Candidate gene screen for potential interaction partners and regulatory targets of the Hox gene \textit{labial} in the spider \textit{Parasteatoda tepidariorum}

Christoph Schomburg\textsuperscript{1} \cdot Natascha Turetzek\textsuperscript{2} \cdot Nikola-Michael Prpic\textsuperscript{1}

Received: 30 November 2019 / Accepted: 31 January 2020 / Published online: 8 February 2020

Abstract
The Hox gene \textit{labial} (\textit{lab}) governs the formation of the tritocerebral head segment in insects and spiders. However, the morphology that results from \textit{lab} action is very different in the two groups. In insects, the tritocerebral segment (intercalary segment) is reduced and lacks appendages, whereas in spiders the corresponding segment (pedipalpal segment) is a proper segment including a pair of appendages (pedipalps). It is likely that this difference between \textit{lab} action in insects and spiders is mediated by regulatory targets or interacting partners of \textit{lab}. However, only a few such genes are known in insects and none in spiders. We have conducted a candidate gene screen in the spider \textit{Parasteatoda tepidariorum} using as candidates \textit{Drosophila melanogaster} genes known to (potentially) interact with \textit{lab} or to be expressed in the intercalary segment. We have studied 75 \textit{P. tepidariorum} genes (including previously published and duplicated genes). Only 3 of these (\textit{proboscipedia-A} (pb-A) and two paralogs of \textit{extradenticle} (exd)) showed differential expression between leg and pedipalp. The low success rate points to a weakness of the candidate gene approach when it is applied to lineage specific organs. The spider pedipalp has no counterpart in insects, and therefore relying on insect data apparently cannot identify larger numbers of factors implicated in its specification and formation. We argue that in these cases a de novo approach to gene discovery might be superior to the candidate gene approach.

Keywords Hox gene \cdot \textit{Labial} \cdot Tritocerebral segment \cdot Arthropod head evolution \cdot Head segmentation

Introduction
One of the main reasons for the evolutionary success of arthropods is the diversity of their appendages (Williams and Nagy 2001; Angelini and Kaufman 2005a). The serially homologous segments of these animals can bear appendages adapted for different purposes (reviewed in Prpic and Damen 2008; Jockusch 2017). While these structures share many developmental features, such as their regionalisation into proximal, median and distal domains, the evolution and development of different morphologies between species and among serially homologous appendages along the body axis of a species is not yet well understood (Angelini and Kaufman 2005a). One of the basic decisions for the development of particular appendage morphologies is the positional information derived from the expression of Hox genes along the antero-posterior axis. For instance, the presence of \textit{Antennapedia} (\textit{Antp}) discriminates between antennal and walking leg fate in the vinegar fly \textit{Drosophila melanogaster} (Casares and Mann 2001), and the absence of \textit{Ultrabithorax} (\textit{Ubx}) in the isopod crustacean \textit{Parhyale hawaiensis} leads to the transformation of gnathopods into a second pair of maxillipods (Liubicich et al. 2009). Recent studies have begun to reveal some of the target genes that are regulated by the Hox genes to achieve specific morphologies. For example, the...
development of different leg morphologies in water striders is achieved by the interaction of Ubx and the gene encoding the Gamma interferon-inducible thiol reductase (Gilt) as a new target gene, which leads to the morphological changes needed for the jumping escape reflex (Armisén et al. 2015).

We study appendage development and appendage-type specification and diversification in the spider Parasteatoda tepidariorum. The spider body is divided into two parts, an anterior prosoma and a posterior opisthosoma. Spiders possess three types of prosomal appendages with different morphologies: the chelicerae, the pedipalps and the walking legs. Pedipalps and walking legs are morphologically very similar and also the expression of appendage patterning genes between pedipalps and walking legs resemble each other closely (reviewed in Pechmann et al. 2010). The main morphological difference between pedipalps and walking legs is that the pedipalps are usually smaller and shorter than the legs, and always lack one of the distal segments, the metatarsus. In addition, pedipalps possess specific modifications, i.e. the gnathendite at the base that is used for food processing, and the bulb at the tip (of the adult male pedipalp only) that is used as a copulatory organ.

The pedipalpal segment is the tritocerebral segment of the spiders and is therefore homologous to the insect intercalary segment (reviewed in Angelini and Kaufman 2005b). Previous work has shown that the positional information of the pedipalp segment and its pair of appendages in *P. tepidariorum* is provided by the expression of the Hox gene *labial-1* (*lab-1*) (Pechmann et al. 2015). RNA interference with the *lab-1* gene in *P. tepidariorum* leads to the loss of the pedipalps and an increase in cell death in the remaining tissue of the pedipalpal segment (Pechmann et al. 2015). The resulting reduced and appendage-less pedipalp segment very much resembles the insect intercalary segment, not only morphologically, but also in terms of segmental gene expression (Pechmann et al. 2015). Thus, *lab-1* is required for normal pedipalp segment formation, and an insect-like reduced segment results from impaired function of *lab-1*. An obvious hypothesis is therefore that the normal insect intercalary segment is produced by an evolutionary loss of the function of *labial* orthologs in insects, but this is not the case. In fact, *labial* orthologs in insects are required for intercalary segment development as well, and loss of *labial* orthologs of several insects leads to increased cell death in the intercalary segment and adjacent head segments, very similar to the effect of *lab-1* loss in spiders (Merrill et al. 1989; Posnien and Bucher 2010; Schaeper et al. 2010). The general roles of *labial* orthologs are therefore very similar in insects and spiders, yet the segmental morphology orchestrated by this Hox gene in the two groups is very different. It is likely that divergent co-factors of *labial* orthologs or divergent regulatory targets downstream of *labial* orthologs are responsible for these differences, but in spiders no such factors have been identified yet. Therefore, we have compiled a list of candidate genes, which are either known or predicted interaction partners of *lab* in *D. melanogaster* or are expressed in the intercalary segment during *D. melanogaster* embryonic development. We have then systematically identified homologs of these genes in *P. tepidariorum* and have analysed their expression in embryos of *P. tepidariorum*. We reasoned that genes that are specifically expressed in the pedipalpal segment of *P. tepidariorum* and/or be involved in producing the specific morphology of the pedipalp appendages.

### Materials and methods

#### Identification of candidate genes

A list of candidate genes was retrieved from The Drosophila Interactions Database (DroID) (Murali et al. 2011), containing all (known and putative) interaction partners of *lab*, as well as genes, which are expressed in the intercalary segment of *D. melanogaster* embryos, according to FlyBase (FB2014_03) (Thurmond et al. 2019).

Protein sequences of genes of interest from *D. melanogaster* were subjected to similarity search via BLAST (Altschul et al. 1990; Boratyn et al. 2013) against the *P. tepidariorum* transcriptome (Posnien et al. 2014). The BLAST parameters were: matrix: BLOSUM62, word size 6; cut-off E-value: 10-1, maximum of 50 hits per sequence. These hits were then reduced, so that only the sequence with the highest blast score per locus was kept. These sequences were subsequently used in a second BLAST-search against the RefSeq (O’Leary et al. 2016) and UniProt (UniProt Consortium T 2019) databases (“back-BLAST”). All sequences retained from the original BLAST search and the back-BLAST search were then aligned using clustalOmega (Sievers and Higgins 2014) and a phylogenetic tree was inferred using FastTree (Price et al. 2009) (results not shown).

Since these trees contained a huge amount of genes, we decided to take from each tree the branch with our gene of interest and its neighbouring branch and infer a new maximum likelihood tree with these sequences (see Table S2 for the substitution model chosen by the program and amount of samples used in the analysis), using MrBayes (Ronquist and Huelsenbeck 2003; Ronquist et al. 2012).

