

RESEARCH

Open Access



Single-cell RNA sequencing of mid-to-late stage spider embryos: new insights into spider development

Brenda I. Medina-Jiménez^{1*}, Graham E. Budd¹ and Ralf Janssen^{1*}

Abstract

Background The common house spider *Parasteatoda tepidariorum* represents an emerging new model organism of arthropod evolutionary and developmental (EvoDevo) studies. Recent technical advances have resulted in the first single-cell sequencing (SCS) data on this species allowing deeper insights to be gained into its *early* development, but mid-to-late stage embryos were not included in these pioneering studies.

Results Therefore, we performed SCS on mid-to-late stage embryos of *Parasteatoda* and characterized resulting cell clusters by means of *in-silico* analysis (comparison of key markers of each cluster with previously published information on these genes). *In-silico* prediction of the nature of each cluster was then tested/verified by means of additional *in-situ* hybridization experiments with additional markers of each cluster.

Conclusions Our data show that SCS data reliably group cells with similar genetic fingerprints into more or less distinct clusters, and thus allows identification of developing cell types on a broader level, such as the distinction of ectodermal, mesodermal and endodermal cell lineages, as well as the identification of distinct developing tissues such as subtypes of nervous tissue cells, the developing heart, or the ventral sulcus (VS). In comparison with recent other SCS studies on the same species, our data represent later developmental stages, and thus provide insights into different stages of developing cell types and tissues such as differentiating neurons and the VS that are only present at these later stages.

Keywords Single-cell sequencing, Spider development, Nervous system, Genetic fingerprint, *Parasteatoda tepidariorum*

Introduction

Arthropod EvoDevo studies, i.e. the investigation of the development and the evolution of arthropods, still relies mainly on data from a single model organism, the vinegar fly *Drosophila melanogaster* (reviewed in e.g. [15].

Drosophila development is in many aspects derived however, and the processes and underlying genetic networks that lead to the development from the fertilized egg to the imago (the adult form) cannot easily be compared with the development of other arthropods (e.g. [38]). The introduction of new model organisms, and especially the gaining of deeper insights into existing model organisms that represent less derived modes of development is thus of the greatest interest for arthropod EvoDevo research.

The common house spider *Parasteatoda tepidariorum* (formerly known as *Achaearanea tepidariorum*) represents an emerging new model organisms that is often used for comparative developmental studies and

*Correspondence:

Brenda I. Medina-Jiménez

brenda.medina@geo.uu.se

Ralf Janssen

ralfjanssen@geo.uu.se

¹ Department of Earth Sciences, Palaeobiology, Uppsala University, Villavägen 16, 75236 Uppsala, Sweden



© The Author(s) 2024. **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/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

the study of arthropod evolution in general (reviewed in [70, 139]). In the last two decades, many standard and advanced molecular biological methods have been established for this species, leading to a large number of comparative studies. Most of these studies, however, have relied on classic candidate gene approaches in which genes that are known to play a certain role during *Drosophila* development (or other established model organisms) are investigated in non-model organisms or emerging model organisms including *Parasteatoda* (e.g. [87, 138, 162, 163]).

In order to investigate genes (and thus gene regulatory networks) beyond the classic candidate gene approach, genomics and embryonic and tissue-specific transcriptomics have been applied in the recent past, leading to the investigation of complete gene families, orphan genes, fast evolving genes and tissue-specific genes (e.g. [20, 77, 89, 142, 155]). A relatively recent new technical innovation now allows investigation of the transcriptome of every cell of a given organism or a developmental stage of a given organism separately. This so-called single-cell

sequencing (SCS) technology can thus give much more detailed insights into gene transcription leading to the identification of genetic fingerprints that are specific for a certain cell type, or the elucidation of trajectories (consecutive developing stages) of differentiating cells (e.g. reviewed in [59, 114]).

Two very recent studies applying SCS on embryos of *Parasteatoda* covered the early stages of spider development. The first paper [2] investigated the so-called germ-disc stage of spider development (stage 5, staging after Mittmann and Wolff [130]) (Fig. 1). Early during development, a spherical disc forms that in a subsequent step transforms into a bilaterally-symmetrical germ band (e.g. [130]). The second paper [111] investigated the earliest stages of germ band development (stages 7–9). These stages include the posterior addition of segments, the onset of nervous system development and the beginning of limb bud formation (Fig. 1).

In our study, we investigated developmental stages 10–12. These stages are characterized by for example the end of segment addition, the outgrowth and development

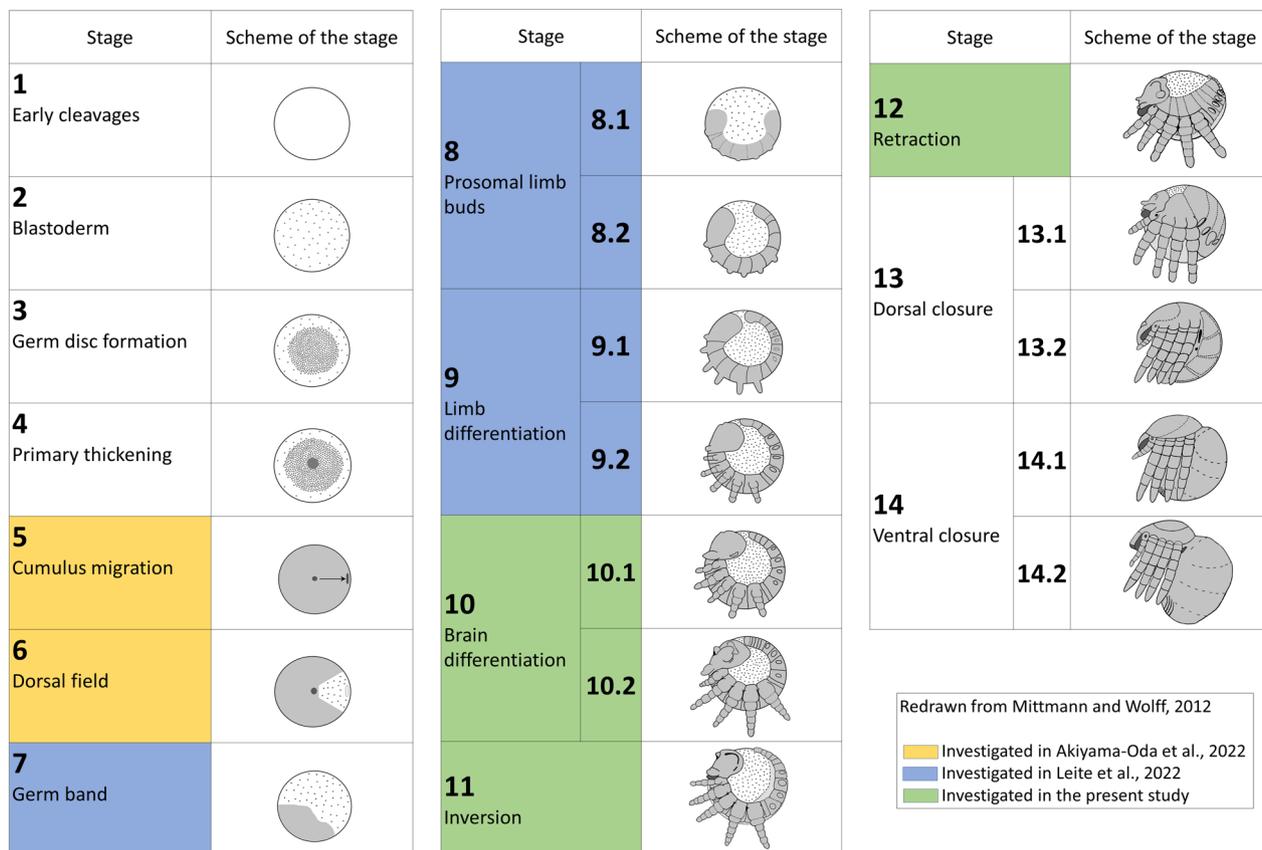


Fig. 1 Stages of spider development. Redrawn after Mittmann and Wolff [130]. The 14 previously described developmental stages of the spider *Parasteatoda tepidariorum* are shown. Stages covered by SCS-analysis of Akiyama-Oda et al. [2] is highlighted in orange. Stages covered by Leite et al. [111] are highlighted in blue, and the stages addressed in our study are highlighted in green

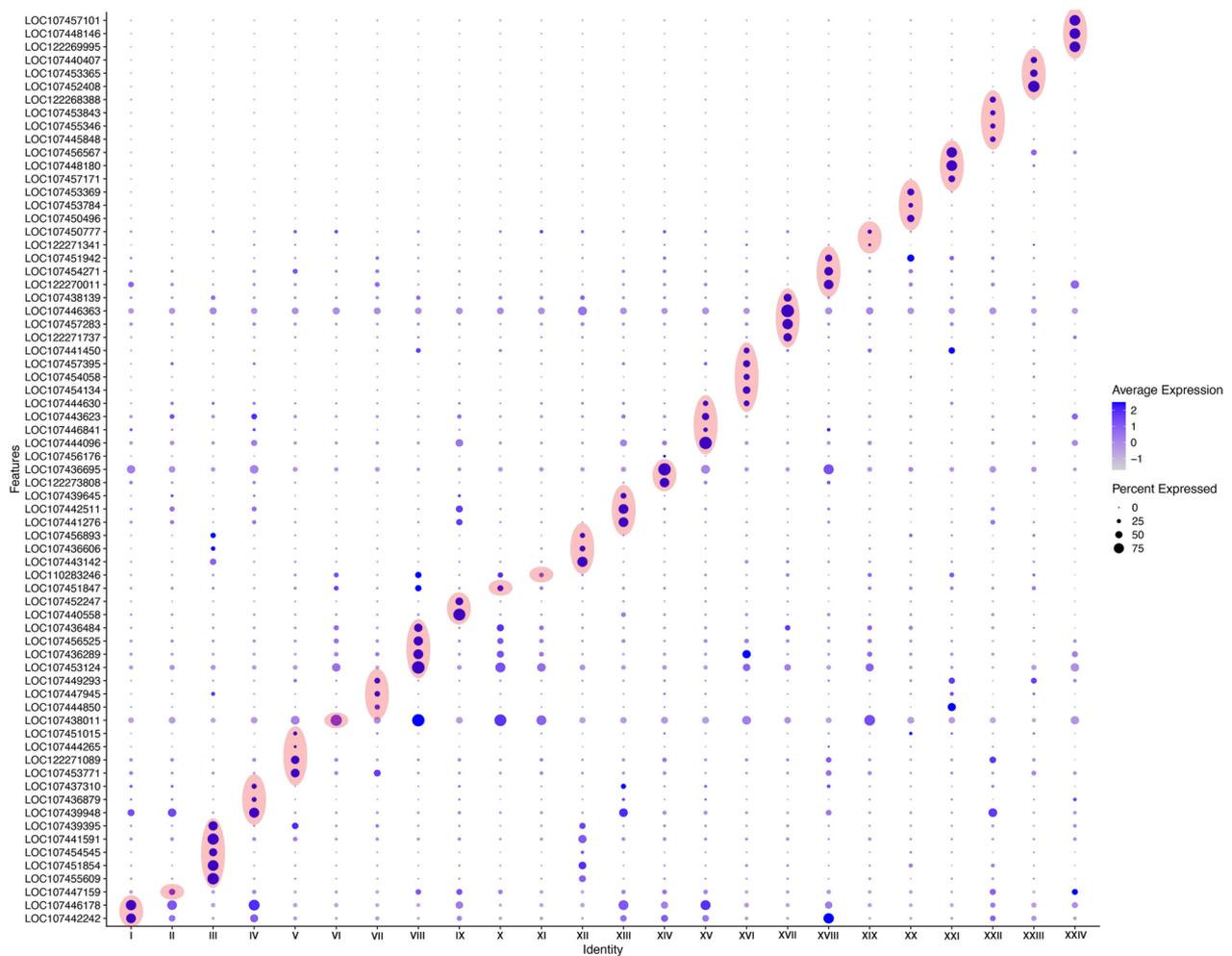


Fig. 2 Dotplot showing all genes (Y axis) per cluster (X axis) for which whole mount *in situ* hybridisation was conducted in this study

of the appendages, nervous system differentiation, heart formation, dorsal closure and ventral splitting of the germ band (Fig. 1). In summary, developmental stages 10–12 likely represent a higher degree of cell-type specification and organogenesis compared with the previous studies on earlier developmental stages. The previous work on the earlier stages of development and our data complement each other and thus each contribute to a larger project driven by the international spider research community to cover all developmental stages of spider development by means of SCS [2, 111]. Our data contribute to this goal by delivering new insights into the differentiation of the spider’s central nervous system (CNS), the patterning of the appendages (including the highly-derived opisthosomal (posterior) appendages), morphogenic movements of cells, and organogenesis. During the course of our investigation, we first characterized the 24

identified cell clusters by means of *in-silico* analysis (literature analysis) using previously known information on marker genes of each cluster from other arthropods and, whenever available, also spiders including *Parasteatoda*. In order to test and verify the *in-silico* based predictions, we then investigated the embryonic expression patterns of 68 genes representing prominent markers of each cell cluster that (for the most) have not been studied previously in *Parasteatoda* (Fig. 2).

Methods

Embryonic tissue dissociation, cell capture, cDNA library preparation, and single-cell RNA sequencing were performed at the Department of Developmental Biology and Gene Core facilities of the European Molecular Biology Laboratory (EMBL), Heidelberg, Germany.

Embryonic tissue dissociation, cell capture, cDNA library preparation, and single-cell mRNA sequencing

In a watch glass, stage 10–12 embryos of *Parasteatoda tepidariorum* were submerged in phosphate buffered saline (PBS), and the vitelline membranes were removed using tweezers. As much yolk as possible was removed manually with tweezers and a pipette to prevent it from clogging the device used for later cell capture. The tissue-suspension was transferred into 1 ml of PBS and collected in a 1.5 ml low protein binding tube. To reduce the number of fat-droplets, the tissue-suspension was shaken on a shaking-platform at 1100 rpm for 5 min, then 100 µl of surface liquid collecting the fat-droplets was discarded and replaced with fresh PBS. To get rid of any remaining fat-droplets, the suspension was centrifuged at 800 rpm for 4 min, the supernatant was discarded, and the pellet was re-suspended in PBS. The latter centrifugation step was repeated one more time.

Cells were disaggregated (cell dissociation) by intermittent pipetting and swirling of the remaining embryonic tissue in an enzymatic mixture of 2% pronase and 2% sodium thioglycolate in PBS. The dissociation progress was constantly monitored under a microscope. When cell aggregates were no longer visible, the suspension was filtered using a 40 µm cell strainer to remove remaining tissue clumps, yolk granules and debris (e.g. strings of leaked DNA). This cell suspension was then centrifuged at 1100 rpm for 5 min to collect intact cells. The supernatant containing leaked RNA molecules and remaining small pieces of debris was removed. Finally, the cell suspension was filtered through a 20 µm cell strainer twice, and collected into a new 1.5 ml low protein binding tube. Cells were counted using a hemocytometer and cell viability was assayed using Fluorescein Diacetate and Propidium Iodide. A detailed tissue dissociation protocol is available in the supplementary data (Supplementary File 1).

The cell suspension was loaded on a 10X Genomics Chromium Controller™, and cell capture was performed following the Chromium Next GEM Single Cell 3' version 3.1 protocol, a process by which transcripts from every cell are labelled with a barcoded oligonucleotide. RT-PCR allowed the synthesis of cDNA from each cell's transcriptome. cDNA quality was checked via electrophoresis on an agarose 2% E-gel precast system (Invitrogen, Cat. No. G402022) (Supplementary Fig. 1). Subsequently, a cDNA library was prepared for sequencing. Paired-end sequencing was performed on an Illumina NextSeq500 platform. Sequenced short reads were de-multiplexed and converted into compressed "(fastq)"-formatted files.

Quality control and mapping of raw data

The bioinformatics tool FastQC (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) was used to check

the quality of the sequencing. A report is available in the supplementary data (Supplementary File 2).

The *Parasteatoda* genome (version Ptep 3.0) and corresponding annotation were downloaded from NCBI (<https://www.ncbi.nlm.nih.gov/genome/?term=Parasteatoda+tepidariorum>).

High levels of mitochondrial transcripts indicate cell stress, which would render the cells unfit for downstream analysis. To identify and remove such cells from our data, a nearly complete mitochondrial genome was assembled and concatenated to the genome prior to mapping. For its assembly, raw DNA sequencing reads (SRR891584) were used. These reads come from the BioProject PRJNA167405 conducted to assemble the *Parasteatoda* genome [166]. The mitochondrial genes present in this dataset were searched for and assembled using GetOrganelle v1.7.3.3, a toolkit for de novo assembly of organelle genomes (<https://github.com/Kinggerm/GetOrganelle>), with kmer sizes 21, 45, 65, 85, 105. The scaffold was annotated using the MITOS server [17], and this information was added to the gene coordinates file (in GTF format) of the genome. The assembled mitochondrial scaffold ("txt") is available in the supplementary data (Supplementary File 3).

Mapping of the reads to the reference genome was done using the Cell Ranger 'mkref' and 'count' pipelines from the Single Cell Software Suite offered by 10×Genomics (version 6.0.2). The output included a web summary in ".html" format (Supplementary File 3), and a folder containing three compressed files in ".gz" format: a list of putative genes and cells ("features.tsv" and "barcodes.tsv", respectively), and a matrix that displays the number of unique transcripts (UMIs) per cell and per gene ("matrix.mtx"). This information is available in a compressed format ".zip" in the supplementary data (Supplementary File 3).

Data processing and downstream analysis

The gene expression matrix was loaded into Seurat ver. 4.1.0 [156] to be processed for downstream analysis (i.e. dimensional reduction, clustering of cells by differential gene expression, identification of cluster markers). In order to filter out low quality cell barcodes (i.e. background RNA encapsulated in droplets, and cells undergoing stress), a subset was made that consisted of cells that express between 250 and 2500 genes, and contained between 1400 and 3500 transcripts, and in which less than 5 percent of transcripts were representing mitochondrial genes. Normalization and variance stabilization of the dataset were performed with the R package SCTransform [30, 61] following the 'glmGamPoi' method. A principal component analysis (PCA) was run for linear dimensional reduction. To visualize the data, 50 principal components (PCs) were selected for Uniform Manifold and Projection (UMAP) analysis. To

construct a Shared Nearest Neighbours (SNN) graph, 50 PCs were used. Cells were clustered together using the Seurat function “FindClusters” at a resolution of 2. The Seurat function ‘FindAllMarkers’ was used to identify the differentially expressed genes (DEGs) to be used for cluster identification. Only the DEGs that are upregulated (only.pos=TRUE), and those that are expressed in at least 1% of the cells in a cluster (min.pct=0.01) were considered. For more specific details regarding the parameters followed for cell filtering, clustering, and marker selection, we provide an R notebook file (“.Rmd”) and R object (“.rds”) containing our data analysis, as well as a gene annotation table (“.tsv”). These files are available in the supplementary data (Supplementary File 4). The top most differentially expressed markers per cluster were selected for *in silico* cell type identification (i.e. by literature review). A spreadsheet containing the markers per cluster is available in the supplementary data (Supplementary File 5). Beyond the *in-silico* analysis we also conducted additional whole-mount *in-situ* hybridization (WISH) experiments to a) verify *in-silico* based cluster-identification of marker genes for which WISH data were not available for *Parasteatoda*, and b) to investigate the spatial expression of marker genes for which comparative data from spiders and other organisms were not available at all, or not in sufficient extent and quality (and thus were not conclusive in the *in-silico* analysis).

Gene amplification, probe synthesis, *in-situ* hybridization, nuclear staining, and data documentation

Total RNA was extracted from a mix of embryos of different developmental stages using TRIZOL (Invitrogen, Cat. No. 15596029). mRNA was isolated from this total RNA using the Dynabeads mRNA Purification Kit (Invitrogen, Cat. No. 61006). cDNA was synthesized from mRNA using RevertAid H Minus First Strand cDNA Synthesis Kit (Thermo scientific, Cat. No. K1631). For fragments of most genes, we performed an initial polymerase chain reaction (PCR), and a subsequent nested (or semi-nested) PCR with a set of internal primers (primer sequences are listed in Supplementary File 5). All backward primers were equipped with a 5′-T7-RNA promotor sequence (gggTAA TACGACTCACTATAG) [37]). The three extra guanines serve as protection against degradation of the PCR fragments prior to probe synthesis. PCRs were purified using the QIAGEN PCR purification kit (QIAGEN, Cat. No. 28104). The purified PCRs directly served as templates for subsequent probe synthesis with T7 RNA polymerase (ROCHE, Cat. No. 10881767001). Synthesized probes were

purified using the QIAGEN RNeasy Kit (QIAGEN, Cat. No. 74104). Whole mount *in-situ* hybridization was performed as previously described [85]. Whenever indicated, for better display, embryos were incubated in glycerol and flat-mounted. Appendages were dissected from glycerol-incubated embryos. SYBR Green (incubation of stained embryos in 1:10,000 SYBR Green in phosphate buffered saline with 0.1% Tween-20 (PBST-0.1%) for 20–30 min) was used to better visualize the morphology of embryos. Stained embryos were photographed under a MZ-FLIII Leica dissection microscope equipped with a Leica DC490 digital camera and an external UV-light source. Linear adjustments on colour, contrast and brightness were performed applying the image-processing software Adobe Photoshop 2022.

