analysis of representative arthropod *ftz* and *Antp* sequences supports a hypothesis that the *Tetranychus ftz* and *Antp* paralogs emerged via lineage-specific duplications (Additional file 8: Figure S2). The percent identity for the *Tetranychus* duplicate *ftz* and *Antp* orthologs is consistent with another well-known tandem gene duplication located in an arthropod Hox cluster; e.g., *Tribolium zerknüllt* and *zerknüllt-2* share 54 % nucleotide identity and 37 % protein identity, respectively. Alignment of the *Tetranychus ftz* and *Antp* amino acid sequences suggests either partial duplications of the genes including the homeodomain-containing region, or deletions occurred post-duplication, upstream of the homeodomain in *Antp2440* and *ftz2530* as the putative coding sequence for both the *ftz* and *Antp* duplicates is substantially smaller (Additional file 8: Figure S2). Additionally, comparison of RNAseq profiles across four developmental stages (RNASeq from embryonic, larval, nymphal, and adult) reveal both *Antp2440* and *ftz2530* have markedly lower levels of expression than their respective paralogs (Additional file 9: Table S6). The scorpion genome revealed the presence of two complete Hox clusters (Fig. 1) [70]. Surveys of the remaining non-insect genomes sampled revealed no apparent Hox gene duplications.

Individual Hox gene characteristics

To further explore variation in Hox gene genomic complexes, we characterized the structure of the *Tetranychus*, *Daphnia*, and *Ixodes* Hox genes and compared them to *Anopheles*, *Drosophila*, *Bombyx*, *Tribolium*, *Nasonia*, *Apis*, *Strigamia*, *Metaseiulus*, *Mesobuthus*, and *Mus musculus* (Fig. 2). In general, we find chelicerates to have

reduced amounts of coding sequence, intronic regions, and untranslated regions (Fig. 2). There is little variation in the coding sequence of myriapod, crustacean, and insect Hox genes. Compared to insects, chelicerates, *Strigamia*, and *Daphnia* contain reduced Hox gene intronic regions (Fig. 2). Similarly, where there are available data on the 5′ and 3′ untranslated regions (UTRs), the UTRs of chelicerate and *Daphnia* Hox genes are reduced compared to insects. For example, the average total intron length and intron number for Hox genes in *Tetranychus* are 5.32 kb and 2.0 introns, in *Ixodes* are 1.45 kb and 0.9 introns, and in *Daphnia* are 7.61 kb and 4.0 introns (Fig. 2 and Additional file 10: Table S3). In contrast, the average total intron length and intron number for Hox genes in *Drosophila* are larger at 25.66 kb and 3.8 introns (Fig. 2 and Additional file 10: Table S3). Similarly, the average length of untranslated regions of *Tetranychus* and *Ixodes* Hox genes (0.49 and 0.31 kb, respectively) is reduced compared to *Drosophila* (2.54 kb) and other arthropods (Fig. 2 and Additional file 10: Table S3). The decrease in untranslated regions and introns of individual Hox genes effectively reduces the size of the overall transcription units in *Tetranychus* and *Ixodes* (Fig. 2 and Additional file 6: Table S3). In *Tetranychus*, these features are not specific to the Hox cluster as a reduction of intergenic regions, introns, and untranslated regions is general features of the highly compact *Tetranychus* genome [24]. Nonetheless, the reduction in putative regulatory DNA associated with *Tetranychus* Hox genes has implications for their regulation during development.

Hox gene-associated miRNAs

We identified multiple sequences with homology to the four conserved miRNAs (i.e., *mir-993*, *mir-10*, *mir-iab-4*, and *mir-iab-8*) in *Tetranychus* (3/4), *Mesobuthus* (4/4), *Metaseiulus* (4/4), and *Strigamia* (4/4) (Fig. 1, Additional file 4: Figure S1; Additional file 2: Table S4, Additional file 3: Table S5). There are two sequences in *Tetranychus* separated by less than 500 bp that contain homology to *mir-993* located near *pb* (Fig. 1). As *mir-993* is typically found conserved between *Hox3* and *Dfd*, the position of the *Tetranychus mir-993* homologs adjacent to *Tu-pb* provides additional support for the complete loss of *Hox3* from the genome. In *Tetranychus mir-10* is located upstream of *Tu-Dfd*, differing from its expected conserved location between *Dfd* and *Scr* (Fig. 1). We were unable to locate *mir-iab-4/mir-iab-8* in *Tetranychus*, which typically overlap each other in the region between *abd-A* and *Abd-B* (Fig. 1). The absence of *mir-iab-4/mir-iab-8* from *Tetranychus* further validates the loss of *abd-A* from the *Tetranychus* genome. All four *Mesobuthus* and *Strigamia* miRNAs were found in the expected genomic positions: *mir-10* is located between *Dfd* and

Scr, *mir-iab-4/mir-iab-8* is located between *abd-A* and *Abd-B*. In *Mesobuthus*, *mir-993* lies between *Hox3* and *Dfd*, whereas in *Strigamia*, despite the loss of *Sm-Hox3* from its conserved position between *pb* and *Dfd* (Fig. 1), *mir-993* is located between *pb* and *Dfd* (Fig. 1). *Metaseiulus* contains a unique arrangement of Hox gene-associated miR-NAs; both *mir-993* and *mir-10* are found on non-Hox gene-containing scaffolds, whereas *mir-iab-4/mir-iab-8* is located on the *Dfd*-containing scaffold (Fig. 1).