#### Animal cultivation and gene cloning

All animals used in this study originate from the Göttingen strain of *P. tepidariorum*. The embryos were staged after the embryonic staging table published by Mittmann and Wolff (2012). Total RNA was extracted from a mix of all embryonic...
stages using TRIzol® (Life Technologies, Carlsbad, CA, USA). cDNA was synthesized from total RNA with the SMARTer™ PCR cDNA Synthesis Kit (Clontech, Mountain View, CA, USA). Gene-specific cDNA fragments were amplified with primers (see Table S1) designed with Primer3 (Untergasser et al. 2012) and cloned into the pCR®II vector using the TA Cloning® Kit Dual Promoter (Invitrogen, Life Technologies, Carlsbad, CA, USA), with the exception of exd-2, which had previously been cloned with the same degenerate primers as exd-1.

In situ staining and imaging

In situ hybridization and nuclear staining with SYTOX® Green were performed as described before (Prpic et al. 2008; Pechmann et al. 2009) with minor modifications. We used commercially available blocking reagent from Roche (Basel, Switzerland) (2% in PBST). Images were taken with a Leica M205 FA binocular (Leica Microsystems, Wetzlar, Germany) equipped with a QImaging MicroPublisher 5.0 RTV camera (QImaging, Surrey, Canada) and UV light. Images were corrected for colour values and brightness with Adobe Photoshop image processing software and arranged with Adobe Illustrator (both version CS6).

Results

Compilation of possible interaction partners of labial

We compiled altogether 105 genes in D. melanogaster, which were either annotated to be a potential interaction partner of labial (lab) in D. melanogaster, or which are expressed in the embryonic intercalary segment during D. melanogaster development (Fig. 1). The genes expressed in the embryonic intercalary segment are co-expressed with lab, but this does not mean that they are also necessarily interaction partners of lab. Therefore, we attempted to reduce the initially very high number of candidate genes. We excluded all genes that, based on their annotated GO terms (taken from FlyBase (FB2014_03) (Thurmond et al. 2019)), have a more general role in the cell, that may not be specific to the intercalary segment. We have instead focused on those genes for which the corresponding gene ontology (GO) terms suggested a function as transcription factor, or suggested a role in development, and/or gene expression. This first round of selection excluded the following 21 genes from further analysis: cell division cycle 14 (cdc14), CG1598, CG9356, CG10089, CG12256, CG14512, CG14692, CG31342, CG31609, Double hit (Dhit), Hsc70Ch, Heat-shock-protein-70Bb (Hsp70Bb), karyopherin α1 (Kapa1), metabotropic Glutamate Receptor (mGluR), Msh6, ora transientless (ort), PFTAIRE-interacting factor 2 (Pf2), Proctolin (Proc), scramblase 1 (scrambl1), spellchecker 1 (spell) and Tachykinin (Tk). We further excluded the gene CG7182, since it is predicted to encode a cytosolic ATPase, and we also excluded the two histone-coding genes His2A and His2B, the short neuropeptide F precursor gene (sNPF), and TFIIIB (coding for the general transcription initiator TFIIIB), for being part of the basic cellular machinery. Of the remaining 79 genes, 12 had previously been studied extensively in P. tepidariorum, namely the Hox genes abdominal-A (abd-A), Antennapedia (Antp), Deformed (Dfd), proboscipedia (pb), Sex combs reduced (Scr), and Ubx (Schwager et al. 2017), and the genes dachshund (dac) (Schomburg et al. 2015; Turetzek et al. 2015), Dichaei (D) (Bonatto Paese et al. 2018; Paese et al. 2018), eyeless (ey) (Schomburg et al. 2015), homothorax (hth) (Turetzek et al. 2017), orthodenticle (otd) (Pechmann et al. 2009; Schomburg et al. 2015) and twin of eyeless (toy) (Schomburg et al. 2015). Because a full analysis of these genes has already been published, we have not analysed these genes here again and have instead relied on those previous studies.

Three further genes, brinker (brk), insensitive (insv) and zeste (z), had no BLAST hit in the P. tepidariorum transcriptome, while another three genes (CG31624, fussel (fuss) and glia cells missing (gcm)) did not identify the original query sequence in the back-BLAST analysis. This left 61 genes for further analysis. For 8 of these genes, namely caudal (cad) (Akiyama-Oda and Oda 2006; Schönauer et al. 2016), engrailed (en) (Akiyama-Oda and Oda 2006; Schwager et al. 2009), hairy (h) (Pechmann et al. 2009; Schwager et al. 2009), hunchback (hh) (Schwager et al. 2009), knot/collie (kn/col) (Schaep et al. 2010), short gastrulation (sog) (Akiyama-Oda and Oda 2006), twist (twi) (Akiyama-Oda and Oda 2003; Schwager et al. 2009; Feitosa et al. 2017) and wingless (wg) (Janssen et al. 2010), at least one homolog had previously been studied, but was published without comprehensive sequence analysis. For the genes en and wg we did not identify any additional similar sequences in the transcriptome, indicating that no additional paralogs of these genes are present in the transcriptome (Figs. S25 and S66). However, we identified additional sequences with similarity to the following genes: h (Fig. S32), hb (Fig. S33), kn/col (Fig. S37), sog (Fig. S58) and twi (Fig. S62). The phylogenetic analyses of the sequences similar to h, kn/col and twi suggest that these genes are each present as two or three paralogous genes in the P. tepidariorum genome (two paralogs of h and kn/col, three paralogs of twi). We have identified three sequences with similarity to hh, but only two of these appear to be proper homologs of hh, while the third sequence could not be placed unequivocally in the phylogenetic analysis (Fig. S33). The five sequences identified for sog perfectly match partially overlapping portions of the published sequence (Fig. S9). Therefore, we conclude that they do not represent individual paralogs, but are parts of the same transcript, which was artificially split up in the transcriptome assembly. We could not identify any
| Gene | Chromosome | Start | End | Description |
|------|------------|-------|-----|-------------|
| abd-A | 2 | 108 | 108 | Schuermann et al. 2017 |
| Aop | 2 | 108 | 108 | Schuermann et al. 2017 |
| Arb | 1 | 108 | 108 | Locus: 42097 |
| btd | 0 | 108 | 108 | |
| Carpok | 0 | 108 | 108 | |
| cad | 1 | 108 | 108 | Alyxiana-Olia and Ol2009 Schuermann et al. 2016 |
| cals | 1 | 108 | 108 | Locus: 4958 |
| eone | 1 | 108 | 108 | Locus: 1274 |
| cul14 | 0 | 108 | 108 | |
| chimo | 0 | 108 | 108 | Cytoplasmic ATPase |
| CG7530 | 0 | 108 | 108 | |
| CG7182 | 0 | 108 | 108 | |
| CG9356 | 0 | 108 | 108 | |
| CG10088 | 0 | 108 | 108 | |
| CG12535 | 0 | 108 | 108 | |
| CG145152 | 0 | 108 | 108 | |
| CG14492 | 0 | 108 | 108 | |
| CG14324 | 0 | 108 | 108 | |
| CG161360 | 0 | 108 | 108 | |
| CG161624 | 0 | 108 | 108 | |
| cpio | 1 | 108 | 108 | Locus: 21216 |
| cro | 1 | 108 | 108 | Locus: 19541 |
| CIBP | 1 | 108 | 108 | Locus: 14608/66947 |
| CTCF | 1 | 108 | 108 | Locus: 14442 |
| D | 1 | 108 | 108 | Paeos et al. 2015 |
| de | 1 | 108 | 108 | Locus: f1603/66647 |
| diec | 1 | 108 | 108 | Locus: Schuermann et al. 2017 Tiettsch et al. 2015 |
| Dnt | 2 | 108 | 108 | Locus: 19578 |
| Dnt1 | 2 | 108 | 108 | Schuermann et al. 2017 |
| all | 3 | 108 | 108 | |
| Dsp1 | 1 | 108 | 108 | Locus: 17502/66647 |
| Dr / Mch | 1 | 108 | 108 | Locus: 17553 |
| evB | 3 | 108 | 108 | see nac |
| ems | 3 | 108 | 108 | Locus: 17574 |
| en | 1 | 108 | 108 | Alyxiana-Olia and Ol2009 Schuermann et al. 2017 |
| exd | 2 | 108 | 108 | |
| ey | 2 | 108 | 108 | Schuermann et al. 2017 |
| Eto | 1 | 108 | 108 | Locus: B1 |
| Fex2 | 3 | 108 | 108 | Locus: 1337/166647 |
| Fex2/2ket | 3 | 108 | 108 | Locus: 5350/66647/ |
| fws | 1 | 108 | 108 | |
| gcm | 2 | 108 | 108 | Locus: 109/66647 |
| gap | 2 | 108 | 108 | Locus: 109/66647 |
| h | 1 | 108 | 108 | Locus: 12291/106647 * Pachmann et al. 2009 |
| hbb | 2 | 108 | 108 | Locus: 1087/66647 * Schuermann et al. 2009 |
| His2A | 2 | 108 | 108 | Locus: 1087/66647 * Schuermann et al. 2009 |
| His2B | 2 | 108 | 108 | Locus: 1087/66647 * Schuermann et al. 2009 |
| hkb | 2 | 108 | 108 | Locus: 1087/66647 |
| Hsc70Cm | 2 | 108 | 108 | Locus: 1087/66647 |
| Hsp70Bb | 2 | 108 | 108 | Locus: 1087/66647 |
| nht | 2 | 108 | 108 | Tavernix et al. 2017 |
homolog for cad in the transcriptome, despite the fact that a P. tepidariorum sequence for this gene has been identified previously (Akiyama-Oda and Oda 2006). We therefore performed a BLAST search with the published P. tepidariorum cad sequence against the transcriptome sequence and found 3 very short sequence fragments, which perfectly align with the query sequence, but are apparently too short to surface in the very short sequence fragments, which perfectly align with the cad sequence against the transcriptome sequence and found 3.