Results and discussion

Single-cell mRNA sequencing of stage 10–12 spider embryos

Embryonic tissue dissociation and single cell isolation were conducted using a mix of 50 *Parasteatoda tepidariorum* embryos at stages 10–12. This yielded 70 µl of suspension containing nearly 90% of live cells with a concentration of 920 cells/µl. Single cell mRNA sequencing was performed using equipment and reagents kits provided by the droplet-based technology Chromium Single Cell 3′ Gene Expression™ sold by 10xGenomics®. Following the specifications in the User Guide, 30 000 cells were loaded on the Chromium Controller™. After cell capture, the RT-PCR generated cDNA had a concentration of 1.43 ng/µl. Quality control performed via electrophoresis showed that the cDNA was of good quality, with little degradation of mRNAs (Supplementary Fig. 1). The subsequent cDNA library had a concentration of 37.6 ng/µl. Paired-end sequencing generated 330,047,610 short reads and a sequencing saturation of 11.1%. Quality control via FastQC indicated an overall good base calling process (i.e. high Phred scores for all reads).

Mapping the reads onto the genome, filtering and processing the resulting gene expression matrix using Cell Ranger pipelines and Seurat v4.1.0 yielded a set of 4103 cells expressing 16 669 genes. These cells were grouped into 24 clusters (indicated by Roman Numerals I–XXIV), with 2784 markers in total (Fig. 3; Table 1); note that some genes represent markers for more than one cluster.

Marker genes for each cluster were determined using a “cluster against all-other-clusters” approach, including all genes that were expressed in at least 1% of the cells of their associated cluster and a return threshold *p*-value of 0.01. “Top-marker genes” are defined as genes with the best in-cluster against all-other-clusters values, i.e. genes with the lowest *p*-value.

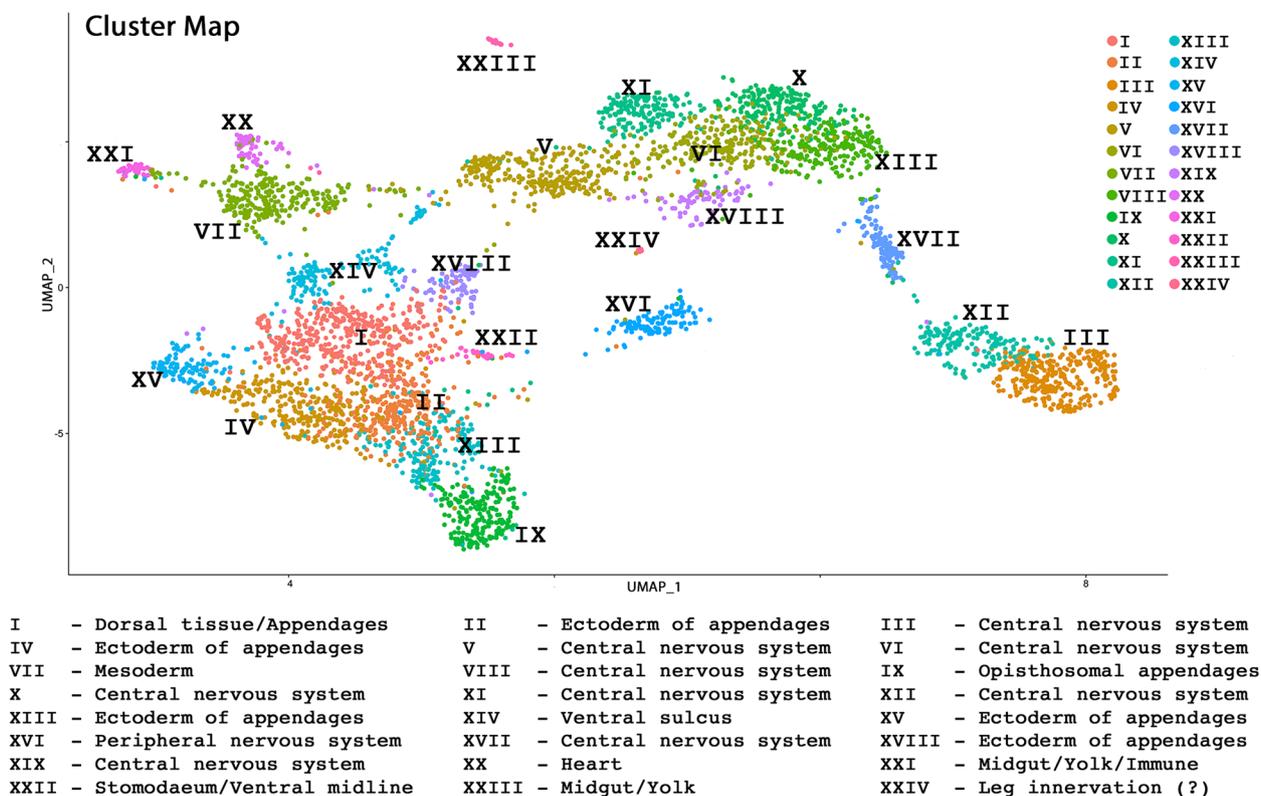


Fig. 3 Integrated data UMAP showing 24 identified clusters. A short description of the nature of each cluster is given below the UMAP data set

Cluster analysis

Cluster-I: Dorsal tissue, and ectoderm patterning of developing appendages (1)

The top markers of Cluster-I have not been investigated in great detail previously, and it is therefore difficult to determine the nature of Cluster-I cells by *in-silico* analysis. Somewhat further down the list of markers, however, we find genes that have previously been studied in *Parasteatoda* and other spiders: *Frizzled-4.1* (*Fz4.1*) (LOC107442148) and *Frizzled-4.2* (*Fz4.2*) (LOC107441380) [82], *optomotor-blind* (*omb*)/*Tbx3* (LOC107450980) [81], *decapentaplegic* (*dpp*) (LOC107441097) [149], and *irx2* (LOC107439315) [110]. *Fz4.1* and *irx2* are strongly expressed in dorsal tissue of the developing *Parasteatoda* embryo. All of these genes have in common that they are expressed in various patterns in the developing appendages.

We tried to substantiate the suggestion that Cluster-I cells could represent dorsal derivatives of the spider embryo and/or the developing appendages by additional whole-mount *in-situ* hybridization (WISH) experiments choosing markers that are high up in the list, and that are expressed as tissue (cluster) specific as possible. Therefore, we chose *hexosaminidase-1* (*hex1*) (LOC107442242) and *elongation of very long chain fatty acids protein*

(*elovl7*) (LOC107446178). These two genes first are expressed in complementing leg gap-gene like patterns proximally and distally respectively and later in rings in the appendages suggesting a function in joint formation. Importantly, however, both genes are also dominantly expressed in dorsal tissue of both the pro- and opisthosoma, while ventral tissue does not express these genes (Fig. 4A-H and Supplementary Figs. 2 and 3). This pattern is very similar to the patterns of *irx2* and *Fz4.1* in dorsal tissue supporting the suggestion that the Cluster-I represents dorsal tissue.

With respect to the somewhat later expression of Cluster-I genes in what we believe are the developing joints, it is worth mentioning that also the Cluster-I marker *dpp* is expressed in a comparative pattern in the developing joints [149]. It is thus likely that Cluster-I indeed also harbours cells that are involved in spider joint formation.

Cluster-II: Ectoderm patterning of developing appendages (2)

Most top markers of Cluster-II represent uncharacterized genes. Among the characterized markers, however, range a number of genes that have previously been described for *Parasteatoda*, other spiders and/or other arthropods. Notably, a large number of these markers are expressed in the ectoderm of developing appendages,

Table 1 Number of markers and cells per cluster. Clusters listed in Roman numerals

Cluster	No. markers	No. cells
I	247	393
II	57	318
III	409	304
IV	148	289
V	118	281
VI	72	269
VII	110	264
VIII	144	223
IX	221	218
X	106	212
XI	43	187
XII	80	172
XIII	199	168
XIV	46	140
XV	75	122
XVI	114	120
XVII	205	108
XVIII	150	84
XIX	42	79
XX	173	58
XXI	239	41
XXII	151	27
XXIII	361	21
XIV	395	5
	3905	4103

and are involved in appendage development, such as *disconnected (disco)* (aka *basonuclin*) (LOC107447159) (Fig. 4I-L and Supplementary Fig. 4), *spineless2 (ss2)* (LOC107457395) (Fig. 10F-H, cf. Cluster-XVI), *aristaless (al)* (LOC107448374), *AP2.2* (LOC107443623) (Fig. 9U-X, cf. Cluster-XV) (note that Leite et al. [111] investigated another paralog of this gene, *AP2* (LOC107452006)), *Fz4.1* (LOC107442148), *Fz4.2* (LOC107441380), and *unchoordinated-5.1 (unc5.1)* (LOC107445619) (e.g. [4, 6, 45, 64, 82, 83, 88, 109, 111, 132, 134, 141, 150, 160, 167]).

We therefore believe that Cluster-II cells generally contribute to the developing appendages. At least some of the investigated genes are expressed in the form of rings that could indicate a function in proximal–distal axis patterning and thus joint formation.

Cluster-III: The developing central nervous system (CNS) (1) – Differentiating and differentiated neurons (1)

In-silico analysis of Cluster-III gene markers strongly suggests that these cells represent part of the developing nervous system. Best markers are *sax3* (LOC107455609), and two paralogs of *neural-cadherin (Ncad1)* (LOC107454545) and

Ncad2 (LOC107451854)). Other high-ranked markers are *follistatin-related protein 5 (fstl5)* (LOC107441591), *Dscam-2* (LOC107456604), *pikachurin (pika)* (LOC107439395), and *ELAV-like protein 3 (Elavl3)* (LOC107436216). Of these markers, *sax3*, *pika*, *fstl5*, and *Elavl3* are unique markers for this cell cluster (see Supplementary Fig. 5). *Ncad1*, *Ncad2* and *Dscam2* are markers of both Cluster-III and the closely related Cluster-XII (cf. chapter on Cluster-XII). These genes are all involved in neuron development, neuron differentiation, and neuronal pathfinding, and thus also in differentiated neurons (e.g. [36, 43, 66, 75, 92, 140, 151, 191, 203]). Beyond that, many of the lesser markers of this cluster indeed also represent genes typically expressed in neurons, such as the *Dscams* (e.g. LOC107456604, LOC107444681, LOC107450135, LOC107445852, LOC107452025, LOC107448201, LOC107443039, LOC107455980) (e.g. reviewed in [171], and synaptotagmins (LOC107455395, LOC107449342) (e.g. reviewed in [195]). It is therefore reasonable to suggest that Cluster-III cells represent differentiating and mature neurons.

The expression and function of most of these genes has not been studied in great detail in arthropods other than *Drosophila*, and thus data from spiders are not available. *In-situ* hybridization of some of these genes show, as expected, that they are expressed in the developing central nervous system (CNS), and some are expressed in the peripheral nervous system (PNS) too (Fig. 5A-L and Supplementary Figs. 5–8).

Cluster-IV: Ectoderm patterning of developing appendages (3)

Most of the top markers of Cluster-IV have not been studied in great detail, or not at all. Two of these markers, *cuticular protein 47Ef-like (cp47EF-like)* (LOC107436879) and *chitinase-7 (cht7)* (LOC107437310), are exclusively expressed in the developing joints (Fig. 4M-T and Supplementary Figs. 9 and 10). Another marker, the protease *notopleural (np)* (LOC107439948), is expressed in dorsal tissue and what we believe are the developing joints of the appendages, very similar like the markers of the closely-related Cluster-I and the aforementioned markers of Cluster-IV (Fig. 4U-X and Supplementary Fig. 11). Interestingly, another marker of Cluster-IV is *trachealeless (trh)* (LOC110283103) which has been studied in *Parasteatoda* and other chelicerates such as a harvestman and a scorpion, is also expressed in the form of rings in the developing appendages suggesting a possible role in joint formation [168]. Likewise, *AP2.2* (LOC107443623) (cf. Cluster-XV) is expressed in the form of rings in the appendages (Fig. 9U-X). Other previously investigated and less prominent markers of this cluster (e.g. the two *Fz4* ohnologs (LOC107442148, LOC107441380) [82], *Wnt6*, *Wnt5*, and *Wnt1* (LOC107438387, LOC107445649, and

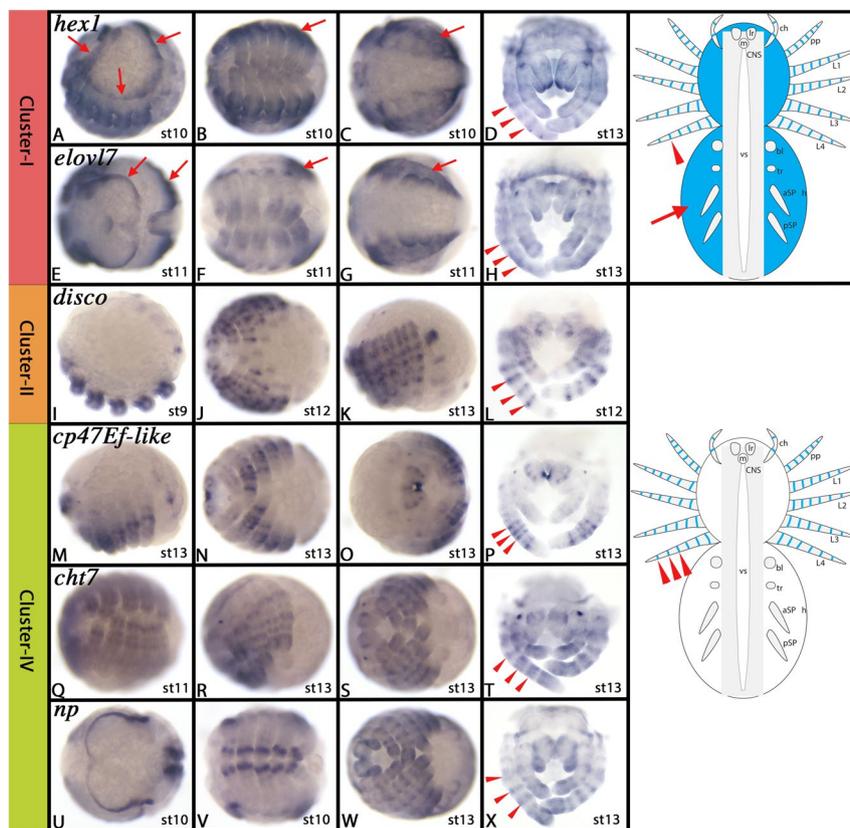


Fig. 4 Gene expression of Cluster-I, -II, and -IV markers. In all panels, anterior is to the left, except panels D, H, L, P, T and X where anterior is up. Panels A, I, K, M and R represent lateral views. Panels B, C, E, F, G, J, N, Q, S, V and W represent ventral views. Panels D, H, L, P, T and X represent anterior ventral views. Schematic drawings in the right column summarize the main gene expression characteristic for a given cell cluster (blue tissue). Arrows and arrowheads in schematic drawings point to comparable expression as marked by the same symbols in the original photographs. Abbreviations of gene names is indicated in the upper left corner. Developmental stages are indicated in the bottom right corner. For further information, see Supplementary Figs. 2–4 and 9–11, and Supplementary File 6. Abbreviations used in schematic drawings: aSP, anterior pair of spinnerets; bl, book lung; ch, chelicera; CNS, central nervous system; h, heart; L1–L4, first to fourth leg; lr, labrum; m, mouth; pp, pedipalp; pSP, posterior pair of spinnerets; tr, trachea; vs, ventral sulcus

LOC107438386) [87], and *unc5* [88] are also expressed in the developing appendages, albeit not in the form of rings. Gene expression analysis of the high-ranked markers of this cluster, however, suggests that these cells could contribute to the development of the joints of the spider appendages.

Cluster-V: The developing central nervous system (CNS) (2) – EMT-like processes in neural precursor determination?

Top markers of this cluster are *neurotactin* (*nrt*) (LOC107441543), *magu* (LOC107453771), *zinc finger protein 395-like* (*zfp395-like*) (LOC122271089), *fasciclin-2* (*fas2*) (LOC107436358), *noggin* (*nog*) (LOC107444265) and *otopetrin-2-like* (*otop2*) (LOC107451015).

In *Drosophila*, *nrt* is involved in morphogenic movements and is expressed in dynamic patterns in both mesodermal and ectodermal tissues including the developing nervous system where it is expressed strongly [39]. The

expression of *Drosophila magu* is not known, but overexpression of *magu* in *Drosophila* causes an elongated life span, especially when it is overexpressed in the nervous system [112]. To our knowledge, there are no data on *zfp395-like*. In *Drosophila*, the cell adhesion molecule *fas2* is an important factor of synaptic growth and maintenance [57, 165]. In vertebrates, *nog1* induces neural tissue development (reviewed in e.g. [98]). These genes have not been studied in spiders prior to this study, except for the recent data on *nog* showing expression in earlier developmental stages [111].

Some other significant markers of Cluster-V, however, have been investigated previously in *Parasteatoda* or other species of spiders. *netrin-2* (*net2*) (LOC107450632) for example is prominently expressed in the developing nervous system, the heart, and the appendages [88, 118], and *snail* (*sna*) (LOC107443696) is prominently expressed in the developing nervous system of the American

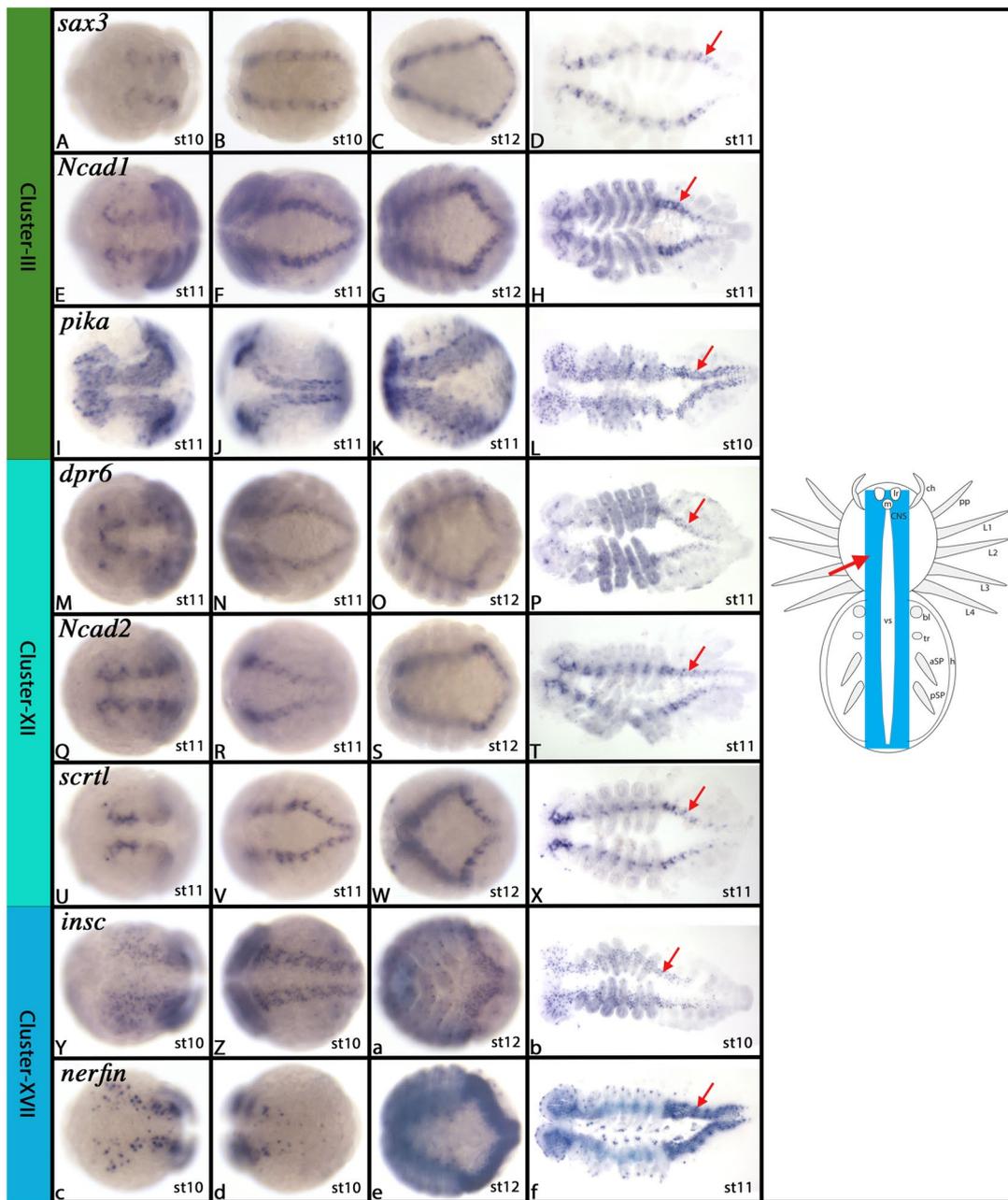


Fig. 5 Gene expression of Cluster-III, XII, and -XVI markers. In all panels, anterior is to the left. All panels represent ventral views. Panels D, H, L, P, T, X, b, and f show flat-mounted embryos. Schematic drawings in the right column summarize the main gene expression characteristic for a given cell cluster (blue tissue). Arrow in schematic drawings point to comparable expression as marked by the same symbol in the original photographs. Abbreviations of gene names is indicated in the upper left corner. Developmental stages are indicated in the bottom right corner. For further information, see Supplementary Figures B5-B8, B28-B31 and B46-B48, and Supplementary File 6. For abbreviations used in schematic drawings see legend of Fig. 4

wandering spider *Cupiennius salei* [177, 193]. In summary, these data suggest that Cluster-V cells are involved in the development of the nervous system. Our gene expression analysis supports this as all investigated genes

of this cluster are expressed *inter alia* in the developing CNS (Fig. 6A-F and Supplementary Figs. 12–15).

