

# Evolutionary emergence of Hairless as a novel component of the Notch signaling pathway

Steven W Miller<sup>1</sup>, Artem Movsesyan<sup>1</sup>, Sui Zhang<sup>1</sup>, Rosa Fernández<sup>2†</sup>, James W Posakony<sup>1\*</sup>

<sup>1</sup>Division of Biological Sciences, Section of Cell and Developmental Biology, University of California, San Diego, La Jolla, United States; <sup>2</sup>Bioinformatics and Genomics Unit, Center for Genomic Regulation, Barcelona, Spain

**Abstract** Suppressor of Hairless [Su(H)], the transcription factor at the end of the Notch pathway in *Drosophila*, utilizes the Hairless protein to recruit two co-repressors, Groucho (Gro) and C-terminal Binding Protein (CtBP), indirectly. Hairless is present only in the Pancrustacea, raising the question of how Su(H) in other protostomes gains repressive function. We show that Su(H) from a wide array of arthropods, molluscs, and annelids includes motifs that directly bind Gro and CtBP; thus, direct co-repressor recruitment is ancestral in the protostomes. How did Hairless come to replace this ancestral paradigm? Our discovery of a protein (S-CAP) in Myriapods and Chelicerates that contains a motif similar to the Su(H)-binding domain in Hairless has revealed a likely evolutionary connection between Hairless and Metastasis-associated (MTA) protein, a component of the NuRD complex. Sequence comparison and widely conserved microsynteny suggest that S-CAP and Hairless arose from a tandem duplication of an ancestral MTA gene.

DOI: <https://doi.org/10.7554/eLife.48115.001>

\*For correspondence:

[jposakony@ucsd.edu](mailto:jposakony@ucsd.edu)

**Present address:** <sup>†</sup>Department of Life Sciences, Barcelona Supercomputing Center, Barcelona, Spain

**Competing interests:** The authors declare that no competing interests exist.

**Funding:** See page 19

**Received:** 03 May 2019

**Accepted:** 23 September 2019

**Published:** 23 September 2019

**Reviewing editor:** Patricia J Wittkopp, University of Michigan, United States

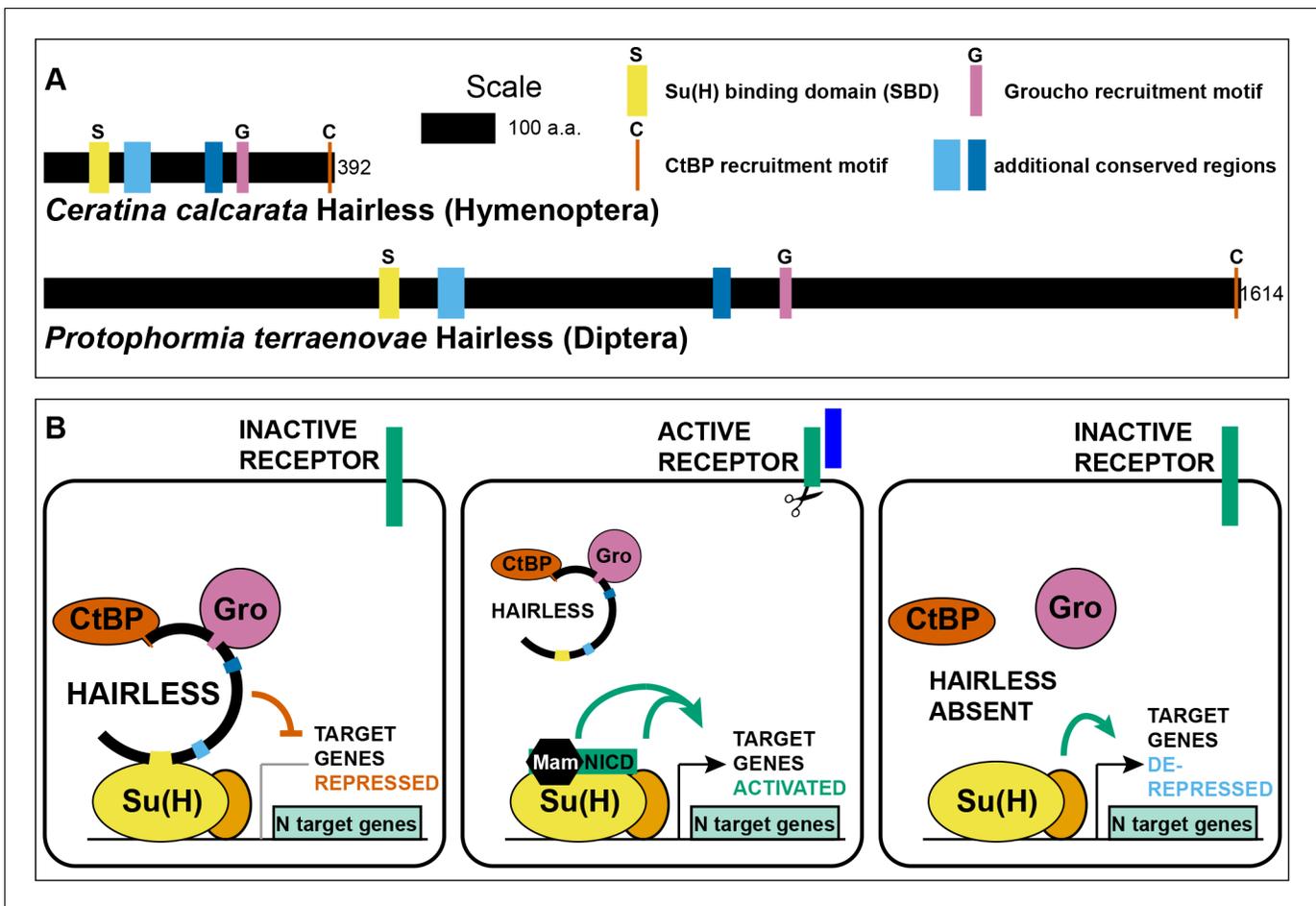
© Copyright Miller et al. This article is distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use and redistribution provided that the original author and source are credited.

## Introduction

A very common paradigm in the regulation of animal development is that DNA-binding transcriptional repressors bear defined amino acid sequence motifs that permit them to recruit, by direct interaction, one or more common co-repressor proteins that are responsible for conferring repressive activity. Two such universal co-repressors are Groucho (Gro) and C-terminal Binding Protein (CtBP).

The ancient and highly conserved transcription factor Suppressor of Hairless [Su(H)] functions at the terminus of the widely utilized Notch cell-cell signaling pathway. Su(H) is converted into an activator by signaling through the Notch receptor, but in the absence of signaling it functions as a repressor. Earlier studies have revealed that in many settings in *Drosophila*, Su(H)'s repressive activity depends on binding to the Hairless protein (**Figure 1**). Hairless includes separate Gro- and CtBP-binding motifs, which permit it to function as an adaptor to bring these two corepressors to Su(H) (**Figure 1B**) (**Barolo et al., 2002a**). Thus, the Su(H)/H partnership in the fly represents a notable exception to the rule of direct co-repressor recruitment.

As genome and transcriptome sequences have become available for more and more insects and other arthropods, we have searched for possible Hairless orthologs in a wide variety of species, in an attempt to determine the protein's phylogenetic distribution. We have found that Hairless is confined to the Pancrustacea (or Tetraconata), a clade of arthropods that includes the Crustacea and Hexapoda (**Misof et al., 2014; Kjer et al., 2016**). While this indicates that Hairless was gained at least 500 Mya, it also raises the question of how Su(H) in other protostomes acquires repressive activity.



**Figure 1.** Hairless mediates indirect recruitment of co-repressor proteins to Su(H). (A) Diagram denoting locations of conserved domains and motifs within Hairless, and illustrating extreme size differences of the protein in different species. Shown is Hairless from the carpenter bee *Ceratina calcarata* and the blowfly *Protophormia terraenovae* (Hase et al., 2017), with scale and protein sizes indicated. (B) Summary of Hairless's known mode of action (Lai, 2002; Maier, 2006) as an adaptor protein that recruits the global co-repressors C-terminal Binding Protein (CtBP) and Groucho (Gro) to Suppressor of Hairless [Su(H)], the transducing transcription factor for the Notch (N) cell-cell signaling pathway; adapted from Figure 6 of Barolo et al. (2002a). In the absence of signaling through the Notch receptor (left), Su(H) acts as a repressor of Notch target genes, despite the presence of transcriptional activator proteins (orange oval). Upon activation of the Notch receptor (middle), Su(H), in a complex with the receptor's intracellular domain (NICD) and the co-activator Mastermind (Mam), functions to activate transcription of pathway target genes in cooperation with other transcriptional activators. In the absence of Hairless and hence in the absence of Su(H)'s repressive activity (right), the partner transcription factors are often sufficient to activate expression of target genes in a signal-independent manner (Barolo and Posakony, 2002b).

DOI: <https://doi.org/10.7554/eLife.48115.002>

The following figure supplement is available for figure 1:

**Figure supplement 1.** Graph showing predicted disordered regions in *Drosophila melanogaster* Hairless, generated by DISOPRED3 (Buchan et al., 2013; Jones and Cozzetto, 2015).

DOI: <https://doi.org/10.7554/eLife.48115.003>

Here we present evidence that direct co-repressor recruitment by Su(H) is likely to be ancestral in the protostomes. We show that Su(H) in a broad range of protostomes, including arthropods, molluscs, and annelids, bears both a short linear motif that mediates binding of CtBP and a novel motif for direct recruitment of Gro. Thus, the evolutionary appearance of Hairless has permitted the replacement of an ancient and predominant regulatory mechanism (direct co-repressor recruitment) with a novel one (indirect recruitment).

What can we learn about the evolutionary history of Hairless? While Hairless itself is found only in the Pancrustacea, we show that the genomes of Myriapods and Chelicerates encode a protein with clear sequence and functional similarities to Hairless. These proteins include a motif that strongly

resembles the Su(H)-binding domain of Hairless, and we demonstrate that this motif from the house spider *Parasteatoda tepidariorum* does indeed bind Su(H). In addition, these Myriapod and Chelicerate proteins also include one or more canonical motifs for recruitment of CtBP. Accordingly, we designate these factors as 'Su(H)-Co-repressor Adaptor Proteins' (S-CAPs).

Finally, further sequence analyses, along with the discovery of conserved microsynteny, have provided substantial evidence that Hairless and the S-CAPs are likely to be homologous and that they arose from a duplication of the gene encoding Metastasis-associated (MTA) protein, a component of the nucleosome remodeling and deacetylase (NuRD) complex.

An intriguing question in evolutionary biology concerns the path by which a particular clade has escaped a strongly selected character that has been conserved for hundreds of millions of years. We believe that our study has yielded valuable insight into both the emergence of an evolutionary novelty and its replacement of an ancestral paradigm.

## Results

### Hairless is present only in the Pancrustacea

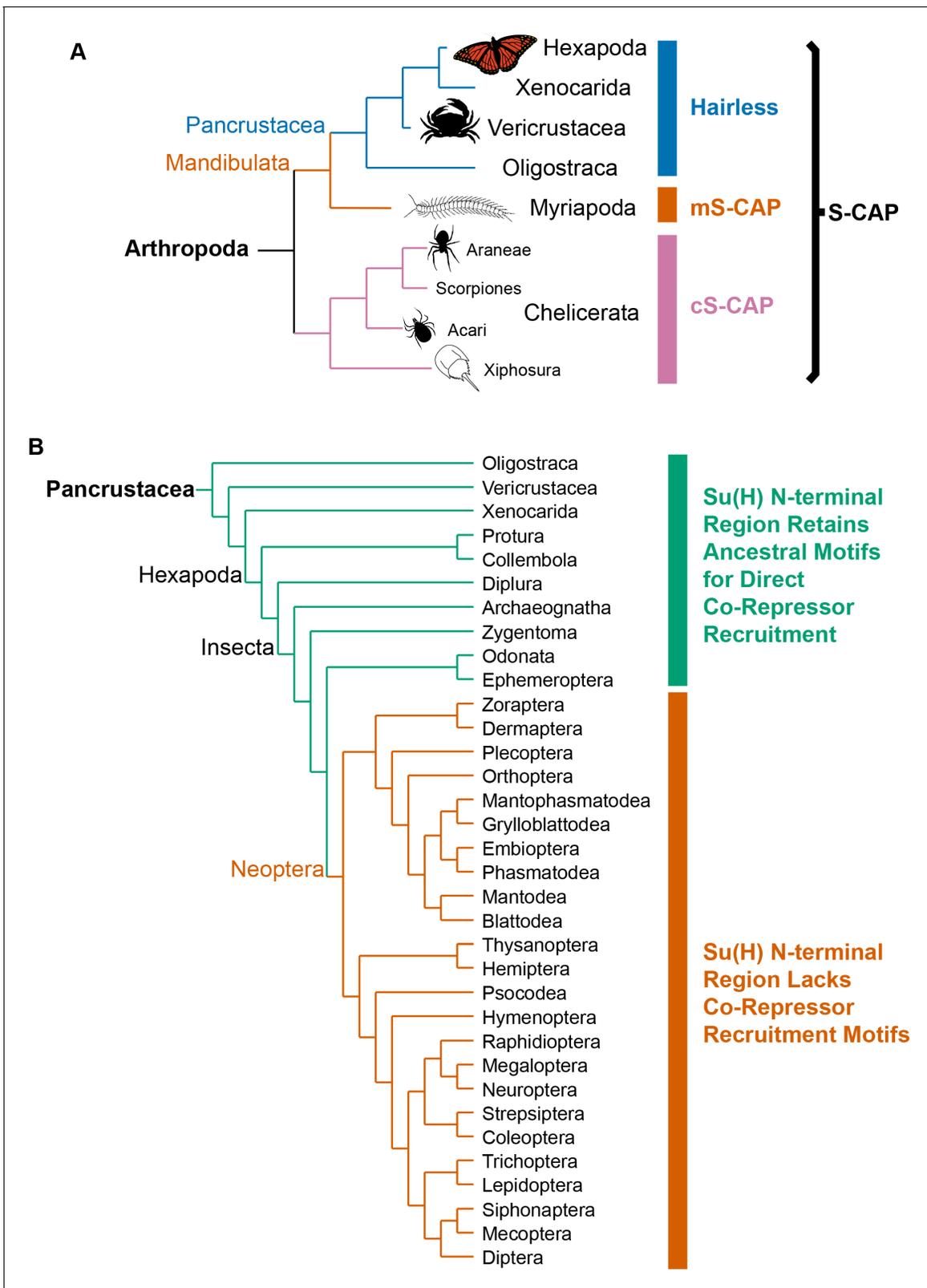
We have conducted extensive BLAST searches of genome and transcriptome sequence data for a wide variety of metazoa in an attempt to define the phylogenetic distribution of Hairless. We find that Hairless as originally described (*Bang and Posakony, 1992; Maier et al., 1992; Maier et al., 2008*) is confined to the Pancrustacea (or Tetraconata), and occurs widely within this clade, including the Hexapoda, Vericrustacea, and Oligostraca (*Figure 2A*). By contrast, no evidence for a true Hairless gene has been detected in either Myriapods or Chelicerates, even in cases where substantially complete genome sequence assemblies are available.