Expression of candidate genes: previously published genes with known or new paralogs

First, we used previously published expression data for the genes on the candidate gene list to assess possible differential expression of the genes in the pedipalps or the pedipalpal segment compared with the walking leg segments and their appendages. One paralog of the Hox gene ph, namely pb-A, is strongly expressed in the pedipalpal segment, but only weakly expressed in the walking leg segments (Schwager et al. 2017). The gene thus shows an expression pattern very similar to lab-1 itself (Pechmann et al. 2015). The other Hox genes on the candidate gene list are not expressed in the pedipalpal segment at all, but are expressed in more posterior body segments (Schwager et al. 2017). This suggests that these Hox genes do not normally interact with lab-1 in the pedipalpal segment, but of course this does not rule out the possibility that they may act as repressors of lab-1 in other segments. Based on previously published accounts (see references in previous chapter and Fig. 1) of their expression pattern, the genes dac, D, ey, hth, old, toy, cad, en, sog and wg (and their paralogs, if any) do not show apparent differential expression between the pedipalpal and the neighbouring walking leg bearing segments. The same is true for the genes h, hb, kn/coll and twi, but in these four cases we have identified additional new putative paralogous genes in the genome of P. tepidariorum that have not been studied in the initial published accounts. The additional paralog of h (Locus_1 125317/166847) is expressed ubiquitously (Fig. S1). The newly identified paralog of hb (Locus_1 68341/166847) is expressed weakly in the pro-neural clusters during stages 10 and 11 (Fig. S5a, a’, b). The newly identified paralog of kn/coll (Locus 24477) is only expressed at stage 12 as three spots in the segments L4 to O2 (white arrowheads in Fig. 2k, k’). It is not expressed in the pedipalpal segment or the head in general (Fig. 2j) and is therefore unlikely to interact with lab-1 in the pedipalpal segment. The newly identified additional paralogs of twi (Locus_1 61174/166847 and Locus_10831) could not be studied further, because we failed to clone fragments of these transcripts from the cDNA preparations used for molecular cloning.

Ubiquitously expressed genes or genes without detectable embryonic expression

Of the genes we have analysed, a relatively large proportion is expressed ubiquitously. The homologs of cap ‘n’ collar (cnc)
Locus_1274 (Fig. S14), C-terminal binding protein (CtBP) (Locus_1 460/166847) (Fig. S17), CTCF (Locus_14442) (Fig. S18), daughterless (da) (Locus_1 9632/166847) (Fig. S19), dorsal (dl) (Locus_15237, Locus_32713) (Fig. S21), Drop (Dr) (Locus_17634) (Fig. S21), Dorsal switch protein 1 (Dsp1) (Locus_1 97552/166847) (Fig. S22), empty spiracles (ems) (Locus_21143) (Fig. S24), Enhancer of zeste (E(z)) (Locus_81) (Fig. S27), Fasciclin-2 (Fas2) (Locus_23706) (Fig. S28), hairy (h) (Locus_1 123,107/166847) (Fig. S32), huckebein (hkb) (Locus_26294) (Fig. S34), Medea (Med) (Locus_2595) (Fig. S42), Polycomb (Pc) (Locus_2368) (Fig. S48), polyhomeotic proximal (ph-p) (Locus_17041) (Fig. S50), pleiohomeotic (pho) (Locus_5955, Locus_10241) (Fig. S49), senseless (sens) (Locus_25081) in the head lobes. White arrowheads in h, i point to expression near the stomodeum. Black arrowheads in h, h” point to expression in the head lobe. White arrowheads in k, k’ denote segmental expression. All embryos are shown with anterior to the top, except for k, which is a ventral aspect of the opisthosoma, and k’, which is a lateral aspect with anterior to the right. Abbreviations: Ch = chelicera, Pp = pedipalp, L = walking leg, O = opisthosomal segment
Scm-related gene containing four mbt domains (Sfmbt) (Locus_1 107630/166847) (Fig. S57), Serrate (Ser) (Locus_8656) (Fig. S56), trithorax (trx) (Locus_690) (Fig. S61) and ventral nerve chord defective (vnd) (Locus_16018) (Fig. S63) show no separate expression domains in the embryos, but uniform and ubiquitous expression throughout the embryo (Figs. S1 and S2). Moreover, Dr (Locus_28432) (Fig. S21), maf-S (Locus_2680) (Fig. S41) and pax3/7
(Locus_13554) (Fig. S31), a homolog of the D. melanogaster genes paired (prd) and gooseberry (gsb), did not display any detectable staining in the embryonic stages analysed (Figs. S1 and S2). Please note, however, that during the screening process no in situ hybridisation experiments have been repeated and, therefore, a cautionary note concerning the ubiquitously expressed and non-expressed genes is appropriate: all probes were subjected to the same quality checks (size control with electrophoresis and quantity control with spectrophotometry). However, we did not control for digoxigenin labelling efficiency during probe synthesis and we therefore cannot guarantee that every probe (especially in the cases where no expression was detected) worked properly.

**Expression in the pre-cheliceral lobes**

Several of the analysed genes show prominent expression in the pre-cheliceral lobes. The two isolated paralogs of no ocelli (noc)/elbow (elB) (Locus_13902, Locus_1 91018/166847; Figs. S23 and S44) are expressed in two spots on either side of the head lobes (white arrowheads in Fig. 2a, b, c, d). While the sequence from Locus_13902 shows no further expression domains, its paralog is expressed ubiquitously in the limb buds at stage 8 (Fig. 2c′). At stage 11, additional expression is seen in the chelicerae (black arrowhead in Fig. 2d′), as well as a segmental expression in several pro-neural clusters (white arrowheads in Fig. 2d′), and proximal expression domains in all appendages (black arrowheads in Fig. 2d′).

The homolog of reversed polarity (repo) (Locus_20045; Fig. S52) shows only one expression domain in an anterior-lateral position on either side of the head lobes (arrowheads in Fig. 2e, f).

The tailless (tll) gene with only one homolog in P. tepidariorum (Locus_135; Fig. S60) shows no expression at stage 8 (Fig. 2g, g′). From stage 10 on tll is expressed in one domain on either side of the stomodeum (white arrowheads in Fig. 2h, i). An additional expression domain is present at stage 10, at the lateral rim of the head lobes (black arrowheads in Fig. 2h, h′), which vanishes again by stage 11 (Fig. 2i, i′).