Discussion

Changes in the genomic organization of arthropod Hox genes

To understand the conservation of Hox gene clustering throughout the arthropod phylum, we performed an arthropod-wide comparison of the genomic structure of Hox genes from representative arthropod genome sequences. We analyzed and compared the genomic organization of Hox genes from several insects, a crustacean, a myriapod, four chelicerates, and the vertebrate *Mus musculus*. Our data suggest the constraints maintaining Hox genes in a tightly clustered genomic complex have been lost during arthropod evolution (Fig. 1), although the sample size remains relatively small (e.g., there are an estimated 2–10 million arthropod species [71]). Excluding arthropods with Hox genes organized into distinct genomic complexes (e.g., *Drosophila*, *Bombyx*, and *Tetranychus*), the variation in clustering of Hox genes is mainly attributed to the amount of intervening space between Hox genes, not to a lack of Hox gene coding sequence, intron, and UTR lengths. For example, intervening space accounts for 0.25 Mb (72 %) of the *Daphnia* Hox gene cluster. In contrast, intervening space accounts for 3.45 Mb (99 %) of the Hox gene cluster in *Ixodes*. There does not appear to be a phylogenetic trend for proportion of intervening sequence. Furthermore, large gaps in the genomic arrangement of Hox genes have independently occurred in insects and chelicerates with sequenced genomes to date (Fig. 1). These large breaks do not appear at consistent locations within the cluster (Fig. 1). For example, the genomic split within the *Tetranychus* Hox genes represents a novel arrangement of Hox genes compared to those in dipteran and lepidopteran genomes.

The predatory mite *Metaseiulus occidentalis* was reported to have a disintegrated arrangement of Hox genes in the genome—each Hox gene is localized to a separate genomic scaffold, that if contiguous on a single chromosome would span approximately 12.1–14.2 Mb [72] (Fig. 1). While *Metaseiulus*, *Tetranychus*, and *Ixodes* are all are members of the chelicerate subclass Acari, *Metaseiulus* is more closely related to *Ixodes* (both are members of the Parasitiformes superorder, and

Tetranychus is a member of the Acariformes superorder). The *Ixodes* Hox genes were uncovered on a single large scaffold, but the Hox genes span 3.5 Mb of the genome and none of the genes are tightly clustered in this region (Fig. 1). Thus, these data suggest that compared to other arthropods, there may have been fewer constraints on mites and ticks to maintain the paradigmatic clustered arrangement of Hox genes.

While the mechanisms that constrain Hox genes into genomic clusters during evolution are not well understood, they are believed to be tied to their temporal collinearity in vertebrates [73]. The evolution of rapid development and simultaneous appearance of segments in flies (22 h [74]) has been proposed as a key factor in the loss of temporal collinearity and thus disintegration of Hox gene clusters [9]. Although it not known whether *Tetranychus* Hox genes are temporally collinear, *Tetranychus* embryonic development is also relatively short (~39 h [75]). The development of *Bombyx mori*, however, contradicts this model. While *Bombyx* have a split arrangement of Hox genes, *Bombyx* embryos take 10 days to complete embryogenesis [76]. Interestingly, in several insects and in *Strigamia*, expression of the posterior class Hox genes is temporally collinear [77–80]. Taken together, these data, or more so, lack of data, reveal that the temporal expression of Hox genes in arthropods is under-appreciated and under-studied. It is possible that shifts in the temporal collinearity of Hox gene expression, like anteroposterior shifts in expression domains, should also be considered as a driver of morphological change.

Conservation of miRNAs in arthropod Hox clusters

Four miRNAs located in close proximity to Hox genes are found conserved throughout Arthropoda—*mir-993*, *mir-10*, *mir-iab-4*, and *mir-iab-8*. To date there has only been one arthropod identified that contains duplicated copies. *Bombyx* has two copies of *mir-993*—*mir-993a* is present in the Hox cluster in the expected position between *pb*/*Hox3* and *Dfd*, whereas *mir-993b* is located on a genomic scaffold separate from the scaffolds that contain the Hox genes [51]. Similarly, *Tetranychus* contains two copies of *mir-993*, although both are in the expected position downstream of *pb* (Fig. 1). However, the *mir-993* duplications in both *Bombyx* and *Tetranychus* likely represent lineage-specific duplications, as they do not cluster together in a phylogenetic analysis (Additional file 4: Figure S1). We also found *Tetranychus mir-10* has translocated from a conserved position between *Dfd* and *Scr* to upstream of *Tu-Dfd* (Fig. 1). *Strigamia* contains a single copy of the four conserved Hox gene-associated miRNAs, and we did not find any evidence for translocation (Fig. 1). We speculate that the losses of *mir-iab-4* and *mir-iab-8* from the *Tetranychus* genome reflect the
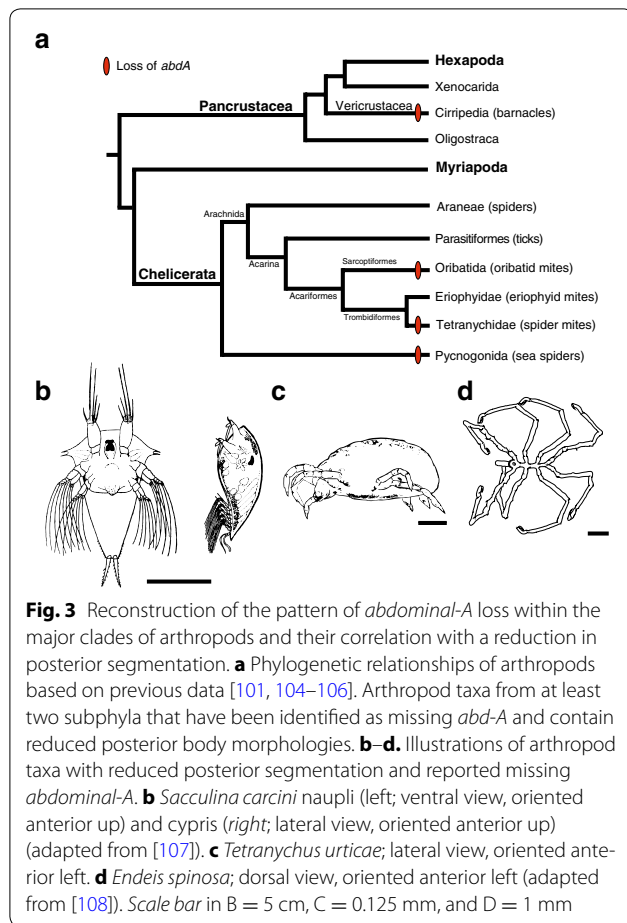
loss of the nearby Hox gene *abd-A*. Notably, in *Drosophila*, *mir-iab-4* appears to be transcriptionally co-regulated with *abd-A* and *mir-iab-8* plays a regulatory role in the expression of *abd-A* and *Ubx* and maintaining posterior segment identities during early patterning [53, 81]. In *Tetranychus*, the loss of the region containing *abd-A*, *mir-iab-4*, and *mir-iab-8* may have contributed to the reduced posterior patterning observed (described in further detail below). However, an examination of the expression and function of *mir-iab-4* and *mir-iab-8* in more phylogenetically basal arthropods, such as chelicerates, remains to be determined.