Unlike the expression of other CNS-Cluster genes (e.g. clusters VI, X, XII, and XVII), markers of Cluster-V are also expressed in other tissues such as the visceral

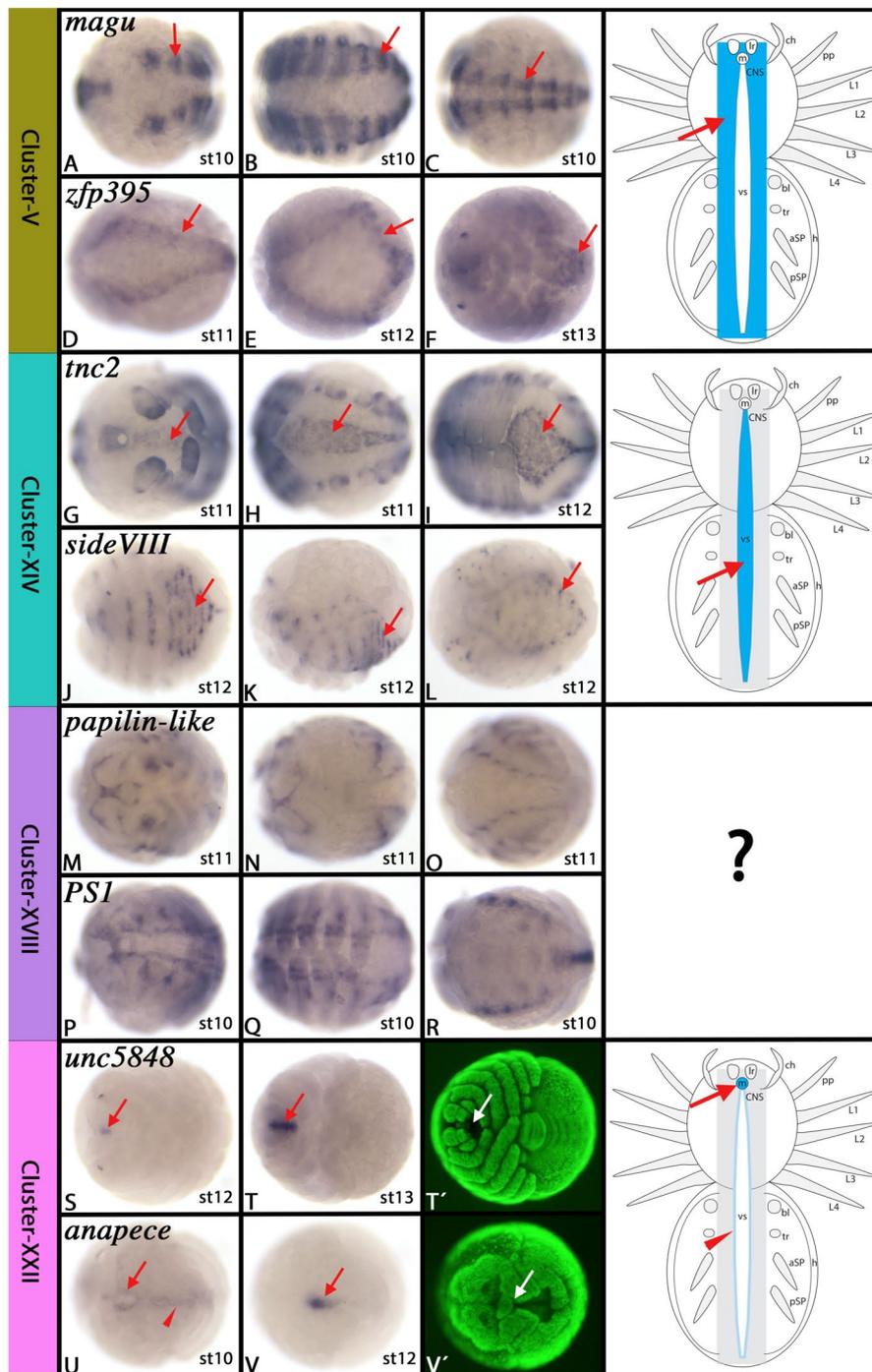


Fig. 6 Gene expression of Cluster-V, -XIV, -XVIII and -XXII markers. In all panels, anterior is to the left. All panels represent ventral views, except panels K (lateral view), N (dorsal view), and V/V' (anterior view). T' and V' represent SYBR-green staining of the embryos shown in panels T and V. Schematic drawings in the right column summarize the main gene expression characteristic for a given cell cluster (blue tissue). Arrows and arrowheads in schematic drawings point to comparable expression as marked by the same symbol in the original photographs. For Cluster-XVIII, it was not possible to determine a specific tissue marked by these genes. Abbreviations of gene names is indicated in the upper left corner. Developmental stages are indicated in the bottom right corner. For further information, see Supplementary Figs. 12–15, 35–37, 49–51, and 60–63, and Supplementary File 6. For abbreviations used in schematic drawings see legend of Fig. 4

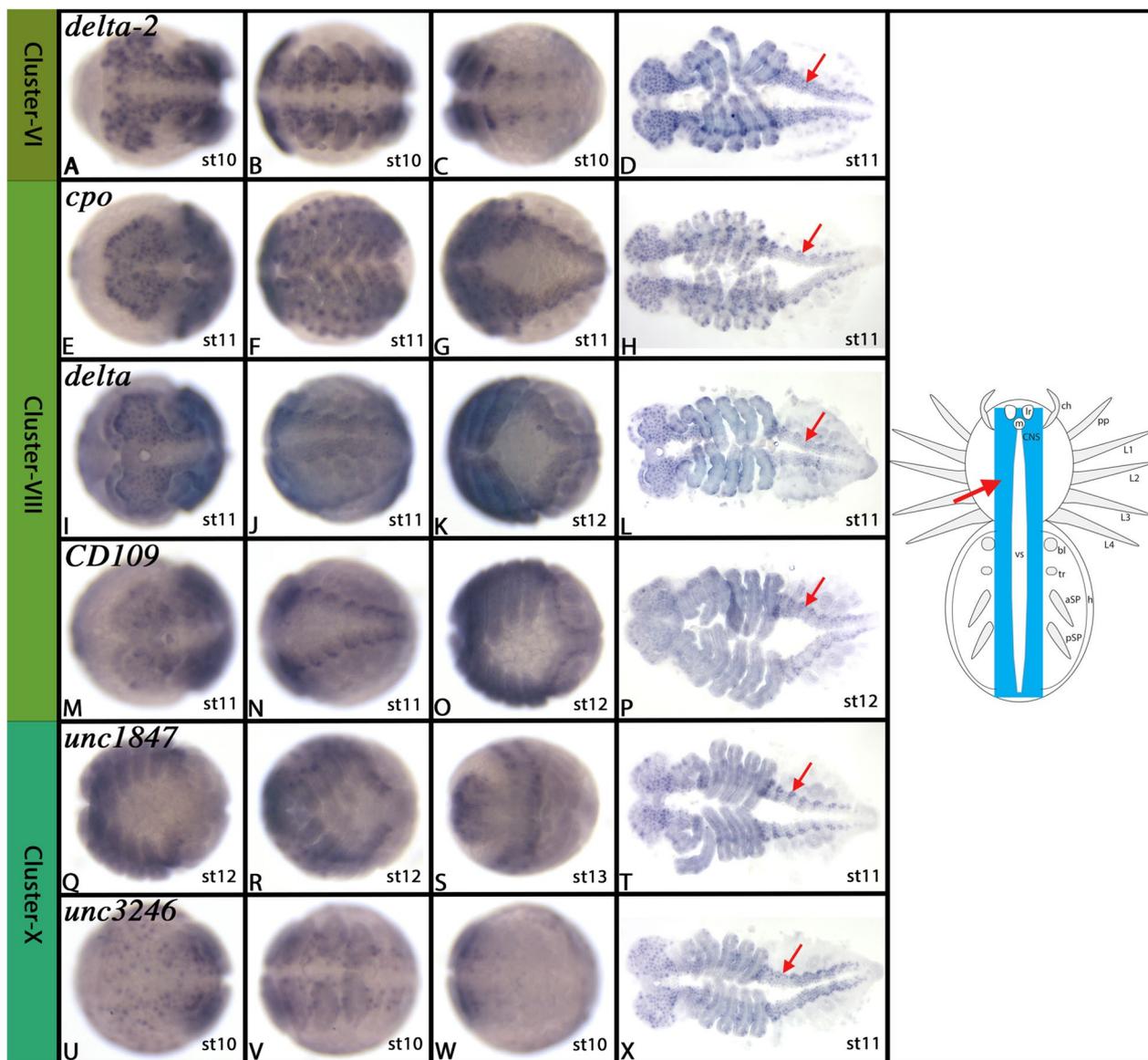


Fig. 7 Gene expression of Cluster-VI, -VIII, and -X markers. In all panels, anterior is to the left. All panels represent ventral views. Panels D, H, L, P, T and X represent flat-mounted embryos. Schematic drawings in the right column summarize the main gene expression characteristic for a given cell cluster (blue tissue). The arrow in the schematic drawing points to comparable expression as marked by the same symbol in the original photographs. Abbreviations of gene names is indicated in the upper left corner. Developmental stages are indicated in the bottom right corner. For further information, see Supplementary Figs. 16, 20–23, 26, and 27, and Supplementary File 6. For abbreviations used in schematic drawings see legend of Fig. 4

mesoderm, the heart and the dorsal field. These aspects of expression are very similar to the pattern of Cluster-VII (discussed below) and may hint to a function in visceral mesoderm development and epithelial-to-mesenchymal transition (EMT). Although Cluster-V cells still show a neural fingerprint, this fingerprint hints to a “near” transition of these cells into a possible mesenchymal nature. Alternatively, Cluster-V cells represent another developmental state of CNS development in which future neural

precursors are selected and delaminate from the “neural epithelium”, in an EMT-related process (e.g. [7]).

Cluster-VI: The developing central nervous system (CNS) (3) – Early differentiating neural cells (1)

Top markers of Cluster-VI are *delta-2* (LOC107438011), and the uncharacterized gene *unc1852* (LOC107451852). *delta* is a well-known neurogenic gene (e.g. [60, 190]). Two other genes that are markers of Cluster-VI, *cyclin*

D (*cycD*) (LOC107457517) and *cyclin dependent kinase 1* (*CDK1*) (LOC107436497) are known factors of cell cycle regulation and have been reported to be highly expressed in the developing nervous system (e.g. [58, 71, 86]). Further down the list of markers, we also find the genes *dachshund* (*dac*) (LOC107453438) and *Nkx6.2* (LOC107450777) (the latter has recently been studied by [111]). We confirm in this study that *Nkx6.2* (Fig. 10Q–T, cf. Cluster-XIX) and *delta-2* (Fig. 7A–D and Supplementary Fig. 16) are expressed exclusively in the developing nervous system of the spider. Previous studies have shown that *dac* is expressed in the developing nervous system of *Parasteatoda* and other spiders [144, 149, 187]. These data and the fact that Cluster-VI is encircled by other cell clusters that likely represent cells of the developing spider nervous system (cf. clusters V, VIII, X, XI, and XIX), strongly suggest that Cluster-VI represents a subtype of cells of the developing nervous system. Notably, none of the detected Achaete-Scute complex (ASH-C) genes is expressed in Cluster-VI, but are expressed in the neighbouring clusters XIII, X and XVII (Supplementary File 5), suggesting that Cluster-VI cells do not represent the earliest steps of CNS development.

Cluster-VII: The developing mesoderm (1) – EMT-like processes in visceral mesoderm development?

In-silico analysis of Cluster-VII markers gives some insight into the possible nature of the cells represented by this cluster. Top markers are *integrin alpha-PS2* (*PS2*) (LOC107444850), *collagen alpha-1(I) chain* (LOC107442626), its paralog (LOC107443413), several other collagen genes (LOC107442286, LOC107441250, LOC107442289), several laminin genes (LOC107439460, LOC107452433, LOC107448892), *papilin* (LOC107447945), and a *fibrinogen-like* gene (LOC107449293).

Many of these genes represent cell adhesion molecules needed for cell migration as present during epithelial-to-mesenchymal transition (EMT) (reviewed in e.g. [14, 73]). These and other genes such as *ECM protein 3-like* (LOC122271235) and *ECM organizing protein Fras1* (LOC107444379) are components of the extracellular matrix (ECM) that is used as a substrate for cell migration (reviewed in e.g. [181]). Interestingly, in *Drosophila*, the visceral mesoderm serves as a substrate for many populations of migrating cells (e.g. [22, 153]), thus connecting cell migration, visceral mesoderm and the ECM. Among the integrins, *PS2*, the top marker of this cluster, is specifically expressed in the visceral mesoderm (e.g. [21, 128]). Another group of genes that are among the top markers of this cluster are the fibroblast growth factor receptors (FGFRs) (LOC107437526, LOC107438841, LOC107445728) which are involved in the development of the visceral mesoderm, somatic muscles and the heart

(e.g. [122, 172]). *papilin* is an important factor of basement membrane development of somatic and visceral muscles, interacts with collagen genes, and is generally involved in cell movement [28, 97], reviewed in [48]. Finally, we also found *FoxF1* (LOC107456534) to be a marker of this cell cluster, a gene that is known to be involved in the development of the visceral mesoderm in *Drosophila* (e.g. [78, 146]). A recent study has shown that in *Parasteatoda*, *FoxF1* is almost exclusively expressed in the visceral mesoderm of the opisthosoma [89].

We performed *in-situ* hybridization experiments with top markers of this cluster and show that they all are expressed *inter alia* in the visceral mesoderm of the trunk, the mesoderm of the appendages, and the developing heart (Fig. 8A–F and Supplementary Figs. 17–19). It is thus likely that Cluster-VII genes represent cells that are in the process of EMT and that are developing towards becoming cells of the visceral mesoderm (cf. Cluster-V).

Cluster-VIII: The developing central nervous system (CNS) (4) – Early differentiating neuronal cells

The top marker of this cluster is *couch potato* (*cpo*) (LOC107453124), a gene that in *Drosophila* is involved in the development of neuronal precursors, and the sensory nervous system (e.g. [16, 54]). Other top markers are *Chronophage* (*Cph*) (LOC107438166), the two ohnologs of *delta* (LOC107456525, LOC107438011), *CD109 antigen* (*CD109*) (LOC107436484) and *epidermal growth factor receptor* (*egfr*) (LOC107446048). *Cph* has recently been identified as a temporal switch of nerve cell subtype specification [51]. One of the two paralogs of *Parasteatoda delta* (LOC107456525) has been described previously, and it has been shown to be expressed in the developing nervous system around stage 9 (later stages were not presented in this study) [138]. In another spider, the American wandering spider *Cupiennius salei*, both ohnologs of *delta* have been investigated, and it has been shown that both are expressed early in the developing nervous system [176]. In *Drosophila*, *delta* represents a neurogenic gene, and is thus involved in cell fate determination within the developing nervous system (reviewed in e.g. [136]). *CD109* is involved in *Drosophila* septate junction formation, but to our knowledge there is no data about its potential function during nervous system development [10]. Finally, *egfr* represents an important factor of nervous system development as EGFR-signalling interacts with TOR-signalling and thereby contributes to neuronal differentiation [93, 94]. *In-silico* analysis thus suggests that Cluster-VIII cells represent nerve cells that are in the (relatively) early process of differentiation. We investigated the embryonic gene expression profiles of *cpo*, *delta*, *delta-2*, *CD109*, and an uncharacterized marker of this cluster (*unc6289*)

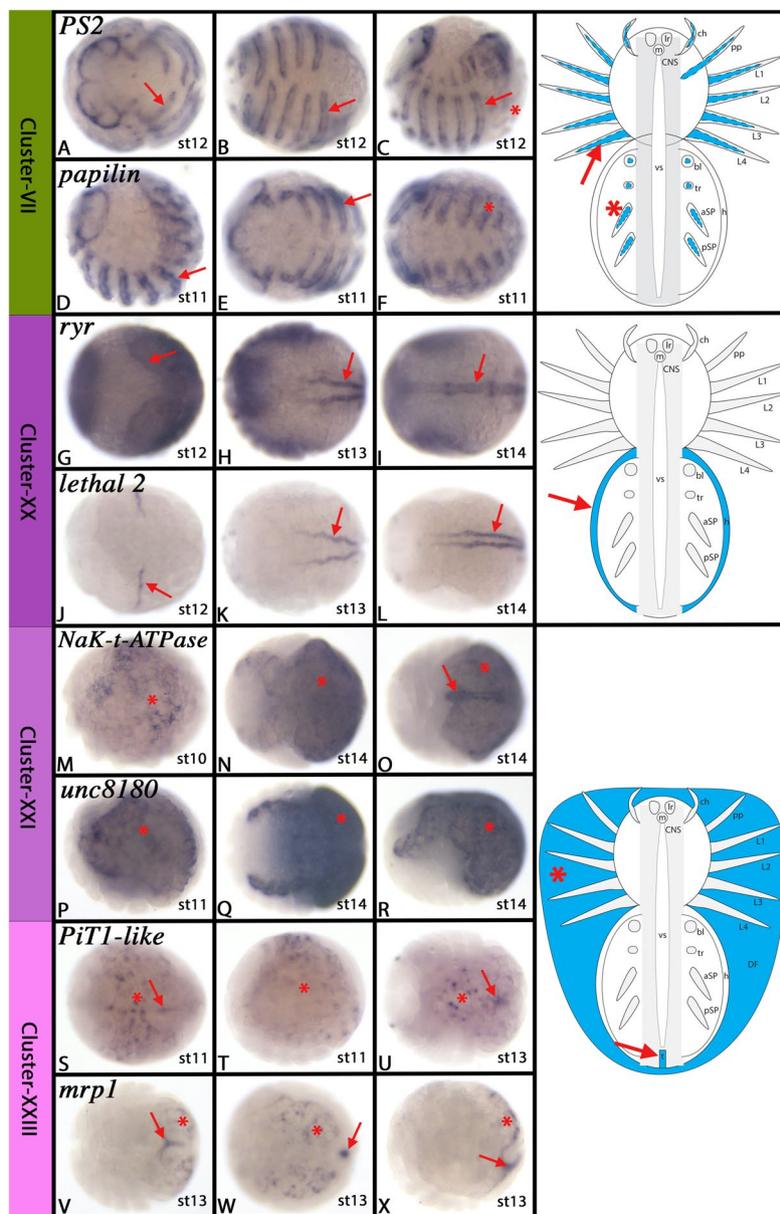


Fig. 8 Gene expression of Cluster-VII, -XX, -XXI and -XXIII markers. In all panels, anterior is to the left. Panels A, B, E, F, U and V represent ventral views. Panels C, D, M, P, R, T and X represent lateral views. Panels G-L, N, O, Q, S and W represent dorsal views. Schematic drawings in the right column summarize the main gene expression characteristic for a given cell cluster (blue tissue). Arrows and asterisks in schematic drawings point to comparable expression as marked by the same symbol in the original photographs. Abbreviations of gene names is indicated in the upper left corner. Developmental stages are indicated in the bottom right corner. For further information, see Supplementary Figs. 17, 18, 54–56, 57–59, and 64–66, and Supplementary File 6. For abbreviations used in schematic drawings see legend of Fig. 4

(LOC107436289). These genes all are expressed almost exclusively in the developing nervous system, thereby supporting the conclusion drawn from *in-silico* analysis (Fig. 7E-P and Supplementary Figs. 20–23).

