Insect science

Insect Science (2016) 23, 335–352, DOI 10.1111/1744-7917.12327

REVIEW

Distal-less homeobox genes of insects and spiders: genomic organization, function, regulation and evolution

Bin Chen[1,†], William H. Piel[2,3,†] and Antónia Monteiro[2,3,†]

[1]*Institute of Entomology and Molecular Biology, College of Life Sciences, Chongqing Normal University, Chongqing 401331, P.R. China;* [2]*Yale-NUS College, Singapore and* [3]*Department of Biological Sciences, National University of Singapore, Singapore*

Abstract The *Distal-less* (*Dll*) genes are homeodomain transcription factors that are present in most Metazoa and in representatives of all investigated arthropod groups. In *Drosophila*, the best studied insect, *Dll* plays an essential role in forming the proximodistal axis of the legs, antennae and analia, and in specifying antennal identity. The initiation of *Dll* expression in clusters of cells in mid-lateral regions of the *Drosophila* embryo represents the earliest genetic marker of limbs. *Dll* genes are involved in the development of the peripheral nervous system and sensitive organs, and they also function as master regulators of black pigmentation in some insect lineages. Here we analyze the complete genomes of six insects, the nematode *Caenorhabditis elegans* and *Homo sapiens*, as well as multiple *Dll* sequences available in databases in order to examine the structure and protein features of these genes. We also review the function, expression, regulation and evolution of arthropod *Dll* genes with emphasis on insects and spiders.

Key words *Distal-less*; evolution; function; genomic organization; regulation

Introduction

The *Distal-less* (*Dll*) gene is a homeodomain transcription factor, named *Dlx* in vertebrates and *Dll* in all other metazoans (Zerucha & Ekker, 2000). *Dll* is expressed in representatives of onychophorans and many arthropod groups including chelicerates, myriapods, crustaceans and insects (Williams & Nagy, 1996; Panganiban *et al.*, 1997; Popadic *et al.*, 1998; Scholtz *et al.*, 1998; Thomas & Telford, 1999; Mittmann & Scholtz, 2001; Pechmann *et al.*, 2010). *Dll* has its earliest origin in metazoan anteroposterior head axis patterning, and subsequently it was likely co-opted for proximodistal patterning of body appendages in arthropods, including serial homologous and non-homologous appendages (Lemons *et al.*, 2010).

Correspondence: Bin Chen, Institute of Entomology and Molecular Biology, College of Life Sciences, Chongqing Normal University, University Town, Chongqing 401331, P.R. China. Tel: +86 23 65918391; fax: +86 23 65910315; email: c_bin@hotmail.com, bin.chen@cqnu.edu.cn

[†]These authors contributed equally.

It is in this role of appendage development that *Dll* is best known – a role found to occur in virtually all arthropod species where *Dll* expression is known (Robledo *et al.*, 2002). That the evolution of appendages is central to the diversification and success of arthropods means that *Dll* likely has played a key role in the success of this large group (Pechmann *et al.*, 2010). Such appendages include the legs, antennae, wing, analia and mouthparts in insects (Casci, 2002; Lin *et al.*, 2014), coxapophyses and ocularia in harvestmen (Sharma *et al.*, 2013), mouthparts, legs and spinnerets in spiders (Abzhanov *et al.*, 1999; Prpic & Damen, 2004; Pechmann & Prpic, 2009), nasi in termites (Toga *et al.*, 2012), and horns in beetles (Moczek & Nagy, 2005). *Dll* has also been involved in the development of the peripheral nervous system and sensitive organs (Casci, 2002; Plavicki *et al.*, 2012), in butterfly eyespots (Carroll *et al.*, 1994; Brakefield *et al.*, 1996; Beldade *et al.*, 2002), and in melanin synthesis in the wings of both flies and butterflies (Arnoult *et al.*, 2013; Monteiro *et al.*, 2013). Vertebrate *Dlx* genes are involved in a variety of other developmental processes ranging from neurogenesis to hematopoiesis (Depew *et al.*, 1999;

Shimamoto *et al.*, 2000). Earlier reviews of vertebrate *Dlx* genes and *Drosophila melanogaster Dll* (Panganiban, 2000; Zerucha & Ekker, 2000; Panganiban & Rubenstein, 2002) did not include much information from other protostomes.

Here we review information pertaining to *Dll* across all available protostome systems with emphasis on insects and spiders, including recent *Drosophila* work. We start by analyzing the complete genomes of seven protostome species and a deuterostome outgroup (humans), and many *Dll* sequences available in the databases in order to examine evolution of copy number and gene sequence. We also describe the *Dll* gene structure and its main sequence features, and summarize the function, expression and regulation of protostome *Dll* genes, mainly in insects and spiders.

Genomic organization and gene structure

We investigated the complete genome of seven protostome species (six insects and one nematode) as well as a deuterostome outgroup, humans, and determined that there is only one *Dll* copy in protostomes (Table 1). For comparison, the basal chordate amphioxus also has a single *Dlx* gene (Holland *et al.*, 1996), whereas the more derived urochordate, *Ciona intestinalis*, and the vertebrates, mouse and humans, have two and six *Dlx* genes, respectively (Digregorio *et al.*, 1995; Zerucha & Ekker, 2000). The six vertebrate *Dlx* genes can be grouped into two clades (*Dlx1/4/6* and *Dlx2/3/5*) based on sequence similarity, and are organized as three gene pairs (*Dlx1/2*, *Dlx3/4* and *Dlx5/6*) with each pair closely linked on the genome (see Table 1, [Zerucha & Ekker, 2000]). Two additional *Dlx* genes have been identified in zebrafish, which are not linked to each other and do not appear to exist in other vertebrates such as mammals (Stock *et al.*, 1996).

Primary messenger RNA (mRNA) transcript sizes (including introns) of insect *Dll* genes (19 773–72 999 bp) are much larger than that of human *Dlx* (3313–5396 bp) and *Caenorhabditis elegans Dll* (3303 bp) genes (Table 1). Correspondingly, the exon number (5–7 exons), amino acid length (312–356 amino acid [aa]) and molecular mass (31.74–38.36 kDa) of insect *Dll* genes are also larger than those of human *Dlx* (3 exons, 240–328 aa and 26.26–34.24 kDa) and *C. elegans Dll* (5 exons, 273 aa and 30.16 kDa) genes. Interestingly, one of the central introns of insect *Dll* genes spans more than half of the gene's primary transcript size and is largely responsible for the observed difference in gene size relative to the human and *C. elegans* homologues (Fig. 1). This large intron is the third intron in *D. melanogaster*, *Anopheles gambiae*, *Aedes aegypti* and *Apis mellifera* and the second intron in *Bombyx mori* and *Tribolium castaneum*.

Two possible *Dll* splicing patterns, RA and RB, were detected in the genomes of the Diptera, Lepidoptera and Hymenoptera species investigated, whereas a single splice variant was found for *T. castaneum* and *C. elegans* (Table 1, Fig. 1). The RA splicing pattern is a modification of the RB splice variant where an extra exon encoding a conserved VWPAV motif is added to the N-terminal end of the RA protein. These two splice variants match all *Dll* mRNA sequences available in the databases for these insect orders except that of the lepidopteran *Precis coenia* (AF404110.1), which has 27 unique residues at the N-terminal end instead of the common VWPAV motif. All the alternatively spliced protein sequences share a common NPS(G)LL(V)T motif at their C-terminus, whereas *T. castaneum* and *C. elegans* lack this motif. By comparison, vertebrate *Dlx* genes produce multiple transcripts by alternative transcription initiation (e.g. *Dlx1*) (McGuinness *et al.*, 1996) as well as alternative splicing (e.g. *Dlx4* and *Dlx5*) (Liu *et al.*, 1997).

The *Dll*/*Dlx* genes in the eight genomes investigated have different sized exons and introns; however, the homeodomain is commonly split between codons 43 and 44 by the largest intron in insects, *C. elegans* (1414 bp), and humans (431–2019 bp) (Fig. 1). This shared splice site suggests the presence of an ancient homologous intron across all these species. The two homeodomain sections encode 43 and 17 aa in their anterior and posterior exons, respectively, with the exception of *C. elegans* whose anterior 43 aa homeodomain section is additionally interrupted by a short intron and split into 19 and 24 codons. The first two exons of *D. melanogaster*, *An. gambiae*, *Ae. aegypti* and *A. mellifera Dll* genes, and the unique first exon of *B. mori* and *T. castaneum* encode the NM-1 and NM-2 regions of *Dll* (Fig. 2A). These regions are conserved regions, which will be described more fully later. The third exon of *D. melanogaster*, *An. gambiae*, *Ae. aegypti* and *A. mellifera* and the second exon of *B. mori* and *T. castaneum* are homologous and display the highest degree of sequence similarity across these species, perhaps because these exons include the largest portion of the homeodomain. Hereafter exons are quite varied in number and sequence. The last exon of the *Dll*-RB splice variant in all insect species but *T. castaneum*, encodes the conserved motif NPS(G)LL(V)T at the C-terminal end, whereas the *Dll*-RA variant has an additional exon encoding the conserved motif VWPAV in all insects examined (Fig. 2A).

Table 1 *Dll* genes present on the genomes of eight species.

Gene/splicing	Accession number						
	mRNA	Amino acid.	Chromosomal location†	Exon no.‡	Amino acid (length)	Mass (kDa)	Genome size (bp)¶
Drosophila melanogaster							
Dll-RA	NM_079133.1	NP_523857.1	2R+ 60E2	7(7)	327	35.260	20334
Dll-RB	NM_166689.1	NP_726486.1	2R+ 60E2	6(6)	322	34.71	19773
Anopheles gambiae							
Dll-RA	XM_308706.1	XP_308706.1	2L–	7(7)	300	32.29	48680
Dll-RB	Newly predicted	Newly predicted	2L–	6(6)	295	31.74	48506
Aedes aegypti							
Dll-RA	Newly predicted	Newly predicted	N/A	7(7)	301	32.43	72999
Dll-RB	Newly predicted	Newly predicted	N/A	6(6)	296	31.88	72825
Bombyx mori							
Dll-RA	Newly predicted	Newly predicted	N/A	7(7)	356	38.36	43839
Dll-RB	Newly predicted	Newly predicted	N/A	6(6)	351	37.81	43539
Apis mellifera							
Dll-RA	Newly predicted	Newly predicted	Group13 +	6(6)	349	36.28	69484
Dll-RB	Newly predicted	Newly predicted	Group13 +	5(5)	344	35.73	68848
Tribolium castaneum							
Dll	AF317551.1	AAG39634.1	Contig1484-, Contig1541+	5(6)	312	34.40	61104
Caenorhabditis elegans							
Dll (*Ceh-43*)	NM_065503.2	NP_497904.1	3–	5(5)	273	30.16	3303
Homo sapiens							
Dlx1	NM_178120.4	NP_835221.2	2+ q31.1	3(3)	255	27.32	4195
Dlx2	NM_004405.3	NP_004396.1	2– q31.1	3(3)	328	34.24	3313
Dlx3	NM_005220.2	NP_005211.1	17– q21.33	3(3)	287	31.74	5220
Dlx4	NM_138281.1	NP_612138.1	17+ q21.33	3(3)	240	26.26	5396
Dlx5	NM_005221.5	NP_005212.1	7– q21.3	3(3)	289	31.54	4442
Dlx6	NM_005222.2	NP_005213.2	7+ q21.3	3(3)	265	29.39	5063

†+, forward strand; –, reverse strand; both preceded by chromosomal name or sequence group/contig, and followed by location in the cytological map. N/A indicates unavailability of the chromosomal information.