Hox gene duplications

Despite all extant arthropods containing the same basic set of Hox genes [1, 12], there are an increasing number of examples of lineage-specific Hox gene duplications. Duplications of Hox genes have been previously observed via PCR surveys in several non-insect arthropods, e.g., *Scr*, *Dfd*, and *Ubx*, in the spider *Cupiennius salei* [82] and *Dfd* in the centipede *Pachymerium ferrugineum* [83]. Additionally, PCR and genomic surveys in horseshoe crabs suggest the entire Hox gene complex has undergone several duplications, potentially via whole genome duplication [84, 85], and similar large-scale duplications of Hox genes have been observed in scorpions [70, 86] (Fig. 1). Where available, analysis of gene expression data reveals the duplicated genes to have overlapping, but distinct spatial domains [82, 86]. We also observed Hox gene duplications in *Tetranychus* for *ftz* and *Antp*. Whether the *Tetranychus* duplicated Hox genes also contain overlapping, but distinct spatial domains of expression awaits examination. However, both *Tetranychus* duplicated Hox genes are truncated in length suggesting potential pseudogenization of the loci. Consistent with a model of pseudogenization, both *Tetranychus* duplicated Hox genes have lower RNASeq expression profiles compared to their respective paralogs (Additional file 9: Table S6), typical of a gene duplication event in which one duplicate is free to lose function retained by the other duplicate [87–89].

Loss of *abd-A* and evolution of arthropod posterior segmentation

In several arthropods there is a correlation between a reduced number of posterior segments and an absence of the posterior class Hox gene *abd-A*. *abd-A* was not identified in PCR surveys of Hox genes in three cirripede crustaceans (*Sacculina carcini*, *Elminius modestus*, and *Trypetesa lampas*) [34–36] and two chelicerates, including the pycnogonid *Endeis spinosa* [33] and the oribatid mite *Archegozetes* [11] (Fig. 3). The addition of *Tetranychus* to this list suggests that this correlation has

Fig. 3 Reconstruction of the pattern of *abdominal-A* loss within the major clades of arthropods and their correlation with a reduction in posterior segmentation. **a** Phylogenetic relationships of arthropods based on previous data [101, 104–106]. Arthropod taxa from at least two subphyla that have been identified as missing *abd-A* and contain reduced posterior body morphologies. **b–d.** Illustrations of arthropod taxa with reduced posterior segmentation and reported missing *abdominal-A*. **b** *Sacculina carcini* naupli (left; ventral view, oriented anterior up) and cypris (*right*; lateral view, oriented anterior up) (adapted from [107]). **c** *Tetranychus urticae*; lateral view, oriented anterior left. **d** *Endeis spinosa*; dorsal view, oriented anterior left (adapted from [108]). *Scale bar* in B = 5 cm, C = 0.125 mm, and D = 1 mm

emerged independently four times during arthropod evolution, once, in the lineage leading to the cirripedes, and three times within Chelicerata (Fig. 3). Although both *Tetranychus* and the oribatid mite are members of the acariformes lineage of chelicerates, the oribatid mites group more closely with the eriophyids, a group of mites that possess two pairs of legs and an elongated opisthosoma comprised of an uneven number of dorsal and ventral cuticular annuli that are not regarded as segments [90–92]. Based on the opisthosomal morphology of eriophyids, it might be expected that *abd-A* is present within this taxa and that loss of *abd-A* has occurred independently at least three times in the Chelicerata. However, a PCR survey of Hox genes and a molecular analysis of posterior body segments in eriophyids with a segmental marker such as *engrailed* have yet to be completed.

What role might the loss of *abd-A* have in the evolution of trunk morphology? It is possible that the loss of *abd-A* has no role in reducing the number of posterior segments and is merely a function of relaxed selection due to overlapping function with other Hox genes, e.g., *Ubx*

and *Abd-B*. This may be similar to the loss, or derivation, of *Hox3* in insects [93]. Alternatively, the loss of *abd-A* could be a key evolutionary event underlying the reduction in the number of posterior segments. In vertebrate somitogenesis, which is superficially similar to arthropod sequential segmentation, termination of segmentation is promoted by the onset of expression of the vertebrate *Abdominal-B* homologs *Hoxb13* and *Hoxc13* [94, 95]. This late onset of expression is entwined with the temporal collinearity of vertebrate Hox genes, in which dynamic shifts in the three-dimensional chromatin arrangements within the Hox cluster occur during anteroposterior patterning [96, 97]. Again, while it is not well known whether arthropod Hox genes exhibit similar temporal controls overall, *Abd-B* is known to have a late onset of expression during segmentation in the apterygote *Thermobia domestica* and the orthopteran *Schistocerca gregaria* [77, 78]. In addition, during pupation *Drosophila Abd-B* contributes to a reduction of posterior segments in a sex-specific manner [98, 99]. Taken together these data suggest that in arthropods that display reduced posterior segmentation, the loss of *abd-A* may have contributed to a change in the temporal chromatin dynamics of Hox clusters, transcriptional regulation of *Abd-B*, and an early termination of segmentation.

Conclusions

The current data, from both genomic studies and PCR surveys, remain consistent with the idea that an ancestral arthropod had ten Hox genes. However, the accumulating non-insect Hox data raises questions as to whether ancestrally multiple Hox clusters were present and subsequently lost, and to what degree the ancestral arthropod clusters were dispersed within the genome. Vertebrate Hox genes are tightly clustered on the same chromosome with virtually no non-Hox genes interspersed within the cluster. Arthropod Hox genes are located on the same chromosome and in the same order and transcriptional orientation as their vertebrate orthologs, but they show varying degrees of dispersion. The Hox clusters in the genomes of *Daphnia* and *Strigamia* span relatively small genomic regions, while all chelicerates surveyed to date show a significantly more dispersed Hox cluster configuration, with the extreme example of the "atomized" cluster of the predatory mite. Within the insects surveyed here, there is an evident trend toward a greater dispersion of the Hox cluster within more derived species. Taken together, these data suggest the constraints maintaining Hox genes in a genomic cluster in arthropods have been relaxed in comparison with vertebrates and may play a functional role in the reduction of posterior body plans.