Cluster-IX: Opisthosomal appendages

The two top markers of this cluster represent the uncharacterized genes *unc0558* (LOC107440558) and *unc2247* (LOC107452247). These genes, and other markers of this cluster do not allow any *in-silico* prediction. These two

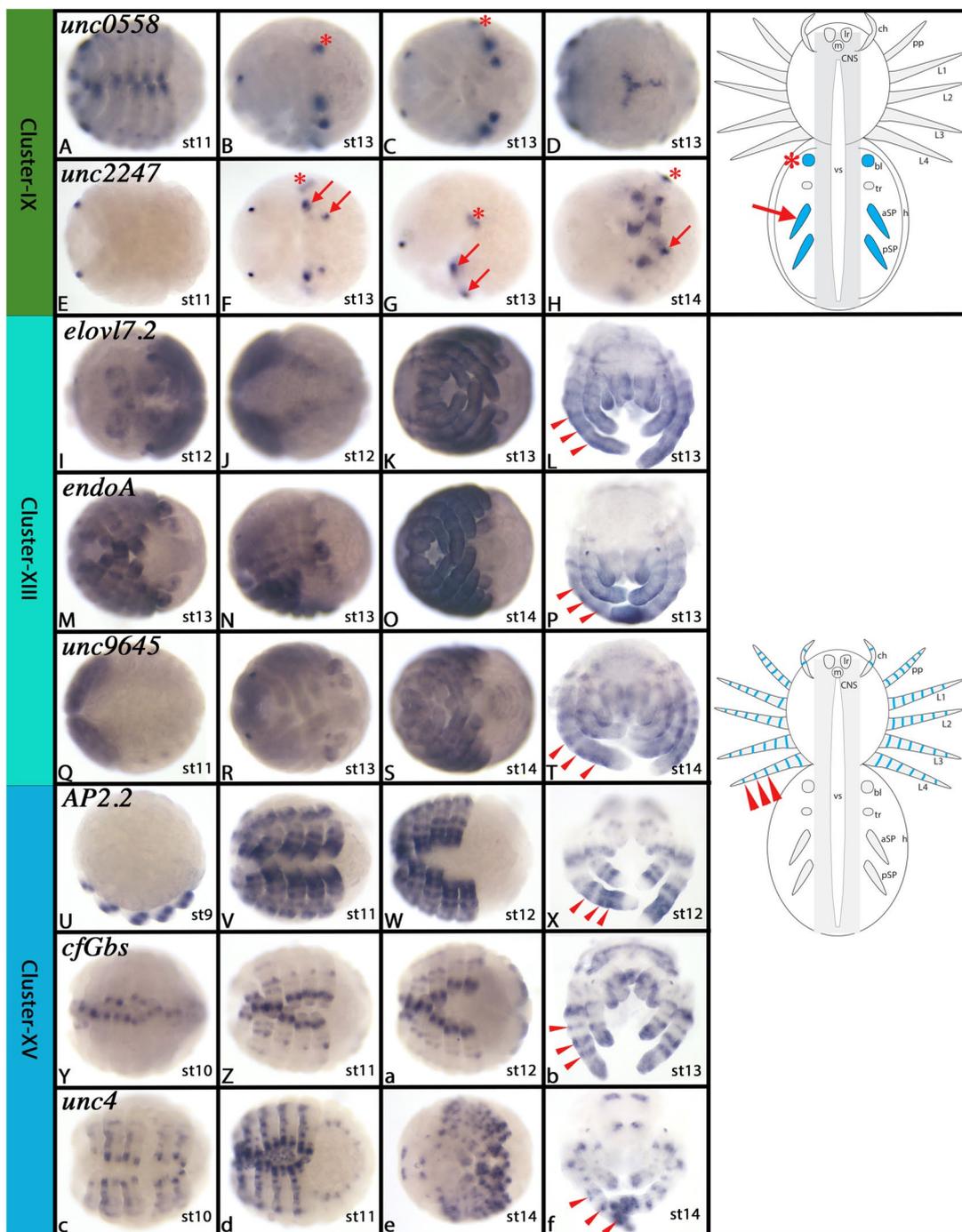


Fig. 9 Gene expression of Cluster-IX, -XIII, and -XV markers. In all panels, anterior is to the left. Panels A, C, E, F, H, I-K, M, O, Q-S, V, W, Y-a, and c-e represent ventral views. Panels B, G, and N represent lateral views. Panel D represents dorsal view. Panels L, P, T, X, b, and f represent anterior views on flat-mounted head regions. Schematic drawings in the right column summarize the main gene expression characteristic for a given cell cluster (blue tissue). Arrows, arrowheads and asterisks in schematic drawings point to comparable expression as marked by the same symbol in the original photographs. Abbreviations of gene names is indicated in the upper left corner. Developmental stages are indicated in the bottom right corner. For further information, see Supplementary Figs. 24, 25, 32–34, and 38–41, and Supplementary File 6. For abbreviations used in schematic drawings see legend of Fig. 4

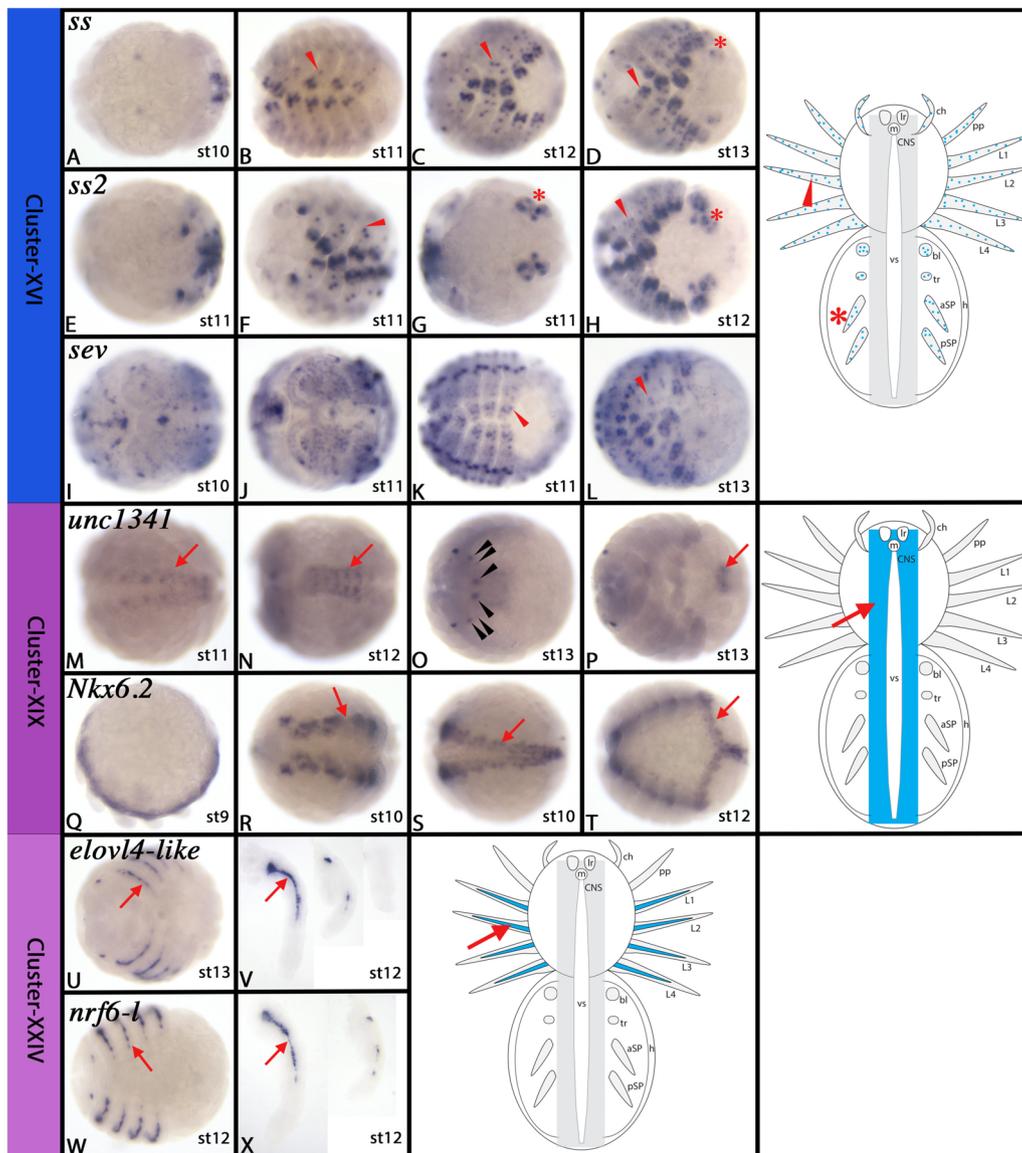


Fig. 10 Gene expression of Cluster-XVI, -XIX, and -XXIV markers. In all panels, anterior is to the left. Panels A-M, P, R-U, and W represent ventral views. Panel N represents dorsal view. Panel O represents anterior view. Panel Q represents lateral view. Panels V and X show isolated walking legs (left), pedipalps (middle) and chelicerae (right), lateral views. Schematic drawings in the right column summarize the main gene expression characteristic for a given cell cluster (blue tissue). Arrows, arrowheads and asterisks in schematic drawings point to comparable expression as marked by the same symbol in the original photographs. Black arrowheads in panel O point to expression in the developing eyes. Abbreviations of gene names is indicated in the upper left corner. Developmental stages are indicated in the bottom right corner. For further information, see Supplementary Figs. 42–45, 52, 53, and 67–69, and Supplementary File 6. For abbreviations used in schematic drawings see legend of Fig. 4

genes are expressed in the developing spinnerets and book lungs (Fig. 9A-H and Supplementary Figs. 24 and 25). Cluster-IX could thus represent cells of the highly-modified opisthosomal appendages. Another previously studied gene of this cluster is one of the two ohnologs of *tracheless* (*trh*) (LOC107455153) which is expressed

more specifically in the developing tracheae of spiders than its paralog (LOC110283103) which itself is expressed in the developing tracheae, but also strongly in legs and dorsal tissue (a marker of clusters IV, XIII, XV, and XVIII) (Zhang (née Turetzek) 2016).

Cluster-X: The developing central nervous system (CNS) (5) – Early differentiating neural cells (2)

One of the two ohnologs of the neurogenic gene *delta*, *delta-2* (LOC107438011) represents the strongest marker of this cluster, followed by *CD109* (OC107436484) (cf. Cluster-VI). Other high-ranked markers are the uncharacterized genes *unc1847* (LOC107451847) (cf. clusters VI and XIII) and *unc3246* (LOC110283246) (cf. clusters VI, VIII and XI), and the *histone-lysine N-methyltransferase MECOM* (LOC107436282) (cf. clusters V, VI, VIII, XI, XVII, and XIX). Cluster-X cells thus clearly share a large number of markers with its neighbouring clusters VI, VIII and XI. One potentially crucial difference between Cluster-X and clusters VI and VIII could be the presence of *cyclin-dependent kinase inhibitor 1 (CKII)* (LOC107443706) which likely interacts with *cycD* (another high-ranked marker of Cluster-X) (LOC107457517) (cf. clusters VI and VIII) [170]. Clusters VI, VIII and X could thus represent cells in a different stage of cell cycle (see distribution of *cycD*, *CDKI*, *CKII* genes) in the developing nervous system (reviewed in [117]). Interestingly, *CKII* is also involved in neural stem cell differentiation and neuron induction by silencing *Sox2* (= *SoxN*), a top marker of cluster-XI [113, 126], reviewed in [117]. Unfortunately, we do not have gene expression data on *cycD*, *CDKI* and *CKII*, but it has recently been proposed that *cycD* and *CDKI* (and other *cycs* and *CDKs*) are expressed in waves of expression in the nervous system of the onychophoran *Euperipatoides kanangrensis* [86]. Cluster-X or clusters VI, X and VIII could thus represent early differentiating neural cells, and Cluster-XI, in which *SoxN* (LOC107457313) is highly expressed, could represent neural precursors (discussed below). In line with this assumption is the fact that the proneural gene *achaete scute homolog 1 (ash1)* (LOC107451231) is also among the higher-ranked markers of Cluster-X. In *Drosophila*, *achaete-scute* complex genes are expressed in segregating neuroblasts (neuronal precursors), but not in neurons [25, 174, 175]. Interestingly, however, studies in insects other than *Drosophila*, such as the beetle *Tribolium castaneum*, and in crustaceans, have shown that the function of *ash* as the first gene to be expressed in developing neuroblasts is not conserved, but instead there seems to be more flexibility in the patterning of the early nervous system (e.g. [188, 194]). What all these studies show, however, is that *ash* is an early factor of nervous system development.

In-situ hybridization analysis of previously unstudied genes that are among the top markers of this cluster reveals that they, as expected, are expressed predominantly in the developing nervous system (Fig. 7Q-X and Supplementary Figs. 26 and 27) (cf. Cluster-VIII for *CD109* expression and Cluster-VI for *delta-2* expression).

Cluster-XI: The developing central nervous system (CNS) (6) – Neural precursors

Two highly-ranked markers of Cluster-XI are *SoxN* (LOC107457313), a well-known factor of stem cell maintenance in the developing nervous system [113, 126], reviewed in [47], and *snail (sna)* (LOC107443696), an important factor of stem cell maintenance in the developing nervous system (e.g. [9, 27]). Both genes are prominently expressed in the developing CNS of spiders [20, 177, 193]. Most interestingly, both genes represent proneural genes that link neural progenitor (neuroblasts in *Drosophila*) selection with epithelial-to-mesenchymal transition (EMT) [7] (cf. *sna* as a marker of Cluster-V). Another highly-ranked marker of this cluster is another potential proneural gene in *Drosophila*, *gooseoid (gsc)* (LOC107439785) [62]. *gsc* is involved in neuroblast delamination in vertebrates (e.g. [96]). The homeobox gene *Dbx* (LOC122269629) is expressed in the developing nervous system of *Parasteatoda* [110], and in vertebrates, its two paralogs are expressed in neural progenitors (e.g. [148]). In *Drosophila*, the single *Dbx* gene is expressed in interneurons that form from a subset of neural progenitor cells (neuroblasts) [106]. Finally, the trunk-determining zinc finger transcription factor *tiptop/teashirt (tio/tsh)* (LOC107441771) is expressed in the CNS of the trunk of *Parasteatoda* [125, 129]. In *Drosophila*, *tio* and *tsh* are expressed in the trunk CNS as well, and it has been shown that both genes can mediate the survival of neurons [107, 135]. Given the genetic fingerprint of Cluster-XI it is possible that these cells represent neural precursor cells.

Cluster-XII: The developing central nervous system (CNS) (7) – Differentiating and differentiated neurons (2)

Cluster-XII is closely related to Cluster-III. Top markers of Cluster-XII are the uncharacterized gene *unc3142* (LOC107443142), *defective proboscis extension response 6 (dpr6)* (LOC107436606), *neural-cadherin 2 (Ncad2)* (LOC107451854), *collier (col)* (aka *knot*) (LOC107437107) and a *scratch-like (scrtl)* gene (LOC107456893).

col has previously been studied in spiders, and indeed a wide range of arthropods, showing that it has a conserved role in nervous system development and neuronal differentiation [13, 35, 41, 158]. Among arthropods, *scrtl* has only been investigated in *Drosophila* (CG1130) where it is expressed in most (or all) neural precursors promoting neuronal development [154]. Information on *Drosophila scrtl* (CG12650) is restricted

to gene expression data, showing that it is expressed in the developing CNS (BDGP in situ homepage: <https://insitu.fruitfly.org/cgi-bin/ex/report.pl?ftype=3&fext=LP01683>). *dpr6* and *Ncad2* are also key-factors of *Drosophila* nervous system development that are involved in the self-organization of neurons and the specification of neuronal subtypes (e.g. [3, 19, 34, 186]). Together, this strongly suggests that Cluster-XII represents cells of the developing nervous system, and indeed, *in-situ* hybridization experiments confirm this suggestion: In *Parasteatoda*, *unc3142*, *scrtl*, *dpr6* and *Ncad2* all are expressed exclusively in the CNS (Fig. 5M–X and Supplementary Figs. 28–31). Beyond that, Cluster-XII is like its neighbouring cluster (Cluster-III) characterized by a significant number of Dscams (LOC107450135, LOC107445852, LOC107444681, LOC107456604). In summary, it is therefore likely that Cluster-XII cells represent differentiating and differentiated neurons (cf. Cluster-III).

Cluster-XIII—Ectoderm patterning of developing appendages (4)

Among the markers of Cluster-XIII are very few previously investigated genes. Top markers are *elongation of very long chain fatty acids protein 7.2 (elovl7.2)* (LOC107441276) and *endochitinase A (endoA)* (LOC107442511). We investigated the *in-situ* hybridization patterns of these two markers and another top marker, the uncharacterized gene *unc9645 (LOC107439645)*. All three marker genes have in common that they are expressed in the form of rings in the developing appendages, indicating a function in appendage ectoderm patterning and thus likely also joint formation (Fig. 9I–T and Supplementary Figs. 32–34).

Cluster-XIV: The ventral sulcus (VS)

The top markers of Cluster-XIV represent two Tenectin-like genes, *Tenectin1 (Tnc1)* (LOC107436695) and *Tenectin1 (Tnc2)* (LOC122273808). Additional top markers of this cluster are a *sidestep (side)/hemicentin-2* like gene hereafter called *sidestep VIII (sideVIII)* (LOC107456176), *netrin-1* (LOC107455212) and *slit* (LOC107443293 & LOC122270376). *netrin-1* has previously been investigated in spiders. It is expressed *inter alia* in the ventral sulcus (VS) (*aka* the ventral midline epithelium [119]) where it is likely involved in axon guidance [88, 118]. *slit* genes encode ligands of the Roundabout receptors and are thus important key players in axon guidance as well (reviewed in [199]). Likewise, *hemicentin-2/sidestep* functions as an important guidance cue for growing axons (e.g. [173]). Finally, also *Tnc* is involved in nervous system development in *Drosophila* where it is expressed in both longitudinal and commissural axon tracts that span

the ventral midline [52, 192]. We therefore assume that Cluster-XIV cells represent cells of the VS, a suggestion that is backed-up by *in-situ* hybridization that confirmed that top markers of this cluster (*Tnc1*, *Tnc2*, and *sideVIII*) indeed are expressed in either the complete VS, or, like *netrin-1*, in transverse stripes spanning the VS (Fig. 6G–L and Supplementary Figs. 35–37). Another high-ranked marker of this cluster, *vitamin K-dependent protein C (vitK-C)* (LOC107451660) has recently been studied by Leite et al. [111] showing that it is exclusively expressed in the ventral midline region prior to ventral splitting, data on developmental stages that possess the split ventral midline and thus the VS are unfortunately not shown in their paper.

Cluster-XV: Ectoderm patterning of developing appendages (5)

Top markers of Cluster-XV are the uncharacterized gene *unc4096 (LOC107444096)*, *clotting factor G subunit (cfGbs)* (LOC107446841), the transcription factor *AP2.2* (LOC107443623) and the homeobox containing gene *unc-4 like* (LOC107444630) (*Drosophila* DPHD-1). Information about the two top markers *unc4096* and *cfGbs* is not available. *AP2*, however, has been investigated in some detail. In *Drosophila*, this gene is expressed *inter alia* in the leg and antennal discs [12, 99]. A previous study in another spider, *Cupiennius salei*, has shown that *AP2* is expressed in the appendages and functions during leg development [150]. The *unc-4 like* gene has been studied in *Drosophila* where it is expressed in the developing nervous system, in a segment-polarity like pattern in the epidermis, and in the eye-antennal disc [179].

A closer look at the Cluster-XV markers also reveals a large number of genes that are known factors of arthropod (and indeed spider) appendage development, such as *clawless/C15* (LOC107451627) Zhang [201] 2016, *dally* (LOC107446074) [72], *distal-less* (LOC107450100) [1, 145, 164], *dachsous* (LOC107449611) [124], the Hox gene *Deformed* (LOC107444120) [166], *unc5.1* (LOC107445619) [88], *optomotor-blind* (LOC107450980) [81], *SP6/9* (LOC107448645) [102] and *Fz4.2* (LOC107441380) [82]. *In-situ* hybridization of the top markers in this cluster revealed that these genes all are expressed strongly in the ectoderm of the development appendages, many of them in the form of rings, indicating a role in proximal–distal axis patterning and joint formation (Fig. 9U–f and Supplementary Figs. 38–41).