The enormous variation in the size of the Hairless protein in various Pancrustacean clades is worthy of note (*Figure 1A*). The known extremes are represented by the Diplostracan (shrimp) *Eulimnadia texana* (343 aa) (*Baldwin-Brown et al., 2018*) and the Dipteran (fly) *Protophormia terraenovae* (1614 aa) (*Hase et al., 2017*), a 4.7-fold difference. There is a broad tendency for the size of the protein to be relatively stable within an order (*Supplementary file 1*). Thus, as noted previously (*Maier et al., 2008*), the Hymenoptera generally have a small Hairless (of the order of 400 aa; see *Figure 1A*), while the Diptera typically have a much larger version (of the order of 1000 aa or more). Notable exceptions to this pattern of uniformity are aphids, where Hairless is typically ~900 aa compared to ~400 aa in other Hemiptera, and chalcid wasps, where the protein is over 500 aa instead of the Hymenoptera-typical ~400 aa noted above (*Supplementary file 1*). Smaller Hairless proteins typically retain all five conserved motifs/domains characteristic of this factor (*Maier et al., 2008*), while the regions that flank and lie between these sequences are reduced in size (*Figure 1A; Supplementary file 2*).

### A known CtBP-binding motif is present in the non-conserved N-terminal region of Su(H) in a wide variety of protostomes

The apparent confinement of the Hairless co-repressor adaptor protein to the Pancrustacea raises the question of the mechanism(s) by which Su(H) in other protostomes might recruit co-repressor proteins to mediate its repressor function. Of course, other protostomes need not utilize the Gro and CtBP co-repressors for this purpose; different co-repressors might substitute. Nevertheless, we first sought to identify known binding motifs for Gro and CtBP in Su(H) from arthropods lacking Hairless. As shown in *Table 1*, we found a canonical CtBP recruitment motif of the form P $\Phi$ D $\Phi$ S (where  $\Phi$  = I, L, M, or V) in predicted Su(H) proteins from a variety of Myriapods and Chelicerates, including the centipede *Strigamia maritima*, the tick *Ixodes scapularis*, the spider *Parasteatoda tepidariorum*, the horseshoe crab *Limulus polyphemus*, and the scorpion *Centruroides sculpturatus*. These motifs are all located in the non-conserved N-terminal region of Su(H) (*Supplementary file 3*).

Extending this sequence analysis to other protostome phyla led to the finding that a similar P $\Phi$ D $\Phi$ S motif occurs in the N-terminal region of Su(H) from a large number of molluscs and annelids, as well as from multiple Nemertea, Brachiopoda, Phoronida, and monogonont rotifers, and also from some flatworms (*Table 1*). It is notable, by contrast, that we do not find CtBP-binding motifs present in Su(H) from nematodes. Nevertheless, given the broad phylogenetic distribution of the



**Figure 2.** Phylogenetic distribution of *Hairless* and related S-CAP proteins. (A) Based on extensive BLAST searches of available genome and transcriptome assemblies, orthologs of the canonical *Hairless* gene are found only in the Pancrustacea (blue bar), while orthologs of a gene that encodes the related S-CAP protein are found in the Myriapods (mS-CAP, red bar) and Chelicerates (cS-CAP, pink bar). We suggest S-CAP as a suitable umbrella nomenclature for this gene family (black bracket). Tree adapted from Figure 2 of **Regier et al. (2010)**. (B) Consistent with the presence of *Hairless* in Pancrustacea, Hexapoda, and Insecta, the presence of Su(H) N-terminal motifs is retained in these groups (green bar). In contrast, the absence of Su(H) N-terminal motifs is observed in Neoptera (orange bar). *Figure 2 continued on next page*

Figure 2 continued

Hairless as an adaptor protein, Su(H) in most insect orders (the Neoptera clade) has lost the ancestral short linear motifs that mediate direct recruitment of the CtBP and Gro co-repressor proteins (red bar). However, in the Crustacea, Collembola, Diplura, and a subset of Insecta, the ancestral recruitment motifs have been retained in Su(H), despite the presence of Hairless (see **Table 1** and **Supplementary file 3**). Tree adapted from **Misof et al. (2014)** and **Kjer et al. (2016)**.

DOI: <https://doi.org/10.7554/eLife.48115.004>

P $\phi$ D $\phi$ S motif in Su(H) from both Ecdysozoa and Lophotrochozoa, our observations strongly suggest that direct recruitment of CtBP by Su(H) is ancestral in the protostomes.

To verify that the shared P $\phi$ D $\phi$ S motif in protostome Su(H) proteins can indeed mediate direct recruitment of CtBP, we carried out an in vitro pulldown assay using GST-tagged *Drosophila* CtBP (bound to Glutathione Sepharose beads) and a His-tagged fragment of *Strigamia maritima* Su(H) (**Figure 3A**). We found that the two proteins do interact directly and robustly, in a manner that is dependent on the integrity of the PVDLS motif in *Strigamia* Su(H).

### A novel conserved motif in protostome Su(H) binds the Gro co-repressor

In addition to a P $\phi$ D $\phi$ S CtBP-binding motif, we have found that Su(H) from a wide variety of protostomes includes a novel motif similar to GSLTPDKV (**Table 1**). Where present, this sequence typically lies a short (but variable) distance C-terminal to the P $\phi$ D $\phi$ S motif, also within the non-conserved N-terminal region of the protein (**Supplementary file 3**). The GSLTPDKV motif is particularly prevalent in Su(H) from the Trochozoa, which includes annelids, sipunculans, molluscs, nemerteans, brachiopods, and phoronids (**Kocot et al., 2017**). Among the Ecdysozoa, it appears consistently in Su(H) from Crustacea and Myriapoda, and in small subsets of both Hexapoda (Ephemeroptera, Odonata, Zygentoma, Archaeognatha, Diplura, and Collembola) and Chelicerata [harvestmen (Opiliones) and Scorpiones]. The motif is absent from Su(H) in all other insect orders, and we have not found it so far in Su(H) from nematodes, flatworms, rotifers, or tardigrades; it is, however, found in the onychophoran *Euperipatoides kanangrensis* (**Table 1**). Perhaps surprisingly, the motif is present in Su(H) from the acorn worms *Saccoglossus kowalevskii* and *Ptychodera flava* (**Simakov et al., 2015**), which are hemichordates (deuterostomes).

Using an in vitro pulldown assay, we tested the possibility that the GSLTPDKV motif mediates binding of the Gro co-repressor (**Figure 3B**). Indeed, we find that GST-tagged Gro protein interacts strongly with a His-tagged protein bearing this motif at its C-terminus, and that this binding is abolished when the motif is replaced by alanine residues. We conclude that Su(H) from a broad range of protostomes is capable of directly recruiting both CtBP and Gro (**Table 1**), and that this capacity is hence very likely to be ancestral in this clade.

### Retention of the hybrid state: Species that have both Hairless and the co-repressor-binding motifs in Su(H)

The evolutionary emergence of Hairless as an adaptor protein capable of mediating the indirect recruitment of both Gro and CtBP to Su(H) might be expected to relieve a selective pressure to retain the ancestral Gro- and CtBP-binding motifs in Su(H) itself. And indeed, we find that Su(H) from multiple insect orders comprising the Neoptera lacks both of these sequences (**Figure 2B**). Strikingly, however, we have observed that Crustacea and a small group of Hexapoda retain both traits (**Figure 2B**). Thus, multiple representatives of the Branchiopoda, Malacostraca, and Copepoda, along with Ephemeroptera, Odonata, Zygentoma, Archaeognatha, Diplura, and Collembola, have both a canonical Hairless protein (including its Gro- and CtBP-binding motifs) and Gro- and CtBP-binding motifs within Su(H). These clades, then, appear to have retained a 'hybrid intermediate' state (**Baker et al., 2012**) characterized by the presence of both co-repressor recruitment mechanisms.

### Myriapods and Chelicerates encode a protein with similarity to Hairless

While canonical Hairless proteins are confined to the Pancrustacea, we have discovered that the genomes of Myriapods and Chelicerates nevertheless encode a protein with intriguing similarities to

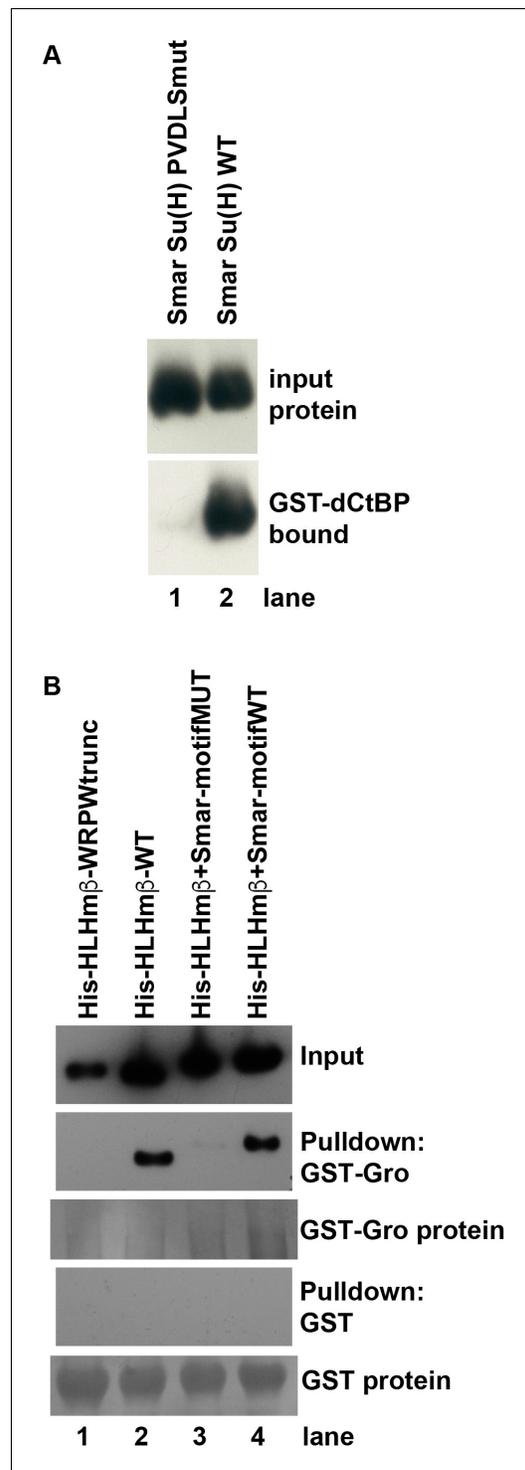
**Table 1.** Co-repressor recruitment motifs in protostome Su(H) proteins.