**Segmentally repeated expression**

We found several of the candidate genes to display a prominent segmental expression. The homolog of crocodile (croc) (Locus_16541; Fig. S29) is expressed in the pre-cheliceral lobes in a v-shape pattern and in later stages surrounds the stomodeum on either side with an additional later domain towards the edge of the head lobes (white arrowheads in Fig. 3a, c, d, e). We also observed a transient segmental expression at stage 9 (arrowheads in Fig. 3b), which vanishes again at stage 10 (Fig. 3c′, d′).

One of the identified homologs of Dr (Locus_11933; Fig. S21) is expressed in broad segmental stripes at stage 6 (Fig. 30), which at stage 8 become a segmentally repeated pattern at the base of the appendages (arrowheads in Fig. 3g, g′, g″), while the striped pattern in the opisthosomal segments splits up along the midline in more anterior segments (arrowheads in Fig. 3g″). At stage 12, Dr expression is present in every segment on the ventral side near the bases of the appendages and strongly in the former segment addition zone, as well as in the neural precursor groups along the body axis (arrowheads in Fig. 3h, h′, h″, h‴).

Two of the identified paralogs of ems (Fig. S25) also show segmental expression. Locus_15791 is weakly expressed at stage 8, in a pattern of segmentally repeated stripes in the opisthosomal segments (Fig. 3i″), and at stage 12 expression is only present in the most posterior tip of the germ band (Fig. 3j″). Locus_16046 is expressed in two v-shaped domains in the pre-cheliceral lobes at stage 8 (black arrowhead in Fig. 3k). This domain splits up into four distinct spots in later stages, which form a line across the head lobes (black arrowheads in Fig. 3l, l′). This transcript has additional expression domains in segmentally repeated stripes around the anterior-median quarter of the limb buds, which in later stages are located in several pro-neural clusters in every segment (white arrowheads in Fig. 3k, k′, k″, k‴).

Locus_23671, which we identified as a homolog of Fas2 (Fig. S28), shows no expression at stage 8 (Fig. 4a). In later stages, staining is present in two domains on either side of the head lobes (black arrowheads in Fig. 4b, c), as well as in segmentally repeated expression domains along the body, and also in the pro-neural clusters (white arrowheads in Fig. 4b, b′, c, c′).

The gene pax3/7 (Locus_17531), which is a homolog of the D. melanogaster genes prd and gsb (Fig. S31), is expressed in segmentally repeated stripes throughout the embryonic stages analysed. At stage 6, it is expressed in the most posterior segments, emerging from the segment addition zone, with broader expression in the more recently formed segments (arrowheads in Fig. 4d, d′). In later stages, this gene is expressed in every segment (Fig. 4e, e′, e″, f, f′, f″).

Finally, the single P. tepidariorum homolog of the two sloppy paired genes in D. melanogaster (slp1 and slp2; Fig. S29), is expressed in segmental stripes, which also include the neural precursor clusters along the body axis in the stages analysed (white arrowheads in Fig. 4g, g′, g″, h, h′, h″, i, i′, i″, j, j′, j″). Furthermore, we observed expression around the median sulcus at stage 8 (red arrowheads in Fig. 4g), two broader domains in the lateral portion of the head lobes (black arrowheads in Fig. 4g, h, i, j) and a domain in the anterior region of the pre-cheliceral lobes at stage 11 and 12 (Fig. 4i, j).

**Genes predominantly expressed in the nervous system**

Six of the analysed genes indicated a role in the development of the nervous system by their expression patterns. The newly
identified paralog of *hb* (Locus_1 68,341/166847; Fig. S33) is expressed weakly in the pro-neural clusters during stages 10 and 11 (Fig. 5a, a′, b).

One of the paralogs of *hkb* (Locus_1422; Fig. S34) shows a segmental expression pattern, starting at stage 8 in the base of the appendages (arrowheads in Fig. 5c, d, d′). At stage 11, this expression is then present in a subset of pro-neural clusters in each segment (arrowheads in Fig. 5e, e′, e″).

The homolog of *klumpfuss* (*klu*) (Locus_5949; Fig. S36) shows no expression at stage 8 (Fig. 5f, f′), but at stage 11 is weakly expressed in all pro-neural clusters of the embryo and expression near the lateral rim of the head lobes. The red arrowhead in g points to expression at the anterior end of the median sulcus near the stomodeum. All embryos are shown with anterior to the top, except for d, which is a lateral aspect with anterior to the right, and d′, e′, f′, g′, h′, i′, j′, k′ which are ventral aspects of the opisthosoma. Abbreviations: Ch = che-licera, Pp = pedipalp, L = walking leg, O = opisthosomal segment

![Fig. 4](image-url) Genes with segmentally repeated expression patterns II. Expression of *Fas2* locus_23671 (a–c′), *pax3/7* locus_17531 (d–f″) and *slp* locus_18540 (g–j″). Arrowheads in a–c′ point to separate expression patches in the head segments. Arrowheads in d–f″ denote diverse segmental patches/stripes of the dynamic expression profile. White arrowheads in g–j″ point to segmental expression patches including a single patch in the head lobes. The black arrowheads in g, h, i, j point to...
the sensory organ precursors at the tips of the appendages (Fig. 5g, g′, g″, g‴).

The homolog of Krueppel (Kr) (Locus_6712; Fig. S38) is expressed ubiquitously in the appendages at stages 8 and 10 (black arrowheads in Fig. 5h, h′, i, i′). There is additional ubiquitous expression at stage 8 from the L3 segment towards the posterior end of the germ band (white arrowhead in Fig. 5h′). At stage 10, Kr also stains a large
number of pro-neural clusters along the antero-posterior axis (Fig. 5i, i', i'').

At earlier stages, the homolog of ovo (Locus_13054; Fig. S47) showed no staining in the entire embryo, by stage 10 it was expressed in a subset of pro-neural clusters in the prosomal segments (Fig. 5k, k'), and a single cluster in each opisthosomal segment (white arrowheads in Fig. 5k''). Finally, the homolog of vnd (Locus_12534; Fig. S63) appears as one spot of expression on either side in the centre of the pre-cheliceral lobes (white arrowheads in Fig. 5m, n), as well as a median stripe of expression along the ventral midline, which at stage 10 splits up on either side of the median sulcus (black arrowheads in Fig. 5m, m').

Extradenticle is differentially expressed between pedipalps and legs

The gene extradenticle (exd) is present as 10 sequences in the transcriptome of P. tepidariorum (Fig. S26). However, these sequences map to only two predicted sequences (Figs. S6 and S7). We therefore assumed that there are two paralogous loci in the genome of P. tepidariorum, one of which is the previously published exd-1 (Khadjeh et al. 2012) and the other (exd-2) we newly describe here. The exd-1 gene is expressed at the base of the appendages throughout embryonic development (black arrowheads in Fig. 5a–i).

Additionally, exd-1 is expressed in the region of the labrum (white arrowheads in Fig. S3b, c), in a distinct ring in the pedipalps and legs (white arrowheads in Fig. S3b', c', d', brackets in Fig. 6a–i). Additionally, exd-1 is expressed in the region of the labrum (white arrowheads in Fig. S3b, c), in a distinct ring in the pedipalps and legs (white arrowheads in Fig. S3b', c', d', brackets in Fig. 6a–i). The second exd gene, exd-2, is expressed at the base of the chelicerae (white arrowheads in Fig. S4a, a', b, b', brackets in Fig. 6j–l), and as a faint stripe along the embryo at the base of the appendages (Fig. S4a'), which becomes more prominent in

![Fig. 5](image-url) Genes with expression in the developing nervous system. Expression of hb locus_1_68341 (a, b), hkb locus_1422 (c–e), klu locus_5949 (f, g), Kr locus_13054 (j, k) and vnd locus_12534 (m, n). White arrowheads in c', d', e', e'' point to segmentally reiterated expression in the ventral nervous system. Black arrowheads in h, h' denote ubiquitous expression in the appendages. White arrowhead in h' points to the anterior border of posterior expression in the nervous system. White arrowheads in k'' denote expression in proneural clusters. White arrowheads in m, n point to expression in the developing brain. Black arrowheads in m–n denote expression along the ventral midline. All embryos are shown with anterior to the top, except for e', g', g'', i', i'' which are ventral aspects of the opisthosa. Abbreviations: Ch = chelicera, Pp = pedipalp, L = walking leg, O = opisthosomal segment.