‡Total number of exons in coding regions. The numbers in brackets include exons in untranslated regions.

¶The size corresponds to the primary mRNA transcript including introns.

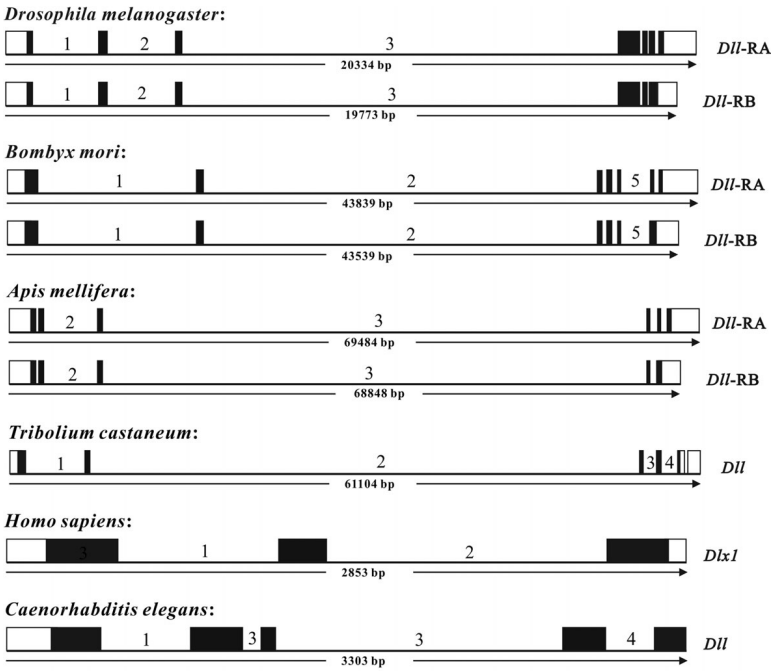

Fig. 1 Scaled intron/exon structures of *Dll/Dlx* genes. The filled rectangles depict protein-coding sequences, and unfilled rectangles represent untranslated regions. Introns are the open spaces between these rectangles with intron numbers partially marked.

Protein sequence features

The *Dll* homeodomain is invariable across all arthropod sequences available in the databases and only 20 of the 60 residues vary across all available animal sequences (Fig. 2A). The homeodomain codes for a homeobox that binds DNA through a helix-turn-helix structure (HTH, Fig. 2A), usually in the regulatory region of *Dll*/*Dlx* target genes. There are four additional motifs that were identified in this study to be relatively conserved across arthropods, NM-1 and NM-2 located at the N-terminus, and HD-5′ and HD-3′ connected to the homeodomain at its 5′- and 3′-ends, respectively. NM-1 is 22 aa long in *D. melanogaster*, with eight identical aa throughout Insecta. NM-2 is 36 aa long, with three and 20 identical aa throughout Arthropoda and Insecta, respectively. HD-5′ is approximately 13 aa long, with three and five identical aa throughout Metazoa and Arthropoda, respectively. HD-3′ is 12 aa long, with three identical aa throughout Arthropoda. The functions of these motifs are still unknown, but likely to mediate interactions with other proteins.

All Dll/Dlx proteins possess two conserved tryptophan (Trp) residues that are C-terminal to the homeodomain. The first (not shown) is typically followed by an aspartic acid, whereas the second (Fig. 2A) is followed by a tyrosine (Panganiban & Rubenstein, 2002). Dll/Dlx proteins lack Trp N-terminal of their homeodomains.

Dll expression and function

Anteroposterior head axis patterning

Early *Dll* expression in cnidarians (Ryan *et al.*, 2007), and in particular the head regions of mollusks (Lee & Jacobs, 1999), hemichordates (Lowe *et al.*, 2003) and spiders (Pechmann *et al.*, 2011) suggests that an anteroposterior axis patterning role existed ancestral to arthropods and that this role was later co-opted into a proximodistal patterning role for the serial homologous appendages (Lemons *et al.*, 2010). Remarkably, RNA interference (RNAi) suppression of early *Dll* expression in spiders shows a novel gap gene-like function with the loss of the first or the first and second walking leg body segments, and corresponding legs (Pechmann *et al.*, 2011) (Figure 4).

Legs

Dll is activated in the thoracic limb primordia of the fly embryo soon after gastrulation, and is one of the earliest known markers of these primordia (Cohen *et al.*, 1989; Cohen, 1990). *Dll* is first expressed in all six thoracic primordia at late embryonic stage 10 and is expressed in the entire limb thoracic primordium at stage 11 (Co-

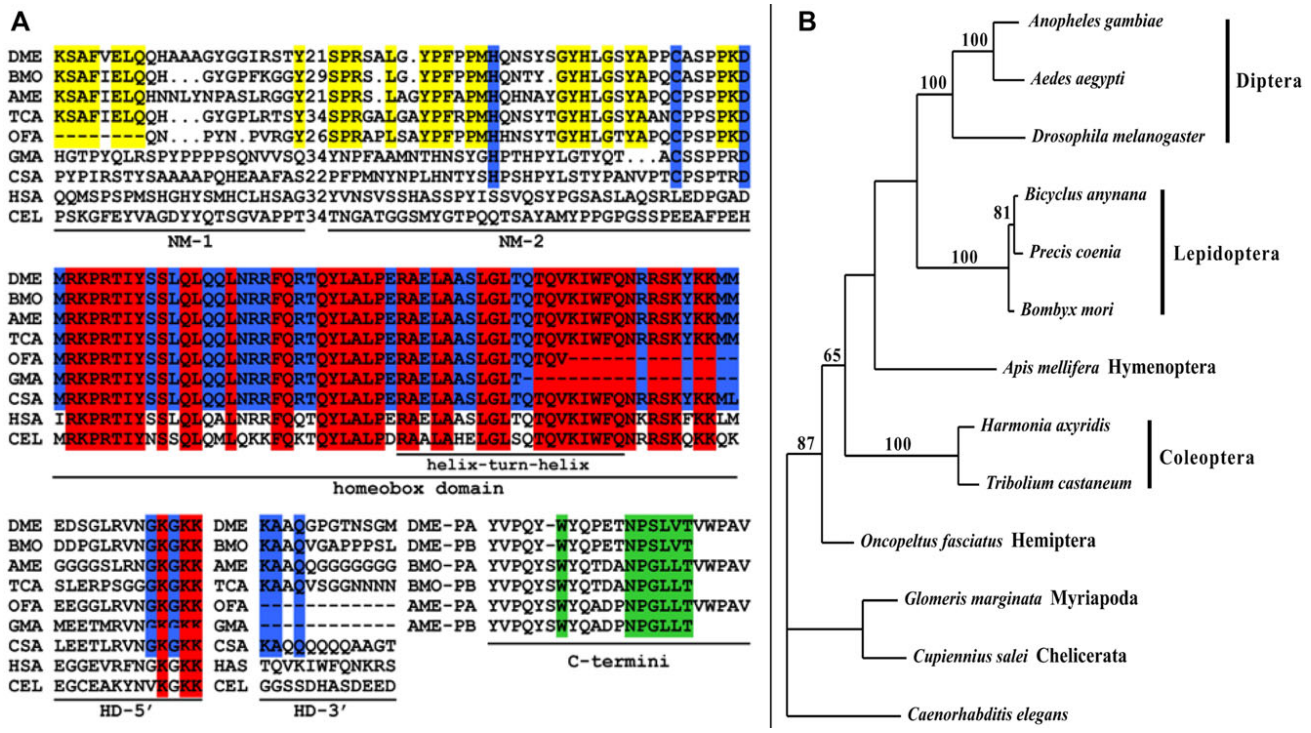

Fig. 2 Analysis of the *Dll* sequences of representative insects. (A) Alignments of conserved motifs. NM-1 and NM-2 are two N-terminal motifs separated by a number of residues not shown, and HD-5′ and HD-3′ are the motifs connected to the homeobox at its 5′- and 3′-end. The shaded residues are identical either throughout Metazoa (red), Arthropoda (blue) or Insecta (yellow). The conserved residue tryptophan and the motif NPS(G)LL(V)T in C-termini are marked in green. Dots and dashes represent gaps and unavailability in individual sequences. AME, *Apis mellifera*; BMO, *Bombyx mori*; CEL, *Caenorhabditis elegans*; CSA, *Cupiennius salei*; DME, *Drosophila melanogaster*; GMA, *Glomeris marginata*; HAS, *Homo sapiens*; OFA, *Oncopeltus fasciatus*; TCA, *Tribolium castaneum*. (B) The most parsimonious *Dll* gene tree inferred from conserved protein sequences with maximum parsimony phylogenetic analysis, which was conducted using PAUP* v4.0b10 (Swofford, 2002). The bootstrap percentages of 1000 replicates are shown on the branches where they exceed 50%.

hen *et al.*, 1989; Cohen, 1990). The second thoracic limb primordium gives rise to three structures: the leg, wing and a larval sensory organ called the Keilin's organ. At stage 12, *Dll* expression becomes restricted to the center of the primordia of the combined disc for wing, leg and Keilin's organ. The central cells of the combined disc were thought to contribute to the distal leg but it was later proposed that they give rise to the Keilin's organ (Fig. 3A) (Kubota *et al.*, 2003; Bolinger & Boekhoff-Falk, 2005). The Keilin's organ precursor cells express *Dll*, *Cut* and *Couch potato* (*Cpo*) but lack the leg imaginal determinant *Escargot* (*Esg*). Distal leg disc cells are marked by the coexpression of *Dll* and *Esg*, whereas proximal cells express only *Esg* (Fig. 5B) (Bolinger & Boekhoff-Falk, 2005). *Dll* expression is lost from medial leg cells either before or during the second larval instar (Weigmann & Cohen, 1999). Late in development, during the third instar, *Dll* expression is activated in a proximal ring corresponding to the trochanter where it seems to play an important role in preventing mixing of proximal and medial cells (Wu & Cohen, 1999) and is expressed at low levels in the developing femur (Weigmann & Cohen, 1999). In the late pupal stage, *Dll* is expressed in the distal trochanter and tibia, and in all tarsal segments (Figs. 3B, 4) (Dong *et al.*, 2002; Panganiban & Rubenstein, 2002).

Across the insects, with the exception of dipterans that have limbless larvae, limb primordia established during embryogenesis give rise to well-formed larval limbs. Before limb primordia are morphologically discernible, *Dll* expression in *Tribolium* closely resembles that of *Drosophila*. Limb development in *Drosophila* is not visible from the outside through the larval stages until the pupal stage when a fly pupal leg resembles an embryonic leg in *Tribolium* (Beermann *et al.*, 2001). Similar patterns of *Dll* expression have been observed in several insect embryos, including lepidopterans (Zheng *et al.*, 1999), grasshoppers (Jockusch *et al.*, 2000) and crickets (Niwa *et al.*, 1997).