Cluster-XVI: The peripheral nervous system (PNS)

Two top markers of this cluster represent ohnologs of the aryl hydrocarbon receptor *spineless (ss)* (LOC107454134, LOC107457395). In *Drosophila*, *ss* causes the transformation of the distal antenna into a distal leg [24, 44], but

beyond that *ss* also is involved in sensory neuron development [40, 100, 116]. Both functions appear to be conserved in other insects [182]. Other top markers of this cluster are a *VEGF receptor like* gene (LOC107454058), *eagle* (LOC107450205), *sevenless (sev)* (LOC107441450) and *Pax2* (LOC107444558). In *Drosophila*, *VEGF receptor-like* is involved in mechanical nociception and normal axon branching, is expressed in sensory neurons, and overexpression leads to mechanical hypersensitivity [108, 120]. *eagle* represents a steroid receptor that is *inter alia* involved in nervous system development [69, 121]. *Drosophila sev* is a well-known factor of photoreceptor specification (reviewed in e.g. [183]). Finally, *Pax2* has been shown to be expressed in the developing peripheral nervous system of *Parasteatoda* [79]. In summary, this suggests that cells of Cluster-XVI represent the developing peripheral nervous system. *In-situ* hybridization experiments of the aforementioned genes supports this suggestion as they are all expressed in distinct cells or groups of cells in the periphery of the developing embryo such as the appendages that display a large number of sensory structures (e.g. [11, 55, 79, 178] (Fig. 10A-L and Supplementary Figs. 42–45). This is further supported by two genes that have been investigated as top markers of other clusters, *unc-4 like* (LOC107444630) (Fig. 9c-f, cf. Cluster-XV) and the uncharacterized gene *unc6289* (LOC107436289) (Supplementary Fig. 23, cf. Cluster-VIII). Both are expressed in similar patterns in the appendages suggesting a function during peripheral/sensory nervous system development.

Cluster-XVII: The developing central nervous system (CNS) (8) – Early differentiating neurons

The top markers of Cluster-XVII are *inscuteable (insc)* (LOC107457283 & LOC122271737), *nervous fingers (nerfin)* (LOC107446363), *prospero (pros)* (LOC107448306), and (less high-ranked) *brain tumor (brat)* (LOC107438139). *insc* is a neural precursor gene in *Drosophila* that is involved in defining neuroblast identity [5, 104]. Beyond that, *insc* is involved in controlling asymmetric cell division during nervous system development where it acts via *miranda (mira)* upstream of *pros* [76, 169]. Interestingly, *pros*, another confirmed factor of spider and arthropod nervous system development [23, 101, 185, 193], is one of the highest-ranking markers of Cluster-XVII as well. In *Drosophila*, *pros* functions as a factor of early neural differentiation by repressing neural stem cell markers such as *snail (sna)* (e.g. [200]). *brat* is involved in the regulation and distinction of intermediate progenitor cells from mitotically active neuroblasts, again interacting with the scaffolding protein *mira* [18]. Finally, *nerfin* prevents reversion and dedifferentiation of neurons into neural stem cells [53,

196]. *In-situ* hybridization of *insc*, *nerfin*, *pros* (in *Cupiennius salei*), and *brat* reveals relatively early expression in the developing CNS suggesting that the cells represented by Cluster-XVII are early differentiating neurons (Fig. 5Y-f and Supplementary Figs. 46–48).

Cluster-XVIII: Ectoderm patterning of developing appendages (6) – EMT-like processes in the appendage epithelium

Among the top markers are *papilin-like* (LOC122270011), a paralog of the aforementioned *papilin* (LOC107447945) (cf. Cluster-VII), *integrin alpha-PS1 (PS1)* (LOC107454271) (cf. Cluster-V), and *nord* (LOC107451942).

papilin-like is an extracellular matrix protein encoding gene that is a likely factor of basement membrane development and cell movement (reviewed in [48]). Comparing the expression patterns of *papilin-like* and *papilin* in *Parasteatoda* (cf. Cluster-VII) reveals only little similarity (Figs. 6M-O and 8D-F). This suggests that these two paralogs have undergone neo- or sub-functionalization, and thus represent markers of different but related cell clusters (clusters XVIII and VII respectively). Unlike *PS2* that is expressed in the visceral mesoderm (discussed above, cf. Cluster-VII), *PS1* is expressed in epithelial cells and mediates cell migration [189]. Notably, a number of laminin-subunit encoding genes (LOC107448892, LOC107452433), the putative ligands of the integrins [67, 127], and a collagen gene (LOC107442289), an interaction partner of *papillin* [97], are also among the strongest markers of Cluster-XVIII, and so is the extracellular matrix (ECM) protein encoding gene *nord* (LOC107451942) [198]. The genetic fingerprint of Cluster-XVIII is thus very similar to that of Cluster-VII, suggesting that its cells are actively involved in cell migration and thus likely in EMT-related processes.

In-situ hybridization of Cluster-XVIII marker genes, however, is not conclusive. All investigated top markers appear to have quite complex and in many aspects different expression patterns, and also the markers of Cluster-XVIII that have been studied previously in *Parasteatoda* such as *Wnt16* (LOC107457243) [87], *dally* (LOC107446074) [72], *FoxB* (LOC107443349) [65], and *hex1* (cf. Cluster-I) show a variety of different expression pattern (Fig. 6M-R and Supplementary Figs. 49–51). What these genes have in common, however, is expression in the developing appendages. It is thus likely that Cluster-XVIII represents cells that undergo an EMT-like process in the developing appendages.

Cluster-XIX: The developing central nervous system (CNS) (9) – The sensory nervous system of the head

Many of the top markers of Cluster-XIX have previously been studied in *Parasteatoda*. Among those are *irx-4*

(LOC107456088) [110], *six3* (LOC107436457) [111, 157], *orthodenticle (otd/otx)* (LOC107457564) [163] and *eyes-absent (eya)* (LOC107452693) [163]. Although these genes are predominantly expressed in the developing nervous system of the head, including the developing eyes, they are also expressed (albeit at a lower level) in the developing CNS of the trunk [110, 157, 163]. Expression data of *irx-4* in later developmental stages are not shown in Leite et al. [110], but the putative homologs of *irx-4* in *Drosophila*, *araucan (ara)* and *caupolican (caup)* both are expressed in the sensory nervous system of the head including the developing eyes (e.g. [56, 147]). The second highest-ranked marker of this cluster is a *scavenger receptor class B member 1-like* gene (LOC122271437) that shows weak sequence similarity with *Drosophila croquemort* and *Sensory neuron membrane protein 1 (Snmp-1)*, a gene that is involved in the function of sensory pheromone receptors [91]. We investigated the embryonic expression of the uncharacterized gene *unc1341* (LOC122271341) and the homeobox gene *Nkx6.2* (LOC107450777) (cf. Cluster-VI) and show that both genes are expressed in the CNS including part of the most anterior region of the CNS (Fig. 10M-T and Supplementary Figs. 52 and 53). Notably, within the developing brain, *unc1341* is specifically expressed in the developing eyes (Fig. 10M-P). Cells representing Cluster-XIX thus clearly contribute to the developing CNS, possibly representing developing sensory structures in the head.

Cluster-XX: The developing heart

The top marker of this cluster is a *ryanodine receptor (ryr)* (LOC107450496) encoding gene that is typically found in muscle tissue (e.g. [63, 180]). The second-best marker is a *lethal 2* like gene (LOC107453784) and the third-best marker is *rho-associated protein kinase 2 (rapk2)* (LOC107453369). Other high-ranked markers are two *titin/twitchin/bent* ohnologs (LOC107453137, LOC107453633), *filamin* (LOC107438849), *myosin heavy chain* (LOC107457063), *myocardin-related transcription factor* (LOC107443971) and *kon-tiki (kon)* (LOC107436488), all of which are involved in muscle and heart development (e.g. [105, 123, 133, 161, 202]). A lower-ranked marker, *myocyte enhancer factor 2 (Mef2)* (LOC107445920) is also expressed in the developing heart of *Parasteatoda* [111] and another spider [80]. *In-silico* analysis thus clearly suggests that Cluster-XX cells represent developing muscle tissue, including the developing heart. We show by *in-situ* hybridization analysis that all three previously unstudied top markers also are expressed almost exclusively in the heart, strongly suggesting that Cluster-XX cells indeed represent the developing heart (Fig. 8G-L and Supplementary Figs. 54–56).

Cluster-XXI: Midgut development, yolk metabolism, hematopoiesis, and immune response

Among the top markers of Cluster-XXI are few previously investigated genes. The top markers are an *aquaporin-7-like (aqp7l)* gene (LOC107457171), the uncharacterized gene *unc8180* (LOC107448180), and a *Na⁺/K⁺ transporting ATPase subunit (NaK-t-ATPase)* (LOC107456567). The majority of markers appear to represent enzymes (e.g. *alkaline phosphatase* (LOC107438694) and *snake venom 5'-nucleotidase* (LOC107439410)), transporters (e.g. the aforementioned *NaK-t-ATPase* and *organic cation transporter protein* (LOC107457207)), and channel-forming proteins (e.g. the aforementioned aquaporin gene and *apolipoprotein* (LOC107450568)), suggesting that these cells are involved in yolk metabolism and uptake. Another group of Cluster-XXI marker genes is involved in the vertebrate immune response (e.g. the aforementioned *snake venom 5'-nucleotidase*, *equilibrative nucleoside transporter 1* (LOC107445032) and *venom phosphodiesterase 2* (LOC107442458)) suggesting that this may be another function of Cluster-XXI cells (e.g. [74, 95, 159]). Indeed, in arthropods, the so-called extraembryonic tissues (*sensu lato*) are involved in providing immune response (recently reviewed in e.g. [143, 184]). *In-situ* hybridization shows that the three top markers of this cluster all exclusively are expressed in cells underlying the germ band (and possibly also the dorsal field and ventral sulcus), and thus the cells/tissue that connects the developing embryo (*sensu stricto*) with the yolk (Fig. 8M-R and Supplementary Fig. 57–59). Interestingly, we also find a midgut marker of *Parasteatoda* in this cluster, the GATA transcription factor *serpent (srp)* (LOC107456523), supporting that these cells may also contribute to midgut development that appears to go hand in hand with nutrition uptake in spiders [46]. Beyond that, hematopoiesis, a process that is closely linked with innate immune defence, depends on the interplay of the aforementioned GATA transcription factor *srp* with the 'friend of GATA' factor *U-shaped (Ush)* (e.g. [49, 50, 152]), another unique marker of this cell cluster (LOC107440842).

Cluster-XXII: The stomodaeum and the ventral midline

The top gene markers of this cluster are a number of uncharacterized genes (e.g. *unc 5848* (LOC107445848), *unc3843* (LOC107453843), *unc5174* (LOC107455174) and *atrial natriuretic peptide-converting enzyme (anapece)* (LOC107455346). The only previously investigated spider genes that represent markers of Cluster-XXII are the two forkhead domain transcription factors *FoxC/crocodile (croc)* (LOC107456536) and *FoxA/forkhead (fkh)* (LOC107452746), *six3* (LOC107450741), and *visual system homeobox (vsx)*

(LOC122268388). In *Parasteatoda* and other spiders, these genes are expressed in the stomodaeum and along the split ventral midline [84, 89, 111, 163]. Expression of *unc5848*, *unc3843*, and *anapece*, however, shows that these top markers are expressed exclusively in the stomodaeum, or strongly in the stomodaeum and faintly at the edges of the split ventral midline (Fig. 6S-V and Supplementary Figs. 60–63). It appears thus that Cluster-XXII predominantly represents cells of the stomodaeum and possibly also cells along the ventral midline.

Cluster-XXIII: Midgut development and yolk metabolism

Cluster-XXIII forms a well-separated small group of cells. The best marker of this cluster is the uncharacterized gene *unc7981* (LOC107437981). The second-best marker is the Na^+/H^+ -exchanger beta (*nhe2*) (LOC107452408). This gene is expressed and functions in the epithelium of the mammalian gut (e.g. [33, 42]. To the best of our knowledge, there are no data on the

expression or function of this gene in any arthropod species. Other markers of this cluster are the *Na-dependent phosphate transporter 1A (PiT1-like)* (LOC107453365) and the *multidrug resistance-associated protein 1 (mrp1)* (LOC107440407). The former gene is involved in the uptake of inorganic phosphate, displays kidney functions, and is often expressed in the intestine (e.g. [68, 131]. *Mrp1* represents a universal transporter, including the transport of lipid derivatives (reviewed in [31, 32]. *In-silico* analysis thus suggests that these cells are involved in metabolic processes, possibly including the uptake of nutrition (yolk metabolism), and the development of the midgut, similar to the predicted functions of Cluster-XXI cells. Detection of gene expression supports this assumption as all of these genes are expressed exclusively in cells of (or underneath) the dorsal field, underneath the embryo proper, and in the tail region that appears to be a key-connective tissue between the yolk and the embryo proper (Fig. 8S-X and Supplementary Fig. 64–66) (cf.

Table 2 Cluster overview

Cluster	Mega Cluster	Title	Tissue
I	A	Dorsal tissue, and ectoderm patterning of developing appendages (1)	Ectoderm
II	A	Ectoderm patterning of developing appendages (2)	Ectoderm
III	B	The developing central nervous system (CNS) (1) – Differentiating and differentiated neurons (1)	CNS
IV	A	Ectoderm patterning of developing appendages (3)	Ectoderm
V	B/C	The developing central nervous system (CNS) (2) – EMT-like processes in neural precursor determination?	CNS
VI	B	The developing central nervous system (CNS) (3)—Early differentiating neural cells (1)	CNS
VII	C	The developing mesoderm (1) – EMT-like processes in visceral mesoderm development?	EMT (Mesoderm)
VIII	B	The developing central nervous system (CNS) (4)—Early differentiating neurons?	CNS
IX	--	Opisthosomal appendages	Ectoderm
X	B	The developing central nervous system (CNS) (5): Early differentiating neural cells (2)	CNS
XI	B	The developing central nervous system (CNS) (6)—Neural precursors	CNS
XII	B	The developing central nervous system (CNS) (7) – Differentiating and differentiated neurons (2)	CNS
XIII	A	Ectoderm patterning of developing appendages (4)	Ectoderm
XIV	B/C	The ventral sulcus (VS)	Ventral Sulcus
XV	A	Ectoderm patterning of developing appendages (5)	Ectoderm
XVI	B	The peripheral nervous system (PNS)	PNS
XVII	B	The developing central nervous system (CNS) (8)—Early differentiating neurons	CNS
XVIII	A/C	Ectoderm patterning of developing appendages (6): EMT-like processes in the appendage epithelium	EMT (Ectoderm)
XIX	B	The developing central nervous system (CNS) (9) – The sensory nervous system of the head	CNS
XX	--	The developing heart	Mesoderm
XXI	--	Midgut development, yolk metabolism, hematopoiesis, and immune response	Yolk + Blood cells /Endoderm
XXII	--	The stomodaeum and the ventral midline	Ectoderm
XXIII	--	Midgut development and yolk metabolism	Yolk cells /Endoderm
XXIV	B	A mini-cluster consisting of only 5 cells that suggests a function in leg-innervation	PNS

Mega Cluster A: Ectoderm patterning of developing appendages

Mega Cluster B: The developing central nervous system (CNS)

Mega Cluster C: Epithelial-to-mesenchymal transition (EMT)

Cluster-XXI). The fact that Cluster-XXI and -XXIII cells are separated into two clusters suggest that they represent different cell populations of this tissue. In this context, it is interesting to note that one endodermal midgut gene, *serpent* (*srp*) (LOC107456523), is expressed in Cluster-XXI cells, but another midgut marker, *hepatocyte nuclear factor 4* (*hnf4*) (LOC107439273) is expressed in Cluster-XXIII cells [46].

Cluster-XXIV: A mini-cluster consisting of only 5 cells that suggests a function in leg-innervation

Some of the top markers of this cluster are *RYamide receptor-like* (LOC107443440), *elongation of very long chain fatty acids protein 4-like* (*elovl4-l*) (LOC122269995), *nose resistant to fluoxetine protein 6-like* (*nrf6-l*) (LOC107448146) and *prostaglandin reductase-3* (*ptgr3*) (LOC107457101).

In-silico analysis of these genes is difficult because there is little or no information about the function of these genes in *Drosophila* and other arthropods. Only the *RYamide receptor-like* gene hints to a function in nervous system development. RYamide peptides have been identified in at least arthropods and tardigrades [26, 103], and represent the orthologs of lophotrochozoan *Luqin* genes [90]. RYamides/Luqins are involved in a wide range of physiological function, including feeding behaviour and locomotion (e.g. [115]). We failed to PCR-amplify *RYamide receptor-like*, but the other aforementioned markers all are expressed exclusively inside the developing legs, and with lower intensity the developing pedipalps (Fig. 10U-X and Supplementary Figs. 67–69), suggesting a function in locomotion, possibly innervation of these appendages.

Mega-clusters and sequencing depth

Overall, the achieved SCS data appear to be of acceptable quality, i.e. most of the predicted cell clusters indeed represent specific tissues of the developing spider embryos such as the developing heart, the ventral sulcus or the peripheral nervous system (PNS). These are also the clusters that are most clearly separated from other clusters, and thus are most well-defined. Other clusters appear to be sub-clusters of larger clusters, often part of what we would like to call mega-clusters (Mega-Cluster-A (Ectoderm patterning of developing appendages; clusters I, II, IV, XIII, XV, and XVIII), Mega-Cluster-B (The developing central nervous system (CNS); clusters III, V, VI, VIII, X, XI, XII, XIV, XVI, XVII, XIX, and XXIV), and Mega-Cluster-C (Epithelial-to-mesenchymal transition (EMT)); clusters V, VII, XIV, and XVIII); note that some clusters such as Cluster-XVIII appear to be part of two mega-clusters

(summarized in Table 2). Naturally, the underlying data that subdivide those closely-related and physically connected (in the UMAP) clusters is less strong (i.e. there are less many markers that are specifically expressed in cells of each given cluster, and that are expressed in more vs less cells of a given cluster). The mega-clusters could represent the actual situation in developing embryos in terms that many related cell types or developmental stages of the same cell type indeed express very similar sets of marker genes. Since our SCS data stem from embryos and thus developing tissues and differentiating cell types, this assumption is not unlikely.

Alternatively, however, mega-clusters may represent “artefacts” as a result of the shortcomings of our analysis. These cells may indeed be much more different from each another, and thus the corresponding clusters could be much more defined than shown in our analysis. A possible reason for such a scenario could be the relatively low sequencing saturation in our analysis (ca. 11%). Low sequencing depth could have led to the lack of detection of more specific, but less numerous transcripts.

A bias towards ectodermal tissues?

We recognized that the majority of all detected tissues, and thus dissociated and captured cells represent ectodermal derivatives, while less mesodermal, and few endodermal cells were detected in our analysis. This could either represent the natural correlation of ectodermal vs mesodermal vs endodermal cells in the developing spider embryo at the investigated stages, or it could be a bias introduced by our methodology, i.e. dissociation, cell capture and downstream analysis. Additionally, we likely removed a fraction of the cells that cover the yolk outside the embryo proper (i.e. cells of the ventral sulcus and the dorsal field) prior to dissociation and cell capture.

We assume, however, that ectodermal cells represent the majority of cells in embryos of the stages 10–12, followed by mesodermal cells and finally endodermal cells. We base this assumption on the fact that most tissues indeed are ectodermal, followed by mesodermal tissues/cells. This has also been shown previously by the expression profiles of typical ectodermal, mesodermal and endodermal markers (e.g. [46, 87, 88, 110, 111]). The study by Feitosa et al. [46], for example, has shown that only relatively few cells of the so called dorsal extraembryonic tissue (the dorsal field in this study) likely contribute to the later (endodermal) midgut [46]. Therefore, we believe that the distribution of captured ectodermal, mesodermal and endodermal cells likely represents the true distribution of these cell types in developing spider embryos of stage 10–12. The recent SCS-study by Leite et al. [111] on earlier stages of *Parasteatoda* further support this assumption.

Whole genome duplication in the spider: How reliable is “in-silico” analysis?

One part of the analysis of cell markers is what we call “in-silico” analysis, the comparison of known data from other arthropods, usually the vinegar fly *Drosophila melanogaster*. One potential problem of this kind of comparative analysis, however, is the fact that many genes are duplicated in *Parasteatoda* (e.g. [87, 110, 166], which likely represents the result of a whole genome duplication (WGD) in this species that dates back to a WGD in the last common ancestor of Arachnopolmonata (e.g. [137, 166]. In-silico analysis therefore often compares one (of two) paralogs/ohnologs of a given marker gene of a particular cell cluster in *Parasteatoda* with the one ortholog in *Drosophila* (or other previously investigated organisms). Only in a few cases, both *Parasteatoda* ohnologs of a given gene (if present as two ohnologs) represent markers of the same cell cluster and are ranked similarly as markers of this cluster. Examples are the two *spineless* ohnologs in Cluster-XVI, or the two *delta* ohnologs in Cluster-VIII.