![Fig. 6](image-url) Expression of exd-1 and exd-2 in different appendage types. Expression of exd-1 (a–i) and exd-2 (j–r) in chelicera (top row), pedipalp (center row) and walking leg (bottom row) at early limb formation (first and fourth column), midterm (second and fifth row) and towards the end of inversion (third and sixth row). In all panels, brackets denote continuous extension of expression from the proximal end towards distal, and arrowheads point to locally elevated expression levels (“rings”). Abbreviations: Ch, chelicera; Pp, pedipalp.
the pedipalps as embryonic development progresses (compare pedipalpal expression in Fig. 6n, o to Fig. 6q, r). The exd-2 gene is also expressed in a ring-shaped domain in the appendages (black arrowheads in Fig. S4b, c′, d′, e′ and Fig. 6m–r), which is slightly fainter in the pedipalps than in the legs (red arrowhead in Fig. S4c′). Moreover, exd-2 is expressed in a domain surrounding the pre-cheliceral lobes and the labrum at the edge of the non-neurogenic ectoderm. This domain continues along the lateral edges of the embryo, which represents future dorsal tissue (red and grey arrowheads in Fig. S4c, d, e).

**Discussion**

**Basal cellular and developmental processes**

The ubiquitous expression of many of the analysed *P. tepidariorum* genes supports a conserved function of these genes in basal cellular and developmental processes. In *D. melanogaster* and other animal species, many genes that are involved in chromatin remodelling or the maintenance of a certain transcriptional state (*CiBP, E(z), Dsp1, Pc, ph-p, pho, sfmbt, try*) (Courey and Jia 2001; Schüttengruber et al. 2007) are expressed ubiquitously. This is possible, because spatial specificity of their function is conferred by other means, e.g. local events of chromatin modification or spatially restricted binding partners (Cao et al. 2002; Czermin et al. 2002). Homologs of these genes are expressed ubiquitously in *P. tepidariorum* as well, suggestive of a conserved function of these genes in insects and spiders. The same is true for *CTCF*, which in *D. melanogaster* is an ubiquitous transcription factor that blocks enhancers (Moon et al. 2005; Mohan et al. 2007). The homolog of *Med* also shows ubiquitous expression in *P. tepidariorum* as well as in *D. melanogaster* (Tomancak et al. 2002, 2007). This is not unexpected for a gene that is involved in the transduction of TGF-beta-like signals (Das et al. 1998; Marquez et al. 2001) that rely on the correct spatio-temporal activation by factors from the outside rather than the specific expression of effectors themselves.

As in *D. melanogaster* (Cronmiller and Cummings 1993), *da* is expressed ubiquitously in *P. tepidariorum*. However, since the molecular mechanisms of the sex-determination system in spiders are largely unknown, no conclusion can be drawn yet on a conserved role of *da* in *P. tepidariorum*.

In *D. melanogaster*, *dl* establishes dorso-ventral patterning of the body axis through a nuclear concentration gradient of the protein (Anderson et al. 1985; Rushlow et al. 1989; Lall and Patel 2001) and is also involved in the immune response after injury (Lemaître et al. 1995). The specification of the dorso-ventral axis in *P. tepidariorum* has been shown to be facilitated by the migration of cells of the cumulus, and thus by a mechanism apparently not dependent on dorsal function (Akiyama-Oda and Oda 2006).

**Head and brain development**

The *D. melanogaster* paralogs of *noc/elB* are expressed in the region of the developing brain, and also appear to be involved in the development of appendages (Cheah et al. 1994; Weihe et al. 2004). In *P. tepidariorum*, the orthologs of *noc/elB* do not seem to be involved in the development of the appendages, but the patch of expression in the pre-cheliceral lobes suggests a conserved role in brain formation. The role of *tll* seems to be partly conserved with respect to the development of the optic system (Rudolph et al. 1997), while terminal structures (Casanova 1990; Moran and Jiménez 2006) appear not to be under its influence in *P. tepidariorum*, at least in those embryonic stages studied in the present work. This is consistent with previous findings that the terminal system is not entirely conserved among arthropods (Duncan et al. 2013).

Both, *cnc* and *kn/col* are involved in the segmentation of head structures in *D. melanogaster* (Veraksa et al. 2000; Peel 2004; Ntini and Wimmer 2011). Previous studies have shown that the expression patterns of both genes are not conserved among all arthropods. The eponymous cap and collar expression (distal “cap” and a “collar” around the stomodeum) is restricted to the Mandibulata clade, while expression in chelicerates and onychophorans is ubiquitous (Sharma et al. 2014; Janssen et al. 2011a; Hunnekuhl and Akam 2017; Janssen 2017). The head-specific expression of *kn/col* has been shown to be conserved in insects and myriapods, but not in other arthropod groups (Schaeper et al. 2010; Janssen et al. 2011b). The newly identified second *P. tepidariorum* homolog of *kn/col* is not expressed in the head either, thus supporting the previous conclusion that the role of *kn/col* in head segmentation is restricted to insects and myriapods.

The *P. tepidariorum* homolog of *repo* is expressed in patches in the pre-cheliceral lobes. This indicates a role in brain development, but provides no evidence for a conserved role in glial cell maintenance known from *D. melanogaster* (Halter et al. 1995).

**Aspects of segmentation**

The anterior expression domains of *croc* are conserved among insects and myriapods (Birkan et al. 2011; Janssen et al. 2011a). Our results in *P. tepidariorum* indicate that the expression (and by inference also the role) of *croc* in head development is conserved in all arthropods. In *P. tepidariorum* *croc* also shows segmental expression in the mesoderm in the earlier stages, which correlates to mesodermal expression reported in *D. melanogaster* (Tomancak et al. 2002, 2007).
The segmentally iterated expression of ems paralogs in P. tepidariorum suggests that the role in the segmental development of the nervous system is conserved, but no paralog appears to be involved in identity specification of anterior segments as known in D. melanogaster (Schöck et al. 2000; Peel 2004).

A role of pax3/7 in segmentation appears to be conserved in P. tepidariorum, although pax3/7 is not expressed in a pair-rule pattern, like prd in D. melanogaster, but is expressed in segmentally repeated stripes. This difference has been shown for several pair-rule genes in spiders, especially members of the Pax group III (Damen et al. 2005; Schoppmeier and Damen 2005) and has been hypothesized to be the ancestral condition in arthropod segmentation (Peel et al. 2005). Based on the expression pattern in P. tepidariorum, pax3/7 seems to have a function in the establishment of newly formed segments in the segment addition zone, and a role in the maintenance of mature segments, indicated by the segmental expression in later stages.

The sloppy-paired genes in D. melanogaster are pair-rule genes involved in trunk segmentation (Cadigan et al. 1994) and the specification of head segments (Grossniklaus et al. 1994). The segmental expression of slp in P. tepidariorum suggests that a function in segmentation might be conserved, although slp seems to specify one segment at a time and not in a pair-rule fashion.

**Similarities in neurogenesis between spiders and insects**

Neuroblasts in D. melanogaster develop from a field of cells in the neuroectoderm, which expresses several proneural genes, such as achaete, scute and lethal of scute (Skeath and Thor 2003). While the establishment of neural precursors by achaete-scute genes and the subsequent lateral inhibition in these groups by Delta / Notch signalling is regarded to be generally conserved in spiders (Stollewerk 2002; Stollewerk et al. 2003), the mechanisms of separating neural precursors are different. While in D. melanogaster single neuroblasts delaminate from the neuro-ectoderm, in spiders there are several rounds of delamination of a larger group of cells, called neural progenitor groups (Stollewerk et al. 2003).