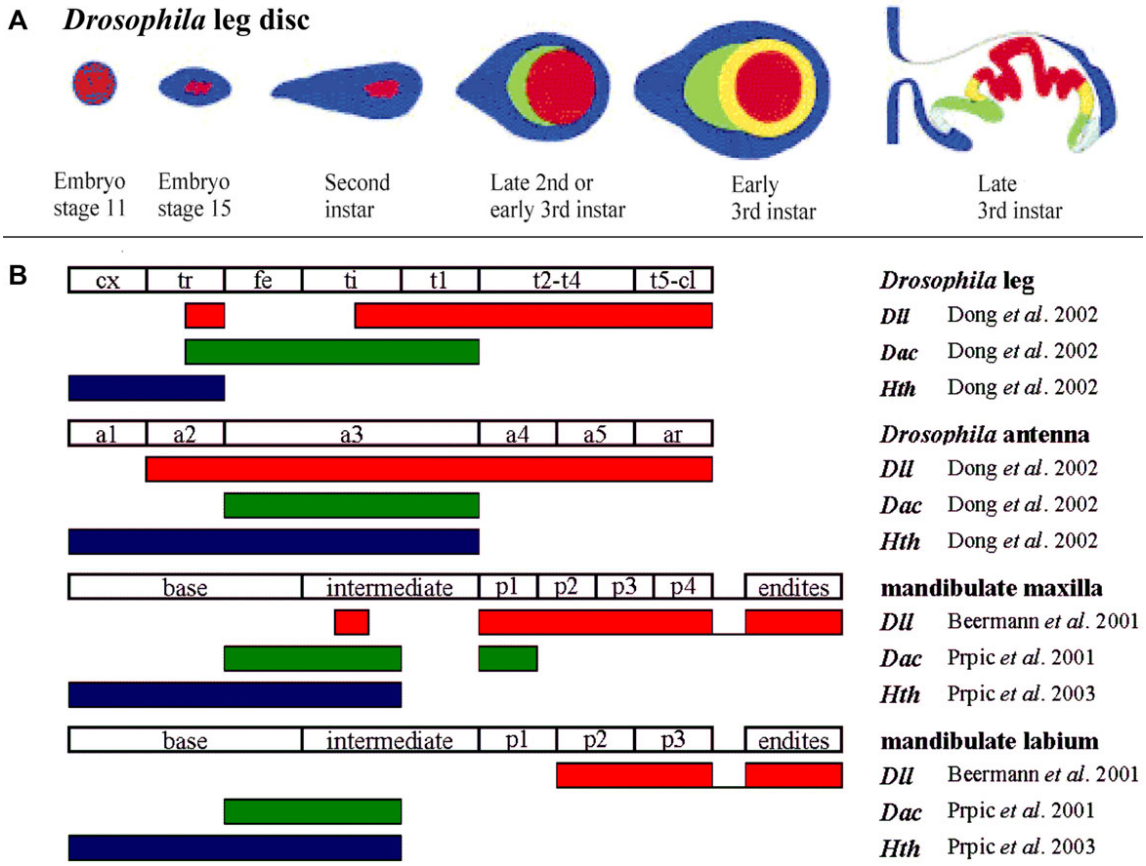

Fig. 3 Expression patterns of domain genes *Dll*, *Dac* and *Hth* along the PD axis of appendages. (A) Modified from Inoue *et al.* (2002). *Dll* is expressed in the entire leg primordium in the first to third thoracic segments at stage 11 of *Drosophila* embryonic development, and restricted in the central domain at stage 15. In the late second or early third instar the leg disc is defined by three discrete domains of gene expressions: distal (*Dll*, red), medial (*Dac*, green) and proximal (*Hth*, dark blue). Following this, *Dll* and *Dac* expressions overlap in a broad domain (yellow). By the late third instar, *Dll* expression is reactivated in a proximal ring corresponding to the trochanter domain (grey), where all three proteins overlap (Postleth & Schneide, 1971; Dong *et al.*, 2001; Inoue *et al.*, 2002). (B) *Dll*, *Dac* and *Hth* expression in the antenna, leg, maxilla and labium at a late disc developmental stage. cx, coxa; tr, trochanter; fe, femur; ti, tibia; t1-t5, tarsi 1–5; cl, claw; a1–a5, antennal segments 1–5; ar, arista; cl, claw; p1–p5, palps 1–5.

Despite some discrepancies in the early leg bud, expression patterns of *Dll* in its later stages are basically consistent between dipterans, orthopterans (Jockusch *et al.*, 2000; Inoue *et al.*, 2002), hemipterans (Rogers *et al.*, 2002), coleopterans (Beermann *et al.*, 2001), myriapods (Prpic & Tautz, 2003), chelicerates (Panganiban *et al.*, 1995; Thomas & Telford, 1999; Abzhanov & Kaufman, 2000; Schoppmeier & Damen, 2001) and crustaceans (Panganiban *et al.*, 1995; Popadic *et al.*, 1998; Abzhanov & Kaufman, 2000). These data suggest a conserved role for *Dll* in patterning the distal part of all arthropod limbs (Fig. 4).

In the abdominal prolegs of the butterfly *Junonia coenia*, the expression of *Dll* is restricted to a proximal ring, and the more distal expression is lacking or highly reduced in levels (Panganiban *et al.*, 1994) (Fig. 4). However, *Dll* is not expressed in the prolegs of sawfly larvae *Neodiprion abietis*, *Diprion similis* and *Athalia rosae* (Suzuki & Palopoli, 2001; Oka *et al.*, 2010). In two crustaceans, *Triops* and *Nebalia*, *Dll* is not expressed prior to branching in all limbs, nor is it expressed in every limb branch, which postulates that *Dll* only plays a role in establishing a single proximal-distal (PD) patterning axis within these branched limbs (Williams, 1998). This suggests that sawfly prolegs may derive from these branch elements but not from the main PD axis (Oka *et al.*, 2010). In addition, the distinct *Dll* expression patterns suggest that larval prolegs have evolved independently in Hymenoptera and Lepidoptera (Suzuki & Palopoli, 2001).

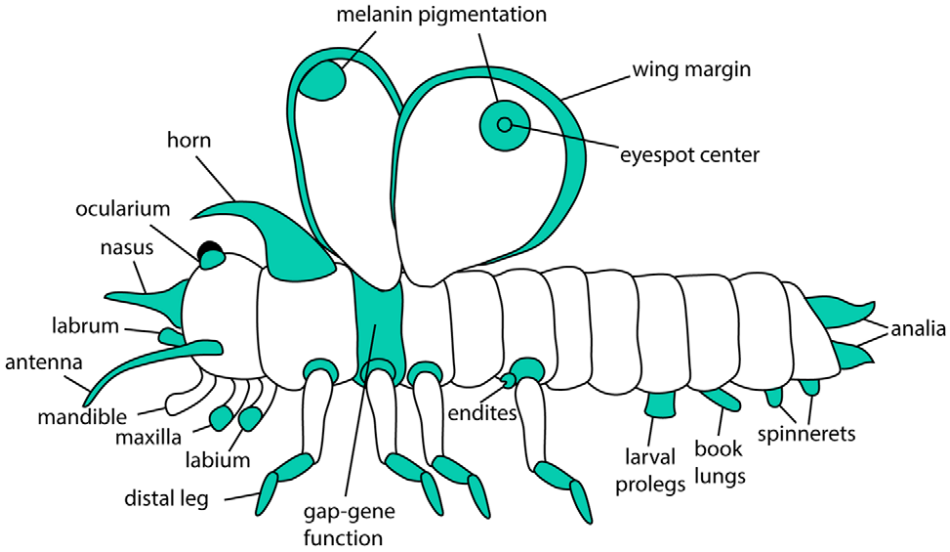

Fig. 4 Expression domains and function of *Dll* in a chimeric arthropod with insect and spider features. *Dll* functions as a gap gene in spiders. It is a positive regulator of horn size in beetles, nasus size in termites and eyespot size in butterflies. *Dll* is required for the development of the analia and all ventral appendages except mandibles (early in development), distal limbs (later in development), and it specifies antennal identity. *Dll* is required for ocularium development in harvestmen, and endites, book lungs, and spinnerets in spiders. *Dll* is also expressed in prolegs in Lepidoptera but it is still unclear whether it has a function in the development of this trait. *Dll* is required for wing margin development in flies, and for the development of multiple parts of the nervous system and sensory cells (not shown). Head appendage nomenclature used is that of insects.

In arthropods, investigation of *Dll* function has been largely focused on *D. melanogaster*. *Dll* is required for limb outgrowth and for differentiation of distal limb structures. *Dll*-null mutants die as embryos because they lack the rudimentary larval limbs (Cohen & Jurgens, 1989). Clones of homozygous Dll^{-} cells are incapable of contributing to any structure but the leg coxal segment, whereas those of Dll^{+} develop normally (Lindsley & Zimm, 1992). Hypomorphic fly *Dll* alleles have been reported to result in various transformations, malformations, fusions and/or reductions of legs (in the ventral thoracic disc, trochanter, femur, tibia, tarsus and joint) (Dong *et al.*, 2000, 2002; Beermann *et al.*, 2001; Dworkin, 2005). Ectopic *Dll* expression in the proximal region of ventral appendages induces nonautonomous duplication of legs by the activation of *Wingless* (*Wg*) and *Decapentaplegic* (*Dpp*) (Gorfinkiel *et al.*, 1997). The requirement for *Dll* in the fly femur and most of the tibia is lost by about the early third instar, and the distal tibia and the tarsus remain the only regions where *Dll* function is required late in development (Campbell & Tomlinson, 1998).

In the arachnid *Cupiennius*, the silencing of *Dll* by RNAi leads to missing distal portions of limbs but to normal proximal parts (Schoppmeier & Damen, 2001). However, *Dll* is also expressed in the spider's palpal coxae, and may be responsible for the endites that extend from them (Pechmann & Prpic, 2009; Pechmann *et al.*, 2010), especially seeing as *Dll* silencing in harvestmen results in the loss of coxapophyses (Sharma *et al.*, 2013). In *Onthophagus* beetles, *Dll* down-regulation in the last larval instar led to loss or fusion of distal appendage regions in pupal and adult legs, as well as mouthparts and antennae (Moczek & Rose, 2009). In *T. castaneum*, disruptions of *Dll* function in earlier larval instars led to more severe leg disruptions, suggesting a role for this gene in maintaining the integrity of the whole larval leg (Suzuki *et al.*, 2009). In the crustacean *Parhyale hawaiensis*, *Dll* small short interfering RNA (siRNA) injections into embryos led to hatchings with truncated appendages (Liubicich *et al.*, 2009).