For all *Parasteatoda* genes, however, that are present in the form of two ohnologs and that are not both markers of a given cell cluster, this may cause a problem. We know that many ohnologs have undergone sub-functionalization and neo-functionalization. It is thus possible that a marker, if this particular *Parasteatoda* ohnolog has been the subject of neo-functionalization, may have a completely different function than its ortholog in other organisms. Likewise, spatial sub-functionalization (expression in different tissues) and temporal sub-functionalization (expression in the same tissue/cell type at different time points during development) may lead to misinterpretation of the function of a given marker gene. This problem can only be addressed beyond doubt by comprehensive WISH analysis of every considered marker (as partially done in this study).

Cell-type vs developmental stage of a cell

Interpretation of the obtained data must consider that an identified cell cluster may either represent a certain cell type (as defined by a cell-type specific genetic fingerprint), or a developmental stage of a certain cell type that therefore may also be represented by other, typically accompanying cell cluster(s). Possible examples are represented by the mega-clusters. Some clusters of Mega-Cluster B (e.g. clusters VI, VIII, X and XI) all harbour cells of the early developing CNS, representing either different (differentiating) cell types, different developmental stages of the same general cell type (that at all times during their development are represented by a more or less stable and specific genetic fingerprint), or cells that represent the same future cell type and/or developmental stage (of that cell type), but show variation (heterogeneity) in their expression profile (e.g. [29]. Indeed, suggestions have been made how to overcome this apparent problem [8], but we currently do

not have the tools to apply them to our current data set. Thus, although we can make predictions on this matter, we do not believe that it is currently possible to distinguish the one from the other beyond considerable doubt.

Our data compared to previously published SCS-projects in *Parasteatoda tepidariorum*

Two previous studies on SCS in *Parasteatoda* embryos covered the developmental stages 5 [2] and stages 7–9 [111]. At stage 5, a germ disc and a cumulus have formed. The cumulus represents later endomesodermal cells and the disc represents the ectodermal tissues of the developing germ band. Additional endodermal and mesodermal tissue come from the rim of the disc (e.g. [138, 197]. The centre of the disc represents the later posterior of the embryo and its segment addition zone (SAZ), and the periphery of the disc represents the anterior of the later germ band (e.g. [2]. Cells from outside the disc represent the yolk and part of the later-forming dorsal field (DF). At this point, the embryo consists of approximately 2000 cells representing mainly undifferentiated cells of the three germ layers [2]. Already at this stage of development, the genetic fingerprint of endodermal cells appears to be quite different from the ectoderm and mesoderm, the latter two which form clusters that are in close proximity to each another (i.e. having related genetic fingerprints) [2]. Interestingly, the study by Akiyama-Oda and colleagues presents two endodermal cell clusters, one single mesodermal cell cluster, and a much larger number of ectodermal cell clusters that represent different regions along the centre-to-periphery-axis of the disc [2].

The study by Leite et al. [111] targets the stages 7–9 each represented by separated dissociation and SCS experiments that were then combined into a single SCS-atlas. Given the more advanced developmental stages that are addressed in this study compared to the study by Akiyama-Oda et al. [2], there are considerably more and more distinct cell clusters [111]. Like in our study, Leite et al. [111] also identified 23 distinct cell clusters, but the nature of these clusters is in many cases quite different from the 23 cell clusters identified in our study. Therefore, it is not possible to directly compare the clusters recovered in Leite et al. [111] with the clusters that we identified in our study. Indeed, there are obvious explanations for these differences. One main difference between the two studies is the higher number of CNS-representing clusters in our analysis which may easiest be explained by the more differentiated (and differentiating) state of the central nervous system in stage 10–12 vs stage 7–9 embryos, where in the latter, development and differentiation of the CNS may only just have begun [111]. Some clusters identified by Leite et al. [111] like the segment addition zone (SAZ) and the segment maturation zone (SMZ) have not been identified in our study, most easily explained by the lack or miniaturization of such tissues in later stage embryos (e.g. [130]. On the other hand, Leite

et al. [111] did not detect cells specific for the ventral sulcus (VS), simply because this structure has not yet formed in the developmental stages studied by them. Another eye-striking difference in the outcome of the previous studies compared with our study is the apparent lack of EMT-related cell types in their studies. This, so we believe, can also be explained by the stages of the sequenced embryos because main morphogenetic events may take place at later developmental stages and after the development of the complete AP axis, rather than early during the processes of axis formation, germ band formation, and segment addition. Finally, Leite et al. [111] also identify a cluster representing putative stem cells, but we do not find a corresponding cell population in our study, most likely because the number of stem cells is low in the later and thus more differentiated developmental stages that we investigated. The paper by Leite et al. [111] uses Hox genes as regional markers and defines some of the recovered clusters according to the expression/distribution of Hox genes. In this way, they for example discriminate between the first two pairs of legs (L1 and L2) and the remaining legs, L3 and L4. Likewise they identify a cluster representing the pedipalp-bearing segment (or merely the developing pedipalp). We did not do this in our analysis because we believe that regional markers are potentially misleading. It is unlikely, for example, that the cells that build L1 and L2 are fundamentally different than those involved in the development of L3 and L4, especially because these segments and their appendages all form from the early germ disc, and do not represent segments/appendages that are partially from the germ disc, and partially added sequentially from the SAZ. Using regional markers such as *labial (lab)* to discriminate the pedipalp-bearing segments from the leg-bearing segments, however, makes sense if the genes involved in pedipalp development indeed differ somewhat from those in leg development.

Mega-Cluster-A: Ectoderm patterning of developing appendages

According to our data and subsequent analysis, six clusters clearly are associated with the ectoderm of developing appendages (i.e. clusters I, II, IV, IX, XIII, and XV). Although both *in-silico* analysis of known genes representing markers of these clusters and accompanying *in-situ* hybridization analysis show that these cells come from the ectoderm of developing appendages, characterization of each cluster turned out to be rather complicated. Firstly, the detected patterns in the appendages are complex and dynamic making it difficult to find common patterns of cells of a given cluster. Secondly, even top markers of certain clusters such as Cluster-XIII showed to be expressed in several tissues further complicating analysis. It is obviously necessary to further investigate these cell clusters, either by applying an improved SCS data set, or by more detailed *in-situ* hybridization studies either comparing marker gene expression in the

appendages of exactly staged embryos, or applying double (or multi) staining. We assume that such data could reveal the interplay(s) of genes involved in appendage patterning.

An interesting finding of our study is the presence of multiple genes that are expressed like genes that are involved in joint formation, i.e. in neatly spaced rings along the proximal–distal (PD) axis of the appendages. The number of cell clusters, and the accompanying large number of markers, that are involved in PD-appendage patterning and (likely) also joint formation is remarkably high, possibly reflecting the complexity of the crucial process of leg-patterning and joint development. Our data thus provide further insight and the possibility to study multiple new genes that likely are involved in these processes.

Another interesting outcome of our study is the finding that Cluster-IX cells are apparently specific for the heavily modified opisthosomal appendages of spiders, the breathing organs (tracheae and book lungs) and the possibly most characteristic feature of spiders, the spinnerets. This may allow a comparison between the genetic networks involved in the developing of the more basal locomotory prosomal legs and the derived appendages of the opisthosoma.

Mega-Cluster-B: The developing CNS

Several of the identified cell clusters represent cells of the developing CNS. Interestingly, the genetic fingerprint of these clusters also enables us to predict their specific identity beyond their mere involvement in nervous system development. We identify proneural genes predominantly in Cluster-XI, suggesting that these cells represent neural progenitors. The adjacent clusters VI and X both likely represent early differentiating neural cells based on the presence of the neurogenic *delta* genes, the proneural gene *ash* that in *Drosophila* is involved in neuroblast segregation (Cluster-X), and a cassette of nervous system cell cycle controlling factors that also block the action of the proneural gene *SoxN* found in Cluster-XI (neural progenitors). Adjacent to clusters VI and X lies Cluster-VIII which possibly represents early differentiating neuronal cells including neurons. Although the cluster is marked by a number of the aforementioned markers of clusters VI and X it also expresses markers such as *cph* that is involved in the early differentiation of neural cell types, and *cpo* that *inter alia* is involved in neuronal precursor development. Cells of Cluster-VIII thus likely represent a later step in nervous system development and differentiation than cells of clusters VI, X and XI. Closely connected to Cluster-VIII lies Cluster-XVII which, based on its genetic fingerprint, appears to represent early differentiating neurons: *insc* is involved in neuronal progenitor determination, *pros* inhibits neural stem cell development and initiates neuronal differentiation, *brat* defines intermediate neuronal progenitors from primary neuronal progenitors, and *nerfn* blocks reversion and dedifferentiation of neurons

into neural precursors. In the periphery of the CNS mega-cluster we find two closely-connected clusters, Cluster-III and Cluster-XII, that very likely represent developing and/or mature neurons because they both express genes that are key-factors of neuron development, neuron differentiation, and neuronal pathfinding. Compared to Cluster-XII, Cluster-III expresses more neuron-specific markers such as the *Dscams*. It may thus be that Cluster-XII indeed represents mostly late developing neurons, and Cluster-III mostly mature neurons.

Remarkably, we can follow CNS development from neural progenitor determination (Cluster-XI), over the early differentiation of these cells (clusters VI, X and VIII) towards the determination of neuronal precursors and early developing neurons (Cluster-XVII) up to the late stages of neuron development and finally mature neurons (clusters III and XII).

A second branch of neuronal cells is represented by clusters XVI, XIX, XXIV, and possibly even XXII. Closest to clusters VI and X, the early differentiating neural cells, lies Cluster-XIX which we believe represents mainly developing sensory nerve cells of the head. Close by lies the mini-cluster Cluster-XXIV that only contains 5 cells. These cells, however, appear to be part of the appendage-innervation system. Cluster-XVI is represented by cells that express typical markers of the peripheral nervous system such as *Pax2*, *sev* and *eagle*. Finally, Cluster-XXII that represents cells of the developing stomodaeum (and ventral midline?) may also represent a neuronal cell type. While the developmental trajectory of Cluster-XI to Cluster-III appears to represent the development from neuronal precursors towards mature neurons, the possible trajectory of Cluster-XI to Cluster-XXII appears less straightforward and thus may need further investigation, probably with the help of an improved data set.

A third short trajectory of CNS-related clusters appears to be the connection of Cluster-XI to its neighbouring Cluster-V. Both clusters express *sna*, a gene that in *Drosophila* regulates the delamination of neuroblasts (neuronal precursors) and that is thus involved in EMT-like processes during neurogenesis [7]. Other genes of Cluster-V indeed are involved in EMT and the expression patterns of many markers of Cluster-V shows that they are not only expressed in the CNS but also various other tissues, likely as a result of the delamination and morphogenic movements of these cells. The trajectory from Cluster-XI to Cluster-V and the involvement of EMT-like processes in CNS development is thus closely related to other cell types such as mesodermal cells (Cluster-VII) or the ectoderm of the developing appendages (Cluster-XVIII) that share an EMT-related genetic fingerprint (discussed below).

Mega-Cluster-C: Epithelial-to-mesenchymal transitions (EMTs)

Several of the identified gene clusters are marked by an EMT-related genetic fingerprint. The main difference between these cells is the underlying second genetic fingerprint that is for example neuronal (clusters V and XIV) (discussed above), mesodermal (Cluster-VII), or appendage-ectoderm specific (Cluster-XVIII). We believe that the reason for detecting a large number of cells that show an EMT-related genetic fingerprint is likely correlated with the relatively late developmental stages we investigated in this study. This goes in line with the other *Parasteatoda* SCS studies that investigate earlier developmental stages and that did not specifically identify cells that are undergoing EMT [2, 111]. Indeed, processes like CNS differentiation, the formation of neuronal networks, and the formation of the mesoderm and mesodermal organs such as the heart most certainly require EMT and morphogenic movement.

Future perspectives

The recent SCS-based papers published on the development of *Parasteatoda* clearly highlight the interest in this model system and the interest in SCS data [2, 111], this study). In the future, the existing data including our own will, we hope, be supplemented and extended towards earlier as well as later developmental stages, including nymphs, juveniles and adults of both sexes, to develop a comprehensive overview over cell types and their developmental trajectories in this model spider. Detailed knowledge about the changes of the genetic fingerprints of developing cell types and the identification of definite cell types will significantly improve our knowledge about the development of spiders. This will include the identification of novel genes, and the untangling of new and conserved gene regulatory networks (GRNs). These data can then be used to study newly identified marker genes (or complete GRNs) in classic candidate gene approaches beyond spiders, possibly enabling us to reconstruct the development and evolution of Arthropoda as a whole. The current study aims to contribute to this goal.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-023-09898-x>.

Additional file 1.

Additional file 2.

Additional file 3.

Additional file 4.

Additional file 5.

Additional file 6.

Additional file 7.

Additional file 8.
Additional file 9.
Additional file 10.
Additional file 11.
Additional file 12.
Additional file 13.
Additional file 14.
Additional file 15.
Additional file 16.
Additional file 17.
Additional file 18.
Additional file 19.
Additional file 20.
Additional file 21.
Additional file 22.
Additional file 23.
Additional file 24.
Additional file 25.
Additional file 26.
Additional file 27.
Additional file 28.
Additional file 29.
Additional file 30.
Additional file 31.
Additional file 32.
Additional file 33.
Additional file 34.
Additional file 35.
Additional file 36.
Additional file 37.
Additional file 38.
Additional file 39.
Additional file 40.
Additional file 41.
Additional file 42.
Additional file 43.
Additional file 44.
Additional file 45.
Additional file 46.
Additional file 47.
Additional file 48.
Additional file 49.
Additional file 50.
Additional file 51.
Additional file 52.
Additional file 53.
Additional file 54.
Additional file 55.

Additional file 56.
Additional file 57.
Additional file 58.
Additional file 59.
Additional file 60.
Additional file 61.
Additional file 62.
Additional file 63.
Additional file 64.
Additional file 65.
Additional file 66.
Additional file 67.
Additional file 68.
Additional file 69.
Additional file 70.
Additional file 71.
Additional file 72.
Additional file 73.
Additional file 74.
Additional file 75.

Acknowledgements

We are grateful to the members of the Developmental Biology department (Arendt Lab) and GeneCore at the European Molecular Biology Laboratory (EMBL) in Heidelberg, Germany for providing facilities, equipment and assistance during the single cell experiments. Special thanks to Jacob M. Musser (Yale University), Daniel Leite, Laura Piovani (University College London), Alba Almazán (Institut de Génomique Fonctionnelle de Lyon), Periklis Paganos (Stazione Zoologica Anton Dohrn) and Åsa Björklund (National Bioinformatics Infrastructure Sweden) for their guidance during the process of data analysis.

The data handling for this project was enabled by resources provided by the Swedish National Infrastructure for Computing (SNIC) at UPPMAX, partially funded by the Swedish Research Council under grant agreement No. 2018-05973.

Authors' contributions

BIMJ, RJ and GEB designed the outline of the project. BIMJ performed all steps of the single-cell RNA sequencing experiments and the subsequent in-silico analysis of the data. BIMJ performed most of the in-situ hybridization experiments. RJ validated the in-situ hybridization experiments and performed additional in-situ hybridization experiments. BIMJ and RJ analysed and discussed the results of the experiments and wrote the first draft version of the manuscript. RJ created the figures. All authors worked on the final version of the manuscript. All authors read and approved the final manuscript.

Funding

Open access funding provided by Uppsala University. This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No. 766053.

Availability of data and materials

The data generated and analysed during this study are included in this published article and its supplementary information files. Raw sequencing reads in fastq.gz format are available from the corresponding authors on reasonable request.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 29 June 2023 Accepted: 12 December 2023

Published online: 07 February 2024

References

- Abzhanov A, Kaufman TC. Homologs of *Drosophila* appendage genes in the patterning of arthropod limbs. *Dev Biol.* 2000;227(2):673–89.
- Akiyama-Oda Y, Akaiwa T, Oda H. Reconstruction of the global polarity of an early spider embryo by single-cell and single-nucleus transcriptome analysis. *Front Cell Dev Biol.* 2022;10:933220.
- Aksu M, Seiradake E. DIPPing into the Fly visual system. *Neuron.* 2018;100(6):1270–2.
- Ahn Y, Zou J, Mitchell PJ. Segment-specific regulation of the *Drosophila* AP-2 gene during leg and antennal development. *Dev Biol.* 2011;355(2):336–48.
- An H, Ge W, Xi Y, Yang X. Inscuteable maintains type I neuroblast lineage identity via numb/notch signaling in the *Drosophila* larval brain. *J Genet Genomics.* 2017;44(3):151–62.
- Angelini DR, Kikuchi M, Jockusch EL. Genetic patterning in the adult capitata antenna of the beetle *Tribolium castaneum*. *Dev Biol.* 2009;327(1):240–51.
- Arefin B, Parvin F, Bahrapour S, Stadler CB, Thor S. *Drosophila* neuroblast selection is gated by notch, snail, SoxB, and EMT gene interplay. *Cell Rep.* 2019;29(11):3636–3651.e3.
- Arendt D, Bertucci PY, Achim K, Musser JM. Evolution of neuronal types and families. *Curr Opin Neurobiol.* 2019;56:144–52.
- Ashraf SI, Ip YT. The snail protein family regulates neuroblast expression of inscuteable and string, genes involved in asymmetry and cell division in *Drosophila*. *Development.* 2001;128(23):4757–67.
- Bätz T, Förster D, Luschnig S. The transmembrane protein macroglobulin complement-related is essential for septate junction formation and epithelial barrier function in *Drosophila*. *Development.* 2014;141(4):899–908.
- Barth FG. Spider mechanoreceptors. *Curr Opin Neurobiol.* 2004;14(4):415–22.
- Bauer R, McGuffin ME, Mattox W, Tainsky MA. Cloning and characterization of the *Drosophila* homologue of the AP-2 transcription factor. *Oncogene.* 1998;17(15):1911–22.
- Baumgardt M, Miguel-Aliaga I, Karlsson D, Ekman H, Thor S. Specification of neuronal identities by feedforward combinatorial coding. *PLoS Biol.* 2007;5(2):e37.
- Barczyk M, Carracedo S, Gullberg D. Integrins. *Cell Tissue Res.* 2010;339(1):269–80.
- Beckingham KM, Armstrong JD, Texada MJ, Munjaal R, Baker DA. *Drosophila melanogaster*—the model organism of choice for the complex biology of multi-cellular organisms. *Gravit Space Biol Bull.* 2005;18(2):17–29.
- Bellen HJ, Kooyer S, D'Evelyn D, Pearlman J. The *Drosophila* couch potato protein is expressed in nuclei of peripheral neuronal precursors and shows homology to RNA-binding proteins. *Genes Dev.* 1992;6(11):2125–36.
- Bernt M, Donath A, Jühling F, Externbrink F, Florentz C, Fritsch G, Pütz J, Middendorff M, Stadler PF. MITOS: improved de novo metazoan mitochondrial genome annotation. *Mol Phylogenet Evol.* 2013;69(2):313–9.
- Betschinger J, Mechtler K, Knoblich JA. Asymmetric segregation of the tumor suppressor brat regulates self-renewal in *Drosophila* neural stem cells. *Cell.* 2006;124(6):1241–53.
- Biersmith B, Liu ZC, Bauman K, Geisbrecht ER. The DOCK protein sponge binds to ELMO and functions in *Drosophila* embryonic CNS development. *PLoS ONE.* 2011;6(1):e16120.
- Bonatto Paese CL, Leite DJ, Schönauer A, McGregor AP, Russell S. Duplication and expression of sox genes in spiders. *BMC Evol Biol.* 2018;18(1):205.
- Boube M, Martin-Bermudo MD, Brown NH, Casanova J. Specific tracheal migration is mediated by complementary expression of cell surface proteins. *Genes Dev.* 2001;15(12):1554–62.
- Bradley PL, Myat MM, Comeaux CA, Andrew DJ. Posterior migration of the salivary gland requires an intact visceral mesoderm and integrin function. *Dev Biol.* 2003;257(2):249–62.
- Broadus J, Doe CQ. Evolution of neuroblast identity: seven-up and prospero expression reveal homologous and divergent neuroblast fates in *Drosophila* and *Schistocerca*. *Development.* 1995;121(12):3989–96.
- Burgess EA, Duncan I. Direct control of antennal identity by the spineless-aristapedia gene of *Drosophila*. *Mol Gen Genet.* 1990;221(3):347–57.
- Cabrera CV, Martinez-Arias A, Bate M. The expression of three members of the achaete-scute gene complex correlates with neuroblast segregation in *Drosophila*. *Cell.* 1987;50(3):425–33.
- Caers J, Verlinden H, Zels S, Vandersmissen HP, Vuerinckx K, Schoofs L. More than two decades of research on insect neuropeptide GPCRs: an overview. *Front Endocrinol (Lausanne).* 2012;3:151.
- Cai Y, Chia W, Yang X. A family of snail-related zinc finger proteins regulates two distinct and parallel mechanisms that mediate *Drosophila* neuroblast asymmetric divisions. *EMBO J.* 2001;20(7):1704–14.
- Campbell AG, Fessler LI, Salo T, Fessler JH. Papilin: a *Drosophila* proteoglycan-like sulfated glycoprotein from basement membranes. *J Biol Chem.* 1987;262(36):17605–12.
- Cembrowski MS, Menon V. Continuous variation within cell types of the nervous system. *Trends Neurosci.* 2018;41(6):337–48.
- Choudhary S, Satija R. Comparison and evaluation of statistical error models for scRNA-seq. *Genome Biol.* 2022;23(1):27.
- Cole SP. Targeting multidrug resistance protein 1 (MRP1, ABCC1): past, present, and future. *Annu Rev Pharmacol Toxicol.* 2014;54:95–117.
- Cole SP. Multidrug resistance protein 1 (MRP1, ABCC1), a "multitasking" ATP-binding cassette (ABC) transporter. *J Biol Chem.* 2014;289(45):30880–8.
- Collins JF, Honda T, Knobel S, Bulus NM, Conary J, DuBois R, Ghishan FK. Molecular cloning, sequencing, tissue distribution, and functional expression of a Na⁺/H⁺ exchanger (NHE-2). *Proc Natl Acad Sci U S A.* 1993;90(9):3938–42.
- Cosmanescu F, Katsamba PS, Sergeeva AP, Ahlsen G, Patel SD, Brewer JJ, Tan L, Xu S, Xiao Q, Nagarkar-Jaiswal S, Nern A, Bellen HJ, Zipursky SL, Honig B, Shapiro L. Neuron-subtype-specific expression, interaction affinities, and specificity determinants of DIP/DPR cell recognition proteins. *Neuron.* 2018;100(6):1385–1400.e6.
- Crozatier M, Vincent A. Control of multidendritic neuron differentiation in *Drosophila*: the role of Collier. *Dev Biol.* 2008;315(1):232–42.
- Daiber T, VanderZwan-Butler CJ, Bashaw GJ, Evans TA. Conserved and divergent aspects of Robo receptor signaling and regulation between *Drosophila* Robo1 and *C. elegans* SAX-3. *Genetics.* 2021;217(3):iy018.
- David R, Wedlich D. PCR-based RNA probes: a quick and sensitive method to improve whole mount embryo in situ hybridizations. *Bio-techniques.* 2001;30(4):769–72 (74).
- Davis GK, Patel NH. Short, long, and beyond: molecular and embryological approaches to insect segmentation. *Annu Rev Entomol.* 2002;47:669–99.
- de la Escalera S, Bockamp EO, Moya F, Piovant M, Jiménez F. Characterization and gene cloning of neurotactin, a *Drosophila* transmembrane protein related to cholinesterases. *EMBO J.* 1990;9(11):3593–601.
- Deak II. Demonstration of sensory neurones in the ectopic cuticle of spineless-aristapedia, a homoeotic mutant of *Drosophila*. *Nature.* 1976;260(5548):252–4.
- Demilly A, Simionato E, Ohayon D, Kerner P, Garcès A, Vervoort M. Coe genes are expressed in differentiating neurons in the central nervous system of protostomes. *PLoS ONE.* 2011;6(6):e21213.