Additional files

Additional file 1: Table S1. Gene accession numbers or genomic locations of Hox genes used in this analysis.

Additional file 2: Table S4. Positive BLAST hits and genomic locations of conserved Hox gene-associated miRNAs in *Mesobuthus*, *Metaseiulus*, *Tetranychus*, and *Strigamia*.

Additional file 3: Table S5. Number of BLAST hits, alignments, and sequences considered for conserved Hox gene-associated miRNAs in *Mesobuthus*, *Metaseiulus*, *Tetranychus*, and *Strigamia*.

Additional file 4: Figure S1. Maximum likelihood phylogenetic tree of arthropod Hox gene-associated miRNAs. Abbreviations used: *Dm – Drosophila melanogaster, Ag – Anopheles gambiae, Tc – Tribolium castaneum, Nv – Nasonia vitripennis, Ap – Apis mellifera, Bm – Bombyx mori, Dp – Daphnia pulex, Sm – Strigamia maritima, Mm – Mesobuthus martensii, Is – Ixodes scapularis, Mo – Metaseiulus occidentalis*, and *Tu – Tetranychus urticae*.

Additional file 5. Muscle alignment of arthropod miRNAs *mir-993*, *mir-10*, and *mir-iab-4/8*. Abbreviations used as in Additional file 4: Figure S1.

Additional file 6: Table S2. Calculated spatial relationship of Hox genes.

Additional file 7. Global alignment of *Tetranychus fushi tarazu* and *Antennnapedia* nucleotide and protein sequences.

Additional file 8: Figure S2. Maximum likelihood phylogenetic tree of *Antp* (A) and *ftz* (B). Phylogenetic trees were constructed with PhyML [64] using full length protein sequences aligned with MUSCLE. PhyML parameters were set as the following: WAG amino acid substitution model, proportion of invariable sites estimated, and the number of categories of substitution rate = 4. Statistical support was provided by approximate likelihood ratio tests based on a Shimodaira-Hasegawa-like procedure, with the scores shown in the tree.

Additional file 9: Table S6. RNASeq expression data for *Tetranychus urticae Antennapedia* and *fushi tarazu*. Data taken from embryonic (embryo_techrep_#), larval (larvae_techrep_#), nymphal (nymph_techrep_#), and adult (adult_techrep_#) developmental stages, values are reads per kilobase per million mapped reads (RPKM).

Additional file 10: Table S3. Measurements of the size of Hox gene coding and non-coding regions.

Abbreviations

lab: labial; *pb*: proboscipedia; *zen*: zerknüllt; *Dfd*: Deformed; *Scr*: Sex combs reduced; *ftz*: fushi tarazu; *Antp*: Antennapedia; *Ubx*: Ultrabithorax; *abd-A*: abdominal-A; *Abd-B*: Abdominal-B; *miR*, miRNA: microRNA; *iab*: infra-abdominal; PCR: polymerase chain reaction; BLAST: basic local alignment search tool; MFE: minimum free energy; bp: base pairs; kb: kilobase pairs; Mb: megabase pairs; Gb: gigabase pairs; CDS: coding sequence; UTR: untranslated region; *M. musculus*: *Mus musculus*; Dm: *Drosophila melanogaster*; *A. gambiae*, Ag: *Anopheles gambiae*; *T. castaneum*, Tc: *Tribolium castaneum*; *N. vitripennis*, Nv: *Nasonia vitripennis*; *A. mellifera*, Ap: *Apis mellifera*; *B. mori*, Bm: *Bombyx mori*; *D. pulex*, Dp: *Daphnia pulex*; *S. maritima*, Sm: *Strigamia maritima*; *M. martensii*, Mm: *Mesobuthus martensii*; *I. scapularis*, Is: *Ixodes scapularis*; *M. occidentalis*, Mo: *Metaseiulus occidentalis*; *T. urticae*, Tu: *Tetranychus urticae*.

Authors' contributions

RMP participated in the conception and design of the study, performed data analysis, and wrote the manuscript. MG participated in data discussion and critical revisions of the manuscript. LMN participated in the conception and design of the study, data discussion and contributed to writing the manuscript. All authors read and approved the final manuscript.

Author details

[1] Department of Molecular and Cellular Biology, University of Arizona, Tucson, AZ 85721, USA. [2] Department of Biology, University of Western Ontario, London, ON N6A 5B7, Canada. [3] Universidad de la Rioja, 26006 Logroño, Spain. [4] Present Address: Division of Maternal-Fetal Medicine, Department of Obstetrics and Gynecology, Baylor College of Medicine, Houston, TX 77030, USA.