Species	CtBP motif	Gro motif	Source
<i>Ecdyonurus insignis</i>	YPDNHPVDLSSPRPH	APMIPGSLTPDKMNGEHPHHG	GCCL01029953.1 ( <i>Simon et al., 2018</i> )
<i>Calopteryx splendens</i>	YTDNHPVDLSSPRPP	HHMIPGSLTPDKMNGEHPAMH	LYUA01002621.1 ( <i>Ioannidis et al., 2017</i> )
<i>Atelura formicaria</i>	YPDNHPVDLSSPRPQ	PHMIPGSLTPDKMNGEHPHHS	GAYJ02050375.1 ( <i>Misof et al., 2014</i> )
<i>Machilis hrabei</i>	YPDNHPVDLSSPRPH	PHMLPGSLTPDKMNGEHPHHG	Scaffold 1 ( <i>i5K Consortium, 2013</i> )
<i>Catajapyx aquilonaris</i>	STANNPVDLSSPRGS	APMIPGSLTPDKVNGEHHSHH	JYFJ02000853.1 ( <i>i5K Consortium, 2013</i> )
<i>Holacanthella duospinosa</i>	VPNSNPVDLSNPSPS	SNFVPGSLSPPERMNGNDPSLL	NIPM01000059.1 ( <i>Wu et al., 2017</i> )
<i>Pollicipes pollicipes</i>	YPDNHPVDLSSPRPE	GPLIAGSLTPDKLGAELGLHA	GGJN01104381.1 (unpublished)
<i>Hyalella azteca</i>	SLGHRPVDLSQAPSP	AAMLAGSLTPDKLNSDPQQQQ	NW_017238139.1 ( <i>i5K Consortium, 2013</i> )
<i>Eurytemora affinis</i>	SETSAPVDLSAPRPN	YGMLPGSLTPDKLNGDHCSFG	NW_019396104.1 ( <i>i5K Consortium, 2013</i> )
<i>Triops cancriformis</i>	HPEARPVDLSSSRLL	YHSSSLTTPDKVNVGDSNSQ	BAYF01001879.1 ( <i>Ikeda et al., 2015</i> )
<i>Argulus siamensis</i>	YPENNPVDLSNSRTG	SPMIPGSLTPDKMNGEHPHGH	JW959185.1 ( <i>Sahoo et al., 2013</i> )
<i>Strigamia maritima</i>	FADNHPVDLSNSHRG	SHMIAGSLTPDKVNGEHHGHL	JH430541.1 ( <i>Chipman et al., 2014</i> )
<i>Sigmoria lator munda</i>	TNENHPVDLSSSHRS	SHMIPGSLTPDKGNAEHSLSH	( <i>Rodriguez et al., 2018</i> )
<i>Metaseiulus occidentalis</i>	GADRKPLDMSAAHRS		NW_003805473.1 ( <i>Hoy et al., 2016</i> )
<i>Ixodes scapularis</i>	QAAGAPVDMSSHPAR		NW_002722632.1 ( <i>Gulia-Nuss et al., 2016</i> )
<i>Parasteatoda tepidariorum 1</i>	VIDSHPVDLSSPKPS		NW_018383625.1 ( <i>Schwager et al., 2017</i> )
<i>Parasteatoda tepidariorum 2</i>	RYEGRPVDLSSPRPN		NW_018370942.1 ( <i>Schwager et al., 2017</i> )
<i>Limulus polyphemus 1</i>	PYDGHVPVDLSNQRPD		NW_013671976.1 ( <i>Battelle et al., 2016</i> )
<i>Limulus polyphemus 2</i>	TYESHVPVDLSNQRPD		NW_013676581.1 ( <i>Battelle et al., 2016</i> )
<i>Centruroides sculpturatus</i>	GYESSPVDLSSHSRV	MQLISGSMTSHDKVNGDQHSLSG	NW_019384406.1 ( <i>Schwager et al., 2017</i> )
<i>Euperipatoides kanangrensis</i>	NSYDNPVDLSSHRSS	QQILPGSLGPSDKVNGDLVSLA	LN881712.1 (unpublished)
<i>Naineris dendritica</i>	DPNGHPVDLSHSRHI	PHMIHGSLTPPDRVNGEPGSGSL	( <i>Andrade et al., 2015</i> )
<i>Platynereis dumerilii</i>	MASENPVDLSSRHVG	GNHFPGTLTPDKLNGDHNAHH	KP293861.1 ( <i>Gazave et al., 2017</i> )
<i>Nephasoma pellucidum</i>	AGYETPVDLSSPRPC	SHLIPGSLTPDKINGEGITTS	( <i>Lemer et al., 2015</i> )
<i>Owenia sp.</i>	QPYENPVDLSSRRHIK	AHLIPGSLTPDKINGDMVTMA	( <i>Andrade et al., 2015</i> )
<i>Octopus bimaculoides</i>	NGFDNPMDLSNGKVV	HLMPAGSLTPDKISGDSISMA	NW_014678436.1 ( <i>Albertin et al., 2015</i> )
<i>Crassostrea gigas</i>	GGYENPMDLSSNKPG	SHIVAGSLTPPEKINGDPGAMA	NW_011936122.1 ( <i>Zhang et al., 2012</i> )
<i>Lottia gigantea</i>	AGVENPVDLNNGRIS	SHLFTGSLTPPEKPNGDLVPMMS	NW_008708401.1 ( <i>Simakov et al., 2013</i> )
<i>Notospermus geniculatus</i>	VQYDNPIDLNSRLEG	NHMIPGSLTPDKVNGDMVPLP	GFY01035878.1 ( <i>Luo et al., 2018</i> )
<i>Malacobdella grossa</i>	LHYDNPDLTLNRLDE	GSGIAGSMTPPDGGKGNLDLQ	( <i>Whelan et al., 2014</i> )
<i>Lingula anatina</i>	GGYENPMDLSRRTEM	AHMIPGNLTPDKVNGEMVPMA	GDJY01029776.1 ( <i>Luo et al., 2015</i> )
<i>Phoronis australis</i>	QHDNRPMDLSSRGQH	SHLIAGSLTPDKVNGDVVSMA	GFSC01078935.1 ( <i>Luo et al., 2018</i> )
<i>Procotyla fluviatilis</i>	ETLFEPLDLRSPIGV		GAKZ01044347.1 (unpublished)
<i>Brachionus koreanus</i>	AKDETPIDLSSKSKK		GBXV02009219.1 ( <i>Lee et al., 2015</i> )
<i>Xenoturbella bocki</i>	KRYSAPLNLTVDHDC	DVRVLGRLTPDKQHVNDVGA	( <i>Brauchle et al., 2018</i> )

Shown are alignments of short linear amino acid motifs (bold) in the N-terminal region of Su(H) proteins that mediate direct recruitment of the co-repressors CtBP and Gro. Column at right shows the source of the corresponding sequence data, with accession numbers and publication citations indicated.

DOI: <https://doi.org/10.7554/eLife.48115.005>

Hairless. Most notable is the presence of a motif that strongly resembles the ‘Su(H)-binding domain’ (SBD) of Hairless, which mediates its high-affinity direct interaction with Su(H) (*Figure 1; Figure 4A*). We will refer to these proteins as ‘S-CAPs’; the basis for this designation will be made clear in forthcoming figures. We note that the occurrence of this protein in the centipede *Strigamia maritima* has also recently been reported by *Maier (2019)*. In the Pancrustacea, the N-terminal and C-terminal halves of the Hairless SBD are encoded by separate exons (*Figure 4B*). Strikingly, the related motif in Myriapod and Chelicerate S-CAPs is likewise encoded by separate exons, with exactly the same



**Figure 3.** Direct binding of co-repressor proteins by Su(H) from the centipede *Strigamia maritima*. (A) The PVDLS motif in the N-terminal region of Su(H) from the centipede *Strigamia maritima* directly binds *Drosophila* CtBP. A His-tagged 116-aa segment of the *Strigamia* Su(H) protein, bearing a PVDLS recruitment motif for CtBP, binds strongly to GST-dCtBP (WT, lane 2). Mutation of the motif to alanines (AAAAA) abolishes

Figure 3 continued on next page

splice junction as in Hairless (**Figure 4B**). We believe that this is highly unlikely to be coincidental, and is instead strongly suggestive of an evolutionary relationship between Hairless and S-CAPs.

A recent structural analysis of the Su(H)-Hairless protein complex identified several residues in the Hairless SBD that are involved in binding to the C-terminal domain (CTD) of Su(H) (Yuan *et al.*, 2016) (**Figure 4A**). These include four hydrophobic amino acids in the main body of the SBD (L235, F237, L245, and L247; these are highlighted in red in **Figure 4A**). Note that the Myriapod and Chelicerate S-CAP motifs share these same residues. In addition, a tryptophan (W258) C-terminal to the main body of the Hairless SBD also participates in binding to Su(H) (**Figure 4A**). Myriapod and Chelicerate S-CAPs all include a tryptophan residue at a similar position C-terminal to the main SBD-like domain (**Figure 4A**). Moreover, this particular W residue in both Hairless and the S-CAPs is followed by a hydrophobic residue, typically V or I. These sequence features, we suggest, is further strong evidence of a common ancestry for the respective segments of Hairless and S-CAPs.

A third structural similarity between Hairless and S-CAPs is the presence in the latter of one or more short linear motifs capable of binding the CtBP co-repressor (**Figure 5A**). These motifs typically reside in the C-terminal half of the S-CAPs, superficially resembling the C-terminal location of Hairless's CtBP recruitment motif.

A table listing representative examples of Myriapod and Chelicerate S-CAPs is provided as **Supplementary file 4**, and an annotated FASTA file of their amino acid sequences is included as **Supplementary file 5**. It is important to note that we have not found non-Hairless S-CAPs in the Pancrustacea.

### Spider S-CAP binds to *Drosophila* Su(H)

Given the clear sequence similarity between the Hairless SBD and the SBD-like motif in Myriapod and Chelicerate S-CAPs, we investigated whether the latter motif is likewise capable of mediating direct binding to Su(H). As noted above, the Hairless SBD interacts specifically with the CTD of Su(H). Since this domain in Su(H) is very highly conserved throughout the Bilateria and Cnidaria, we thought it reasonable to utilize *Drosophila* Su(H) for this binding assay. As shown in **Figure 4C**, we find that a 200-amino-acid segment of S-CAP from the spider *Parasteatoda tepidariorum* binds directly to *Drosophila* Su(H) in vitro. This

Figure 3 continued

this interaction (PVDLSmut, lane 1). The results shown in this panel have been replicated in eight additional experiments, utilizing three independent isolations of GST-CtBP protein from bacterial cultures. (B) The conserved GSLTPPDVKV motif in the N-terminal region of *Strigamia* Su(H) directly binds *Drosophila* Gro. His-tagged E(spl)mβ-HLH protein, which bears a C-terminal WRPW motif that recruits Gro, is used as a binding control. Wild-type (WT) HLHmβ binds GST-Gro (lane 2), while a truncated version of the protein lacking the WRPW motif (lane 1) fails to bind. A synthetic version of HLHmβ in which the WRPW motif has been replaced by the wild-type GSLTPPDVKV motif also binds GST-Gro efficiently (lane 4), while a mutant version in which GSLTPPDVKV is replaced by alanines (AAAAAAAAA) shows extremely weak binding (lane 3). No binding of any of the His-tagged proteins to GST alone is observed, even with substantially greater amounts of GST compared to GST-Gro. The results shown in this panel have been replicated in seven additional experiments, utilizing five independent isolations of GST-Gro protein from bacterial cultures.

DOI: <https://doi.org/10.7554/eLife.48115.006>

domains: BAH (Bromo-Adjacent Homology), ELM2 (Egl-27 and MTA1 homology), SANT (Swi3, Ada2, N-CoR, and TFIIIB), and GATA-like zinc finger (Millard et al., 2014) (Figure 5B). Of these, the ELM2 and SANT domains are retained at the N-terminal end of Chelicerate S-CAPs (Figure 5B; Figure 5—figure supplement 1A). This is highly likely to have functional significance, as the ELM2 and SANT domains of MTA proteins work together to recruit and activate the histone deacetylases HDAC1 and HDAC2 (Millard et al., 2013). Further suggesting homology between Chelicerate S-CAPs and MTAs is the observation that their shared ELM2 and SANT domains are each encoded by two exons with exactly the same splice junction (Figure 5C).

It is noteworthy that, despite sharing the SBD-like and CtBP recruitment motifs of Chelicerate S-CAPs, the available Myriapod S-CAP protein sequences lack the N-terminal ELM2 and SANT homologies with MTA proteins (Figure 5B). Consistent with this, the SBD motif in Myriapod S-CAPs lies much closer to the protein's N terminus than the SBD motif in Chelicerate S-CAPs, suggesting that simple loss of the ELM2/SANT-encoding exons might underlie this difference between the two S-CAP clades. Likewise, Hairless proteins are devoid of clear similarities to MTAs.

In addition to their SBD and ELM2/SANT domains, Chelicerate S-CAPs share a third region of homology that lies between the ELM2 and SANT sequences (Figure 5—figure supplement 1A). This region is absent from both Hairless and the Myriapod S-CAPs. Conversely, Myriapod S-CAPs include a segment of sequence similarity that is not found in either Hairless or Chelicerate S-CAPs (Figure 5—figure supplement 1B).

### Conserved microsynteny between MTA and S-CAP/Hairless genes

Our analysis of the genomic locations of genes encoding MTA proteins in Arthropoda, Hairless in Pancrustacea, and S-CAPs in Myriapods and Chelicerates has yielded the surprising finding that proximate or near-proximate linkage between MTA and Hairless genes or between MTA and S-CAP genes is broadly conserved among arthropods (Figure 6; Supplementary file 1; Supplementary file 4). Thus, in the centipede *Strigamia maritima*, the gene encoding S-CAP lies immediately upstream of that encoding MTA, in the same orientation (Figure 6; Supplementary file 4). A similar linkage relationship between S-CAP and MTA genes is seen in many arachnids, including the spiders *Nephila clavipes* (Supplementary file 4) and *Parasteatoda tepidariorum* (Figure 6; Supplementary file 4) and the mites *Achipteria coleoptrata* and *Sarcoptes scabiei* (Supplementary file 4). Likely due to

interaction depends strictly on the integrity of the five residues that in Hairless have been shown to contact the Su(H) CTD (highlighted in red in Figure 4A).