Phenotypic analysis indicates that distal leg (and antennal) structures are more sensitive to changes in *Dll* levels than are medial structures. Based on this, it was anticipated that *Dll* would be expressed in a graded manner along the developing PD axis, with the highest levels present distally (Cohen *et al.*, 1989). However, neither RNA *in situ* data nor antibody stainings have provided convincing support for this view. It may be that there is a shallow *Dll* gradient, not readily observed using standard techniques. Alternatively, there may be a gradient early in development that has disappeared before the third instar, which is the stage most commonly analyzed. Yet

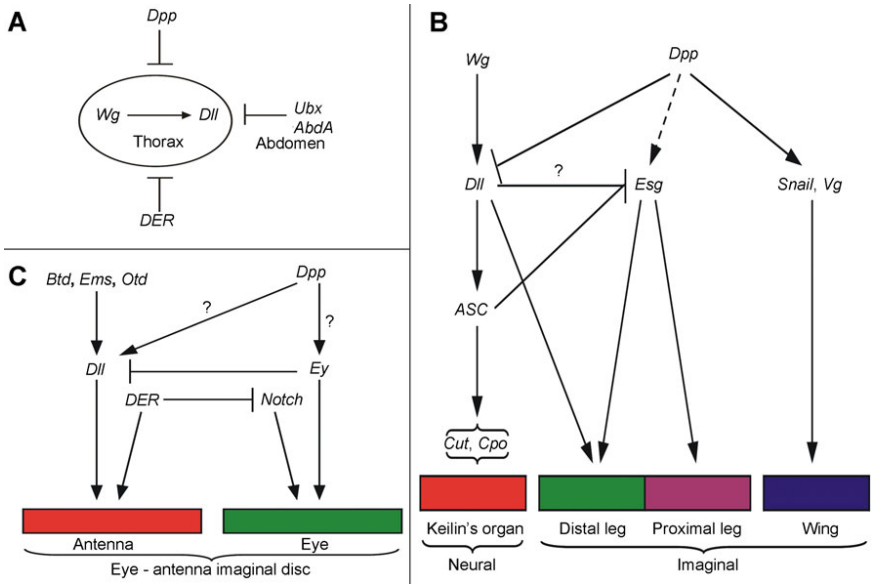

Fig. 5 The role of *Dll* in the genetic pathway of imaginal disc differentiation and subsequent subdivision in *D. melanogaster*. A. Modified from Panganiban (2000). *Dll* is activated by *Wg* in the embryonic thorax, and repressed by *Dpp* dorsally and *DER* ventrally. *Dll* is not expressed in the abdomen because of repression by *Ubx* and *AbdA*. B. Modified from Bolinger and Boekhoff-Falk (2005). *Wg* and *Dpp* serve key roles in the subdivision of thoracic limb primordium through determinants such as *Dll* and *Esg* (Cohen, 1990b; Goto & Hayashi, 1997). The Keilin's organ (red) is specified by *Dll* and by the downstream *acuaete-scute* (*ASC*) complex. The specification involves both the activation of the neural genes *Cut* and *Cpo* and the repression of *Est* (Bolinger & Boekhoff-Falk, 2005). High *Dpp* activity (solid arrow) is required for specification of the wing disc and lower (dashed arrow) for specification of the leg disc (Raz & Shilo, 1993; Goto & Hayashi, 1997; Kubota *et al.*, 2000). Imaginal components are patterned by *Esg* (leg disc) and Snail and *Vg* (wing disc, blue), and *Dll* is a determinant for distal leg (green) but not for proximal leg (purple) (Hayashi *et al.*, 1993; Fuse *et al.*, 1996; Bolinger & Boekhoff-Falk, 2005). C. Subdivision of eye-antenna imaginal disc by *Eyeless* (*Ey*) – *Dll* selectors. *Dll* and *Ey* are expressed throughout the combined eye-antenna disc early in fly development (Kumar & Moses, 2001). Beginning late in the second-larval stage, *Dll* expression becomes confined to the antenna portion (Diaz-Benjumea *et al.*, 1994) and *Ey* to the eye portion of the disc (Halder *et al.*, 1998; Quiring *et al.*, 1994). *Dpp* controls the subdivision (Kenyon *et al.*, 2003), and *Der* and *Notch* participate in this subdivision through delimiting *Dll* and *Ey* expression to the antenna and eye portions, respectively (Kumar & Moses, 2001). Embryos mutant for *Btd*, *Ems* and *Otd* lack *Dll* expression in the primordium of the antenna (Inoue *et al.*, 2002), which indicates that these genes regulate *Dll*.

another possibility is that other factors present in a graded manner in the developing limbs differentially sensitize cells to homogeneous levels of Dll protein. For instance, it is possible that Dll cooperates with the graded Dpp and Wg signals to achieve differential target gene regulation along the limb (Panganiban, 2000).

Wings

Although *Dll* is expressed early in the imaginal discs of *Drosophila* wings and halteres, its function is not required for their formation (Weihe *et al.*, 2004). Subsequently during embryogenesis *Dll* expression is lost from the wing and haltere discs (Kaphingst & Kunes, 1994), but is reactivated along the presumptive wing margin at some point before the third and final larval instar (Weigmann & Cohen, 1999) (Fig. 4). However, this expression is modified in halteres (Weatherbee *et al.*, 1999). Dll^{17} affects the wing margin (including hairs and bristles) and vein differentiation in the vicinity of the margin during the third larval stage (Gorfinkiel *et al.*, 1997; Campbell & Tomlinson, 1998). Ectopic expression of *Dll* in the third larval wing discs leads to the formation of distal leg elements. This effect is likely due, in part, to repression of the wing selector gene *vestigial* (*vg*) by ectopic *Dll* (Gorfinkiel *et al.*, 1997).

Wing eyespots and melanic spots

Dll is also expressed along the margin of both butterfly and moth wings (Carroll *et al.*, 1994; Kango-Singh *et al.*, 2001; Reed & Gilbert, 2004; Reed & Serfas, 2004;

Monteiro *et al.*, 2006). In derived lepidopteran lineages, *Dll* is additionally expressed along intervenous stripes that, in species with border eyespots, become enlarged at the proximal end and resolved into a circular pattern. This group of cells map to the center of the future border eyespots (Brakefield *et al.*, 1996; Reed & Serfas, 2004) (Fig. 4). In saturniid moths with more centrally located discal-cell eyespots *Dll* is also expressed in the future eyespot centers, which later are intercepted by a cross-vein (Monteiro *et al.*, 2006). This cross-vein expression of *Dll* is also seen in nymphalid butterflies that do not carry discal-cell eyespots (Reed & Gilbert, 2004). In the pupal stage of nymphalid butterflies, *Dll* expression extends from the center of the future eyespot pattern to a disc of cells that maps to one of the concentric rings of colored scales of the adult eyespots where black pigmented scales, likely containing melanin, later appear (Brakefield *et al.*, 1996; Brunetti *et al.*, 2001) (Fig. 4). Polymorphic markers at the *Dll* locus in a nymphalid butterfly were associated with the size of the eyespots, suggesting a role for this gene in the control of eyespot size (Beldade *et al.*, 2002). This role was supported with functional studies where the over-expression of *Dll* during late larval development led to both larger eyespots as well as additional eyespots, and ectopic expression during the early pupal stage led to patches of black pigmentation (Monteiro *et al.*, 2013). These experiments indicated that *Dll* is a positive regulator of eyespot size, as well as a master control gene promoting black pigmentation. The latter function of *Dll* is also present in *D. biarmipes* where *Dll* expression is associated with a black spot of melanization on the tip of the wing and where *Dll* down-regulation removes the spot and over-expression induces melanization across the whole wing (Arnoult *et al.*, 2013) (Fig. 4).

Mouthparts

In the fly embryo, shortly after blastoderm formation (at stage 11), *Dll* is expressed in the precursors of the labrum, maxillae and labium (Cohen *et al.*, 1989; Cohen & Jurgens, 1989; Cohen, 1990), and *Dll* is required for the development of the mouthparts, for example, clypeolabrum, maxillary palps and labial palps (Cohen & Jurgens, 1989) (Fig. 4). *Dll* expression is lost from presumptive proximal cells in all these structures during either embryogenesis or the first larval instar (Panganiban, 2000).

Dll is also expressed in the labrum of representatives of Chelicerata, Myriapoda, Crustacea and Insecta (Panganiban *et al.*, 1995; Popadic *et al.*, 1998; Abzhanov *et al.*, 1999; Thomas & Telford, 1999; Beermann *et al.*, 2001; Schoppmeier & Damen, 2001; Urbach & Technau, 2003), and *Dll* RNAi embryos of a spider lack a labral structure (Schoppmeier & Damen, 2001) (Fig. 4). However, *Dll* appears to lack an AP axis patterning role in the mouthparts of myriapods (Prpic & Tautz, 2003). In the primitive mandibulate insect mouthparts, *Dll* is expressed and functionally required in the palps and the medial endites of maxillae and labium (Fig. 3B) (Abzhanov & Kaufman, 2000; Beermann *et al.*, 2001). No *Dll* expression was found in the mandibles of insects (Panganiban *et al.*, 1994) or other arthropods (Scholtz *et al.*, 1998) (Fig. 4). Moreover, levels of *Dll* expression can influence the length of the gnathal appendages during larval development in Hemiptera (Angelini & Kaufman, 2005).

Analia

Dll is expressed in the genital discs and is required in the formation of the analia (dorsal and ventral anal plates) (Gorfinkiel *et al.*, 1999) (Fig. 4). The lack of *Dll* function in the anal primordia transforms the anal tissue into hindgut by the extension of the *even-skipped* (*eve*) domain, whereas ectopic *Dll* represses *eve* expression and hindgut formation in the fly (Gorfinkiel *et al.*, 1999). *Dll* is required for the development of both anal plates in males but only for the dorsal anal plate in females, supporting the idea that the analia arise from two primordia (Gorfinkiel *et al.*, 1999). Fly Dll^{2} somatic clones do not develop anal plates in males, or dorsal anal plates in females (Gorfinkiel *et al.*, 1999). So far there is no report for *Dll* involvement in genital and hindgut primordia development and their morphogenesis.

Antennae

Similar to its expression in the mouthparts, *Dll* is also expressed in the precursors of the antennae in the fly embryo at stage 11 (Cohen *et al.*, 1989; Cohen, 1990). In the late pupal stage, *Dll* is expressed in the antennae from distal segment 2 through to the tip, or arista (Figs. 3B, 4) (Panganiban & Rubenstein, 2002).

Ectopic *Dll* expression in the head-eye region leads to ectopic antennae (Gorfinkiel *et al.*, 1997), whereas clones of homozygous Dll^{-} cells are incapable of contributing to any structure but the first antennal segment (Lindsley & Zimm, 1992). Hypomorphic fly *Dll* alleles result in various transformations, malformations, fusions and/or reductions of antennae (in segments 2 and 3, and arista) (Dong *et al.*, 2000, 2002; Dworkin, 2005).

Dll also plays an essential role in antennal identity. *Dll* and *Eyeless* (*Ey*), are initially co-expressed in the fly eye-antennal disc (Fig. 5C) (Curtiss *et al.*, 2002), but after the input of *Dpp* that appears first in eye and later in the antennal portion of the disc, *Ey* and *Dll* become restricted to either the eye or the antennal disc, respectively, and function as selector genes (Kenyon *et al.*, 2003). *Dll* can induce ectopic antennae in the eye part of the eye-antenna disc (Dong *et al.*, 2000). In addition, both *Dll* and *Homothorax* (*Hth*) function together in specifying antennae from legs. Hypomorphic alleles of either of these genes lead to antenna-to-leg transformation (Dong *et al.*, 2000), and ectopic co-expression of *Dll* and *Hth* can induce antennal differentiation in the leg, head and genital discs (Dong *et al.*, 2000).

Novelties in Chelicerates

Chelicerates have separately evolved the suppression of legs in the abdomen relative to insects and crustaceans (Khadjeh *et al.*, 2012), and early expression of *Dll* correlates with the development of book lungs in arachnids, endites on palps, and of spinnerets in spiders (Abzhanov *et al.*, 1999; Pechmann & Prpic, 2009; Pechmann *et al.*, 2010) (Fig. 4). Whether the spinnerets are homologous with the wings of insects or the gills of crustaceans (Damen *et al.*, 2002) or whether they are serial homologs with legs (Pechmann & Prpic, 2009) is not yet clear, but it is clear that they are novelties, since the ancestors to spiders lacked spinneret-like appendages.

Horns, nasus and ocularium

Dll, together with other appendage patterning genes, is also expressed in horn primordia of *Onthophagus* horned beetles suggesting that these novel structures may have co-opted the limb developmental network in a novel location on the head (Moczek & Nagy, 2005; Moczek *et al.*, 2006; Monteiro & Podlaha, 2009) (Fig. 4). Interestingly, Dll accumulates in distal regions of the male pupal horn, but not in corresponding regions of the female. *Dll* down-regulation in late larval instars led to shorter horns (Moczek & Rose, 2009).