References

1. Akam M, Averof M, Castelli-Gair J, Dawes R, Falciani F, Ferrier DEK. The evolving role of Hox genes in arthropods. Development. 1994;1994:209–15.
2. Wang BB, Müller-Immergluck MM, Austin J, Robinson NT, Chisholm A, Kenyon C. A homeotic gene cluster patterns the anteroposterior body axis of *C. elegans*. Cell. 1993;74:29–42.
3. Garcia-Fernàndez J, Holland PWH. Archetypal organization of the amphioxus Hox gene cluster. Nature. 1994;370:563–6.
4. McGinnis W, Krumlauf R. Homeobox genes and axial patterning. Cell. 1992;68:283–302.
5. Duboule D, Morata G. Colinearity and functional hierarchy among genes of the homeotic complexes. Trends Genet. 1994;10:358–64.
6. Duboule D, Dollé P. The structural and functional organization of the murine HOX gene family resembles that of Drosophila homeotic genes. EMBO J. 1989;8:1497–505.
7. Seo H-C, Edvardsen RB, Maeland AD, Bjordal M, Jensen MF, Hansen A, Flaat M, Weissenbach J, Lehrach H, Wincker P, Reinhardt R, Chourrout D. Hox cluster disintegration with persistent anteroposterior order of expression in *Oikopleura dioica*. Nature. 2004;431:67–71.
8. Albertin CB, Simakov O, Mitros T, Wang ZY, Pungor JR, Edsinger-Gonzales E, Brenner S, Ragsdale CW, Rokhsar DS. The octopus genome and the evolution of cephalopod neural and morphological novelties. Nature. 2015;524:220–4.
9. Hoy MA, Waterhouse RM, Wu K, Estep AS, Ioannidis P, Palmer WJ, Pomerantz AF, Simão FA, Thomas J, Jiggins FM, Murphy TD, Pritham EJ, Robertson HM, Zdobnov EM, Gibbs RA, Richards S. Genome sequencing of the phytoseiid predatory mite *Metaseiulus occidentalis* reveals completely atomised Hox genes and super-dynamic intron evolution. Genome Biol Evol 2016;evw048. doi:10.1093/gbe/evw048.
10. Janssen R, Eriksson BJ, Tait NN, Budd GE. Onychophoran Hox genes and the evolution of arthropod Hox gene expression. Front Zool. 2014;11:22.
11. Cook CE, Akam M, Smith ML, Telford MJ, Bastianello A. Hox genes and the phylogeny of the arthropods. Curr Biol. 2001;11:759–63.
12. Hughes CL, Kaufman TC. Hox genes and the evolution of the arthropod body plan1. Evol Dev. 2002;4:459–99.
13. Powers TP, Hogan J, Ke Z, Dymbrowski K, Wang X, Collins FH, Kaufman TC. Characterization of the Hox cluster from the mosquito *Anopheles gambiae* (Diptera: Culicidae). Evol Dev. 2000;2:311–25.
14. Devenport MP, Blass C, Eggleston P. Characterization of the Hox gene cluster in the malaria vector mosquito, *Anopheles gambiae*. Evol Dev. 2000;2:326–39.
15. Brown SJ, Fellers JP, Shippy TD, Richardson EA, Maxwell M, Stuart JJ, Denell RE. Sequence of the *Tribolium castaneum* homeotic complex: the region corresponding to the *Drosophila melanogaster* antennapedia complex. Genetics. 2002;160:1067–74.
16. Ferrier DEK, Akam M. Organization of the Hox gene cluster in the grasshopper, *Schistocerca gregaria*. Proc Natl Acad Sci USA. 1996;93:13024–9.
17. Kaufman TC, Lewis R, Wakimoto B. Cytogenetic analysis of chromosome 3 in *Drosophila melanogaster*: the homoeotic gene complex in polytene chromosome interval 84a-B. Genetics. 1980;94:115–33.