Given the presence of one or more CtBP recruitment motifs in the Myriapod and Chelicerate S-CAP proteins (Figure 5A), along with the ability of their SBD-like domains to bind Su(H) (Figure 4C), we have designated these as 'Su(H)-Co-repressor Adaptor Proteins' (S-CAPs).

### Chelicerate S-CAP proteins are related to Metastasis-associated (MTA) proteins

In addition to their similarities to Hairless, the S-CAP proteins of Chelicerates include two regions with strong sequence homology to the Metastasis-associated (MTA) protein family, which is highly conserved among Metazoa. The MTA proteins play an important role in transcriptional regulation via their function as core components of the nucleosome remodeling and deacetylase (NuRD) complex (Allen et al., 2013; Burgold et al., 2019). The N-terminal half of MTAs includes four well-defined functional



Figure 4 continued

highlighted in bold. Pairwise amino acid sequence identities within the motifs are indicated by vertical lines; conservative substitutions are indicated by + signs. Amino acids in *Hairless* that have been shown to make direct contact with Su(H) [including the non-contiguous tryptophan (W) residue] (Yuan et al., 2016) are highlighted in red. Hydrophobic residues nearly always found immediately adjacent to the W are underlined. Species names are as follows: Smar (*Strigamia maritima*); Hsub (*Hydroshendyla submarina*) (Fernández et al., 2016); Agir (*Anopsobius giribeti*) (Fernández et al., 2016); Onosp (*Onomeris* sp.) (Rodríguez et al., 2018); SIm (*Sigmoria latior munda*) (Rodríguez et al., 2018); Ptep (*Parasteatoda tepidariorum*); Hruf (*Hypochthonius rufulus*) (Bast et al., 2016); Ssca (*Sarcoptes scabiei*); Vjac (*Varroa jacobsoni*) (Techer et al., 2019); Iper (*Ixodes persulcatus*); Mmar (*Mesobuthus martensii*) (Cao et al., 2013); Lpol (*Limulus polyphemus*). We note that Maier (2019) has previously described the presence of the SBD-like element in the *Strigamia maritima* sequence. (B) SBD motifs in both *Hairless* and S-CAP proteins (red) are encoded in two exons with the same splice junction (indicated by /; see gray highlight). Pairwise amino acid sequence identities within the motifs are indicated by vertical lines; conservative substitutions are indicated by + signs. Species names as in A, except for Isca (*Ixodes scapularis*). (C) Spider S-CAP protein binds directly to *Drosophila* Su(H) in vitro. In all panels, lanes 1–4 represent the indicated His-tagged segments of wild-type *Drosophila* (Dmel) *Hairless* (lane 1); Dmel *Hairless* bearing alanine substitutions for each of five SBD residues shown to contact Su(H) (lane 2); wild-type S-CAP from the spider *Parasteatoda tepidariorum* (Ptep) (lane 3); Ptep S-CAP bearing the same five alanine substitutions (lane 4). Input levels of these His-tagged proteins for each experiment are shown in the respective ‘input’ panels. Remaining two panels show the results of pull-down assays using Sepharose beads bearing only GST (left side) or GST-Su(H) (right side). Left: No binding of the His-tagged proteins to GST alone is observed. Right: Wild-type Dmel *Hairless* binds efficiently to GST-Su(H) (lane 1); this interaction is severely reduced by the introduction of the five alanine substitutions (lane 2). Wild-type Ptep S-CAP likewise binds to GST-Su(H) (lane 3), while no binding is observed with the alanine-substitution mutant (lane 4); the same result is obtained even when the amount of input Ptep S-CAPs (wild-type and mutant) is increased by a factor of 10 (lanes 3 and 4, 10X). Amounts of GST and GST-Su(H) on the beads are shown in the Coomassie stains below the corresponding pull-down lanes. The results shown in this panel have been replicated in two additional experiments, including one utilizing new isolations of GST-Su(H) and His-tagged proteins.

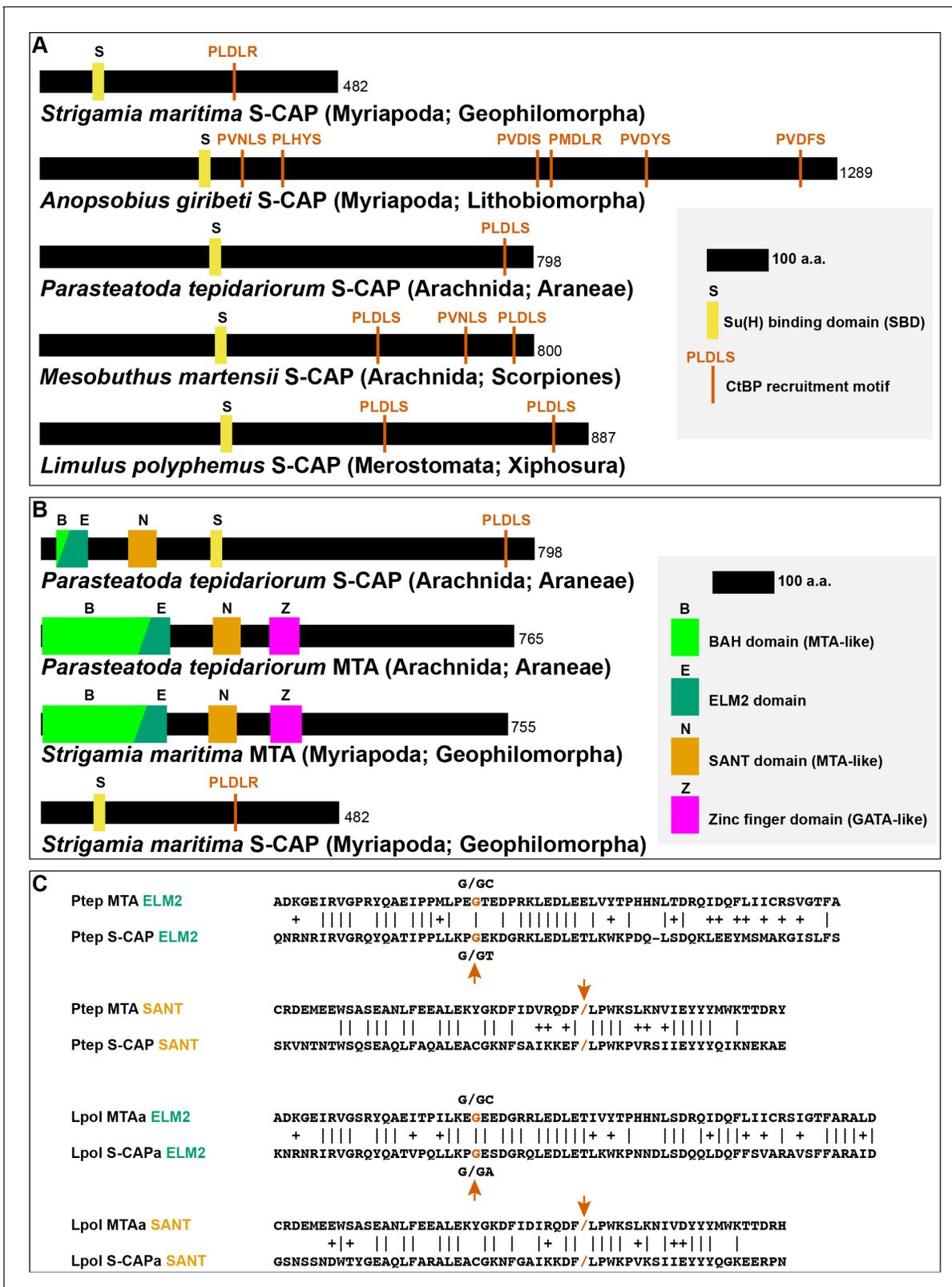
DOI: <https://doi.org/10.7554/eLife.48115.007>

least in part to its history of whole-genome duplication (Nossa et al., 2014; Kenny et al., 2016), the horseshoe crab *Limulus polyphemus* (representing the Merostomata/Xiphosura) has three paralogous copies of this same S-CAP-MTA linkage pairing (Supplementary file 4). Some exceptions to this pattern do exist. In the genomes of the mites *Metaseiulus occidentalis* (Supplementary file 4) and *Varroa destructor* (Techer et al., 2019), for example, the genes encoding S-CAP and MTA are far separated from each other.

Close, typically adjacent, linkage between *Hairless* and MTA genes is likewise widely observed in the genomes of Pancrustacea. Among the Hexapoda, this pattern can be found in many different orders (Supplementary file 1), including Diptera, Lepidoptera, Coleoptera (Figure 6), Hymenoptera (Figure 6), Psocodea, Hemiptera (Figure 6), Thysanoptera, Blattodea, Orthoptera, Odonata, and Collembola. Among the Vericrustacea, adjacent linkage of *Hairless* and MTA is seen in the shrimp *Triops cancriformis* (Notostraca) (Supplementary file 1). Nevertheless, exceptions are readily found, even within the same orders as above (Supplementary file 1). Examples include *Drosophila melanogaster*, *Ceratitis capitata*, and *Lucilia cuprina* (Diptera; Supplementary file 1), *Bicyclus anynana* (Lepidoptera), *Anoplophora glabripennis*, *Dendroctonus ponderosae*, and *Nicrophorus vespilloides* (Coleoptera), and *Cimex lectularius* (Hemiptera; Supplementary file 1).

Interestingly, in some instances *Hairless*/MTA microsynteny is preserved, but the genes' relative orientation is different (Figure 6; Supplementary file 1). Thus, in the aphids — in contrast to other Hemiptera — MTA lies downstream of *Hairless*, but in the opposite orientation (Figure 6). In the beetle *Harmonia axyridis* (Coleoptera), MTA lies upstream of *Hairless* (Figure 6).

Despite the multiple instances in which it has been lost, we believe that the most parsimonious interpretation of our analysis is that close linkage between MTA and S-CAP/*Hairless* genes is ancestral in the respective taxa (Myriapods/Chelicerates and Pancrustacea). We leave for the Discussion our proposed interpretation of the evolutionary significance of this adjacency.



**Figure 5.** Sequence characteristics of S-CAP proteins in Myriapods and Chelicerates. (A) Diagrams of representative examples of Myriapod and Chelicerate S-CAP proteins, denoting locations of SBD motifs and CtBP recruitment motifs. Scale and protein sizes are indicated. (B) Chelicerate, but not Myriapod, S-CAP proteins share N-terminal ELM2 and SANT domains with an MTA zinc-finger protein from the same species. Scale and protein sizes are indicated. (C) Shared ELM2 and SANT domains in Chelicerate MTA and S-CAP proteins are encoded in two exons with the same splice

Figure 5 continued on next page

Figure 5 continued

junction (indicated by /; red arrows). Pairwise amino acid sequence identities within the motifs are indicated by vertical lines; conservative substitutions are indicated by + signs. Species names as in **Figure 4A**.

DOI: <https://doi.org/10.7554/eLife.48115.008>

The following figure supplement is available for figure 5:

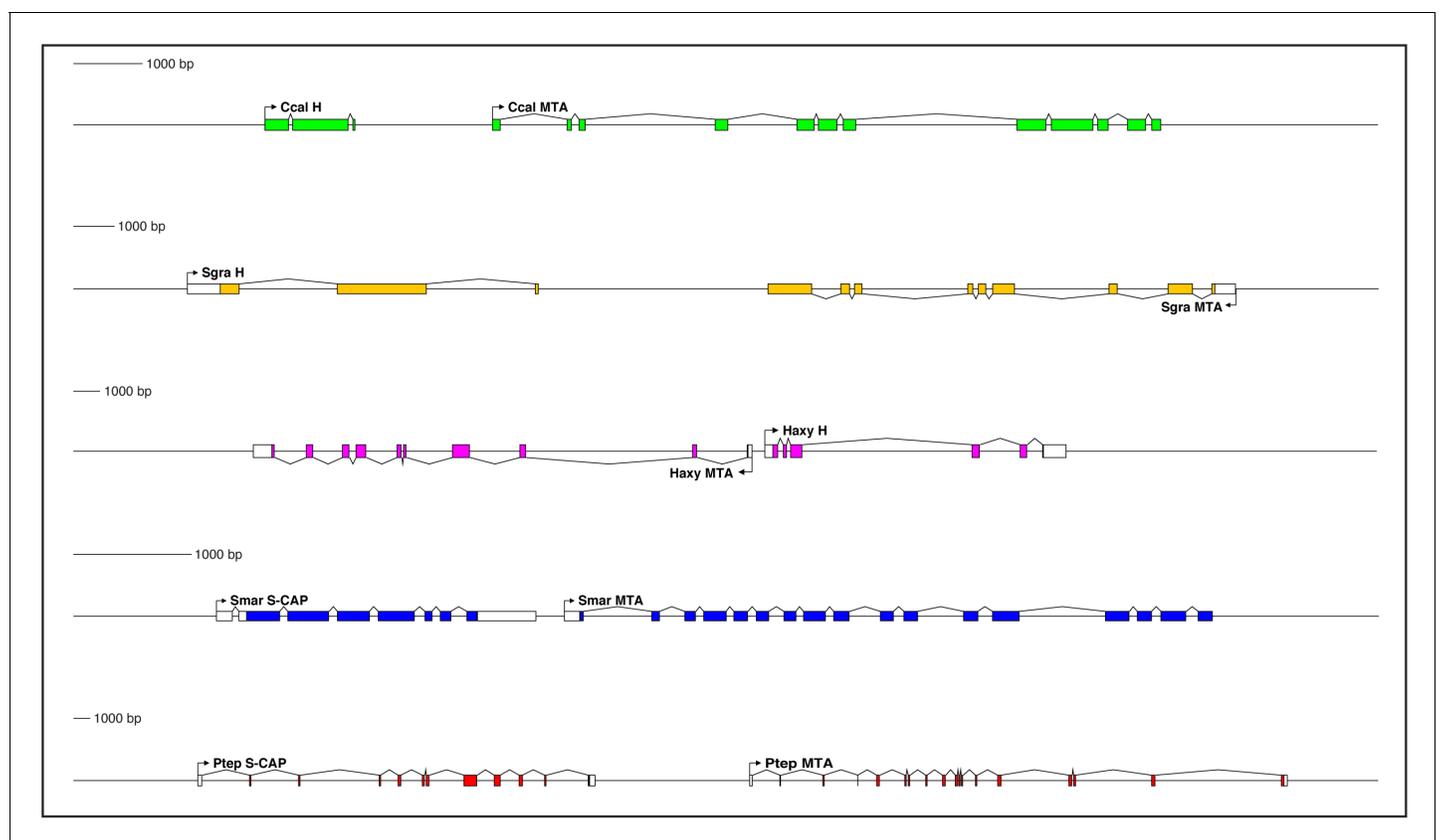
**Figure supplement 1.** Alignments of sequence regions shared by representative S-CAP proteins from (A) Chelicerates and (B) Myriapods.