Dll is also expressed in the primordia of the nasus, a novel horn-like frontal projection that is present in termite soldiers, and *Dll* down-regulation represses nasus growth (Toga *et al.*, 2012) (Fig. 4). *Dll* is expressed in the ocularium of harvestman and its down-regulation leads to the loss of this head protuberance beneath the eyes (Sharma *et al.*, 2013) (Fig. 4).

Nervous system

Dll is required for the formation of parts of the central and peripheral nervous systems. *Dll* is expressed in some brain precursor cells at stage 15 of the fly embryo, and in both the optic lobe neurons of the brain and in the glial cells of the ventral nerve cord at a late stage (Kaphingst & Kunes, 1994). The proximal ring of *Dll* expression in the fly leg correlates with the location of a group of campaniform sensillae in the adult, and fly embryos that are *Dll* null lack certain sensilla, including Keilin's organs and antennal, maxillary, labial and labral sense organs. All of these sense organs are thought to correspond to vestiges of the distal sensilla of rudimentary larval appendages (Cohen & Jurgens, 1989; Lindsley & Zimm, 1992). Various fly *Dll* alleles also affect the development of the mechanosensory bristles and macrochaeta over cuticle, the labial, maxillary and antennal sense organs of the head, and Keilin's organs and leg sensilla (bracts, sex combs) (Campbell & Tomlinson, 1998; Dworkin, 2005). Moreover, it has been shown that *Dll* expression along the wing margin is autonomously required for activation of the proneural gene *achaete* in ventral sensory bristles (Campbell & Tomlinson, 1998). Similarly, loss of *Dll* activity in the genital disc of the fly eliminates sensory bristles from the analia and genitalia (Gorfinkiel *et al.*, 1999).

In basal insects such as the silverfish *Lepisma* (Mittmann & Scholtz, 2001), Dll protein accumulation correlates with sensory organs on the mouthparts and terminalia. In the mandibles of this species, specific sensory cells are the only nuclei that stain for Dll. Similarly, in the myriapod *Glomeris* embryonic expression of *Dll* appears in presumptive sensory organs of the maxilla and mandible (Prpic & Tautz, 2003). In the crustaceans *Thamnocephalus* and *Triops*, *Dll* expression appears in cells throughout the limbs at the base of bristle-like setae, which likely have a sensory function (Williams *et al.*, 2002). In the chelicerate *Limulus*, Dll protein correlates with developing mechanoreceptors and sensory neurons in the proximal legs, book gill opercula and dorsal body surface (Mittmann & Scholtz, 2001).

Dll regulation and downstream targets

So far only a few *Dll* regulatory elements have been characterized in the fly. One of the best characterized is enhancer element 208. A 877 bp region within this element, known as Dll304, is sufficient to recapitulate the early expression pattern of *Dll* in the embryonic leg primordium, and is located approximately 12 kb upstream of the

presumed *Dll* promoter region. The remainder of the 208 element directs *Dll* expression in the head (Vachon *et al.*, 1992). In addition to activation functions, the Dll304 region contains two Hox binding sites, Bx1 and Bx2, which repress *Dll* activity in the abdomen. Bx1 is 26 bp long and sufficient to fully repress *Dll*, while BX2 has only a small repressive effect (Vachon *et al.*, 1992; White *et al.*, 2000; Gebelein *et al.*, 2002).

The regulation of *Dll* expression is dynamic and tissue specific. In the fly embryo, *Dll* is activated in the thoracic imaginal primordia by *Wg* and repressed by *Dpp* dorsally and *epidermal growth factor receptor* (*DER*) ventrally (Fig. 5A) (Cohen *et al.*, 1989; Cohen, 1990; Raz & Shilo, 1993; Goto & Hayashi, 1997). Once activated, maintenance of *Dll* expression in the embryo does not require continued Wg signaling (Cohen *et al.*, 1993). The subdivision of the thoracic imaginal primordia is allocated in response to determinants such as *Dll*, *Esg*, *Snail* and *Vg* under the control of *Wg* and *Dpp* (Fig. 5B) (Bolinger & Boekhoff-Falk, 2005). The expression of *buttonhead* (*btd*) and *Sp1* is necessary to allow expression of *Dll* in the leg and sufficient to induce leg identity (Estella *et al.*, 2003). *Dll* is activated by *Wg*, expressed along the wing margin, and together these genes participate in the formation of the wing margin (Tabata & Takei, 2004). *Dll* is also activated by Wg signaling in the optic lobes of the brain (Kaphingst & Kunes, 1994). Expression of *Dll* in the embryonic head has been shown to depend partially on *Wg* and *Engrailed* (*En*) in the gnathal segments (Cohen, 1990), and on *btd*, *empty spiracles* and *orthodenticle* in the antenna (Cohen & Jurgens, 1990). The homeodomain of *Ey* is able to downregulate the expression of *Dll*, which is required during endogenous eye development (Fig. 5C) (Punzo *et al.*, 2004). In addition, expression of *Dll* in the ventral maxillary segment is dependent upon *Deformed* (*Dfd*) (Ohara *et al.*, 1993). Maintenance and refinement of several *Dll* expression patterns through the larval stages requires cooperative positive inputs from both *Dpp* and *Wg*, as well as autoregulatory inputs from *Dll* itself (Goto & Hayashi, 1997; Lecuit & Cohen, 1997). For example, the activation of *Dll* for analia development is dependent on the combined action of *Wg* and *Dpp* (Gorfinkiel *et al.*, 1999). In the larval leg and antennal imaginal discs, both *Wg* and *Dpp* are required for the maintenance of *Dll* expression (Diaz-Benjumea *et al.*, 1994; Lecuit & Cohen, 1997). *Dpp* is expressed dorsally along the anterior-posterior axis, whereas *Wg* is expressed ventrally along each segment's anterior-posterior compartment boundary (Blair, 1995; Held, 1995). Thus, only cells in the center of the disc are exposed to high levels of both *Dpp* and *Wg*. As the discs grow, cells at the periphery of the *Dll* expression domain may continue to be exposed to high Dpp or high Wg, but not both, and therefore stop expressing *Dll*. At some point during the second larval instar, *Dll* expression becomes independent of *Dpp* and *Wg* (Lecuit & Cohen, 1997), probably due to autoregulation (Gorfinkiel *et al.*, 1997; Weigmann & Cohen, 1999). *Dll* itself is also able to induce Wg and Dpp signals, which in turn induce *Dll* expression nonautonomously (Gorfinkiel *et al.*, 1997).

Gradients of the morphogens Dpp and Wg initiate the PD organization of the fly leg by activating *Dll* and repressing *dachshund* (*dac*) and *homothorax* (*hth*) in the center of the disc, and by allowing the activation of *dac* while repressing *hth* medially (Lecuit & Cohen, 1997; Abu-Shaar & Mann, 1998). This creates three domains, distal (center of the disc), medial and proximal, that are specified respectively by the expression of *Dll*, *dac* and *hth* (Fig. 3A). Additionally, duplication of *dac* in spiders created a paralog, *dac2*, which is responsible for an additional leg segment, the patella (Turetzek *et al.*, 2016). Mutually antagonistic interactions between the genes expressed in the proximal and medial and between medial and distal domains maintain the domain identity in the leg (Dong *et al.*, 2001). *Antennapedia* (*Antp*) (together with *Dll*) represses distal expression of *Hth* in the leg, precluding the overlap of *Dll* and *Hth* and thereby preventing antennal differentiation (Dong *et al.*, 2000). As in the leg, *Dll* and *hth* are required to specify the distal and proximal domains of the antenna (Dong *et al.*, 2000). However, *dac*, required for the patterning of the medial leg (Mardon *et al.*, 1994), has a different function in the antenna. These three genes extensively overlap in expression in the antenna and there is no mutual antagonism between *Dll* and *hth* (Dong *et al.*, 2001). Instead of *dac*, *spalt* (*sal*) functions in the specification of the medial domain in antennae (Dong *et al.*, 2000), while *aristaless* (*al*) and *bric à brac* (*bab*) are required for the patterning of the tarsus (Campbell & Tomlinson, 1998). Mutually repressive interactions are also required to separate domains along the PD axis of the fly wing (Dong *et al.*, 2001), but these involve *hth* and *vestigial* (*vg*).

The role of *Dll* and that of other genes in the genetic network for the specification of limb axes has probably been extensively modified across arthropods. An indication of network evolution is the different requirements found for *Wg* in limb axis specification in *Drosophila*, *Tribolium*, *Oncopeltus* and *Gryllus* (Angelini & Kaufman, 2005). Obviously, more sampling of taxonomic groups and more functional studies with other members of the network will be necessary for a complete understanding of the evolutionary history of limb axis specification.

In the fly and the moth *Manduca*, *Dll* expression and limb formation are repressed in the abdomen by the Hox proteins Ultrabithorax (Ubx) and Abdominal A (AbdA)

(Fig. 5A) (Vachon *et al.*, 1992; Zheng *et al.*, 1999). Spiders also prevent limb formation in the abdomen, but have evolved this convergently: *Antp*, and both *Antp* and *Ubx*, in the first and second abdominal segments respectively, each suppressing *Dll* expression (Khadjeh *et al.*, 2012). In insects, these bithorax-complex proteins are known to repress *Dll* expression by binding to a small number of specific sites in a minimal *cis*-regulatory enhancer element (Vachon *et al.*, 1992). However, *Dll* repression is absent in the abdomens of species that carry appendages such as in myriapods and crustaceans (Averof & Cohen, 1997; Grenier *et al.*, 1997), and in the first abdominal segments (A1) of the beetle *Tribolium* (Lewis *et al.*, 2000) and the grasshopper *Schistocerca* (Lewis *et al.*, 2000). In *Tribolium*, Abd-A and Ubx diverge in function relative to limb repression. Whereas Abd-A represses early expression of *Dll* in the embryonic abdomen, Ubx, expressed in A1, appears to allow A1 appendage growth (Lewis *et al.*, 2000). Thus, the repression of *Dll* by one or more *Hox* genes was apparently acquired progressively within the arthropod lineage.

Dll expression in butterfly hindwings is down-regulated directly or indirectly by *Ubx*. Clones of cells lacking *Ubx* within an eyespot field lead to over-expression of *Dll* and subsequent alterations in the size of the hindwing eyespots relative to their forewing counterparts (Weatherbee *et al.*, 1999). Additionally, ectopic expression of *Ubx* on the pupal wings of *B. anynana* activates the black-scale associated genes *sal* and *Dll*, and leads to the differentiation of black wing scales (Tong *et al.*, 2014).

In butterfly larval wings, *Notch* upregulation is followed by *Dll* up-regulation in an intervenous line of cells, as well as in cells that map to the center of the eyespots (Reed & Serfas, 2004), but functional essays to test the hypothesis of direct regulation of *Dll* by *Notch* are still lacking.

Dll expression in both the anal and genital primordia in the third instar larvae of the fly is induced by the joint activities of caudal (cad) and the hedgehog pathways (Gorfinkiel *et al.*, 1999; Moreno & Morata, 1999).