18. Duncan I. The bithorax complex. Annu Rev Genet. 1987;21:285–319.
19. Yasukochi Y, Ashakumary LA, Wu C, Yoshido A, Nohata J, Mita K, Sahara K. Organization of the Hox gene cluster of the silkworm, *Bombyx mori*: a split of the Hox cluster in a non-Drosophila insect. Dev Genes Evol. 2004;214:606–14.
20. Chai C-L, Zhang Z, Huang F-F, Wang X-Y, Yu Q-Y, Liu B-B, Tian T, Xia Q-Y, Lu C, Xiang Z-H. Insect biochemistry and molecular biology. Insect Biochem Mol Biol. 2008;38:1111–20.
21. Negre B, Ruiz A. HOM-C evolution in Drosophila: Is there a need for Hox gene clustering? Trends Genet. 2007;23:55–9.
22. Lewis EB. A gene complex controlling segmentation in Drosophila. Nature. 1978;276:565–70.
23. Chipman AD, Ferrier DEK, Brena C, Qu J, Hughes DST, Schröder R, Torres-Oliva M, Znassi N, Jiang H, Almeida FC, Alonso CR, Apostolou Z, Aqrawi P, Arthur W, Barna JCJ, Blankenburg KP, Brites D, Capella-Gutiérrez S, Coyle M, Dearden PK, Du Pasquier L, Duncan EJ, Ebert D, Eibner C, Erikson G, Evans PD, Extavour CG, Francisco L, Gabaldón T, Gillis WJ, et al. The first myriapod genome sequence reveals conservative arthropod gene content and genome organisation in the centipede *Strigamia maritima*. PLoS Biol. 2014;12:e1002005.
24. Grbić M, Van Leeuwen T, Clark RM, Rombauts S, Rouzé P, Grbić V, Osborne EJ, Dermauw W, Ngoc PCT, Ortego F, Hernandez-Crespo P, Diaz I, Martinez M, Navajas M, Sucena É, Magalhães S, Nagy LM, Pace RM, Djuranovic S, Smagghe G, Iga M, Christiaens O, Veenstra JA, Ewer J, Villalobos RM, Hutter JL, Hudson SD, Velez M, Yi SV, Zeng J, et al. The genome of *Tetranychus urticae* reveals herbivorous pest adaptations. Nature. 2011;479:487–92.
25. Abzhanov A, Popadic A, Kaufman TC. Chelicerate Hox genes and the homology of arthropod segments. Evol Dev. 1999;1:77–89.
26. Gibson G. Evolution: hox genes and the cellared wine principle. Curr Biol. 2000;10:R452–5.
27. Löhr U, Yussa M, Pick L. Drosophila fushi tarazu. a gene on the border of homeotic function. Curr Biol. 2001;11:1403–12.
28. Alonso CR, Maxton-Kuechenmeister J, Akam M. Evolution of Ftz protein function in insects. Curr Biol. 2001;11:1473–8.
29. Hsia CC, McGinnis W. Evolution of transcription factor function. Curr Opin Genet Dev. 2003;13:199–206.
30. Stauber M, Prell A, Schmidt-Ott U. A single Hox3 gene with composite bicoid and zerknullt expression characteristics in non-Cyclorrhaphan flies. Proc Natl Acad Sci USA. 2002;99:274–9.
31. Stauber MM, Jäckle HH, Schmidt-Ott UU. The anterior determinant bicoid of Drosophila is a derived Hox class 3 gene. Proc Natl Acad Sci USA. 1999;96:3786–9.
32. Heffer A, Shultz JW, Pick L. Surprising flexibility in a conserved Hox transcription factor over 550 million years of evolution. Proc Natl Acad Sci USA. 2010;107:18040–5.
33. Manuel M, Jager M, Murienne J, Clabaut C, Guyader HL. Hox genes in sea spiders (Pycnogonida) and the homology of arthropod head segments. Dev Genes Evol. 2006;216:481–91.
34. Gibert JM, Mouchel-Vielh E, Quéinnec E, Deutsch JS. Barnacle duplicate engrailed genes: divergent expression patterns and evidence for a vestigial abdomen. Evol Dev. 2000;2:194–202.
35. Blin M, Rabet N, Deutsch JS, Mouchel-Vielh E. Possible implication of Hox genes *Abdominal-B* and abdominal-A in the specification of genital and abdominal segments in cirripedes. Dev Genes Evol. 2003;213:90–6.
36. Mouchel-Vielh E, Rigolot C, Gibert JM, Deutsch JS. Molecules and the body plan: the Hox genes of Cirripedes (Crustacea). Mol Phylogenet Evol. 1998;9:382–9.
37. Averof M, Akam M. Hox genes and the diversification of insect and crustacean body plans. Nature. 1995;376:420–3.
38. Averof M, Patel NH. Crustacean appendage evolution associated with changes in Hox gene expression. Nature. 1997;388:682–6.
39. Pavlopoulos A, Akam M. Hox gene Ultrabithorax regulates distinct sets of target genes at successive stages of Drosophila haltere morphogenesis. Proc Natl Acad Sci USA. 2011;108:2855–60.
40. Vachon G, Cohen B, Pfeifle C, McGuffin ME, Botas J, Cohen SM. Homeotic genes of the Bithorax complex repress limb development in the abdomen of the Drosophila embryo through the target gene Distal-less. Cell. 1992;71:437–50.
41. Sanchez-Herrero E. Hox targets and cellular functions. Scientifica. 2013;2013:738257.
42. Karch F, Celniker SE, Weiffenbach B, Peifer M, Bender W, Duncan I, Crosby M, Lewis EB. The abdominal region of the bithorax complex. Cell. 1985;43:81–96.
43. Yekta S, Tabin CJ, Bartel DP. MicroRNAs in the Hox network: an apparent link to posterior prevalence. Nat Rev Genet. 2008;9:789–96.
44. Griffiths-Jones S. miRBase: microRNA sequences, targets and gene nomenclature. Nucleic Acids Res. 2006;34:D140–4.
45. Miura S, Nozawa M, Nei M. Evolutionary changes of the target sites of two microRNAs encoded in the Hox gene cluster of Drosophila and other insect species. Genome Biol Evol. 2011;3:129–39.
46. Wheeler BM, Heimberg AM, Moy VN, Sperling EA, Holstein TW, Heber S, Peterson KJ. The deep evolution of metazoan microRNAs. Evol Dev. 2009;11:50–68.
47. Shippy TD, Ronshaugen M, Cande J, He J, Beeman RW, Levine M, Brown SJ, Denell RE. Analysis of the Tribolium homeotic complex: insights into mechanisms constraining insect Hox clusters. Dev Genes Evol. 2008;218:127–39.
48. Werren JH, Richards S, Desjardins CA, Niehuis O, Gadau J, Colbourne JK, The Nasonia Genome Working Group, Beukeboom LW, Desplan C, Elsik CG, Grimmelikhuijzen CJP, Kitts P, Lynch JA, Murphy T, Oliveira DCSG, Smith CD, Zande LVD, Worley KC, Zdobnov EM, Aerts M, Albert S, Anaya VH, Anzola JM, Barchuk AR, Behura SK, Bera AN, Berenbaum MR, Bertossa RC, Bitondi MMG, Bordenstein SR, et al. Functional and evolutionary insights from the genomes of three parasitoid Nasonia species. Science. 2010;327:343–8.
49. Weaver DB, Anzola JM, Evans JD, Reid JG, Reese JT, Childs KL, Zdobnov EM, Samanta MP, Miller J, Elsik CG. Computational and transcriptional evidence for microRNAs in the honey bee genome. Genome Biol. 2007;8:R97.
50. Lemons D, Paré A, McGinnis W. Three Drosophila Hox complex microRNAs do not have major effects on expression of evolutionarily conserved Hox gene targets during embryogenesis. PLoS ONE. 2012;7:e31365.
51. Liu S, Li D, Li Q, Zhao P, Xiang Z, Xia Q. MicroRNAs of *Bombyx mori* identified by Solexa sequencing. BMC Genom. 2010;11:148.
52. Bae E, Calhoun VC, Levine M, Lewis EB, Drewell RA. Characterization of the intergenic RNA profile at *abdominal-A* and *Abdominal-B* in the Drosophila bithorax complex. Proc Natl Acad Sci USA. 2002;99:16847–52.
53. Stark A, Bushati N, Jan CH, Kheradpour P, Hodges E, Brennecke J, Bartel DP, Cohen SM, Kellis M. A single Hox locus in Drosophila produces functional microRNAs from opposite DNA strands. Genes Dev. 2008;22:8–13.
54. Colbourne JK, Pfrender ME, Gilbert D, Thomas WK, Tucker A, Oakley TH, Tokishita S, Aerts A, Arnold GJ, Basu MK, Bauer DJ, Caceres CE, Carmel L, Casola C, Choi JH, Detter JC, Dong Q, Dusheyko S, Eads BD, Frohlich T, Geiler-Samerotte KA, Gerlach D, Hatcher P, Jogdeo S, Krijgsveld J, Kriventseva EV, Kultz D, Laforsch C, Lindquist E, Lopez J, et al. The ecoresponsive genome of *Daphnia pulex*. Science. 2011;331:555–61.
55. Megy K, Emrich SJ, Lawson D, Campbell D, Dialynas E, Hughes DST, Koscielny G, Louis C, MacCallum RM, Redmond SN, Sheehan A, Topalis P, Wilson D. The VectorBase Consortium: VectorBase: improvements to a bioinformatics resource for invertebrate vector genomics. Nucleic Acids Res. 2011;40:D729–34.
56. Sterck L, Billiau K, Abeel T, Rouze P, Van de Peer Y. ORCAE: online resource for community annotation of eukaryotes. Nat Methods. 2012;9:1041.
57. Gulia-Nuss M, Nuss AB, Meyer JM, Sonenshine DE, Roe RM, Waterhouse RM, Sattelle DB, la Fuente de J, Ribeiro JM, Megy K, Thimmapuram J, Miller JR, Walenz BP, Koren S, Hostetler JB, Thiagarajan M, Joardar VS, Hannick LI, Bidwell S, Hammond MP, Young S, Zeng Q, Abrudan JL, Almeida FC, Ayllón N, Bhide K, Bissinger BW, Bonzon-Kulichenko E, Buckingham SD, Caffrey DR, et al. Genomic insights into the *Ixodes scapularis* tick vector of Lyme disease. Nat Commun. 2016;7:10507.
58. Holt RA. The genome sequence of the malaria mosquito *Anopheles gambiae*. Science. 2002;298:129–49.
59. Mita K, Kasahara M, Sasaki S, Nagayasu Y, Yamada T, Kanamori H, Namiki N, Kitagawa M, Yamashita H, Yasukochi Y, Kadono-Okuda K, Yamamoto K, Ajimura M, Ravikumar G, Shimomura M, Nagamura Y, Shin-i T, Abe H, Shimada T, Morishita S, Sasaki T. The genome sequence of silkworm, *Bombyx mori*. DNA Res. 2004;11:27–35.
60. Elsik CG, Tayal A, Diesh CM, Unni DR, Emery ML, Nguyen HN, Hagen DE. Hymenoptera genome database: integrating genome annotations in HymenopteraMine. Nucleic Acids Res. 2015;44:gkv1208–D800.