DOI: <https://doi.org/10.7554/eLife.48115.009>

## Discussion

### The evolution of Hairless represents a shift from the ancestral and dominant paradigm of direct co-repressor recruitment by Su(H)

Our analysis of sequences from a broad range of protostomes strongly suggests that direct recruitment of the CtBP and Gro co-repressors by Su(H) is ancestral in this clade. This is consonant with the fact that direct co-repressor recruitment by DNA-binding repressor proteins in general is a dominant paradigm among Metazoa. This evokes the intriguing question of what might have led to the loss of direct recruitment by Su(H) in the Neoptera (see **Figure 1B**) and its replacement by Hairless-mediated indirect recruitment? Does Hairless provide some advantageous functional capacity? Note that this is not intended to suggest that Hairless must be an evolutionary adaptation per se (**Lynch, 2007**);



**Figure 6.** Genes encoding both Hairless and S-CAP proteins are frequently located immediately adjacent to an MTA gene. Separate scale for each diagram is shown at the left. Three examples are shown for *Hairless*: the carpenter bee *Ceratina calcarata* (Ccal), the wheat aphid *Schizaphis graminum* (Sgra) (QEWZ01001380.1), and the lady beetle *Harmonia axyridis* (Haxy). Note that microsynteny is often preserved even when gene locations and relative orientations are changed. One example each is shown for S-CAP in Myriapods [the centipede *Strigamia maritima* (Smar)] and Chelicerates [the house spider *Parasteatoda tepidariorum* (Ptep)]. See also **Supplementary file 1** and **Supplementary file 4**.

DOI: <https://doi.org/10.7554/eLife.48115.010>

rather, we are asking: What capability might it have conferred that would lead to its retention and the subsequent loss of the recruitment motifs in Su(H)?

One appealing (but of course speculative) possibility is that Hairless may have permitted Su(H) for the first time to recruit both CtBP and Gro simultaneously to the same target genes. As we have noted, the apparently ancestral P $\phi$ D $\phi$ S and G $\phi$ LTPDKV motifs in protostome Su(H) typically lie quite close to each other in the protein's linear sequence (**Supplementary file 3**). CtBP (~400 aa) and Gro (~700 aa) are both large proteins that engage in oligomerization as part of their functional mechanism (**Song et al., 2004; Bhambhani et al., 2011**). It is very unlikely that both could bind stably to DNA-bound Su(H) at the same time. In contrast, the Gro and CtBP recruitment motifs in Hairless are far apart in the linear sequence (**Figure 1A**) and are separated by a region predicted to be largely disordered (**Figure 1—figure supplement 1**). We suggest that this might be compatible with simultaneous recruitment of the two co-repressors.

Whatever may have been the selective forces that led to the loss of direct co-repressor recruitment by Su(H) in the Neoptera and its replacement by Hairless-mediated indirect recruitment, Hairless is a notable evolutionary novelty for having permitted the unusual abandonment of an ancestral and highly conserved paradigm. We suggest that this represents a striking example of 'developmental system drift' (**True and Haag, 2001**), in which a common output (widespread 'default repression' of Notch pathway target genes) is achieved via distinct molecular mechanisms in different species.

### A possible evolutionary pathway for the appearance of Hairless

We have described here several findings that we believe have important implications for an attempt to reconstruct the history of Hairless as an evolutionary novelty. First, we observe that Hairless is apparently confined to the Pancrustacea, wherein it is widely distributed among diverse taxa (**Figure 2A; Supplementary file 1**). Second, we have discovered in the sister groups Myriapoda and Chelicerata a protein (S-CAP) with clear sequence homology to the Su(H)-binding domain (SBD) of Hairless (**Figure 4A**). Significantly, in both Hairless and the S-CAPs these motifs are encoded by contributions from two exons, with the associated splice junction in precisely the same location (**Figure 4B; Supplementary file 4**). Third, we find that S-CAPs in the Chelicerata include in their N-terminal region strong homology to the ELM2 and SANT domains of MTAs, which themselves are highly conserved among Metazoa, and therefore would have been present in the arthropod common ancestor (**Figure 5B,C**). Finally, our analysis indicates that close, usually adjacent, linkage of *Hairless* and *MTA* genes (in the Pancrustacea) and between *S-CAP* and *MTA* genes (in the Myriapoda and Chelicerata) is widespread (**Figure 6; Supplementary file 1; Supplementary file 4**), and hence very likely to be ancestral, in these taxa.

While any attempt to infer the sequence of evolutionary events that led to the appearance of Hairless is necessarily speculative, we believe that the above findings offer substantial support for the following hypothetical pathway. We propose that in a deep arthropod ancestor a tandem duplication of the *MTA* gene occurred. One copy retained the strong sequence conservation (and presumably ancestral function) of metazoan *MTA* genes, while the second copy diverged very substantially, eventually encoding a protein that had lost all but the ELM2 and SANT domains of the *MTA* ancestor. The extensive reconfiguration of this paralog also included the eventual acquisition of the SBD motif and the addition of one or more CtBP recruitment motifs (see **Figure 7** for some possible sources of these components). In the Myriapod lineage, even the ELM2 and SANT domains were eventually lost. In the Pancrustacea, we suggest that this same divergent *MTA* paralog evolved to become *Hairless*. Beyond the alterations described for the Myriapoda, this would have involved the acquisition of sequences encoding additional now-conserved domains and motifs, including the Gro recruitment motif (**Supplementary file 2**). This radical evolutionary transformation resulted in a protein with little or no remaining homology to its *MTA* ancestor, and with an entirely novel regulatory function (**Holland et al., 2017**).

In this context, it is of interest that the *Drosophila* Mi-2/Nurd complex — which includes the *MTA* protein — has recently been shown to engage in direct repression of multiple Notch pathway target genes, independent of both Su(H) and Hairless (**Zacharioudaki et al., 2019**). Whether this activity preceded the emergence of Hairless is unknown, but the possibility that it is in some way connected to Hairless's evolutionary history is indeed intriguing.

## Materials and methods

## Key resources table

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Antibody	anti-HIS G antibody, mouse monoclonal	Invitrogen	CAT#46–1008 (now ThermoFisher CAT#R940-25), RRID:AB_2556557	1:5000 dilution
Antibody	GOAT anti-mouse HRP, polyclonal	Jackson Immuno-research	CAT#115-035-003, RRID:AB_10015289	1:10000 dilution
Recombinant DNA reagent	GST-dCtBP pGEX-5X-3 clone	<b>Nibu et al., 1998</b>		Construct encoding GST-tagged <i>Drosophila</i> CtBP for expression in <i>E. coli</i>
Recombinant DNA reagent	GST-Gro pGEX-KG clone	This paper		Construct encoding GST-tagged <i>Drosophila</i> Groucho for expression in <i>E. coli</i>
Recombinant DNA reagent	GST-Su(H) pGEX-KG clone	<b>Bailey and Posakony, 1995</b>		Construct encoding GST-tagged <i>Drosophila</i> Su(H) for expression in <i>E. coli</i>
Recombinant DNA reagent	HIS-H192-389 WT pRSET-C clone	This paper		HIS-tagged expression construct encoding amino acids 192–389 of <i>Drosophila</i> Hairless, synthesized by GeneWiz, Inc, and codon-optimized for expression in <i>E. coli</i>
Recombinant DNA reagent	HIS-H192-389 5AMUT pRSET-C clone	This paper		HIS-tagged expression construct encoding amino acids 192–389 of <i>Drosophila</i> Hairless with five alanine substitutions, synthesized by GeneWiz, Inc, and codon-optimized for expression in <i>E. coli</i>
Recombinant DNA reagent	HIS-HLHmBetaSmar WT pRSET-C clone	This paper		HIS-tagged expression construct encoding <i>Drosophila</i> HLHmBeta with the last four amino acids (WRPW) replaced with the sequence GSLTPDKV
Recombinant DNA reagent	HIS-HLHmBetaSmar MUT pRSET-C clone	This paper		HIS-tagged expression construct encoding <i>Drosophila</i> HLHmBeta with the last four amino acids (WRPW) replaced with the sequence AAAAAAAAAA
Recombinant DNA reagent	HIS-HLHmBetaWT pRSET-C clone	This paper		HIS-tagged expression construct encoding full-length <i>Drosophila</i> HLHmBeta
Recombinant DNA reagent	HIS-HLHmBetatrunc pRSET-C clone	This paper		HIS-tagged expression construct encoding <i>Drosophila</i> HLHmBeta with the last four amino acids (WRPW) deleted

Continued on next page

Continued

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Recombinant DNA reagent	HIS-PtepSCAP233-432 WT pRSET-C clone	This paper		HIS-tagged expression construct encoding amino acids 233–432 of <i>Parasteatoda tepidariorum</i> S-CAP, synthesized by GeneWiz, Inc, and codon-optimized for expression in <i>E. coli</i>
Recombinant DNA reagent	HIS-PtepSCAP233-432 5AMUT pRSET-C clone	This paper		HIS-tagged expression construct encoding amino acids 233–432 of <i>Parasteatoda tepidariorum</i> S-CAP with five alanine substitutions, synthesized by GeneWiz, Inc, and codon-optimized for expression in <i>E. coli</i>
Recombinant DNA reagent	HIS-SmarSu(H)ex2-3 WT pRSET-C clone	This paper		HIS-tagged expression construct containing exons 2–3 of <i>Strigamia maritima</i> Su(H), synthesized by GeneWiz, Inc, and codon-optimized for expression in <i>E. coli</i>
Recombinant DNA reagent	HIS-SmarSu(H)ex2-3 mut pRSET-C clone	This paper		HIS-tagged expression construct containing exons 2–3 of <i>Strigamia maritima</i> Su(H) with a PVDSL > AAAAAA mutation, synthesized by GeneWiz, Inc, and codon-optimized for expression in <i>E. coli</i>
Recombinant DNA reagent	pGEX-5X-3	Sigma (formerly Amersham; discontinued)	CAT#28-9545-53	
Recombinant DNA reagent	pRSET-C	Invitrogen	CAT#V35120	
Commercial assay or kit	Chem Illumination Reagents	Pierce ECL Western Blotting Substrate	CAT#32209	
Resource, sequence database	NCBI	NCBI	RRID:SCR_006472	
Software, algorithm	NCBI BLAST	NCBI	RRID:SCR_004870	
Software, algorithm	GenePalette	<a href="http://www.genepalette.org">Smith et al., 2017; Rebeiz and Posakony, 2004; http://www.genepalette.org</a>		
Software, algorithm	DNA Strider	<a href="#">Marck, 1988; Douglas, 1995</a>		
Software, algorithm	BlastStation-Local64	TM Software, Inc		

### Sequence searches, analysis, and annotation

Genome and transcriptome sequences encoding Hairless, Suppressor of Hairless, S-CAP, and MTA proteins from a wide variety of species were recovered via BLAST searches, using either the online version at the NCBI website ([Boratyn et al., 2013](#)) or the version implemented by the BlastStation-Local64 desktop application (TM Software, Inc). Sequences were analyzed and annotated using the GenePalette ([Rebeiz and Posakony, 2004; Smith et al., 2017](#)) and DNA Strider ([Marck, 1988; Douglas, 1995](#)) desktop software tools. Analysis of predicted disordered regions in Hairless was

conducted using DISOPRED3 on the PSIPRED server (Buchan et al., 2013; Jones and Cozzetto, 2015).