All of Dll's putative targets in the fly embryo and/or larvae encode transcription factors. In the embryo, Dll activates *disconnected* (*disco*) and *D-Wnt5* in the thorax (Cohen *et al.*, 1991; Emerald *et al.*, 2003), *al* in the antennal and maxillary segments (Panganiban, 2000), and represses *hth* in the head and thorax (Panganiban, 2000). In the leg and/or antenna imaginal discs, Dll activates *al* (the most distal elements of the antenna and leg) (Campbell & Tomlinson, 1998), *bab* (fourth-fifth antennal segments) (Campbell & Tomlinson, 1998), *bar* (fourth through to arista antennal segments), *spineless* (*ss*) (tarsal segments, and second and arista antennal segments), *dac* (third antennal segment), *sal* and *atonal* (*ato*) (second antennal segment) (Dong *et al.*, 2002), *distal antenna* (*dan*) and *distal antenna-related* (*danr*) (distal antennal segments) (Emerald *et al.*, 2003), and *hernandez* (*hern*) and *fernandez* (*fer*) (third antennal segment) (Suzanne *et al.*, 2003). Dll represses *hth* in the leg (Abu-Shaar & Mann, 1998; Wu & Cohen, 1999) but not in the antenna (Dong *et al.*, 2000), and *dac* in the distal leg (Dong *et al.*, 2002). Other genes regulated by Dll include *BarH1/BarH2* (Kojima *et al.*, 2000) and *Notch ligand Serrate* (*Ser*) (Rauskolb, 2001) in the leg disc. In the wing disc, *Dll* regulates expression of *achaete* (*ac*) (Campbell & Tomlinson, 1998) and *bab* (Panganiban, 2000). Dll has low DNA binding site specificity so, in order to activate specific target genes, Dll probably forms complexes with other transcription factors (Panganiban & Rubenstein, 2002). So far, from the four identified candidate cofactors, all are homeodomain proteins: two Hox proteins, Deformed (Dfd) and Antp, and two TALE proteins Exd and Hth (Panganiban, 2000). Dll cooperates with Dfd to establish ventral maxillary identity (Ohara *et al.*, 1993), and may interact with Antp to specify leg identity (Struhl, 1981). Both *hth* and *Dll* are required to establish antennal identity (Casares & Mann, 1998), and are also needed to activate antenna-specific transcription of *sal* (Dong *et al.*, 2000) and probably *dac* (Dong *et al.*, 2002).

Molecular and functional evolution

Our Basic Local Alignment Search Tool (BLAST) analysis shows that *Ceh-43* of *C. elegans* is the homolog of *Dll* and *Dlx* of arthopods and vertebrates, respectively. This *C. elegans* gene is the most similar to *Dll/Dlx*, having 31.3% and 27.7% identity with *Dll-RA* of flys and *Dlx1* of humans, respectively. The presence of a single *Dll*-like gene in *C. elegans* appears to represent the more ancestral state in Metazoa. *Ceh-43* is physically linked to the *C. elegans* Hox cluster and located on the side of the cluster corresponding to the most posteriorly expressed genes (Stock *et al.*, 1996). However, in flies, *Dll* became separated from the *Hox* cluster via a translocation, as these genes do not map to the same chromosomes (Cohen *et al.*, 1989). In the lineage leading to the Deuterostomes, *Dll* was duplicated multiple times to give rise to the six *Dlx* genes in mice and humans and eight *Dlx* genes in zebrafish. Fly and amphioxus *Dll* are most closely related to *Dlx1* (Holland *et al.*, 1996; Stock *et al.*, 1996), suggesting that *Dlx1* may have retained most of the ancestral functions of the vertebrate *Dlx* family of genes. The very large genomic size of *Dll* in the insect genomes analyzed, due for the most part to the extremely long central intron, appears to be a derived feature of the insect lineage.

Evolution of the Dll protein sequence throughout the metazoans is mainly occurring in the C-terminus following the homeodomain motif (Fig. 2A). This part of the protein is mostly varying in length across arthropod lineages. For instance, the C-terminus length in outgroups *C. elegans* and humans, ranges from 95 aa to 65–118 aa (across the six *Dlx* copies), respectively. In the spider *Cupiennius salei*, this same region is 90 aa in length, whereas in the coleopteran, *T. castaneum*, Dll is 115 aa long, and in the more derived Holometabola, Dll is at least 125 aa long.

Using the few but representative full-length Dll sequences available from the National Center for Biotechnology Information (NCBI) we attempted to estimate the phylogeny of the represented taxa using this gene. Using only the conserved Dll protein-coding regions we observed that the clades containing Diptera, Lepidoptera, Coleoptera and Eumetabola are strongly supported monophylies with at least 87% bootstrap values (Fig. 2B). The phylogeny is congruent with the inference of insect relationships based on morphological and other molecular data (Wheeler *et al.*, 2001), but lacks bootstrap support at some nodes. In addition, Myriapoda appears as the sister taxa of Chelicerata instead of Insecta, as expected, but this relationship does not have strong bootstrap support. This analysis suggests that *Dll* alone may be insufficient to resolve phylogenetic relationships among Arthropod lineages.

Dll is likely to have acquired its multiple functional roles in development in a gradual fashion. By analyzing the common roles of Dll and Dlx in protostome and deuterostome development it was proposed that the ancestral *Dll* gene may have functioned first in the developing nervous system of both invertebrates and vertebrates, acquiring roles in appendage development later in evolution, but still before the split of protostomes and deuterostomes (Panganiban & Rubenstein, 2002). According to this hypothesis, *Dll* would initially be involved in patterning structures of the peripheral nervous system when selection for sensory structures to protrude from the body wall to better sample the environment would have modified *Dll*'s role into a gene that promotes outgrowth. Later, because of *Dll*'s pre-existing association with the sensory protrusions, Dll would be a good candidate to co-opt into the PD axis patterning process of limbs. Thus, primitively Dll would have at least two main roles during development, one in the formation of peripheral sensory structures (such as setae) and another in PD axis formation (Williams *et al.*, 2002). Within the insects, this second function appears to have been co-opted into horn development in beetles, whereas a new non-outgrowth function evolved in the patterning of moth discal-cell eyespots and butterfly border eyespots. The origin of the gap gene function of Dll is unclear because this function was discovered in a single spider species. More comparative work will be required here. Also, it is still unclear when Dll's function of regulating melanin synthesis genes evolved. It may have evolved independently in butterfly and fly lineages or in a common ancestor to both lineages. *Dll/Dlx* genes have also acquired functions in patterning other organs/tissues, including the mouthparts, the auditory and olfactory systems, the hematopoietic system, and skeleton and connective tissue systems. However, it is not yet known whether these roles were acquired independently in the protostomes and deuterostomes or whether they also predate the divergence of these animal lineages.

Conclusions

The study of genomic and individual sequences of *Distal-less* showed that there is only one *Dll* copy in protostomes and two *Dll* splicing variants, RA and RB, in most species. The genome sizes of insect *Dll* are much larger than those of human and nematode homologues due, for the most part, to a large intron with a shared splicing site in protostomes and deuterostomes. In addition to the homeodomain, four additional motifs are identified to be relatively conserved across arthropods, NM-1, NM-2, HD-5′ and HD-3. Evolution of the Dll protein sequence throughout metazoans is mainly occurring in the C-terminus following the homeodomain motif.

The data reviewed here indicate that *Dll* functions as a gap gene in spiders, and is required for limb (leg, antenna, mouthparts, annalia) outgrowth and for differentiation of distal limb structures. *Dll* participates in the differentiation of the wing margin (including hairs and bristles) and of veins in the vicinity of the margin, and is a major regulator of melanin synthesis in flies and butterflies. *Dll* also plays a role in eyespot development in butterflies, horn development in beetles, nasi development in termites, and ocularium development in harvestman, outgrowths that are novel traits in the respective lineages. *Dll*, functioning as a selector gene, also plays an essential role in antennal identity. In addition, *Dll* is required for the formation of parts of the central and peripheral nervous systems. *Dll* is activated by a variety of genes, depending on the species and the developmental context, and affects a variety of downstream targets. Its complex regulation is probably due to the gradual evolution of multiple enhancers in its *cis*-regulatory region. Its multiple numbers of targets, so far all transcription factors, are the result of the evolution of Dll binding sites in these genes' *cis*-regulatory domains. Dll acts on the target genes either as a repressor

or activator of gene transcription. Target specificity appears to require the binding of Dll to additional co-factors such as the homeodomain proteins Dfd, Antp, Exd and Hth. The ancestral *Dll* may have functioned first in the developing nervous system, acquiring roles as a gap gene, and in head patterning and appendage development later in evolution.

With only a few exceptions, most of the functional work on *Dll* in arthropods has been performed on *Drosophila*. Further functional work on a variety of other arthropod lineages would be welcome to confirm the inferred ancestral functions for this gene, but mostly to determine the branches on the phylogeny where the new functions evolved. Additionally, as the regulatory code becomes better understood, and further arthropod genomes are sequenced, comparative sequence analyses alone may provide insights into *Dll*'s functional evolution both in the ontogeny of an organism and across the tree of life.

Acknowledgments

This work was supported by Par-Eu Scholars Program, and The National Science Foundation (IOB-0516705), The National Natural Science Foundation of China (31372265), Coordinated Research Project of the International Atomic Energy Agency (18268/R0), National Key Program of Science and Technology Foundation Work of China (2015FY210300), Singapore Ministry of Education (MOE) AcRF and Yale-NUS College grant numbers IG14-SI002, IG15-LR002, R-607-265-052-121, and MOE R-154-000-602-112 and R-154-000-643-112. Wrote the paper: BC, WHP, AM.

Disclosure

The authors declare no conflict of interest.

References

Abu-Shaar, M. and Mann, R.S. (1998) Generation of multiple antagonistic domains along the proximodistal axis during *Drosophila* leg development. *Development*, 125, 3821–3830.

Abzhanov, A. and Kaufman, T.C. (2000) Homologs of *Drosophila* appendage genes in the patterning of arthropod limbs. *Developmental Biology*, 227, 673–689.

Abzhanov, A., Popadic, A. and Kaufman, T.C. (1999) Chelicerate *Hox* genes and the homology of arthropod segments. *Evolution & Development*, 1,77–89.

Angelini, D.R. and Kaufman, T.C. (2005) Insect appendages and comparative ontogenetics. *Developmental Biology*, 286, 57–77.

Arnoult, L., Su, K.F., Manoel, D., Minervino, C., Magrina, J., Gompel, N. and Prud'homme, B. (2013) Emergence and diversification of fly pigmentation through evolution of a gene regulatory module. *Science*, 339, 1423–1426.

Averof, M. and Cohen, S.M. (1997) Evolutionary origin of insect wings from ancestral gills. *Nature*, 385, 627–630.

Beermann, A., Jay, D.G., Beeman, R.W., Hulskamp, M., Tautz, D. and Jurgens, G. (2001) The short antennae gene of *Tribolium* is required for limb development and encodes the orthologue of the *Drosophila* Distal-less protein. *Development*, 128, 287–297.

Beldade, P., Brakefield, P.M. and Long, A.D. (2002) Contribution of *Distal-less* to quantitative variation in butterfly eyespots. *Nature*, 415, 315–317.

Blair, S.S. (1995) Compartments and appendage development in *Drosophila*. *BioEssays*, 17, 299–309.

Bolinger, R.A. and Boekhoff-Falk, G. (2005) *Distab-less* functions in subdividing the *Drosophila* thoracic limb primordium. *Developmental Dynamics*, 232, 801–816.

Brakefield, P.M., Gates, J., Keys, D., Kesbeke, F., Wijngaarden, P.J., Monteiro, A., French, V. and Carroll, S.B. (1996) Development, plasticity and evolution of butterfly eyespot patterns. *Nature*, 384, 236–242.