61. Wang X, Zhang J, Li F, Gu J, He T, Zhang X, Li Y. MicroRNA identification based on sequence and structure alignment. Bioinformatics. 2005;21:3610–4.

62. Ruan M-B, Zhao Y-T, Meng Z-H, Wang X-J, Yang W-C. Conserved miRNA analysis in *Gossypium hirsutum* through small RNA sequencing. Genomics. 2009;94:263–8.

63. Ikeda KT, Hirose Y, Hiraoka K, Noro E, Fujishima K, Tomita M, Kanai A. Identification, expression, and molecular evolution of microRNAs in the "living fossil" *Triops cancriformis* (tadpole shrimp). RNA. 2015;21:230–42.

64. Jiang J, Lv M, Liang Y, Ma Z, Cao J. Identification of novel and conserved miRNAs involved in pollen development in *Brassica campestris* ssp. chinensi s by high-throughput sequencing and degradome analysis. BMC Genom. 2014;15:1.

65. Gruber AR, Lorenz R, Bernhart SH, Neuböck R, Hofacker IL. The Vienna RNA websuite. Nucleic Acids Res. 2008;36(Web Server issue):W70–4.

66. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004;32:1792–7.

67. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol. 2010;59:307–21.

68. Pace RM, Eskridge PC, Grbić M, Nagy LM. Evidence for the plasticity of arthropod signal transduction pathways. Dev Genes Evol. 2014;224:209–22.

69. Nene V, Wortman JR, Lawson D, Haas B, Kodira C, Tu Z, Loftus B, Xi Z, Megy K, Grabherr M, Ren Q, Zdobnov EM, Lobo NF, Campbell KS, Brown SE, Bonaldo MF, Zhu J, Sinkins SP, Hogenkamp DG, Amedeo P, Arensburger P, Atkinson PW, Bidwell S, Biedler J, Birney E, Bruggner RV, Costas J, Coy MR, Crabtree J, Crawford M, et al. Genome Sequence of *Aedes aegypti*, a Major Arbovirus Vector. Science. 2007;316:1718–23.

70. Di Z, Yu Y, Wu Y, Hao P, He Y, Zhao H, Li Y, Zhao G, Li X, Li W, Cao Z. Genome-wide analysis of homeobox genes from *Mesobuthus martensii* reveals Hox gene duplication in scorpions. Insect Biochem Mol Biol. 2015;61:25–33.

71. Ødegaard F. How many species of arthropods? Erwin's estimate revised. Biol J Linn Soc. 2000;71:583–97.

72. Tschopp P, Zákány J, Tarchini B, Spitz F, Duboule D. Uncoupling time and space in the collinear regulation of Hox genes. PLoS Genet. 2009;5:e1000398.

73. Markow TA, Beall S, Matzkin LM. Egg size, embryonic development time and ovoviviparity in Drosophila species. J Evol Biol. 2009;22:430–4.

74. Ferrier DEK, Minguillón C. Evolution of the Hox/ParaHox gene clusters. Int J Dev Biol. 2003;47:605–11.

75. Grbić M, Khila A, Lee K-Z, Bjelica A, Grbić V, Whistlecraft J, Verdon L, Navajas M, Nagy LM. Mity model: *Tetranychus urticae*, a candidate for chelicerate model organism. BioEssays. 2007;29:489–96.

76. Ueno K, Nagata T, Suzuki Y. Roles of homeotic genes in the *Bombyx* body plan. In: Goldsmith MR, Wilkins AS, editors. Molecular model systems in the Lepidoptera. Cambridge: Cambridge University Press; 1995. p. 165–180.

77. Peterson MD, Rogers BT, Popadic A, Kaufman TC. The embryonic expression pattern of labial, posterior homeotic complex genes and the teashirt homologue in an apterygote insect. Dev Genes Evol. 1999;209:77–90.

78. Kelsh R, Dawson I, Akam M. An analysis of *Abdominal-B* expression in the locust Schistocerca gregaria. Development. 1993;117:293–305.

79. He J. Molecular and genetic analysis of the *Abdominal-B* homolog in the beetle *Tribolium Castaneum*. Ph.D. Thesis, Kansas State University, Manhattan. 1996.