## Generation of constructs for GST pulldown experiments

### *Strigamia maritima* Su(H) protein constructs to test CtBP binding

A codon-optimized fragment corresponding to exons 2 and 3 from *S. maritima* Su(H) mRNA was synthesized by Genewiz, Inc, and cloned into pRSET-C using Acc65I and BamHI restriction sites. The CtBP-motif mutant was subsequently generated by overlap extension PCR using the primers HISsmarSUH-f (CGCTGGATCCGCGGCCAGTATGAC), HISsmarSUH-r (CCATGGTACCAGTTATGCGTGGTG), HISsmarSUHctbpm-f (AACCACgCCGcaGcTGcGgCTAACAGCCATCGCGGTGAAGGCGGCCAC), HISsmarSUHctbpm-r (GCTGTTAGcCgCAgCtgCGGcGTGGTTGTGGCGGAAGTGAGGGTTCAG). After sequence confirmation, this fragment was also cloned into pRSET-C using the same enzymes. Binding of these constructs to *Drosophila melanogaster* CtBP was assayed using GST alone and a GST-CtBP fusion protein (Nibu et al., 1998).

### Constructs to test potential Gro-binding motif in *Strigamia maritima* Su(H)

A truncated version of HLHm $\beta$  (HLHm $\beta$ -WRPWtrunc) was amplified from a pRSET-HLHm $\beta$ -WT construct using the primers HISmbeta-f (cgatggatccgaATGGTTCTGGAAATGGAGATGTCCAAG) and HISmbetrunc-r (ccatggatccagTCACATGGGGCCagaggtggagctggcctcgctggcg); a version of HLHm $\beta$  with the WRPW motif replaced with the amino acids GSLTPDKV (HLHm $\beta$ +Smar-motifWT) was amplified from the WT construct with HISmbeta-f and mbetaSmarSuH-r (ccatggatccagTCACAC TTTATCAGGTGGAGTGAGAGAACCCATGGGGCCagaggtggagctggc); and a version of HLHm $\beta$  with the WRPW motif replaced with a stretch of 9 alanine residues (HLHm $\beta$ +Smar-motifMUT) was amplified using HISmbeta-f and mbetaSmarSuHmut-r (ccatggatccagTCA ggctgccgctgaggctgccgctgctgcCATGGGGCCagaggtggagctggc). Each construct was then subsequently cloned into pRSET-C using the restriction enzymes BamHI and Acc65I and sequence verified. Binding of these constructs to *Drosophila melanogaster* Gro was assayed using GST alone and a GST-Gro fusion protein. The latter construct was made by cloning the full-length Gro coding sequence into the pGEX-KG expression vector at the EcoRI and Sall restriction sites: gtggcgac-catcctccaaaatcggatctggttccgctggatccccgggaatttccggtggtggtggaattctaATG...TAAATCCA-CAAAACCATGCAGTTTTTTCATTTTGTAAATAAGCTCGTATAGTTTTTATTACAACATGTTTCGAAATCA TGCaccgggctgcaggaattcgatcaagcttatcgataccgtcgactcgactcaagcttaattcatctgactgactgacgatctg (underlined = pGEX KG vector; uppercase = *gro* cDNA; bold = *gro* start and stop codons; italic = linker)

### S-CAP/Hairless constructs for Su(H) interaction analysis

Codon-optimized fragments from *Drosophila melanogaster* Hairless (residues 192–389), and *Parasteatoda tepidariorum* cS-CAP (residues 233–432) as well as 5-alanine mutant substitutions (Dmel: **GGRLOFFKDGKFILELARSKDGDKSGW** - > **GGRAQAFKDGKFI~~E~~AARSKDGDKSGA**; Ptep: **VG SLKFFLGGRLV~~L~~KLNAQQDGGSGNKCO~~W~~** - > **VGSAKAF~~L~~GGRLVAKANAQQDGGSGNKCO~~A~~**) were synthesized by Genewiz, Inc. Inserts were subsequently cloned into pRSET-C using the restriction enzymes BamHI and Acc65I. Binding of these constructs to *Drosophila melanogaster* Su(H) was assayed using GST alone and a GST-Su(H) fusion protein (Bailey and Posakony, 1995).

GST pulldowns using each of the above constructs were performed as previously described (Fontana and Posakony, 2009).

## Synthesized, codon-optimized sequences

>Smar Su(H)ex2-3 WT (116 aa)

CGCTGGATCCGCGGCCAGTATGACTACCCGCCCGTTAGCCAGCACATACAGCCGCGAGGCC-GACCTGTGGAACGTGAACCTGGCCACCTACAGCAGCGCACCGACCACATGCACCGG TGCAACCCCGGCACCTAGCGTTACCGGTTTCTACGCCAGGCCACCGGCAGCAACAGCG TTAGCCCGAGTAGCGTGAGCCTGACCACCCTGACCCTCACTTCGCCGACAACCACCCGG TGGACCTGAGCAACAGCCATCGCGGTGAAGGCGGCCACCTGGATCTGGTGCCTTCCAGAGC-GACCGCGTGGATGCCTACAAGCACGCCAACGGCCTGAGCGTGCATATCCCGGACCACCACGCA TAACTGGTACCATGG



TGAAAACAGCACCTCTTTAAGCTTCAGCGACGACAATAGCAGCATTAGAGCAGCCCG  
 TGGCAGCGTGATCAGCCGTGGAAACAGAGTCGTCCGCGCCGTGGCATCAGCAAAGAACTGTG  
 TTTATTTTTCCACCGCCCGCGCAATAGTACTGGTCTGTCAGCCTTACGTACCGCAGCCCG-  
 CAAACGTCGTCCGCATGAACCGCTGACCACCAGCGAAGATCAGCAGCCGATC  
 TTTGCCACCGCAATCAAAGCCGAGAACGGTGATGATACTTTAAAAGCCGAAGCAGCCGAATAAC  
 TGGTACCATGG

>Dmel Hairless192-389 5Amut

CGATGGATCCGAGCCGTTGTGGCAGCAGCAGCTGGCACTGCCAAAATCGGCAAAGGCAGCAA  
 TAGCGGTGGTAGCTTTGACATGGGCCGACCCCGATTAGCACCCATGGCAACAACAGCTGGGG  
 TGGTTATGGTGGTCTGTGCCAAGCTTTAAAGACGGCAAGTTCATCGCCGAAGCCGCACGCAG-  
 CAAAGATGGCGACAAAAGCGGTGCCGTGAGCGTGACCCGCAAACCTTTTCGTCCGCCGAG  
 TGCAGCAACCAGCGCAACCGTTACCCCGACCAGCGCAGTTACCACCGCCTACCCGAAAAAC-  
 GAAAAACAGCACCTCTTTAAGCTTTAGCGACGACAACAGCAGCATTAGAGCAGCCCG  
 TGGCAGCGGATCAGCCGTGGAAACAGAGCCGTCTCGTCGCGGCATCAGCAAAGAGCTGTC  
 TTTATCTTTTCATCGCCCGCGCAATAGCACTTTAGGTCTGTCAGCACTGCGCACAGCAGCAGC  
 TAAACGTCGTCCCGCATGAACCGCTGACCACCAGCGAAGACCAGCAGCCGA  
 TTTTGGCACCAGCAATCAAAGCCGAGAACGGCGATGATACTTTAAAAGCAGAAGCAGCCGAA  
 TAACTGGTACCATGG

>Ptep s-CAP233-432 WT

CGATGGATCCGAACCGTGAATACCGAAGATCCGCCGAAGGATAGCATCAACTTTCTGGACCA-  
 CAGCCGCGTGACCGATCCGTGTAGTGCCGCAAGCGAAACCAGCCTGCCGAGGATG  
 TGCCGGCAAACAAGCACCGTGGGCAGCCTGAAATTTTTCTGGGCGGTGCGCTGGTGCTGAAA  
 TTAACGCCAGCAGGATGGCGGCAGCGCAATAAATGCCAGTGGGTGCAGAGCAACGATC  
 TGCCGAAACATAGCAACCATAACAAAAAGATAAACATAAGAAAAAATTTGCACCGTATAGCTA  
 TAGCAGCAGCGGCACTCAGAAACCGCTGAAGAAAGGCGACGATACCAGTGCCGTGCCGGACTG  
 TGATCCGAGCGGCATCAAAAAGCCGCGCCTGAAAGAGTACGAGACCAGCGAGAATAGCGCCC  
 TGGGTCTGCTGTGTGCAGCAGCAGTTGGACCCCGCCGTTGCAGATGGTTCAGGAGAGCA  
 TTGACGTGGACGATACCAGCAGCAAAACCAGCGAGGGCTATATTAGCCCGATCCTGAGCAACAA  
 TAGCCGCACCAGCAAAATCGACACCATCAAGCACGATTTTGGCAGCAACCCGAACACCTAAC  
 TGGTACCATGG

>Ptep s-CAP233-432 5Amut

CGATGGATCCGAACCGTGAACACCGAAGACCCGCCGAAAGATAGCATCAACTTTTTAGACCA  
 TAGCCGCGTGACAGACCCGTGCAGTGCCGCAAGTGAAACCTTTTACCGCAAGATG  
 TGCCGGCAAACCAGCACCGTGGGTAGCGCCAAAGCCTTTCTGGGCGGTCTGTCTGG  
 TGGCCAAAGCCAATGCCAGCAAGATGGTGGTAGTGGTAACAAATGCCAAGCTGTGCAGAG-  
 CAACGATCTGCCGAAACACAGCAATCACAATAAGAAAGACAAACACAAGAAAAAATTTGCCCGG  
 TATAGCTATAGCAGCAGCGGCACCCAGAAACCGCTGAAAAAGGCGATGACACCAGCGCAG  
 TGCCGGATTGCGATCCGAGCGGCATTAAGAAACCGCGTTTAAAGGAGTACGAGACCAGC-  
 GAAAACAGTGCTTTAGGTTTACTGCTGTGCAGCAGCAGTTGGACACCGCCGGTGGCCGATGG  
 TCAAGAAAGTATCGATGTGGACGACACCAGCAGCAAAACCAGCGAAGGCTACATCAGCCCGA  
 TTCTGAGCAACAATAGCCGCACCAGCAAAATTGATACCATTAAACATGATTTTGAAGCAA  
 TCCGAATACCTAACTGGTACCATGG

## Acknowledgements

We are especially grateful to our many colleagues who made their genome and transcriptome sequence assemblies freely available; without their generosity, this study would not have been possible. We thank Scott Rifkin and Mark Rebeiz for helpful discussion and input during the preparation of the manuscript. We also thank the following artists for making available the illustrations shown in **Figure 2A**: crab, by Firkin (<https://openclipart.org/detail/270221/crab-silhouette>); monarch butterfly, by carolemagnet (<https://openclipart.org/detail/263384/monarch-butterfly>); centipede, by Firkin (<https://openclipart.org/detail/261126/centipede-3>); spider, by liftarn (<https://openclipart.org/detail/>)

179190/spider); horseshoe crab, by Gosc (<https://openclipart.org/detail/174556/horseshoe-crab>); tick, by Juhele (<https://openclipart.org/detail/279073/simple-tick-ixodes-ricinus-silhouette>). This work was supported by NIH Grants R01GM046993 and R01GM120377 (to JWP).

---

## Additional information

### Funding

Funder	Grant reference number	Author
National Institute of General Medical Sciences	R01GM046993	James W Posakony
National Institute of General Medical Sciences	R01GM120377	James W Posakony

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

### Author contributions

Steven W Miller, Supervision, Investigation, Visualization, Methodology, Writing—original draft, Writing—review and editing; Artem Movsesyan, Sui Zhang, Investigation, Writing—original draft, Writing—review and editing; Rosa Fernández, Resources, Investigation, Writing—review and editing; James W Posakony, Conceptualization, Data curation, Supervision, Funding acquisition, Investigation, Methodology, Writing—original draft, Writing—review and editing

### Author ORCIDs

Steven W Miller  <https://orcid.org/0000-0001-7610-6336>

James W Posakony  <https://orcid.org/0000-0001-6377-1732>

### Decision letter and Author response

Decision letter <https://doi.org/10.7554/eLife.48115.020>

Author response <https://doi.org/10.7554/eLife.48115.021>

---

## Additional files

### Supplementary files

- Supplementary file 1. Representative catalog of Hairless proteins in the Pancrustacea, selected from a curated collection of approximately 400 full-length sequences. Annotated sequences of the entries in this list are provided in **Supplementary file 2**. In the ‘MTA microsynteny?’ column, + and – in parentheses denote the relative orientation of the *Hairless* and *MTA* genes in the genome. References not in the main reference list are provided in **Supplementary file 6**.

DOI: <https://doi.org/10.7554/eLife.48115.012>

- Supplementary file 2. Annotated FASTA file of sequences of the Hairless proteins included in **Supplementary file 1**. Characteristic conserved domains and motifs are colored as in **Figure 1A**.

DOI: <https://doi.org/10.7554/eLife.48115.013>

- Supplementary file 3. Annotated FASTA file of sequences of the Su(H) proteins included in **Table 1**; species for which a full-length Su(H) sequence is not available are omitted here. Colors denote conserved sequence features. Motifs for direct recruitment of the CtBP and Gro co-repressor proteins (aligned in **Table 1**) are shown in red and green, respectively. Large region highlighted in orange is the highly conserved body of Su(H), extending from ‘LTREAM’ to ‘YTPEP’.