Brunetti, C.R., Selegue, J.E., Monteiro, A., French, V., Brakefield, P.M. and Carroll, S.B. (2001) The generation and diversification of butterfly eyespot color patterns. *Current Biology*, 11, 1578–1585.

Campbell, G. and Tomlinson, A. (1998) The roles of the homeobox genes *aristaless* and *Distal-less* in patterning the legs and wings of *Drosophila*. *Development*, 125, 4483–4493.

Carroll, S.B., Gates, J., Keys, D.N., Paddock, S.W., Panganiban, G.E.F., Selegue, J.E. and Williams, J.A. (1994) Pattern formation and eyespot determination in butterfly wings. *Science*, 265, 109–114.

Casares, F. and Mann, R.S. (1998) Control of antennal versus leg development in *Drosophila*. *Nature*, 392, 723–726.

Casci, T. (2002) Developmental biology – *Distal-less* makes the leap. *Nature Reviews Genetics*, 3, 413–413.

Cohen, B., Simcox, A.A. and Cohen, S.M. (1993) Allocation of the thoracic imaginal primordia in the *Drosophila* embryo. *Development*, 117, 597–608.

Cohen, B., Wimmer, E.A. and Cohen, S.M. (1991) Early development of leg and wing primordia in the *Drosophila* embryo. *Mechanisms of Development*, 33, 229–240.

Cohen, S.M. (1990) Specification of limb development in the *Drosophila* embryo by positional cues from segmentation genes. *Nature*, 343, 173–177.

Cohen, S.M., Bronner, G., Kuttner, F., Jurgens, G. and Jackle, H. (1989) *Distal-less* encodes a homeodomain protein required for limb development in *Drosophila*. *Nature*, 338, 432–434.

Cohen, S.M. and Jurgens, G. (1989) Proximal distal pattern-formation in *Drosophila* — cell autonomous requirement for *Distal-less* gene activity in limb development. *EMBO Journal*, 8, 2045–2055.

Cohen, S.M. and Jurgens, G. (1990) Mediation of *Drosophila* head development by gap-like segmentation genes. *Nature*, 346, 482–484.

Curtiss, J., Halder, G. and Mlodzik, M. (2002) Selector and signalling molecules cooperate in organ patterning. *Nature Cell Biology*, 4, E48–E51.

Damen, W.G., Saridaki, T. and Averof, M. (2002) Diverse adaptations of an ancestral gill: a common evolutionary origin for wings, breathing organs, and spinnerets. *Current Biology*, 12, 1711–1716.

Depew, M.J., Liu, J.K., Long, J.E., Presley, R., Meneses, J.J., Pedersen, R.A. and Rubenstein, J.L.R. (1999) *Dlx5* regulates regional development of the branchial arches and sensory capsules. *Development*, 126, 3831–3846.

Diaz-Benjumea, F.J., Cohen, B. and Cohen, S.M. (1994) Cell interaction between compartments establishes the proximal-distal axis of *Drosophila* legs. *Nature*, 372, 175–179.

Digregorio, A., Spagnuolo, A., Ristoratore, F., Pischetola, M., Aniello, F., Branno, M., Cariello, L. and Dilauro, R. (1995) Cloning of ascidian homeobox genes provides evidence for a primordial chordate cluster. *Gene*, 156, 253–257.

Dong, P.D.S., Chu, J. and Panganiban, G. (2000) Coexpression of the homeobox genes *Distal-less* and *homothorax* determines *Drosophila* antennal identity. *Development*, 127, 209–216.

Dong, P.D.S., Chu, J. and Panganiban, G. (2001) Proximodistal domain specification and interactions in developing *Drosophila* appendages. *Development*, 128, 2365–2372.

Dong, P.D.S., Dicks, J.S. and Panganiban, G. (2002) *Distal-less* and *homothorax* regulate multiple targets to pattern the *Drosophila* antenna. *Development*, 129, 1967–1974.

Dworkin, I. (2005) Evidence for canalization of *Distal-less* function in the leg of *Drosophila melanogaster*. *Evolution & Development*, 7, 89–100.

Emerald, B.S., Curtiss, J., Mlodzik, M. and Cohen, S.M. (2003) *Distal antenna* and *distal antenna* related encode nuclear proteins containing pipsqueak motifs involved in antenna development in *Drosophila*. *Development*, 130, 1171–1180.

Estella, C., Rieckhof, G., Calleja, M. and Morata, G. (2003) The role of *buttonhead* and *Sp1* in the development of the ventral imaginal discs of *Drosophila*. *Development*, 130, 5929–5941.

Fuse, N., Hirose, S. and Hayashi, S. (1996) Determination of wing cell fate by the *escargot* and *snail* genes in *Drosophila*. *Development*, 122, 1059–1067.

Gebelein, B., Culi, J., Ryoo, H.D., Zhang, W. and Mann, R.S. (2002) Specificity of Distalless repression and limb primordia development by abdominal Hox proteins. *Developmental Cell*, 3, 487–498.

Gorfinkiel, N., Morata, G. and Guerrero, I. (1997) The homeobox gene *Distal-less* induces ventral appendage development in *Drosophila*. *Genes & Development*, 11, 2259–2271.

Gorfinkiel, N., Sanchez, L. and Guerrero, I. (1999) *Drosophila* terminalia as an appendage-like structure. *Mechanisms of Development*, 86, 113–123.

Goto, S. and Hayashi, S. (1997) Specification of the embryonic limb primordium by graded activity of Decapentaplegic. *Development*, 124, 125–132.

Grenier, J.K., Garber, T.L., Warren, R., Whitington, P.M. and Carroll, S. (1997) Evolution of the entire arthropod *Hox* gene set predated the origin and radiation of the onychophoran/arthropod clade. *Current Biology*, 7, 547–553.

Halder, G., Callaerts, P., Flister, S., Walldorf, U., Kloter, U. and Gehring, W.J. (1998) Eyeless initiates the expression of both sine oculis and eyes absent during *Drosophila* compound eye development. *Development*, 125, 2181–2191.

Hayashi, S., Hirose, S., Metcalfe, T. and Shirras, A.D. (1993) Control of imaginal cell-development by the *escargot* gene of *Drosophila*. *Development*, 118, 105–115.

Held, L.I. (1995) Axes, boundaries and coordinates – the ABCs of fly leg development. *BioEssays*, 17, 721–732.

Holland, N.D., Panganiban, G., Henyey, E.L. and Holland, L.Z. (1996) Sequence and developmental expression of *AmphiDII*, an amphioxus *Distal-less* gene transcribed in the ectoderm, epidermis and nervous system: Insights into evolution of craniate forebrain and neural crest. *Development*, 122, 2911–2920.

Inoue, Y., Mito, T., Miyawaki, K., Matsushima, K., Shinmyo, Y., Heanue, T.A., Mardon, G., Ohuchi, H. and Noji, S. (2002) Correlation of expression patterns of *homothorax*, *dachshund*, and *Distal-less* with the proximodistal segmentation of the cricket leg bud. *Mechanisms of Development*, 113, 141–148.

Jockusch, E.L., Nulsen, C., Newfeld, S.J. and Nagy, L.M. (2000) Leg development in flies versus grasshoppers: differences in *dpp* expression do not lead to differences in the expression of downstream components of the leg patterning pathway. *Development*, 127, 1617–1626.

Kango-Singh, M., Singh, A. and Gopinathan, K.P. (2001) The wings of *Bombyx mori* develop from larval discs exhibiting an early differentiated state: a preliminary report. *Journal of Biosciences*, 26, 167–177.

Kaphingst, K. and Kunes, S. (1994) Pattern-formation in the visual centers of the *Drosophila* brain – wingless acts via decapentaplegic to specify the dorsoventral axis. *Cell*, 78, 437–448.

Kenyon, K.L., Ranade, S.S., Curtiss, J., Mlodzik, M. and Pignoni, F. (2003) Coordinating proliferation and tissue specification to promote regional identity in the *Drosophila* head. *Developmental Cell*, 5, 403–414.

Khadjeh, S., Turetzek, N., Pechmann, M., Schwager, E.E., Wimmer, E.A., Damen, W.G. and Prpic, N.M. (2012) Divergent role of the *Hox* gene *Antennapedia* in spiders is responsible for the convergent evolution of abdominal limb repression. *Proceedings of the National Academy of Sciences of the United States of America*, 109, 4921–4926.

Kojima, T., Sato, M. and Saigo, K. (2000) Formation and specification of distal leg segments in *Drosophila* by dual *Bar homeobox* genes, *BarH1* and *BarH2*. *Development*, 127, 769–778.

Kubota, K., Goto, S., Eto, K. and Hayashi, S. (2000) EGF receptor attenuates Dpp signaling and helps to distinguish the wing and leg cell fates in *Drosophila*. *Development*, 127, 3769–3776.

Kubota, K., Goto, S. and Hayashi, S. (2003) The role of Wg signaling in the patterning of embryonic leg primordium in *Drosophila*. *Developmental Biology*, 257, 117–126.

Kumar, J.P. and Moses, K. (2001) EGF receptor and notch signaling act upstream of Eyeless/Pax6 to control eye specification. *Cell*, 104, 687–697.

Lecuit, T. and Cohen, S.M. (1997) Proximal-distal axis formation in the *Drosophila* leg. *Nature*, 388(6638), 139–145.

Lee, S.E. and Jacobs, D.K. (1999) Expression of *Distal-less* in molluscan eggs, embryos, and larvae. *Evolution & Development*, 1, 172–179.

Lemons, D., Fritzenwanker, J.H., Gerhart, J., Lowe, C.J. and McGinnis, W. (2010) Co-option of an anteroposterior head axis patterning system for proximodistal patterning of appendages in early bilaterian evolution. *Developmental Biology*, 344, 358–362.

Lewis, D.L., DeCamillis, M. and Bennett, R.L. (2000) Distinct roles of the homeotic genes *Ubx* and *abd-A* in beetle embryonic abdominal appendage development. *Proceedings of the National Academy of Sciences of the United States of America*, 97, 4504–4509.

Lin, X.D., Yao, Y., Jin, M.N. and Li, Q.L. (2014) Characterization of the *Distal-less* gene homologue, *NlDll*, in the brown planthopper, *Nilaparvata lugens* (Stål). *Gene*, 535, 112–118.

Lindsley, D.L. and Zimm, G.G. (1992) *The Genome of Drosophila melanogaster*. San Diego: Academic Press. viii, 1133 pp.

Liu, J.K., Ghattas, I., Liu, S.Y., Chen, S. and Rubenstein, J.L.R. (1997) *Dlx* genes encode DNA-binding proteins that are expressed in an overlapping and sequential pattern during basal ganglia differentiation. *Developmental Dynamics*, 210, 498–512.

Liubicich, D.M., Serano, J.M., Pavlopoulos, A., Kontarakis, Z., Protas, M.E., Kwan, E., Chatterjee, S., Tran, K.D., Averof, M. and Patel, N.H. (2009) Knockdown of *Parhyale Ultrabithorax* recapitulates evolutionary changes in crustacean appendage morphology. *Proceedings of the National Academy of Sciences of the United States of America*, 106, 13892–13896.

Lowe, C.J., Wu, M., Salic, A., Evans, L., Lander, E., Stange-Thomann, N., Gruber, C.E., Gerhart, J. and Kirschner, M. (2003) Anteroposterior patterning in hemichordates and the origins of the chordate nervous system. *Cell*, 113, 853–865.