Interestingly, despite the mechanistical differences of conferring proneural identity between chelicerates and insects, P. tepidariorum and D. melanogaster share many of the genes to mark neural precursor cells in the ectoderm. Paralogs of Dr, ems, Fas2, hb, hkb, klu, Kr, slp and vnd are all expressed in subsets of neural precursor clusters in P. tepidariorum, similar to the expression in neuroblasts in D. melanogaster (Isshiki et al. 2001; Urbach and Technau 2003). Despite this superficial conservation, these genes do not mark separate neural precursors in the same area of their respective segments, but appear only to be expressed as marker genes to set aside the bulk of these cells from the rest of the neurogenic ectoderm. Notable differences concerning these genes include tll, which is expressed in D. melanogaster neuroblasts, but shows no neural expression in P. tepidariorum, and ovo, which marks neural precursors in P. tepidariorum, but not in D. melanogaster.

**extradenticle is differently expressed in pedipalps and legs**

While exd-1 is expressed in the proximal part of the legs up to a prominent ring of expression in the median part, in the pedipalp the whole appendage expresses the transcript. Thus, exd-1 is differentially expressed especially in the distal portion of legs and pedipalps. This expression therefore correlates with the lack of the metatarsus leg segment in the distal part of the pedipalp compared with the legs. The expression of exd-2 at the base of the pedipalps in later embryonic stages correlates with the other morphological peculiarity of the pedipalps, the gnathendite. Only in the pedipalp exd-2 is expressed in a proximal-ventral region, i.e. the area in which the gnathendite develops later. Since both exd paralogs show additional expression domains in the pedipalp, and because exd is implicated as a regulatory target of lab in D. melanogaster, the two exd paralogs are good candidates for genes that mediate the formation of major morphological peculiarities of the pedipalp under the control of lab-1.

**Conclusions**

We have compiled a list of all genes from D. melanogaster that may be regulated by lab in the intercalary segment based on known interactions with lab and/or gene expression in the intercalary segment (the main expression locus of lab). We have then tested whether homologs of these genes from the spider P. tepidariorum are expressed in the pedipalp segment, which is the homolog of the insect intercalary segment and also the main expression locus of the spider lab-1 gene. We have also tested whether the genes are differentially expressed in the pedipalpal segment and the adjacent leg bearing segments. Differential gene expression is a first indication that a gene might be involved in pedipalpal segment or pedipalp appendage specification and may be controlled by lab-1. After having compiled gene expression data from all previously published genes on the candidate gene list (32 genes, including duplicates), plus expression patterns of 43 newly studied genes (including duplicates), we find that three genes show differential expression between the intercalary and walking leg segments, namely exd-1, exd-2 and pb-A. Thus, our screen for potential lab regulatory targets or co-factors in
P. tepidariorum successfully identified 3 interesting candidates that now warrant further study, but the overall success rate is only 4% (3 genes out of 75 studied genes). This rather disappointingly low success rate points to a potential general problem of the candidate gene approach when it is applied to lineage-specific organs or evolutionary novelties. The spider pedipalp has no counterpart in D. melanogaster or other insects, and therefore relying on insect data as a basis for its potential genetic basis apparently cannot identify a larger number of conserved factors implicated in its specification and formation. Based on our results, we suggest that pedipalp-specific factors, and factors that are regulated by lab-1 in P. tepidariorum, should be identified more reliably by de novo gene discovery approaches, for example transcriptome analyses of developing pedipalp tissue and differential transcriptomics of pedipalp tissue versus non-pedipalp tissue. In addition, co-factors and regulatory targets of labial orthologs specific to the spider model might also be identified more reliably via biochemical methods, such as ChIP-Seq and Co-IP pull-down.

Acknowledgements A portion of the experiments reported in this paper have been performed in the lab of Ernst Wimmer at the University of Göttingen (GAU). We thank him and all members of the Wimmer lab for their hospitality and support.

Funding information Open Access funding provided by Projekt DEAL. This work has been funded by the Deutsche Forschungsgemeinschaft (DFG; grant numbers PR1109/4-1, PR1109/6-1 and PR1109/7-1)