DOI: <https://doi.org/10.7554/eLife.48115.014>

- Supplementary file 4. Representative catalog of S-CAP proteins in the Myriapoda and Chelicerata, selected from a curated collection of approximately 50 full-length sequences. Annotated sequences of the entries in this list are provided in **Supplementary file 5**. In the ‘MTA microsynteny?’ column, + and – in parentheses denote the relative orientation of the *S-CAP* and *MTA* genes in the genome. The ‘H SBD splice?’ column indicates whether the Su(H)-binding domain (SBD) in the listed protein is

encoded by two exons with the same splice junction as in *Hairless*. If not, the alternative exon structure is indicated. References not in the main reference list are provided in **Supplementary file 6**.

DOI: <https://doi.org/10.7554/eLife.48115.015>

- Supplementary file 5. Annotated FASTA file of sequences of the S-CAP proteins included in **Supplementary file 4**. Characteristic domains and motifs are colored as follows: Su(H)-binding domain (SBD), orange; CtBP recruitment motifs, red.

DOI: <https://doi.org/10.7554/eLife.48115.016>

- Supplementary file 6. Sequence data references not cited in the main paper text.

DOI: <https://doi.org/10.7554/eLife.48115.017>

- Transparent reporting form

DOI: <https://doi.org/10.7554/eLife.48115.018>

### Data availability

All data generated or analysed during this study are included in the manuscript and supporting files.

## References

- Albertin CB**, Simakov O, Mitros T, Wang ZY, Pungor JR, Edsinger-Gonzales E, Brenner S, Ragsdale CW, Rokhsar DS. 2015. The octopus genome and the evolution of cephalopod neural and morphological novelties. *Nature* **524**:220–224. DOI: <https://doi.org/10.1038/nature14668>, PMID: 26268193
- Allen HF**, Wade PA, Kutateladze TG. 2013. The NuRD architecture. *Cellular and Molecular Life Sciences* **70**: 3513–3524. DOI: <https://doi.org/10.1007/s00018-012-1256-2>
- Andrade SC**, Novo M, Kawauchi GY, Worsaae K, Pleijel F, Giribet G, Rouse GW. 2015. Articulating "archannelids": phylogenomics and annelid relationships, with emphasis on meiofaunal taxa. *Molecular Biology and Evolution* **32**:2860–2875. DOI: <https://doi.org/10.1093/molbev/msv157>, PMID: 26205969
- Bailey AM**, Posakony JW. 1995. Suppressor of *Hairless* directly activates transcription of *Enhancer of split* complex genes in response to Notch receptor activity. *Genes & Development* **9**:2609–2622. DOI: <https://doi.org/10.1101/gad.9.21.2609>
- Baker CR**, Booth LN, Sorrells TR, Johnson AD. 2012. Protein modularity, cooperative binding, and hybrid regulatory states underlie transcriptional network diversification. *Cell* **151**:80–95. DOI: <https://doi.org/10.1016/j.cell.2012.08.018>, PMID: 23021217
- Baldwin-Brown JG**, Weeks SC, Long AD. 2018. A new standard for crustacean genomes: the highly contiguous, annotated genome assembly of the clam shrimp *Eulimnadia texana* reveals HOX gene order and identifies the sex chromosome. *Genome Biology and Evolution* **10**:143–156. DOI: <https://doi.org/10.1093/gbe/evx280>, PMID: 29294012
- Bang AG**, Posakony JW. 1992. The *Drosophila* gene *Hairless* encodes a novel basic protein that controls alternative cell fates in adult sensory organ development. *Genes & Development* **6**:1752–1769. DOI: <https://doi.org/10.1101/gad.6.9.1752>, PMID: 1516831
- Barolo S**, Stone T, Bang AG, Posakony JW. 2002a. Default repression and Notch signaling: *Hairless* acts as an adaptor to recruit the corepressors Groucho and dCtBP to Suppressor of *Hairless*. *Genes & Development* **16**: 1964–1976. DOI: <https://doi.org/10.1101/gad.987402>, PMID: 12154126
- Barolo S**, Posakony JW. 2002b. Three habits of highly effective signaling pathways: principles of transcriptional control by developmental cell signaling. *Genes & Development* **16**:1167–1181. DOI: <https://doi.org/10.1101/gad.976502>, PMID: 12023297
- Bast J**, Schaefer I, Schwander T, Maraun M, Scheu S, Kraaijeveld K. 2016. No accumulation of transposable elements in asexual arthropods. *Molecular Biology and Evolution* **33**:697–706. DOI: <https://doi.org/10.1093/molbev/msv261>, PMID: 26560353
- Battelle BA**, Ryan JF, Kempler KE, Saraf SR, Marten CE, Warren WC, Minx PJ, Montague MJ, Green PJ, Schmidt SA, Fulton L, Patel NH, Protas ME, Wilson RK, Porter ML. 2016. Opsin repertoire and expression patterns in horseshoe crabs: evidence from the genome of *Limulus polyphemus* (Arthropoda: Chelicerata). *Genome Biology and Evolution* **8**:1571–1589. DOI: <https://doi.org/10.1093/gbe/eww100>, PMID: 27189985
- Bhambhani C**, Chang JL, Akey DL, Cadigan KM. 2011. The oligomeric state of CtBP determines its role as a transcriptional co-activator and co-repressor of Wingless targets. *The EMBO Journal* **30**:2031–2043. DOI: <https://doi.org/10.1038/emboj.2011.100>, PMID: 21468031
- Boratyn GM**, Camacho C, Cooper PS, Coulouris G, Fong A, Ma N, Madden TL, Matten WT, McGinnis SD, Merezuk Y, Raytselis Y, Sayers EW, Tao T, Ye J, Zaretskaya I. 2013. BLAST: a more efficient report with usability improvements. *Nucleic Acids Research* **41**:W29–W33. DOI: <https://doi.org/10.1093/nar/gkt282>, PMID: 23609542
- Brauchle M**, Bilican A, Eyer C, Bailly X, Martínez P, Ladurner P, Bruggmann R, Sprecher SG. 2018. Xenacoelomorpha survey reveals that all 11 animal homeobox gene classes were present in the first bilaterians. *Genome Biology and Evolution* **10**:2205–2217. DOI: <https://doi.org/10.1093/gbe/evy170>, PMID: 30102357

- Buchan DW**, Minneci F, Nugent TC, Bryson K, Jones DT. 2013. Scalable web services for the PSIPRED protein analysis workbench. *Nucleic Acids Research* **41**:W349–W357. DOI: <https://doi.org/10.1093/nar/gkt381>, PMID: 23748958
- Burgold T**, Barber M, Kloet S, Cramard J, Gharbi S, Floyd R, Kinoshita M, Ralsler M, Vermeulen M, Reynolds N, Dietmann S, Hendrich B. 2019. The nucleosome remodelling and deacetylation complex suppresses transcriptional noise during lineage commitment. *The EMBO Journal* **38**:e100788. DOI: <https://doi.org/10.15252/embj.2018100788>, PMID: 31036553
- Cao Z**, Yu Y, Wu Y, Hao P, Di Z, He Y, Chen Z, Yang W, Shen Z, He X, Sheng J, Xu X, Pan B, Feng J, Yang X, Hong W, Zhao W, Li Z, Huang K, Li T, et al. 2013. The genome of *Mesobuthus martensii* reveals a unique adaptation model of arthropods. *Nature Communications* **4**:2602. DOI: <https://doi.org/10.1038/ncomms3602>, PMID: 24129506
- Chipman AD**, Ferrier DE, Brena C, Qu J, Hughes DS, Schröder R, Torres-Oliva M, Znassi N, Jiang H, Almeida FC, Alonso CR, Apostolou Z, Aqrabi P, Arthur W, Barna JC, Blankenburg KP, Brites D, Capella-Gutiérrez S, Coyle M, Dearden PK, et al. 2014. The first myriapod genome sequence reveals conservative arthropod gene content and genome organisation in the centipede *Strigamia maritima*. *PLOS Biology* **12**:e1002005. DOI: <https://doi.org/10.1371/journal.pbio.1002005>, PMID: 25423365
- Clarke TH**, Garb JE, Hayashi CY, Arensburger P, Ayoub NA. 2015. Spider transcriptomes identify ancient large-scale gene duplication event potentially important in silk gland evolution. *Genome Biology and Evolution* **7**:1856–1870. DOI: <https://doi.org/10.1093/gbe/evv110>, PMID: 26058392
- Douglas SE**. 1995. DNA strider. An inexpensive sequence analysis package for the Macintosh. *Molecular Biotechnology* **3**:37–45. DOI: <https://doi.org/10.1007/BF02821333>, PMID: 7606503
- Egekwa N**, Sonenshine DE, Garman H, Barshis DJ, Cox N, Bissinger BW, Zhu J, M Roe R. 2016. Comparison of synganglion neuropeptides, neuropeptide receptors and neurotransmitter receptors and their gene expression in response to feeding in *Ixodes scapularis* (Ixodidae) vs. *Ornithodoros turicata* (Argasidae). *Insect Molecular Biology* **25**:72–92. DOI: <https://doi.org/10.1111/imb.12202>, PMID: 26783017
- Fernández R**, Edgecombe GD, Giribet G. 2016. Exploring phylogenetic relationships within Myriapoda and the effects of matrix composition and occupancy on phylogenomic reconstruction. *Systematic Biology* **65**:871–889. DOI: <https://doi.org/10.1093/sysbio/syw041>, PMID: 27162151
- Fontana JR**, Posakony JW. 2009. Both inhibition and activation of Notch signaling rely on a conserved Neuralized-binding motif in Bearded proteins and the Notch ligand Delta. *Developmental Biology* **333**:373–385. DOI: <https://doi.org/10.1016/j.ydbio.2009.06.039>, PMID: 19580805
- Fuzita FJ**, Pinkse MWH, Patane JSL, Juliano MA, Verhaert PDEM, Lopes AR. 2015. Biochemical, transcriptomic and proteomic analyses of digestion in the scorpion *Tityus serrulatus*: insights into function and evolution of digestion in an ancient arthropod. *PLOS ONE* **10**:e0123841. DOI: <https://doi.org/10.1371/journal.pone.0123841>
- Gazave E**, Lemaître QI, Balavoine G. 2017. The Notch pathway in the annelid *Platynereis*: insights into chaetogenesis and neurogenesis processes. *Open Biology* **7**:160242. DOI: <https://doi.org/10.1098/rsob.160242>, PMID: 28148821
- Gulia-Nuss M**, Nuss AB, Meyer JM, Sonenshine DE, Roe RM, Waterhouse RM, Sattelle DB, de la Fuente J, Ribeiro JM, Megy K, Thimmapuram J, Miller JR, Walenz BP, Koren S, Hostetler JB, Thiagarajan M, Joardar VS, Hannick LI, Bidwell S, Hammond MP, et al. 2016. Genomic insights into the *Ixodes scapularis* tick vector of Lyme disease. *Nature Communications* **7**:10507. DOI: <https://doi.org/10.1038/ncomms10507>, PMID: 26856261
- Hase H**, Koukai M, Hamanaka Y, Goto SG, Tachibana S-I, Shiga S. 2017. Transcriptome analysis of the brain under diapause and nondiapause conditions in the blowfly *Protophormia terraenovae*. *Physiological Entomology* **42**:282–289. DOI: <https://doi.org/10.1111/phen.12205>
- Holland PWH**, Marlétaz F, Maeso I, Dunwell TL, Paps J. 2017. New genes from old: asymmetric divergence of gene duplicates and the evolution of development. *Philosophical Transactions of the Royal Society B: Biological Sciences* **372**:20150480. DOI: <https://doi.org/10.1098/rstb.2015.0480>
- Hosono K**, Sasaki T, Minoshima S, Shimizu N. 2004. Identification and characterization of a novel gene family YPEL in a wide spectrum of eukaryotic species. *Gene* **340**:31–43. DOI: <https://doi.org/10.1016/j.gene.2004.06.014>, PMID: 15556292
- Hoy MA**, Waterhouse RM, Wu K, Estep AS, Ioannidis P, Palmer WJ, Pomerantz AF, Simão FA, Thomas J, Jiggins FM, Murphy TD, Pritham EJ, Robertson HM, Zdobnov EM, Gibbs RA, Richards S. 2016. Genome sequencing of the phytoseiid predatory mite *Metaseiulus occidentalis* reveals completely atomized *Hox* genes and superdynamic intron evolution. *Genome Biology and Evolution* **8**:1762–1775. DOI: <https://doi.org/10.1093/gbe/evw048>, PMID: 26951779
- i5K Consortium**. 2013. The i5K initiative: advancing arthropod genomics for knowledge, human health, agriculture, and the environment. *Journal of Heredity* **104**:595–600. DOI: <https://doi.org/10.1093/jhered/est050>, PMID: 23940263
- Ikeda KT**, Hirose Y, Hiraoka K, Noro E, Fujishima K, Tomita M, Kanai A. 2015. Identification, expression, and molecular evolution of microRNAs in the "living fossil" *Triops cancriformis* (tadpole shrimp). *RNA* **21**:230–242. DOI: <https://doi.org/10.1261/rna.045799.114>, PMID: 25525151
- Ioannidis P**, Simao FA, Waterhouse RM, Manni M, Seppely M, Robertson HM, Misof B, Niehuis O, Zdobnov EM. 2017. Genomic features of the damselfly *Calopteryx splendens* representing a sister clade to most insect orders. *Genome Biology and Evolution* **9**:415–430. DOI: <https://doi.org/10.1093/gbe/evx006>, PMID: 28137743
- Jones DT**, Cozzetto D. 2015. DISOPRED3: precise disordered region predictions with annotated protein-binding activity. *Bioinformatics* **31**:857–863. DOI: <https://doi.org/10.1093/bioinformatics/btu744>, PMID: 25391399