80. Brena C, Chipman AD, Minelli A, Akam M. Expression of trunk Hox genes in the centipede *Strigamia maritima*: sense and anti-sense transcripts. Evol Dev. 2006;8:252–65.

81. Karch F, Bender W, Weiffenbach B. abdA expression in Drosophila embryos. Genes Dev. 1990;4:1573–87.

82. Schwager EE, Schoppmeier M, Pechmann M, Damen WGM. Duplicated Hox genes in the spider *Cupiennius salei*. Front Zool. 2007;4:10.

83. Bastianello A. Hox gene sequences from the geophilomorph centipede *Pachymerium ferrugineum* (C. L. Koch, 1835) (Chilopoda: Geophilomorpha: Geophilidae): implications for the evolution of the Hox class genes of arthropods. Mol Phylogenet Evol. 2002;22:155–61.

84. Cartwright P, Dick M, Buss LW. HOM/Hox type homeoboxes in the chelicerate *Limulus polyphemus*. Mol Phylogenet Evol. 1993;2:185–92.

85. Kenny NJ, Chan KW, Nong W, Qu Z, Maeso I, Yip HY, Chan TF, Kwan HS, Holland PWH, Chu KH, Hui JHL. Ancestral whole-genome duplication in the marine chelicerate horseshoe crabs. Heredity. 2016;116:190–9.

86. Sharma PP, Schwager EE, Extavour CG, Wheeler WC. Hox gene duplications correlate with posterior heteronomy in scorpions. Proc R Soc B. 2014;281:20140661.

87. Lynch M. The evolutionary fate and consequences of duplicate genes. Science. 2000;290:1151–5.

88. Lynch M, Force A. The probability of duplicate gene preservation by subfunctionalization. Genetics. 2000;154:459–73.

89. Ohno S. Evolution by gene duplication. Berlin: Springer; 1970.

90. Keifer HH. Eriophyid studies. Bull Dept Agric State Calif. 1938;27:181–206.

91. Fusco G, Minelli A. Arthropod segmentation and tagmosis. In: Minelli A, Boxshall G, Fusco G, editors. Arthropod biology and evolution. Berlin: Springer; 2013. p. 197–221.

92. Lindquist EE, Sabelis MW, Bruin J. Eriophyid mites: their biology, natural enemies and control, vol. 6. Amsterdam: Elsevier; 1996. p. 1–83.

93. Falciani F, Hausdorf B, Schroder R, Akam M, Tautz D, Denell R, Brown SJ. Class 3 Hox genes in insects and the origin of zen. Proc Natl Acad Sci USA. 1996;93:8479–84.

94. Economides KD, Zeltser L, Capecchi MR. Hoxb13 mutations cause overgrowth of caudal spinal cord and tail vertebrae. Dev Biol. 2003;256:317–30.

95. Young T, Rowland JE, van de Ven C, Bialecka M, Novoa A, Carapuco M, van Nes J, de Graaff W, Duluc I, Freund J-N, Beck F, Mallo M, Deschamps J. Cdx and Hox genes differentially regulate posterior axial growth in mammalian embryos. Dev Cell. 2009;17:516–26.

96. Noordermeer D, Leleu M, Splinter E, Rougemont J, De Laat W, Duboule D. The dynamic architecture of Hox Gene clusters. Science. 2011;334:222–5.

97. Soshnikova N, Duboule D. Epigenetic temporal control of mouse Hox genes in vivo. Science. 2009;324:1320–3.

98. Wang W, Kidd BJ, Carroll SB, Yoder JH. Sexually dimorphic regulation of the Wingless morphogen controls sex-specific segment number in Drosophila. Proc Natl Acad Sci USA. 2011;108:11139–44.

99. Wang W, Yoder JH. Hox-mediated regulation of doublesex sculpts sex-specific abdomen morphology in Drosophila. Dev Dyn. 2012;241:1076–90.

100. Wang DY, Kumar S, Hedges SB. Divergence time estimates for the early history of animal phyla and the origin of plants, animals and fungi. Proc R Soc B. 1999;266:163–71.

101. Misof B, Liu S, Meusemann K, Peters RS, Donath A, Mayer C, Frandsen PB, Ware J, Flouri T, Beutel RG, Niehuis O, Petersen M, Izquierdo-Carrasco F, Wappler T, Rust J, Aberer AJ, Aspöck U, Aspöck H, Bartel D, Blanke A, Berger S, Böhm A, Buckley TR, Calcott B, Chen J, Friedrich F, Fukui M, Fujita M, Greve C, Grobe P, et al. Phylogenomics resolves the timing and pattern of insect evolution. Science. 2014;346:763–7.

102. Jeyaprakash A, Hoy MA. First divergence time estimate of spiders, scorpions, mites and ticks (subphylum: Chelicerata) inferred from mitochondrial phylogeny. Exp Appl Acarol. 2008;47:1–18.

103. Hueber SD, Weiller GF, Djordjevic MA, Frickey T. Improving Hox protein classification across the major model organisms. PLoS ONE. 2010;5:e10820.

104. Regier JC, Shultz JW, Zwick A, Hussey A, Ball B, Wetzer R, Martin JW, Cunningham CW. Arthropod relationships revealed by phylogenomic analysis of nuclear protein-coding sequences. Nature. 2010;463:1079–83.

105. Dabert M, Witalinski W, Kazmierski A, Olszanowski Z, Dabert J. Molecular phylogeny of acariform mites (Acari, Arachnida): strong conflict between phylogenetic signal and long-branch attraction artifacts. Mol Phylogenet Evol. 2010;56:222–41.

106. Dunlop JA, Selden PA. Calibrating the chelicerate clock: a paleontological reply to Jeyaprakash and Hoy. Exp Appl Acarol. 2009;48:183–97.

107. Calman WT. The life of Crustacea. London: METHUEN & CO. LTD; 1911.

108. Brenneis G, Arango CP, Scholtz G. Morphogenesis of *Pseudopallene* sp. (Pycnogonida, Callipallenidae) I: embryonic development. Dev Genes Evol. 2011;221:309–28.