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References
Akiyama-Oda Y, Oda H (2003) Early patterning of the spider embryo: a cluster of mesenchymal cells at the cumulus produces Dpp signals received by germ disc epithelial cells. Development 130:1735–1747. https://doi.org/10.1242/dev.00390
Akiyama-Oda Y, Oda H (2006) Axis specification in the spider embryo: dpp is required for radial-to-axial symmetry transformation and sog for ventral patterning. Development 133:2347–2357. https://doi.org/10.1242/dev.02400
Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410. https://doi.org/10.1016/S0022-2836(05)80360-2
Anderson KV, Jurgens G, Nusslein-Volhard C (1985) Establishment of dorsal-ventral polarity in the Drosophila embryo. Genetic studies on the role of the Toll gene product. Cell 22:779–789
Angelini DR, Kaufman TC (2005a) Insect appendages and comparative ontogenetics. Dev Biol 286:57–77. https://doi.org/10.1016/j.ydbio.2005.07.006
Angelini DR, Kaufman TC (2005b) Comparative developmental genetics and the evolution of arthropod body plans. Annu Rev Genet 39:95–119. https://doi.org/10.1146/annurev.genet.39.073003.112310
Arnsen D, Nagui Refki P, Crumière AJJ, Viala S, Toubina W, Khila A (2015) Predator strike shapes antipredator phenotype through new genetic interactions in water striders. Nat Commun 6:8153. https://doi.org/10.1038/ncomms9153
Birkan M, Schaeppe ND, Chipman AD (2011) Early patterning and blastodermal fate map of the head in the milkweed bug Oncopeltus fasciatus. Evol Dev 13:436–447. https://doi.org/10.1115.1525-142X.2011.00497.x
Bonatto Paese CL, Leite DJ, Schönaauer A, McGregor AP, Rüssel S (2018) Duplication and expression of Sox genes in spiders. BMC Evol Biol 18:205. https://doi.org/10.1186/s12862-018-1337-4
Boratyn GM, Camacho C, Cooper PS, Coulouris G, Fong A, Ma N, Madden TL, Matten WWT, McGinnis SD, Merezhuk Y, Raytsev Y, Sayers EW, Tao T, Ye J, Zaret'skaya I (2013) BLAST: a more efficient report with usability improvements. Nucleic Acids Res 41:29–33. https://doi.org/10.1093/nar/gkt282
Cadigan KM, Grossniklaus U, Gehring WJ (1994) Localized expression of sloppy paired protein maintains the polarity of Drosophila parasegments. Genes Dev 8:899–913. https://doi.org/10.1101/gad.8.8.899
Cao R, Wang L, Wang H, Xia L, Erdjument-Bromage H, Tempst P, Jones RS, Zhang Y (2002) Role of histone H3 lysine 27 methylation in Polycomb group silencing. Science 298:1039–1044. https://doi.org/10.1126/science.1076997
Casanova J (1990) Pattern formation under the control of the terminal system in the Drosophila embryo. Development 110:621–628
Casares F, Mann RS (2001) The ground state of the ventral appendage in Drosophila. Science 293:1477–1480. https://doi.org/10.1126/science.1076997
Cheah PY, Meng YB, Yang X, Kimbrell D, Ashburner M, Chia W (1994) A cluster of mesenchymal cells at the cumulus produces Dpp signals received by germ disc epithelial cells. Development 120:1735–1747. https://doi.org/10.1242/dev.00390
Courey AJ, Jia S (2001) Transcriptional repression: the long and the short. Cell 111:185. https://doi.org/10.1016/S0092-8674(02)00975-3
Damen WG, Janssen R, Pripic NM (2005) Pair-rule gene orthologs in spider segmentation. Evol Dev 7:618–628. https://doi.org/10.1111/j.1525-142X.2005.00080.x
Das P, Madzazia LL, Wang H, Finelli AL, Cho S, Smith MM, Paggett RW (1998) The Drosophila gene Medea demonstrates the requirement for different classes of Smads in dpp signaling. Development 125:1519–1528
Duncan EJ, Benton MA, Dearden PK (2013) Canonical terminal patterning is an evolutionary novelty. Dev Biol 377:245–261. https://doi.org/10.1016/j.ydbio.2013.02.010
Feiotos NM, Pechmann M, Schwager EE, Tobias-Santos V, McGregor AP, Daren WG, Nunes da Fonseca R (2017) Molecular control of
gut formation in the spider Parasteatoda tepidariorum. Genesis 55: e23033. https://doi.org/10.1002/dvg.23033
Grossniklaus U, Cadigan KM, Gehring WJ (1994) Three maternal coordinate systems cooperate in the patterning of the Drosophila head. Development 120:3155–3171
Halter DA, Urban J, Rickert C, Ner SS, Ito K, Travers AA, Technau GM (1995) The homeobox gene repo is required for the differentiation and maintenance of glia function in the embryonic nervous system of Drosophila melanogaster. Development 121:317–322. https://doi.org/10.1101/s0209783.3098139
Hunnekuhl VS, Akam M (2017) Formation and subdivision of the head field in the centipede Strigamia maritima, as revealed by the expression of head gap gene orthologues and hedgehog dynamics. Evodevo 8:18. https://doi.org/10.1186/s13227-017-0082-x
Ishiki T, Pearson B, Holbrook S, Doe CQ (2001) Drosophila neuroblasts sequentially express transcription factors which specify the temporal identity of their neuronal progeny. Cell 106:511–521. https://doi.org/10.1101/S0092-8674(01)00465-2
Janssen R (2017b) Comparative analysis of gene expression patterns in the arthropod labrum and the onychophoran frontal appendages, and its implications for the arthropod head problem. Evodevo 8:1. https://doi.org/10.1186/s13227-016-0064-4
Janssen R, Le Gouar M, Poulin F, Bolognesi R, Schwager Janssen R (2017) Comparative analysis of gene expression patterns in the head of Parasteatoda tepidariorum: a case of convergence? BMC Evol Biol 10:375. https://doi.org/10.1186/s1471-2148-10-374
Janssen R, Budd GE, Damen WGM (2011a) Gene expression suggests conservation mechanisms patterning the heads of insects and myiopods. Dev Biol 357:64–72. https://doi.org/10.1016/j.ydbio.2011.05.670
Janssen R, Damen WG, Budd GE (2011b) Expression of collier in the premandibular segment of myiopods: support for the traditional Atelocerata concept or a case of convergence? BMC Evol Biol 11:50. https://doi.org/10.1186/1471-2148-11-50
Jockusch EL (2017) Developmental and evolutionary perspectives on the origin and diversification of arthropod appendages. Integr Comp Biol 57:533–545. https://doi.org/10.1093/icb/icx063
Khad J, Turetzek N, Peichmann M, Schwager EE, Wimmer EA, Damen WG, Ppric NM (2012) Divergent role of the Hox gene Antennapedia in spiders is responsible for the convergent evolution of abdominal limb morphology. Proc Natl Acad Sci 109:4921–4926. https://doi.org/10.1073/pnas.1116421109
Lall S, Patel N (2001) Conservation and divergence in molecular mechanisms of axis formation. Annu Rev Genet 35:407–437. https://doi.org/10.1146/annurev.cellbio.15.1.81
Lemaître B, Meister M, Govind S, Georgell P, Stewart R, Reichart JM, Hoffmann JA (1995) Functional analysis and regulation of nuclear import of dorsal during the immune response in Drosophila. EMBO J 14:536–545
Liubichert DM, Serano JM, Pavlopoulos A, Kontarakis Z, Prokes ME, Kwan E, Chatterjee S, Tran KD, Averof M, Patel NH (2009) Newfeld SJ (2001) Transgenic analysis of the Smad family of transcription factors which specify the temporal identity of their neuronal progeny. Cell 106:511–521. https://doi.org/10.1146/annurev.cellbio.15.1.81
Mittmann B, Wolff C (2012) Embryonic development and staging of the cobweb spider Parasteatoda tepidariorum: aSynopsis: Achaearanea tepidariorum; Araneomorphae; Theridiidae. Dev Genes Evol 222:189–216. https://doi.org/10.1007/s00427-012-0401-0
Mohan M, Bartkum M, Herold M, Philippen A, Heinl N, Bardenhagen I, Leers J, White RAH, Renkawitz-Pohl R, SAumweber H, Renkawitz-R (2007) The Drosophila insulator proteins CTCF and CP190 link enhancer blocking to body patterning. EMBO J 26:4203–4214. https://doi.org/10.1038/sj.emboj.7601851
Moon H, Filippova G, Loukinov D, Pugacheva E, Chen Q, Smith ST, Munhall A, Grewen B, Bartkum M, Arnold R, Burke LJ, Renkawitz-Pohl R, Ohlsson R, Zhou J, Renkawitz-R, Lobanankov V (2005) CTCF is conserved from Drosophila to humans and confers enhancer blocking of the Fab-8 insulator. EMBO Rep 6:165–170. https://doi.org/10.1038/sj.emboj.7400334
Nitin E, Wimmer EA (2011) Second order regulator Collier directly controls intercalary-specific segment polarity gene expression. Dev Biol 360:403–414. https://doi.org/10.1016/j.ydbio.2011.09.035
O’Leary NA, Wright MW, Brister JR, Ciulo S, Haddad D, McVeigh R, Rajput B, Robbertse B, Smith-White B, Ako-Adjei AA, Bradetdin A, Bao Y, Blinkova O, Brover V, Chevermin V, Choi J, Cox E, Ermolaeva O, Farrell CM, Goldfarb T, Gupta T, Haft D, Hatcher E, Hivaia W, Joardar VS, Kodali VK, Li W, Maglott D, Masterson M, McCarvey KM, Murphy MR, O’Neill K, Pujar S, Rangwala SH, Rausch D, Riddle LD, Schoch C, Shkeda A, Storz SS, Sun H, Thibaud-Nissen F, Toulty I, Tully RE, Vatsar AN, Wallin C, Webb D, Wu W, Landrum MJ, Kimichi A, Tatusova T, DiCuccio M, Kitts P, Murphy TD, Pruitt KD (2016) Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. Nucleic Acids Res 44:D733–D745. https://doi.org/10.1093/nar/gkv1189
Paese CLB, Schoenauer A, Leite DJ, Russel S, McGregor AP (2018) A SoxB gene acts as an anterior gap gene and regulates posterior segment addition in a spider. eLife 7:e37567. https://doi.org/10.7554/eLife.37567
Peichmann M, McGregor AP, Schwager EE, Feistosa NM, Damen WGM (2009) Dynamic gene expression is required for anterior regionalization in a spider. Proc Natl Acad Sci 106:1468–1472. https://doi.org/10.1073/pnas.0811501010
Peichmann M, Khadjeh S, Sprenger F, Ppric NM (2010) Pattern mechanisms and morphological diversity of spider appendages and their importance for spider evolution. Arthropod Struct Dev 39:453–467. https://doi.org/10.1016/j.asd.2010.07.007
Peichmann M, Schwager EE, Turetzek N, Ppric NM (2015) Regressive evolution of the arthropod tritocerebral segment linked to functional divergence of the Hox gene labial. Proc R Soc B Biol Sci 282:20151162. https://doi.org/10.1098/rspb.2015.1162
Peel A (2004) The evolution of arthropod segmentation mechanisms. BioEssays 26:1108–1116. https://doi.org/10.1002/bies.20097
Peel AD, Chipman AD, Akam M (2005) Arthropod segmentation: beyond the Drosophila paradigm. Nat Rev Genet 6:905–916. https://doi.org/10.1038/nrg1724
Posnien N, Buchar G (2010) Formation of the insect head involves lateral contribution of the intercalary segment, which depends on Te-labial function. Dev Biol 338:107–116. https://doi.org/10.1016/j.ydbio.2009.11.010
Posnien N, Zeng V, Schwager EE, Peichmann M, Hillbrant M, Keefe JD, Damen WGM, Ppric NM, McGregor AP, Ex bouts CM (2014) A
comprehensive reference transcriptome resource for the common house spider Parasteatoda tepidariorum. PLoS One 9:e104885. https://doi.org/10.1371/journal.pone.0104885

Price MN, Dehal PS, Arkin AP (2009) FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. Mol Biol Evol 26:1641–1650. https://doi.org/10.1093/molbev/msp077