42. Dudeja PK, Rao DD, Syed I, Joshi V, Dahdal RY, Gardner C, Risk MC, Schmidt L, Bavishi D, Kim KE, Harig JM, Goldstein JL, Layden TJ, Ramaswamy K. Intestinal distribution of human Na⁺/H⁺ exchanger isoforms NHE-1, NHE-2, and NHE-3 mRNA. *Am J Physiol*. 1996;271(3 Pt 1):G483–93.
43. Dumstrei K, Wang F, Shy D, Tepass U, Hartenstein V. Interaction between EGFR signaling and DE-cadherin during nervous system morphogenesis. *Development*. 2002;129(17):3983–94.
44. Duncan DM, Burgess EA, Duncan I. Control of distal antennal identity and tarsal development in *Drosophila* by spineless-aristapedia, a homolog of the mammalian dioxin receptor. *Genes Dev*. 1998;12(9):1290–303.
45. Emmons RB, Duncan D, Duncan I. Regulation of the *Drosophila* distal antennal determinant spineless. *Dev Biol*. 2007;302(2):412–26.
46. Feitosa NM, Pechmann M, Schwager EE, Tobias-Santos V, McGregor AP, Damen WGM, Nunes de Fonseca R. Molecular control of gut formation in the spider *Parasteatoda tepidariorum*. *Genesis*. 2017;55(5). <https://doi.org/10.1002/dvg.23033>.
47. Feng R, Wen J. Overview of the roles of Sox2 in stem cell and development. *Biol Chem*. 2015;396(8):883–91.
48. Fessler JH, Kramerova I, Kramerov A, Chen Y, Fessler LI. Papilin, a novel component of basement membranes, in relation to ADAMTS metalloproteases and ECM development. *Int J Biochem Cell Biol*. 2004;36(6):1079–84.
49. Fossett N, Tevosian SG, Gajewski K, Zhang Q, Orkin SH, Schulz RA. The friend of GATA proteins U-shaped, FOG-1, and FOG-2 function as negative regulators of blood, heart, and eye development in *Drosophila*. *Proc Natl Acad Sci U S A*. 2001;98(13):7342–7.
50. Fossett N, Hyman K, Gajewski K, Orkin SH, Schulz RA. Combinatorial interactions of serpent, lozenge, and U-shaped regulate crystal cell lineage commitment during *drosophila* hematopoiesis. *Proc Natl Acad Sci U S A*. 2003;100(20):11451–6.
51. Fox PM, Tang JLY, Brand AH. The *Drosophila* homologue of CTIP1 (Bcl11a) and CTIP2 (Bcl11b) regulates neural stem cell temporal patterning. *Development*. 2022;149(17):dev200677.
52. Fraichard S, Bouge AL, Chaveuil I, Bouhin H. Tenectin, a novel extracellular matrix protein expressed during *Drosophila melanogaster* embryonic development. *Gene Expr Patterns*. 2006;6(8):772–6.
53. Frolidi F, Szuperak M, Weng CF, Shi W, Papenfuss AT, Cheng LY. The transcription factor Nerfin-1 prevents reversion of neurons into neural stem cells. *Genes Dev*. 2015;29(2):129–43.
54. Glasscock E, Tanouye MA. *Drosophila* couch potato mutants exhibit complex neurological abnormalities including epilepsy phenotypes. *Genetics*. 2005;169(4):2137–49.
55. Gold K, Cotton JA, Stollewerk A. The role of notch signalling and numb function in mechanosensory organ formation in the spider *Cupiennius salei*. *Dev Biol*. 2009;327(1):121–31.
56. Gomez-Skarmeta JL, Díez del Corral R, De la Calle-Mustienes E, Ferré-Marcó D, Modolell J. Araucan and caupolican, two members of the novel iroquois complex, encode homeoproteins that control proneural and vein-forming genes. *Cell*. 1996;85(1):95–105.
57. Grenningloh G, Rehm EJ, Goodman CS. Genetic analysis of growth cone guidance in *Drosophila*: fasciclin II functions as a neuronal recognition molecule. *Cell*. 1991;67(1):45–57.
58. Grison A, Atanasoski S. Cyclins, cyclin-dependent kinases, and cyclin-dependent kinase inhibitors in the mouse nervous system. *Mol Neurobiol*. 2020;57(7):3206–18.
59. Grün D, van Oudenaarden A. Design and analysis of single-cell sequencing experiments. *Cell*. 2015;163(4):799–810.
60. Haenlin M, Kunisch M, Kramatschek B, Campos-Ortega JA. Genomic regions regulating early embryonic expression of the *drosophila* neurogenic gene Delta. *Mech Dev*. 1994;47(1):99–110.
61. Hafemeister C, Satija R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol*. 2019;20(1):296.
62. Hahn M, Jäckle H. *Drosophila* goosecoid participates in neural development but not in body axis formation. *EMBO J*. 1996;15(12):3077–84.
63. Hasan G, Rosbash M. *Drosophila* homologs of two mammalian intracellular Ca(2+)-release channels: identification and expression patterns of the inositol 1,4,5-triphosphate and the ryanodine receptor genes. *Development*. 1992;116(4):967–75.
64. Heilig JS, Freeman M, Laverty T, Lee KJ, Campos AR, Rubin GM, Steller H. Isolation and characterization of the disconnected gene of *Drosophila melanogaster*. *EMBO J*. 1991;10(4):809–15.
65. Heingård M, Turetzek N, Prpic NM, Janssen R. FoxB, a new and highly conserved key factor in arthropod dorsal-ventral (DV) limb patterning. *EvoDevo*. 2019;10:28.
66. Hemmati-Brivanlou A, Kelly OG, Melton DA. Follistatin, an antagonist of activin, is expressed in the Spemann organizer and displays direct neuralizing activity. *Cell*. 1994;77(2):283–95.
67. Henchcliffe C, García-Alonso L, Tang J, Goodman CS. Genetic analysis of laminin a reveals diverse functions during morphogenesis in *drosophila*. *Development*. 1993;118(2):325–37.
68. Hernando N, Wagner CA. Mechanisms and regulation of intestinal phosphate absorption. *Compr Physiol*. 2018;8(3):1065–90.
69. Higashijima S, Shishido E, Matsuzaki M, Saigo K. eagle, a member of the steroid receptor gene superfamily, is expressed in a subset of neuroblasts and regulates the fate of their putative progeny in the *Drosophila* CNS. *Development*. 1996;122(2):527–36.
70. Hilbrant M, Damen WG, McGregor AP. Evolutionary crossroads in developmental biology: the spider *Parasteatoda tepidariorum*. *Development*. 2012;139(15):2655–62.
71. Hindley C, Philpott A. Co-ordination of cell cycle and differentiation in the developing nervous system. *Biochem J*. 2012;444(3):375–82.
72. Hogvall M, Budd GE, Janssen R. Gene expression analysis of potential morphogen signalling modifying factors in Panarthropoda. *EvoDevo*. 2018;9:20.
73. Huang H, Wright S, Zhang J, Brekken RA. Getting a grip on adhesion: cadherin switching and collagen signaling. *Biochim Biophys Acta Mol Cell Res*. 2019;1866(11):118472.
74. Huang YS, Tseng WY, Clanchy FIL, Topping LM, Ogbechi J, McNamee K, Perocheau D, Chiang NY, Ericsson P, Sundstedt A, Xue ZT, Salford LG, Sjögren HO, Stone TW, Lin HH, Luo SF, Williams RO. Pharmacological modulation of T cell immunity results in long-term remission of autoimmune arthritis. *Proc Natl Acad Sci U S A*. 2021;118(19):e2100939118.
75. Hunter DD, Manglapus MK, Bachay G, Claudepierre T, Dolan MW, Gesuelli KA, Brunken WJ. CNS synapses are stabilized trans-synaptically by laminins and laminin-interacting proteins. *J Comp Neurol*. 2019;527(1):67–86.
76. Ikeshima-Kataoka H, Skeath JB, Nabeshima Y, Doe CQ, Matsuzaki F. Miranda directs prospero to a daughter cell during *Drosophila* asymmetric divisions. *Nature*. 1997;390(6660):625–9.
77. Iwasaki-Yokozawa S, Nanjo R, Akiyama-Oda Y, Oda H. Lineage-specific, fast-evolving GATA-like gene regulates zygotic gene activation to promote endoderm specification and pattern formation in the Theridiidae spider. *BMC Biol*. 2022;20(1):223.
78. Jakobsen JS, Braun M, Astorga J, Gustafson EH, Sandmann T, Karzynski M, et al. Temporal ChIP-on-chip reveals Biniou as a universal regulator of the visceral muscle transcriptional network. *Genes Dev*. 2007;21:2448–60.
79. Janeschik M, Schacht MI, Platten F, Turetzek N. It takes two: discovery of spider Pax2 duplicates indicates prominent role in chelicerate central nervous system, eye, as well as external sense organ precursor formation and diversification after neo-and subfunctionalization. *Front Ecol Evol*. 2022;10. <https://doi.org/10.3389/fevo.2022.810077>.
80. Janssen R, Damen WG. Diverged and conserved aspects of heart formation in a spider. *Evol Dev*. 2008;10(2):155–65.
81. Janssen R, Feitosa NM, Damen WG, Prpic NM. The T-box genes H15 and optomotor-blind in the spiders *Cupiennius salei*, *Tegenaria atrica* and *Achaearanea tepidariorum* and the dorsoventral axis of arthropod appendages. *Evol Dev*. 2008;10(2):143–54.
82. Janssen R, Schönauer A, Weber M, Turetzek N, Hogvall M, Goss GE, Patel NH, McGregor AP, Hilbrant M. The evolution and expression of panarthropod frizzled genes. *Front Ecol Evol*. 2015;3:96.
83. Janssen R. Gene expression reveals evidence for EGFR-dependent proximal-distal limb patterning in a myriapod. *Evol Dev*. 2017;19(3):124–35.
84. Janssen R, Budd GE. Investigation of endoderm marker-genes during gastrulation and gut-development in the velvet worm *Euperipatoides kanangrensis*. *Dev Biol*. 2017;427(1):155–64.
85. Janssen R, Andersson E, Betnér E, Bijl S, Fowler W, Höök L, Leyhr J, Manelqvist A, Panara V, Smith K, Tiemann S. Embryonic expression patterns

- and phylogenetic analysis of panarthropod sox genes: insight into nervous system development, segmentation and gonadogenesis. *BMC Evol Biol.* 2018;18(1):88.
86. Janssen R, Budd GE. Oscillating waves of fox, cyclin and CDK gene expression indicate unique spatiotemporal control of cell cycling during nervous system development in onychophorans. *Arthropod Struct Dev.* 2021;62:101042.
 87. Janssen R, Pechmann M, Turetzek N. A chelicerate Wnt gene expression atlas: novel insights into the complexity of arthropod Wnt-patterning. *EvoDevo.* 2021;12(1):12.
 88. Janssen R, Budd GE. Expression of netrin and its receptors uncoordinated-5 and frazzled in arthropods and onychophorans suggests conserved and diverged functions in neuronal pathfinding and synaptogenesis. *Dev Dyn.* 2022. <https://doi.org/10.1002/dvdy.459>.
 89. Janssen R, Schomburg C, Prpic NM, Budd GE. A comprehensive study of arthropod and onychophoran Fox gene expression patterns. *PLoS ONE.* 2022;17(7):e0270790.
 90. Jékely G. Global view of the evolution and diversity of metazoan neuropeptide signaling. *Proc Natl Acad Sci U S A.* 2013;110(21):8702–7.
 91. Jin X, Ha TS, Smith DP. SNMP is a signaling component required for pheromone sensitivity in drosophila. *Proc Natl Acad Sci U S A.* 2008;105(31):10996–1001.
 92. Jontes JD. The cadherin superfamily in neural circuit assembly. *Cold Spring Harb Perspect Biol.* 2018;10(7):a029306.
 93. Jussen D, von Hilchen J, Urbach R. Genetic regulation and function of epidermal growth factor receptor signalling in patterning of the embryonic drosophila brain. *Open Biol.* 2016;6(12):16202.
 94. McNeill H, Craig GM, Bateman JM. Regulation of neurogenesis and epidermal growth factor receptor signaling by the insulin receptor/target of rapamycin pathway in *Drosophila*. *Genetics.* 2008;179(2):843–53.
 95. Kabashima K, Nakashima C, Nonomura Y, Otsuka A, Cardamone C, Parente R, De Feo G, Triggiani M. Biomarkers for evaluation of mast cell and basophil activation. *Immunol Rev.* 2018;282(1):114–20.
 96. Kantarci H, Gerberding A, Riley BB. Spemann organizer gene Goosecoid promotes delamination of neuroblasts from the otic vesicle. *Proc Natl Acad Sci U S A.* 2016;113(44):E6840–8.
 97. Kawano T, Zheng H, Merz DC, Kohara Y, Tamai KK, Nishiwaki K, Culotti JGC. *elegans mig-6* encodes papilin isoforms that affect distinct aspects of DTC migration, and interacts genetically with mig-17 and collagen IV. *Development.* 2009;136(9):1433–42.
 98. Kelly OG, Melton DA. Induction and patterning of the vertebrate nervous system. *Trends Genet.* 1995;11(7):273–8.
 99. Kerber B, Monge I, Mueller M, Mitchell PJ, Cohen SM. The AP-2 transcription factor is required for joint formation and cell survival in *Drosophila* leg development. *Development.* 2001;128(8):1231–8.
 100. Kim MD, Jan LY, Jan YN. The bHLH-PAS protein spineless is necessary for the diversification of dendrite morphology of *drosophila* dendritic arborization neurons. *Genes Dev.* 2006;20(20):2806–19.
 101. Klann M, Stollewerk A. Evolutionary variation in neural gene expression in the developing sense organs of the crustacean *Daphnia magna*. *Dev Biol.* 2017;424(1):50–61.
 102. Königsman T, Turetzek N, Pechmann M, Prpic NM. Expression and function of the zinc finger transcription factor Sp6-9 in the spider *Parasteatoda tepidariorum*. *Dev Genes Evol.* 2017;227(6):389–400.
 103. Koziol U. Precursors of neuropeptides and peptide hormones in the genomes of tardigrades. *Gen Comp Endocrinol.* 2018;267:116–27.
 104. Kraut R, Campos-Ortega JA. *inscuteable*, a neural precursor gene of *drosophila*, encodes a candidate for a cytoskeleton adaptor protein. *Dev Biol.* 1996;174(1):65–81.
 105. Kuwahara K, Barrientos T, Pipes GC, Li S, Olson EN. Muscle-specific signaling mechanism that links actin dynamics to serum response factor. *Mol Cell Biol.* 2005;25(8):3173–81.
 106. Lacin H, Zhu Y, Wilson BA, Skeath JB. *dbx* mediates neuronal specification and differentiation through cross-repressive, lineage-specific interactions with *eve* and *hb9*. *Development.* 2009;136(19):3257–66. <https://doi.org/10.1242/dev.037242>.
 107. Laugier E, Yang Z, Fasano L, Kerridge S, Vola C. A critical role of teashirt for patterning the ventral epidermis is masked by ectopic expression of *tiptop*, a paralog of teashirt in *Drosophila*. *Dev Biol.* 2005;283(2):446–58.
 108. Learte AR, Forero MG, Hidalgo A. Gliatrophic and gliatropic roles of PVF/PVR signaling during axon guidance. *Glia.* 2008;56(2):164–76.
 109. Lee KJ, Freeman M, Steller H. Expression of the disconnected gene during development of *Drosophila melanogaster*. *EMBO J.* 1991;10(4):817–26.
 110. Leite DJ, Baudouin-Gonzalez L, Iwasaki-Yokozawa S, Lozano-Fernandez J, Turetzek N, Akiyama-Oda Y, Prpic NM, Pisani D, Oda H, Sharma PP, McGregor AP. Homeobox gene duplication and divergence in arachnids. *Mol Biol Evol.* 2018;35(9):2240–53.
 111. Leite DJ, Schönauer A, Blakeley G, Harper A, Garcia-Castro H, Baudouin-Gonzalez L, Wang R, Sarkis N, Nikola AG, Koka VS, Kenny NJ. An atlas of spider development at single-cell resolution provides new insights into arthropod embryogenesis. *bioRxiv.* 2022.
 112. Li Y, Tower J. Adult-specific over-expression of the *Drosophila* genes *magu* and *hebe* increases life span and modulates late-age female fecundity. *Mol Genet Genomics.* 2009;281(2):147–62.
 113. Li H, Collado M, Villasante A, Matheu A, Lynch CJ, Cañamero M, Rizzoti K, Carneiro C, Martínez G, Vidal A, Lovell-Badge R, Serrano M. p27(Kip1) directly represses Sox2 during embryonic stem cell differentiation. *Cell Stem Cell.* 2012;11(6):845–52.
 114. Li H. Single-cell RNA sequencing in *Drosophila*: technologies and applications. *Wiley Interdiscip Rev Dev Biol.* 2021;10(5):e396.
 115. Li C, Zheng Y, Cong X, Liu H, Storey KB, Chen M. Molecular and functional characterization of the luqin-type neuropeptide signaling system in the sea cucumber *Apostichopus japonicus*. *Peptides.* 2022;155:170839.
 116. Lienhard MC, Stocker RF. The development of the sensory neuron pattern in the antennal disc of wild-type and mutant (*lz3, ssa*) *Drosophila melanogaster*. *Development.* 1991;112(4):1063–75.
 117. Lim S, Kaldis P. Cdks, cyclins and CKIs: roles beyond cell cycle regulation. *Development.* 2013;140(15):3079–93.
 118. Linne V, Stollewerk A. Conserved and novel functions for netrin in the formation of the axonal scaffold and glial sheath cells in spiders. *Dev Biol.* 2011;353(1):134–46.
 119. Linne V, Eriksson BJ, Stollewerk A. Single-minded and the evolution of the ventral midline in arthropods. *Dev Biol.* 2012;364(1):66–76.
 120. Lopez-Bellido R, Puig S, Huang PJ, Tsai CR, Turner HN, Galko MJ, Gutstein HB. Growth factor signaling regulates mechanical nociception in flies and vertebrates. *J Neurosci.* 2019;39(30):6012–30.
 121. Lundell MJ, Hirsh J. *eagle* is required for the specification of serotonin neurons and other neuroblast 7–3 progeny in the *Drosophila* CNS. *Development.* 1998;125(3):463–72.
 122. Mandal L, Dumstrei K, Hartenstein V. Role of FGFR signaling in the morphogenesis of the *Drosophila* visceral musculature. *Dev Dyn.* 2004;231(2):342–8.
 123. Manieu C, Olivares GH, Vega-Macaya F, Valdivia M, Olguín P. Jitterbug/Filamin and myosin-II form a complex in tendon cells required to maintain epithelial shape and polarity during musculoskeletal system development. *Mech Dev.* 2018;154:309–14.
 124. Mao Y, Rauskolb C, Cho E, Hu WL, Hayter H, Minihan G, Katz FN, Irvine KD. *Dachs*: an unconventional myosin that functions downstream of fat to regulate growth, affinity and gene expression in *Drosophila*. *Development.* 2006;133(13):2539–51.
 125. March LE, Smaby RM, Setton EVW, Sharma PP. The evolution of selector gene function: Expression dynamics and regulatory interactions of *tiptop/teashirt* across Arthropoda. *Evol Dev.* 2018;20(6):219–32.
 126. Marqués-Torrejón MÁ, Porlan E, Banito A, Gómez-Ibarlucea E, Lopez-Contreras AJ, Fernández-Capetillo O, Vidal A, Gil J, Torres J, Fariñas I. Cyclin-dependent kinase inhibitor p21 controls adult neural stem cell expansion by regulating Sox2 gene expression. *Cell Stem Cell.* 2013;12(1):88–100.
 127. Martin D, Zusman S, Li X, Williams EL, Khare N, DaRocha S, Chiquet-Ehrismann R, Baumgartner S. wing blister, a new *Drosophila* laminin alpha chain required for cell adhesion and migration during embryonic and imaginal development. *J Cell Biol.* 1999;145(1):191–201.
 128. Martin-Bermudo MD, Alvarez-Garcia I, Brown NH. Migration of the *Drosophila* primordial midgut cells requires coordination of diverse PS integrin functions. *Development.* 1999;126(22):5161–9.
 129. Medina-Jiménez BI, Budd GE, Janssen R. Panarthropod *tiptop/teashirt* and *Spalt* orthologs and their potential role as “trunk”-selector genes. *EvoDevo.* 2021;12(1):7.
 130. Mittmann B, Wolff C. Embryonic development and staging of the cobweb spider *Parasteatoda tepidariorum* C. L. Koch, 1841 (syn.: *Achaearanea tepidariorum*; *Araneomorphae*; *Theridiidae*). *Dev Genes Evol.* 2012;222(4):189–216.