- Kenny NJ, Chan KW, Nong W, Qu Z, Maeso I, Yip HY, Chan TF, Kwan HS, Holland PW, Chu KH, Hui JH. 2016. Ancestral whole-genome duplication in the marine chelicerate horseshoe crabs. *Heredity* **116**:190–199. DOI: <https://doi.org/10.1038/hdy.2015.89>, PMID: 26419336
- Kjer KM, Simon C, Yavorskaya M, Beutel RG. 2016. Progress, pitfalls and parallel universes: a history of insect phylogenetics. *Journal of the Royal Society Interface* **13**:20160363. DOI: <https://doi.org/10.1098/rsif.2016.0363>
- Kocot KM, Struck TH, Merkel J, Waits DS, Todt C, Brannock PM, Weese DA, Cannon JT, Moroz LL, Lieb B, Halanych KM. 2017. Phylogenomics of Lophotrochozoa with consideration of systematic error. *Systematic Biology* **66**:256–282. DOI: <https://doi.org/10.1093/sysbio/syw079>, PMID: 27664188
- Lai EC. 2002. Keeping a good pathway down: transcriptional repression of Notch pathway target genes by CSL proteins. *EMBO Reports* **3**:840–845. DOI: <https://doi.org/10.1093/embo-reports/kvf170>, PMID: 12223465
- Lee B-Y, Kim H-S, Hwang D-S, Won E-J, Choi B-S, Choi I-Y, Park HG, Rhee J-S, Lee J-S. 2015. Whole transcriptome analysis of the monogonont rotifer *Brachionus koreanus* provides molecular resources for developing biomarkers of carbohydrate metabolism. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics* **14**:33–41. DOI: <https://doi.org/10.1016/j.cbd.2015.02.003>
- Lemer S, Kawauchi GY, Andrade SC, González VL, Boyle MJ, Giribet G. 2015. Re-evaluating the phylogeny of Sipuncula through transcriptomics. *Molecular Phylogenetics and Evolution* **83**:174–183. DOI: <https://doi.org/10.1016/j.ympev.2014.10.019>, PMID: 25450098
- Luo YJ, Takeuchi T, Koyanagi R, Yamada L, Kanda M, Khalturina M, Fujie M, Yamasaki SI, Endo K, Satoh N. 2015. The *Lingula* genome provides insights into brachiopod evolution and the origin of phosphate biomineralization. *Nature Communications* **6**:8301. DOI: <https://doi.org/10.1038/ncomms9301>, PMID: 26383154
- Luo YJ, Kanda M, Koyanagi R, Hisata K, Akiyama T, Sakamoto H, Sakamoto T, Satoh N. 2018. Nemertean and phoronid genomes reveal lophotrochozoan evolution and the origin of bilaterian heads. *Nature Ecology & Evolution* **2**:141–151. DOI: <https://doi.org/10.1038/s41559-017-0389-y>, PMID: 29203924
- Lynch M. 2007. The frailty of adaptive hypotheses for the origins of organismal complexity. *PNAS* **104**:8597–8604. DOI: <https://doi.org/10.1073/pnas.0702207104>
- Maier D, Stumm G, Kuhn K, Preiss A. 1992. *Hairless*, a *Drosophila* gene involved in neural development, encodes a novel, serine rich protein. *Mechanisms of Development* **38**:143–156. DOI: [https://doi.org/10.1016/0925-4773\(92\)90006-6](https://doi.org/10.1016/0925-4773(92)90006-6), PMID: 1419850
- Maier D. 2006. *Hairless*: the ignored antagonist of the Notch signalling pathway. *Hereditas* **143**:212–221. DOI: <https://doi.org/10.1111/j.2007.0018-0661.01971.x>, PMID: 17362357
- Maier D, Chen AX, Preiss A, Ketelhut M. 2008. The tiny *Hairless* protein from *Apis mellifera*: a potent antagonist of Notch signaling in *Drosophila melanogaster*. *BMC Evolutionary Biology* **8**:175. DOI: <https://doi.org/10.1186/1471-2148-8-175>, PMID: 18559091
- Maier D. 2019. The evolution of transcriptional repressors in the Notch signaling pathway: a computational analysis. *Hereditas* **156**:5. DOI: <https://doi.org/10.1186/s41065-019-0081-0>, PMID: 30679936
- Marck C. 1988. 'DNA Strider': a 'C' program for the fast analysis of DNA and protein sequences on the Apple Macintosh family of computers. *Nucleic Acids Research* **16**:1829–1836. DOI: <https://doi.org/10.1093/nar/16.5.1829>, PMID: 2832831
- Millard CJ, Watson PJ, Celardo I, Gordiyenko Y, Cowley SM, Robinson CV, Fairall L, Schwabe JW. 2013. Class I HDACs share a common mechanism of regulation by inositol phosphates. *Molecular Cell* **51**:57–67. DOI: <https://doi.org/10.1016/j.molcel.2013.05.020>, PMID: 23791785
- Millard CJ, Fairall L, Schwabe JW. 2014. Towards an understanding of the structure and function of MTA1. *Cancer and Metastasis Reviews* **33**:857–867. DOI: <https://doi.org/10.1007/s10555-014-9513-5>, PMID: 25352341
- Miller SW, Rebeiz M, Atanasov JE, Posakony JW. 2014. Neural precursor-specific expression of multiple *Drosophila* genes is driven by dual enhancer modules with overlapping function. *PNAS* **111**:17194–17199. DOI: <https://doi.org/10.1073/pnas.1415308111>, PMID: 25404315
- Misof B, Liu S, Meusemann K, Peters RS, Donath A, Mayer C, Frandsen PB, Ware J, Flouri T, Beutel RG, Niehuis O, Petersen M, Izquierdo-Carrasco F, Wappler T, Rust J, Aberer AJ, Aspöck U, Aspöck H, Bartel D, Blanke A, et al. 2014. Phylogenomics resolves the timing and pattern of insect evolution. *Science* **346**:763–767. DOI: <https://doi.org/10.1126/science.1257570>, PMID: 25378627
- Nibu Y, Zhang H, Levine M. 1998. Interaction of short-range repressors with *Drosophila* CtBP in the embryo. *Science* **280**:101–104. DOI: <https://doi.org/10.1126/science.280.5360.101>, PMID: 9525852
- Nossa CW, Havlak P, Yue JX, Lv J, Vincent KY, Brockmann HJ, Putnam NH. 2014. Joint assembly and genetic mapping of the Atlantic horseshoe crab genome reveals ancient whole genome duplication. *GigaScience* **3**:9. DOI: <https://doi.org/10.1186/2047-217X-3-9>, PMID: 24987520
- Rebeiz M, Posakony JW. 2004. GenePalette: a universal software tool for genome sequence visualization and analysis. *Developmental Biology* **271**:431–438. DOI: <https://doi.org/10.1016/j.ydbio.2004.04.011>, PMID: 15223345
- Regier JC, Shultz JW, Zwick A, Hussey A, Ball B, Wetzer R, Martin JW, Cunningham CW. 2010. Arthropod relationships revealed by phylogenomic analysis of nuclear protein-coding sequences. *Nature* **463**:1079–1083. DOI: <https://doi.org/10.1038/nature08742>, PMID: 20147900
- Rodriguez J, Jones TH, Sierwald P, Marek PE, Shear WA, Brewer MS, Kocot KM, Bond JE. 2018. Step-wise evolution of complex chemical defenses in millipedes: a phylogenomic approach. *Scientific Reports* **8**:3209. DOI: <https://doi.org/10.1038/s41598-018-19996-6>
- Roxström-Lindquist K, Faye I. 2001. The *Drosophila* gene *Yippee* reveals a novel family of putative zinc binding proteins highly conserved among eukaryotes. *Insect Molecular Biology* **10**:77–86. DOI: <https://doi.org/10.1046/j.1365-2583.2001.00239.x>, PMID: 11240639

- Sahoo PK**, Kar B, Mohapatra A, Mohanty J. 2013. *De novo* whole transcriptome analysis of the fish louse, *Argulus siamensis*: First molecular insights into characterization of Toll downstream signalling molecules of crustaceans. *Experimental Parasitology* **135**:629–641. DOI: <https://doi.org/10.1016/j.exppara.2013.09.018>, PMID: 24090566
- 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, et al. 2017. The house spider genome reveals an ancient whole-genome duplication during arachnid evolution. *BMC Biology* **15**:62. DOI: <https://doi.org/10.1186/s12915-017-0399-x>, PMID: 28756775
- Simakov O**, Marletaz F, Cho SJ, Edsinger-Gonzales E, Havlak P, Hellsten U, Kuo DH, Larsson T, Lv J, Arendt D, Savage R, Osoegawa K, de Jong P, Grimwood J, Chapman JA, Shapiro H, Aerts A, O'tillar RP, Terry AY, Boore JL, et al. 2013. Insights into bilaterian evolution from three spiralian genomes. *Nature* **493**:526–531. DOI: <https://doi.org/10.1038/nature11696>, PMID: 23254933
- Simakov O**, Kawashima T, Marlétaz F, Jenkins J, Koyanagi R, Mitros T, Hisata K, Bredeson J, Shoguchi E, Gyoja F, Yue JX, Chen YC, Freeman RM, Sasaki A, Hikosaka-Katayama T, Sato A, Fujie M, Baughman KW, Levine J, Gonzalez P, et al. 2015. Hemichordate genomes and deuterostome origins. *Nature* **527**:459–465. DOI: <https://doi.org/10.1038/nature16150>, PMID: 26580012
- Simon S**, Blanke A, Meusemann K. 2018. Reanalyzing the Palaeoptera problem - The origin of insect flight remains obscure. *Arthropod Structure & Development* **47**:328–338. DOI: <https://doi.org/10.1016/j.asd.2018.05.002>, PMID: 29763650
- Smith AF**, Posakony JW, Rebeiz M. 2017. Automated tools for comparative sequence analysis of genic regions using the GenePalette application. *Developmental Biology* **429**:158–164. DOI: <https://doi.org/10.1016/j.ydbio.2017.06.033>, PMID: 28673819
- Song H**, Hasson P, Paroush Z, Courey AJ. 2004. Groucho oligomerization is required for repression in vivo. *Molecular and Cellular Biology* **24**:4341–4350. DOI: <https://doi.org/10.1128/MCB.24.10.4341-4350.2004>, PMID: 15121853
- Techer MA**, Rane RV, Grau ML, Roberts JMK, Sullivan ST, Liachko I, Childers AK, Evans JD, Mikheyev AS. 2019. Genomic analyses of sibling honey bee ectoparasitic mite species show divergent strategies of adaptation. *bioRxiv*. DOI: <https://doi.org/10.1101/512988>
- Tribolium Genome Sequencing Consortium**. 2008. The genome of the model beetle and pest *Tribolium castaneum*. *Nature* **452**:949–955. DOI: <https://doi.org/10.1038/nature06784>, PMID: 18362917
- True JR**, Haag ES. 2001. Developmental system drift and flexibility in evolutionary trajectories. *Evolution and Development* **3**:109–119. DOI: <https://doi.org/10.1046/j.1525-142x.2001.003002109.x>, PMID: 11341673
- Whelan NV**, Kocot KM, Santos SR, Halanych KM. 2014. Nemertean toxin genes revealed through transcriptome sequencing. *Genome Biology and Evolution* **6**:3314–3325. DOI: <https://doi.org/10.1093/gbe/evu258>, PMID: 25432940
- Wu C**, Jordan MD, Newcomb RD, Gemmill NJ, Bank S, Meusemann K, Dearden PK, Duncan EJ, Grosser S, Rutherford K, Gardner PP, Crowhurst RN, Steinwender B, Tooman LK, Stevens MI, Buckley TR. 2017. Analysis of the genome of the New Zealand giant collembolan (*Holacanthella duospinosa*) sheds light on hexapod evolution. *BMC Genomics* **18**:795. DOI: <https://doi.org/10.1186/s12864-017-4197-1>, PMID: 29041914
- Yuan Z**, Praxenthaler H, Tabaja N, Torella R, Preiss A, Maier D, Kovall RA. 2016. Structure and function of the Su (H)-Hairless repressor complex, the major antagonist of Notch signaling in *Drosophila melanogaster*. *PLOS Biology* **14**:e1002509. DOI: <https://doi.org/10.1371/journal.pbio.1002509>, PMID: 27404588
- Zacharioudaki E**, Falo Sanjuan J, Bray S. 2019. Mi-2/NuRD complex protects stem cell progeny from mitogenic Notch signaling. *eLife* **8**:e41637. DOI: <https://doi.org/10.7554/eLife.41637>, PMID: 30694174
- Zhang G**, Fang X, Guo X, Li L, Luo R, Xu F, Yang P, Zhang L, Wang X, Qi H, Xiong Z, Que H, Xie Y, Holland PW, Paps J, Zhu Y, Wu F, Chen Y, Wang J, Peng C, et al. 2012. The oyster genome reveals stress adaptation and complexity of shell formation. *Nature* **490**:49–54. DOI: <https://doi.org/10.1038/nature11413>, PMID: 22992520