Prpic NM, Damen WGM (2008) Arthropod appendages: a prime example for the evolution of morphological diversity and innovation. In: Evolving pathways: key themes in evolutionary developmental biology. Cambridge University Press, Cambridge, pp 381–398

Prpic NM, Schopppmeier M, Damen WGM (2008) Collection and fixation of spider embryos. Cold Spring Harb Protoc 3. https://doi.org/10.1101/pdb.prot5067

Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572–1574. https://doi.org/10.1093/bioinformatics/btg180

Ronquist F, Teslenko M, Van Der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP (2012) MrBayes 3.2: efficient baysian phylogenetic inference and model choice across a large model space. Syst Biol 61:539–542. https://doi.org/10.1093/sysbio/sys029

Schaeper ND, Pechmann M, Damen WGM, Prpic NM, Schöck F, Reischl J, Wimmer E, Taubert H, Purnell BA, Jäckle H (2000) Evolutionary plasticity of dachshund expression in early embryonic patterning and brain development. Development 124:4297–4308

Rushlow CA, Han K, Manley JL, Levine M (1989) The graded distribution of the dorsal morphogen is initiated by selective nuclear transport in Drosophila. Cell 59:1165–1177

Schaeper ND, Pechmann M, Damen WGM, Prpic NM, Wimmer EA (2010) Evolutionary plasticity of collar function in head development of diverse arthropods. Dev Biol 344:363–376. https://doi.org/10.1016/j.ydbio.2010.05.001

Schöck F, Rausch J, Wimmer E, Taubert H, Purnell BA, Jäckle H (2000) Phenotypic suppression of empty spiracles is prevented by buttonhead. Nature 405:351–354. https://doi.org/10.1038/35012620

Schomburg C, Turetzek N, Schacht MI, Schneider J, Kirfel P, Prpic NM, Posnien N (2015) Molecular characterization and embryonic origin of the eyes in the common house spider Parasteatoda tepidariorum. Evodevo 6:15. https://doi.org/10.1186/s12227-015-0011-9

Schönauer A, Paese CLB, Hilbrant M, Leite DJ, Schwager EE, Feitosa FC, Speight AD, Akiyama-Oda Y, Esposito L, Bechsgaard J, Bilde T, Buffry AD, Lengyel JA (1997) Complex regulatory region mediating dorsal expression in early embryonic patterning and brain development. Genome Biol 3:research0088.1–0088.14. https://doi.org/10.1186/gb-2002-3-12-research0088

Schomburg C, Turetzek N, Schacht MI, Schneider J, Kirfel P, Prpic NM, Schöffmeier M, Damen WGM, McGregor AP (2016) The Wnt and dachshund gene complex. Development 143:2455–2463. https://doi.org/10.1242/dev.131656

Schwengruber B, Chourrout D, Vervoort M, Leblanc B, Cavalli G (2007) Genome regulation by polycomb and trithorax proteins. Curr Opin Neurobiol 17:169–176. https://doi.org/10.1016/j.conb.2007.07.013

Schwager EE, Pechmann M, Feitosa M, McGregor AP, Damen WGM (2009) lunchback functions as a segmentation gene in the spider Achaearanea tepidariorum. Curr Biol 19:1333–1340. https://doi.org/10.1016/j.cub.2009.06.061

Schwager EE, Sharma PP, Clarke T, Leite DJ, Wierschin T, Pechmann M, Akiyama-Oda Y, Esposito L, Bechsgaard J, Bilde T, Buffry AD, Chao H, Dinh H, Dodapaneni HV, Dugan S, Eibner C, Eutvou CG, Funch P, Garb J, Gonzalez LB, Gonzalez VL, Grifiths-Jones S, Han Y, Hayashi C, Hilbrant M, Hughes DST, Janssen R, Lee SL, Maes I, Murali SC, Muzny DM, Nunes da Fonseca R, Paese CLB, Qu J, Ronshaugen M, Schomburg C, Schönauer A, Stollewerk A, Torres-Oliva M, Turetzek N, vanhournout B, Werren JH, Wolf C, Worley KC, Bucher G, Gibbs RA, Coddington J, Oda H, Stanke M, Ayoub NA, Prpic NM, Flot JF, Posnien N, Richards S, McGregor AP (2017) The house spider genome reveals an ancient whole-genome duplication during arachnid evolution. BMC Biol 15:62. https://doi.org/10.1186/s12862-020-01559-6

Sharma PP, Gupta T, Schrager EE, Wheeler WC, Extavour CG (2014) Subdivision of arthropod cap-n-collar expression domains is restricted to Mandibulata. EvoDevo 5:3. https://doi.org/10.1186/2041-9139-5-3

Sievers F, Higgins DG (2014) Clustal Omega. Curr Protoc Bioinformatics 48:3.13.1–3.13.16. https://doi.org/10.1002/9781118145440.scg013a48

Skeath JB, Thor S (2003) Genetic control of Drosophila nerve cord development. Curr Opin Neurobiol 13:8–15. https://doi.org/10.1016/S0959-4388(03)00007-2

Stollewerk A (2002) Recruitment of cell groups through Delta/Notch signalling during spider neurogenesis. Development 129:5339–5348. https://doi.org/10.1242/dev.00109

Stollewerk A, Tautz D, Weller M (2003) Neurogenesis in the spider: new insights from comparative analysis of morphological processes and gene expression patterns. Arthropod Struct Dev 32:5–16. https://doi.org/10.1016/S1467-8039(03)00041-0

Thurmond J, Goodman JL, Strellets VB, Attrill H, Gramates LS, Marygold SJ, Matthews BB, Millerburn G, Antonazzo G, Trovisco V, Kaufman TC, Calvi BR, FlyBase Consortium (2019) FlyBase 2.0: the next generation. Nucleic Acids Res 47:759–765. https://doi.org/10.1093/nar/gky1003

Tomanick P, Beaton A, Weizsamm R, Kwan E, Shu SQ, Lewis SE, Richards S, Ashburner M, Hartenstein V, Celniker SE, Rubin GM (2002) Systematic determination of patterns of gene expression during Drosophila embryogenesis. Genome Biol 3:research0088.1–0088.14. https://doi.org/10.1186/gb-2002-3-12-research0088

Tomanick P, Berman BP, Beaton A, Weizsamm R, Kwan E, Hartenstein V, Celniker SE, Rubin GM (2007) Global analysis of patterns of gene expression during Drosophila embryogenesis. Genome Biol 8:R145. https://doi.org/10.1186/gb-2007-8-7-r145

Turetzek N, Pechmann M, Schomburg C, Schneider J, Prpic NM (2015) Neofunctionalization of a duplicate dachshund gene underlies the evolution of a novel leg segment in arachnids. Mol Biol Evol 33:109–121. https://doi.org/10.1093/molbev/msv200

Turetzek N, Khadje S, Schomburg C, Prpic NM (2017) Rapid diversification of homothorax expression patterns after gene duplication in spiders. BMC Evol Biol 17:168. https://doi.org/10.1186/s12862-017-1013-0

UniProt Consortium (2019) UniProt: a worldwide hub of protein knowledge. Nucleic Acids Res 47:D506–D515. https://doi.org/10.1093/nar/gky1049

Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG (2012) Primer3-nc. new capabilities and interfaces. Nucleic Acids Res 40:e115. https://doi.org/10.1093/nar/gks596

Urba R, Techau GM (2003) Molecular markers for identified neuroblasts in the developing brain of Drosophila. Development 130:3621–3637. https://doi.org/10.1242/dev.005535

Veraksa A, McGinnis N, Li X, Mohler J, McGinnis W (2000) Cap ‘n’ collar B cooperates with a small MaF subunit to specify pharyngeal development and suppress deformed homeotic function in the Drosophila head. Development 127:4023–4037

Weihe U, Dorfman R, Wernet MF, Cohen SM, Milän M (2004) Proximal subdivision of Drosophila legs and wings: the elbow-no ocelli gene complex. Development 131:767–774. https://doi.org/10.1242/dev.00979

Williams TA, Nagy LM (2001) Developmental modularity and the evolution of arthropod limbs. J Exp Zool 291:241–257. https://doi.org/10.1002/jez.1101

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