131. Miyaji T, Kawasaki T, Togawa N, Motome H, Moriyama Y. Type 1 sodium-dependent phosphate transporter acts as a membrane potential-driven urate exporter. *Curr Mol Pharmacol*. 2013;6(2):88–94.
132. Miyawaki K, Inoue Y, Mito T, Fujimoto T, Matsushima K, Shinmyo Y, Ohuchi H, Noji S. Expression patterns of aristaless in developing appendages of *Gryllus bimaculatus* (cricket). *Mech Dev*. 2002;113(2):181–4.
133. Mogami K, O'Donnell PT, Bernstein SI, Wright TR, Emerson CP Jr. Mutations of the *Drosophila* myosin heavy-chain gene: effects on transcription, myosin accumulation, and muscle function. *Proc Natl Acad Sci U S A*. 1986;83(5):1393–7.
134. Monge I, Mitchell PJ. DAP-2, the *Drosophila* homolog of transcription factor AP-2. *Mech Dev*. 1998;76(1–2):191–doe5.
135. Moran MT, Tare M, Kango-Singh M, Singh A. Homeotic gene *teashirt* (*tsh*) has a neuroprotective function in amyloid-beta 42 mediated neurodegeneration. *PLoS ONE*. 2013;8(11):e80829.
136. Muskavitch MA. Delta-notch signaling and *Drosophila* cell fate choice. *Dev Biol*. 1994;166(2):415–30.
137. Nolan ED, Santibáñez-López CE, Sharma PP. Developmental gene expression as a phylogenetic data class: support for the monophyly of Arachnospulmonata. *Dev Genes Evol*. 2020;230(2):137–53.
138. Oda H, Nishimura O, Hirao Y, Tarui H, Agata K, Akiyama-Oda Y. Progressive activation of Delta-notch signaling from around the blastopore is required to set up a functional caudal lobe in the spider *Achaearanea tepidariorum*. *Development*. 2007;134(12):2195–205.
139. Oda H, Akiyama-Oda Y. The common house spider *Parasteatoda tepidariorum*. *EvoDevo*. 2020;11:6.
140. Ogawa Y, Kakumoto K, Yoshida T, Kuwako KI, Miyazaki T, Yamaguchi J, Konno A, Hata J, Uchiyama Y, Hirai H, Watanabe M, Darnell RB, Okano H, Okano HJ. *Elavl3* is essential for the maintenance of Purkinje neuron axons. *Sci Rep*. 2018;8(1):2722.
141. Oliveira MB, Liedholm SE, Lopez JE, Lochte AA, Pazio M, Martin JP, Mörch PR, Salakka S, York J, Yoshimoto A, Janssen R. Expression of arthropod distal limb-patterning genes in the onychophoran *Euperipatoides kanangrensis*. *Dev Genes Evol*. 2014;224(2):87–96.
142. Panara V, Budd GE, Janssen R. Phylogenetic analysis and embryonic expression of panarthropod *Dmrt* genes. *Front Zool*. 2019;16:23.
143. Panfilio KA, Chuva de Sousa Lopes SM. The extended analogy of extraembryonic development in insects and amniotes. *Phil Trans R Soc*. 2022;B377:20210268. <https://doi.org/10.1098/rstb.2021.0268>.
144. Pechmann M, Prpic NM. Appendage patterning in the South American bird spider *Acanthoscurria geniculata* (araneae: mygalomorphae). *Dev Genes Evol*. 2009;219(4):189–98.
145. Pechmann M, Khadjeh S, Turetzek N, McGregor AP, Damen WG, Prpic NM. Novel function of *distal-less* as a gap gene during spider segmentation. *PLoS Genet*. 2011;7(10):e1002342.
146. Perez Sanchez C, Casas-Tinto S, Sanchez L, Rey-Campos J, Granadino B. *Dm fox F*, a novel *drosophila* fork head factor expressed in visceral mesoderm. *Mech Dev*. 2002;111:163–6.
147. Pichaud F, Casares F. *homothorax* and *Iroquois-C* genes are required for the establishment of territories within the developing eye disc. *Mech Dev*. 2000;96(1):15–25.
148. Pierani A, Moran-Rivard L, Sunshine MJ, Littman DR, Goulding M, Jessell TM. Control of interneuron fate in the developing spinal cord by the progenitor homeodomain protein *Dbx1*. *Neuron*. 2001;29(2):367–84.
149. Prpic NM, Janssen R, Wigand B, Klingler M, Damen WG. Gene expression in spider appendages reveals reversal of *exd/hth* spatial specificity, altered leg gap gene dynamics, and suggests divergent distal morphogen signaling. *Dev Biol*. 2003;264(1):119–40.
150. Prpic NM, Damen WG. Notch-mediated segmentation of the appendages is a molecular phylotypic trait of the arthropods. *Dev Biol*. 2009;326(1):262–71.
151. Punovuori K, Migueles RP, Malaguti M, Blin G, Macleod KG, Carragher NO, Pieters T, van Roy F, Stemmler MP, Lowell S. N-cadherin stabilises neural identity by dampening anti-neural signals. *Development*. 2019;146(21):dev183269.
152. Rehorn KP, Thelen H, Michelson AM, Reuter R. A molecular aspect of hematopoiesis and endoderm development common to vertebrates and *Drosophila*. *Development*. 1996;122(12):4023–31.
153. Reuter R, Grunewald B, Leptin M. A role for the mesoderm in endodermal migration and morphogenesis in *Drosophila*. *Development*. 1993;119(4):1135–45.
154. Roark M, Sturtevant MA, Emery J, Vaessin H, Grell E, Bier E. *scratch*, a pan-neural gene encoding a zinc finger protein related to snail, promotes neuronal development. *Genes Dev*. 1995;9(19):2384–98.
155. Sasaki M, Akiyama-Oda Y, Oda H. Evolutionary origin of type IV classical cadherins in arthropods. *BMC Evol Biol*. 2017;17(1):142.
156. Satija R, Farrell JA, Gennert D, Schier AF, Regev A. Spatial reconstruction of single-cell gene expression data. *Nat Biotechnol*. 2015;33(5):495–502.
157. Schacht MI, Schomburg C, Bucher G. *six3* acts upstream of *foxQ2* in labrum and neural development in the spider *Parasteatoda tepidariorum*. *Dev Genes Evol*. 2020;230(2):95–104.
158. Schaeper ND, Pechmann M, Damen WG, Prpic NM, Wimmer EA. Evolutionary plasticity of *collier* function in head development of diverse arthropods. *Dev Biol*. 2010;344(1):363–76.
159. Schneider E, Winzer R, Rissiek A, Ricklefs I, Meyer-Schwesinger C, Ricklefs FL, Bauche A, Behrends J, Reimer R, Brenna S, Wasielewski H, Lauten M, Rissiek B, Puig B, Cortesi F, Magnus T, Flegert R, Müller CE, Gagliani N, Tolosa E. CD73-mediated adenosine production by CD8 T cell-derived extracellular vesicles constitutes an intrinsic mechanism of immune suppression. *Nat Commun*. 2021;12(1):5911.
160. Schneitz K, Spielmann P, Noll M. Molecular genetics of *aristaless*, a *prd-type* homeo box gene involved in the morphogenesis of proximal and distal pattern elements in a subset of appendages in *Drosophila*. *Genes Dev*. 1993;7(1):114–29.
161. Schnorrer F, Kalchauer I, Dickson BJ. The transmembrane protein *Kontiki* couples to *Dgrip* to mediate myotube targeting in *Drosophila*. *Dev Cell*. 2007;12(5):751–66.
162. Schönauer A, Paese CL, Hilbrant M, Leite DJ, Schwager EE, Feitosa NM, Eibner C, Damen WG, McGregor AP. The Wnt and Delta-Notch signalling pathways interact to direct pair-rule gene expression via caudal during segment addition in the spider *Parasteatoda tepidariorum*. *Development*. 2016;143(13):2455–63.
163. Schomburg C, Turetzek N, Schacht MI, Schneider J, Kirfel P, Prpic NM, Posnien N. Molecular characterization and embryonic origin of the eyes in the common house spider *Parasteatoda tepidariorum*. *EvoDevo*. 2015;6:15.
164. Schoppmeier M, Damen WG. Double-stranded RNA interference in the spider *Cupiennius salei*: the role of *Distal-less* is evolutionarily conserved in arthropod appendage formation. *Dev Genes Evol*. 2001;211(2):76–82.
165. Schuster CM, Davis GW, Fetter RD, Goodman CS. Genetic dissection of structural and functional components of synaptic plasticity. I. Fasciclin II controls synaptic stabilization and growth. *Neuron*. 1996;17(4):641–54.
166. 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, Doddapaneni H, Dugan S, Eibner C, Extavour CG, Funch P, Garb J, Gonzalez LB, Gonzalez VL, Griffiths-Jones S, Han Y, Hayashi C, Hilbrant M, Hughes DST, Janssen R, Lee SL, Maeso 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, Vanthournout B, Werren JH, Wolff 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. The house spider genome reveals an ancient whole-genome duplication during arachnid evolution. *BMC Biol*. 2017;15(1):62.
167. Setton EVW, March LE, Nolan ED, Jones TE, Cho H, Wheeler WC, Extavour CG, Sharma PP. Expression and function of spineless orthologs correlate with distal deutocerebral appendage morphology across Arthropoda. *Dev Biol*. 2017;430(1):224–36.
168. Sharma PP. Chelicerates and the conquest of land: a view of arachnid origins through an evo-devo spyglass. *Integr Comp Biol*. 2017;57(3):510–22.
169. Shen CP, Jan LY, Jan YN. *Miranda* is required for the asymmetric localization of *Prospero* during mitosis in *Drosophila*. *Cell*. 1997;90(3):449–58.
170. Sherr CJ, Roberts JM. CDK inhibitors: positive and negative regulators of G1-phase progression. *Genes Dev*. 1999;13(12):1501–12.
171. Shi L, Lee T. Molecular diversity of *Dscam* and self-recognition. *Adv Exp Med Biol*. 2012;739:262–75.

172. Shishido E, Ono N, Kojima T, Saigo K. Requirements of DFR1/Heartless, a mesoderm-specific *Drosophila* FGF-receptor, for the formation of heart, visceral and somatic muscles, and ensheathing of longitudinal axon tracts in CNS. *Development*. 1997;124(11):2119–28.
173. Sink H, Rehm EJ, Richstone L, Bulls YM, Goodman CS. sidestep encodes a target-derived attractant essential for motor axon guidance in *Drosophila*. *Cell*. 2001;105(1):57–67.
174. Skeath JB, Carroll SB. The achaete-scute complex: generation of cellular pattern and fate within the *Drosophila* nervous system. *FASEB J*. 1994;8(10):714–21.
175. Skeath JB, Doe CQ. The achaete-scute complex proneural genes contribute to neural precursor specification in the *Drosophila* CNS. *Curr Biol*. 1996;6(9):1146–52.
176. Stollewerk A. Recruitment of cell groups through Delta/Notch signaling during spider neurogenesis. *Development*. 2002;129(23):5339–48.
177. Stollewerk A, Tautz D, Weller M. Neurogenesis in the spider: new insights from comparative analysis of morphological processes and gene expression patterns. *Arthropod Struct Dev*. 2003;32(1):5–16.
178. Stollewerk A, Seyfarth EA. Evolutionary changes in sensory precursor formation in arthropods: embryonic development of leg sensilla in the spider *Cupiennius salei*. *Dev Biol*. 2008;313(2):659–73.
179. Tabuchi K, Yoshikawa S, Yuasa Y, Sawamoto K, Okano H. A novel *Drosophila* paired-like homeobox gene related to *Caenorhabditis elegans* unc-4 is expressed in subsets of postmitotic neurons and epidermal cells. *Neurosci Lett*. 1998;257(1):49–52.
180. Takeshima H, Nishi M, Iwabe N, Miyata T, Hosoya T, Masai I, Hotta Y. Isolation and characterization of a gene for a ryanodine receptor/calcium release channel in *Drosophila melanogaster*. *FEBS Lett*. 1994;337(1):81–7.
181. Theocharis AD, Skandalis SS, Gialeli C, Karamanos NK. Extracellular matrix structure. *Adv Drug Deliv Rev*. 2016;97:4–27.
182. Toegel JP, Wimmer EA, Ppirc NM. Loss of spineless function transforms the *Tribolium* antenna into a thoracic leg with pretarsal, tibiotarsal, and femoral identity. *Dev Genes Evol*. 2009;219(1):53–8.
183. Tomlinson A, Mavromatakis YE, Arias R. The role of sevenless in *Drosophila* R7 photoreceptor specification. *Dev Biol*. 2019;454(2):181–9.
184. Treffkorn S, Mayer G, Janssen R. Review of extra-embryonic tissues in the closest arthropod relatives, onychophorans and tardigrades. *Phil Trans R Soc B*. 2022;377:20210270. <https://doi.org/10.1098/rstb.2021.0270>.
185. Tripathi BK, Das R, Mukherjee A, Mutsuddi M. Interaction of spoonbill with prospero in *Drosophila*: Implications in neuroblast development. *Genesis*. 2017;55(9). <https://doi.org/10.1002/dvg.23049>.
186. Trush O, Liu C, Han X, Nakai Y, Takayama R, Murakawa H, Carrillo JA, Takechi H, Hakeda-Suzuki S, Suzuki T, Sato M. N-cadherin orchestrates self-Organization of Neurons within a columnar unit in the *Drosophila* medulla. *J Neurosci*. 2019;39(30):5861–80.
187. Turetzek N, Pechmann M, Schomburg C, Schneider J, Ppirc NM. Neofunctionalization of a duplicate dachshund gene underlies the evolution of a novel leg segment in Arachnids. *Mol Biol Evol*. 2016;33(1):109–21.
188. Ungerer P, Eriksson BJ, Stollewerk A. Neurogenesis in the water flea *Daphnia magna* (Crustacea, Branchiopoda) suggests different mechanisms of neuroblast formation in insects and crustaceans. *Dev Biol*. 2011;357(1):42–52.
189. Urbano JM, Domínguez-Giménez P, Estrada B, Martín-Bermudo MD. PS integrins and laminins: key regulators of cell migration during *Drosophila* embryogenesis. *PLoS ONE*. 2011;6(9):e23893.
190. Vässin H, Bremer KA, Knust E, Campos-Ortega JA. The neurogenic gene Delta of *Drosophila melanogaster* is expressed in neurogenic territories and encodes a putative transmembrane protein with EGF-like repeats. *EMBO J*. 1987;6(11):3431–40.
191. Wang J, Ding M. Robo and Ror function in a common receptor complex to regulate Wnt-mediated neurite outgrowth in *Caenorhabditis elegans*. *Proc Natl Acad Sci U S A*. 2018;115(10):E2254–63.
192. Wang Q, Han TH, Nguyen P, Jarnik M, Serpe M. Tenectin recruits integrin to stabilize Bouton architecture and regulate vesicle release at the *Drosophila* neuromuscular junction. *Elife*. 2018;7:e35518.
193. Weller M, Tautz D. Prospero and snail expression during spider neurogenesis. *Dev Genes Evol*. 2003;213(1):54–66.
194. Wheeler SR, Carrico ML, Wilson BA, Brown SJ, Skeath JB. The expression and function of the achaete-scute genes in *Tribolium castaneum* reveals conservation and variation in neural pattern formation and cell fate specification. *Development*. 2003;130(18):4373–81.
195. Wu X, Hu S, Kang X, Wang C. Synaptotagmins: beyond presynaptic neurotransmitter release. *Neuroscientist*. 2020;26(1):9–15.
196. Xu J, Hao X, Yin MX, Lu Y, Jin Y, Xu J, Ge L, Wu W, Ho M, Yang Y, Zhao Y, Zhang L. Prevention of medulla neuron dedifferentiation by Nerfin-1 requires inhibition of notch activity. *Development*. 2017;144(8):1510–7.
197. Yamazaki K, Akiyama-Oda Y, Oda H. Expression patterns of a twist-related gene in embryos of the spider *Achaearanea tepidariorum* reveal divergent aspects of mesoderm development in the fly and spider. *Zoolog Sci*. 2005;22(2):177–85.
198. Yang S, Wu X, Daoutidou EI, Zhang Y, Shimell M, Chuang KH, Peterson AJ, O'Connor MB, Zheng X. The NDNF-like factor Nord is a hedgehog-induced extracellular BMP modulator that regulates *Drosophila* wing patterning and growth. *Elife*. 2022;11:e73357.
199. Ypsilanti AR, Chedotal A. Roundabout receptors. *Adv Neurobiol*. 2014;8:133–64.
200. Zander MA, Cancino GI, Gridley T, Kaplan DR, Miller FD. The snail transcription factor regulates the numbers of neural precursor cells and newborn neurons throughout mammalian life. *PLoS ONE*. 2014;9(8):e104767.
201. Zhang N (née Turetzek). Developmental Studies of Appendage Patterning and Formation in Spiders. PhD thesis. 2016. Georg-August University Göttingen, Germany. <https://doi.org/10.53846/goediss-5988>
202. Zhang Y, Featherstone D, Davis W, Rushton E, Broadie K. *Drosophila* D-titin is required for myoblast fusion and skeletal muscle striation. *J Cell Sci*. 2000;113(Pt 17):3103–15.
203. Zipursky SL, Sanes JR. Chemoaffinity revisited: dscams, protocadherins, and neural circuit assembly. *Cell*. 2010;143(3):343–53.

Publisher's Note

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