Mardon, G., Solomon, N.M. and Rubin, G.M. (1994) Dachshund encodes a nuclear-protein required for normal eye and leg development in *Drosophila*. *Development*, 120, 3473–3486.

McGuinness, T., Porteus, M.H., Smiga, S., Bulfone, A., Kingsley, C., Qiu, M.S., Liu, J.K., Long, J.E., Xu, D.W. and Rubenstein, J.L.R. (1996) Sequence, organization, and transcription of the *Dlx-1* and *Dlx-2* locus. *Genomics*, 35, 473–485.

Mittmann, B. and Scholtz, G. (2001) *Distal-less* expression in embryos of *Limulus polyplremus* (Chelicerata, Xiphosura) and *Lepisma saccharina* (Insecta, Zygentoma) suggests a role in the development of mechanoreceptors, chemoreceptors, and the CNS. *Development Genes and Evolution*, 211, 232–243.

Moczek, A.P. and Nagy, L.M. (2005) Diverse developmental mechanisms contribute to different levels of diversity in horned beetles. *Evolution & Development*, 7, 175–185.

Moczek, A.P., Rose, D., Sewell, W. and Kesselring, B.R. (2006) Conservation, innovation, and the evolution of horned beetle diversity. *Development Genes and Evolution*, 216, 655–665.

Moczek, A.P. and Rose, D.J. (2009) Differential recruitment of limb patterning genes during development and diversification of beetle horns. *Proceedings of the National Academy of Sciences of the United States of America*, 106, 8992–8997.

Monteiro, A., Chen, B., Ramos, D.M., Oliver, J.C., Tong, X., Guo, M., Wang, W.K., Fazzino, L. and Kamal, F. (2013) *Distal-less* regulates eyespot patterns and melanization in *Bicyclus* butterflies. *Journal of Experimental Zoology Part B Molecular & Developmental Evolution*, 320, 321–331.

Monteiro, A., Glaser, G., Stockslagger, S., Glansdorp, N. and Ramos, D.M. (2006) Comparative insights into questions of lepidopteran wing pattern homology. *BMC Developmental Biology*, 6, 52.

Monteiro, A. and Podlaha, O. (2009) Wings, horns, and butterfly eyespots: how do complex traits evolve? *PLoS Biology*, 7, e37.

Moreno, E. and Morata, G. (1999) Caudal is the *Hox* gene that specifies the most posterior *Drosophile* segment. *Nature*, 400, 873–877.

Niwa, N., Saitoh, M., Ohuchi, H., Yoshioka, H. and Noji, S. (1997) Correlation between *distal-less* expression patterns and structures of appendages in development of the two-spotted cricket, *Gryllus bimaculatus*. *Zoological Science*, 14, 115–125.

Ohara, E., Cohen, B., Cohen, S.M. and Mcginnis, W. (1993) *Distal-less* is a downstream gene of deformed required for ventral maxillary identity. *Development*, 117, 847–856.

Oka, K., Yoshiyama, N., Tojo, K., Machida, R. and Hatakeyama, M. (2010) Characterization of abdominal appendages in the sawfly, *Athalia rosae* (Hymenoptera), by morphological and

gene expression analyses. *Development Genes and Evolution*, 220, 53–59.

Panganiban, G. (2000) *Distal-less* function during *Drosophila* appendage and sense organ development. *Developmental Dynamics*, 218, 554–562.

Panganiban, G., Irvine, S.M., Lowe, C., Roehl, H., Corley, L.S., Sherbon, B., Grenier, J.K., Fallon, J.F., Kimble, J. and Walker, M. *et al*. (1997) The origin and evolution of animal appendages. *Proceedings of the National Academy of Sciences of the United States of America*, 94, 5162–5166.

Panganiban, G., Nagy, L. and Carroll, S.B. (1994) The role of the *Distal-less* gene in the development and evolution of insect limbs. *Current Biology*, 4, 671–675.

Panganiban, G. and Rubenstein, J.L.R. (2002) Developmental functions of the *Distal-less*/*Dlx* homeobox genes. *Development*, 129, 4371–4386.

Panganiban, G., Sebring, A., Nagy, L. and Carroll, S. (1995) The development of crustacean limbs and the evolution of arthropods. *Science*, 270, 1363–1366.

Pechmann, M., Khadjeh, S., Sprenger, F. and Prpic, N.M. (2010) Patterning mechanisms and morphological diversity of spider appendages and their importance for spider evolution. *Arthropod Structure and Development*, 39, 453–467.

Pechmann, M., Khadjeh, S., Turetzek, N., McGregor, A.P., Damen, W.G. and Prpic, N.M. (2011) Novel function of *Distal-less* as a gap gene during spider segmentation. *PLoS Genetics*, 7, e1002342.

Pechmann, M. and Prpic, N.M. (2009) Appendage patterning in the South American bird spider *Acanthoscurria geniculata* (Araneae: Mygalomorphae). *Developmental Genes and Evolution*, 219, 189–198.

Plavicki, J., Mader, S., Pueschel, E., Peebles, P. and Boekhoff-Falk, G. (2012) Homeobox gene *distal-less* is required for neuronal differentiation and neurite outgrowth in the *Drosophila* olfactory system. *Proceedings of the National Academy of Sciences of the United States of America*, 109, 1578–1583.

Popadic, A., Panganiban, G., Rusch, D., Shear, W.A. and Kaufman, T.C. (1998) Molecular evidence for the gnathobasic derivation of arthropod mandibles and for the appendicular origin of the labrum and other structures. *Development Genes and Evolution*, 208, 142–150.

Postlethwait, J.H. and Schneiderman, H.A. (1971) Pattern formation and determination in antenna of homoeotic mutant *Antennapedia* of *Drosophila melanogaster*. *Developmental Biology*, 25, 606–640.

Prpic, N.M. and Damen, W.G. (2004) Expression patterns of leg genes in the mouthparts of the spider *Cupiennius salei* (Chelicerata: Arachnida). *Developmental Genes and Evolution*, 214, 296–302.

Prpic, N.M. and Tautz, D. (2003) The expression of the proximodistal axis patterning genes *Distal-less* and *dachshund* in the appendages of *Glomeris marginata* (Myriapoda: Diplopoda) suggests a special role of these genes in patterning the head appendages. *Developmental Biology*, 260, 97–112.

Punzo, C., Plaza, S., Seimiya, M., Schnupf, P., Kurata, S., Jaeger, J. and Gehring, W.J. (2004) Functional divergence between eyeless and twin of eyeless in *Drosophila melanogaster* (vol. 131, pg 3943, 2004). *Development*, 131, 4635–4635.

Quiring, R., Walldorf, U., Kloter, U. and Gehring, W.J. (1994) Homology of the *Eyeless* gene of *Drosophila* to the small eye gene in mice and aniridia in humans. *Science*, 265, 785–789.

Rauskolb, C. (2001) The establishment of segmentation in the *Drosophila* leg. *Development*, 128, 4511–4521.

Raz, E. and Shilo, B.Z. (1993) Establishment of ventral cell fates in the *Drosophila* embryonic ectoderm requires DER, the EGF receptor homolog. *Genes & Development*, 7, 1937–1948.

Reed, R.D. and Gilbert, L.E. (2004) Wing venation and *Distal-less* expression in *Heliconious* butterfly wing pattern development. *Developmental Genes and Evolution*, 214, 628–634.

Reed, R.D. and Serfas, M.S. (2004) Butterfly wing pattern evolution is associated with changes in a Notch/Distal-less temporal pattern formation process. *Current Biology*, 14, 1159–1166.

Robledo, R.F., Rajan, L., Li, X. and Lufkin, T. (2002) The *Dlx5* and *D1x6* homeobox genes are essential for craniofacial, axial, and appendicular skeletal development. *Genes & Development*, 16, 1089–1101.

Rogers, B.T., Peterson, M.D. and Kaufman, T.C. (2002) The development and evolution of insect mouthparts as revealed by the expression patterns of gnathocephalic genes. *Evolution & Development*, 4, 96–110.

Ryan, J.F., Mazza, M.E., Pang, K., Matus, D.Q., Baxevanis, A.D., Martindale, M.Q. and Finnerty, J.R. (2007) Pre-bilaterian origins of the *Hox* cluster and the *Hox* code: evidence from the sea anemone, *Nematostella vectensis*. *PLoS ONE* 2, e153.

Scholtz, G., Mittmann, B. and Gerberding, M. (1998) The pattern of *Distal-less* expression in the mouthparts of crustaceans, myriapods and insects: new evidence for a gnathobasic mandible and the common origin of Mandibulata. *International Journal of Developmental Biology*, 42, 801–810.

Schoppmeier, M. and Damen, W.G.M. (2001) Double-stranded RNA interference in the spider *Cupiennius salei*: the role of *Distal-less* is evolutionarily conserved in arthropod appendage formation. *Development Genes and Evolution*, 211, 76–82.

Sharma, P.P., Schwager, E.E., Giribet, G., Jockusch, E.L. and Extavour, C.G. (2013) *Distal-less* and dachshund pattern both plesiomorphic and apomorphic structures in chelicerates: RNA interference in the harvestman *Phalangium opilio* (Opiliones). *Evolution & Development*, 15, 228–242.

Shimamoto, T., Ohyashiki, K. and Takeshita, K. (2000) Overexpression of the homeobox gene *DLX-7* inhibits apoptosis by induced expression of intercellular adhesion molecule-1. *Experimental Hematology*, 28, 433–441.

Stock, D.W., Ellies, D.L., Zhao, Z.Y., Ekker, M., Ruddle, F.H. and Weiss, K.M. (1996) The evolution of the vertebrate *Dlx* gene family. *Proceedings of the National Academy of Sciences of the United States of America*, 93, 10858–10863.

Struhl, G. (1981) A homeotic mutation transforming leg to antenna in *Drosophila*. *Nature*, 292, 635–638.

Suzanne, M., Estella, C., Calleja, M. and Sanchez-Herrero, E. (2003) The *hernandez* and *fernandez* genes of *Drosophila* specify eye and antenna. *Developmental Biology*, 260, 465–483.

Suzuki, Y. and Palopoli, M.F. (2001) Evolution of insect abdominal appendages: are prolegs homologous or convergent traits? *Development Genes and Evolution*, 211, 486–492.

Suzuki, Y., Squires, D.C. and Riddiford, L.M. (2009) Larval leg integrity is maintained by *Distal-less* and is required for proper timing of metamorphosis in the flour beetle, *Tribolium castaneum. Developmental Biology*, 326, 60–67.

Tabata, T. and Takei, Y. (2004) Morphogens, their identification and regulation. *Development*, 131, 703–712.

Thomas, R.H. and Telford, M.J. (1999) Appendage development in embryos of the oribatid mite *Archegozetes longisetosus* (Acari, Oribatei, Trhypochthoniidae). *Acta Zoologica*, 80, 193–200.

Toga, K., Hojo, M., Miura, T. and Maekawa, K. (2012) Expression and function of a limb-patterning gene *Distal-less* in the soldier-specific morphogenesis in the nasute termite *Nasutitermes takasagoensis*. *Evolution & Development*, 14, 286–295.

Tong, X.L., Hrycaj, S., Podlaha, O., Popadic, A. and Monteiro, A. (2014) Over-expression of *Ultrabithorax* alters embryonic body plan and wing patterns in the butterfly *Bicyclus anynana*. *Developmental Biology*, 394, 357–366.

Turetzek, N., Pechmann, M., Schomburg, C., Schneider, J. and Prpic, N.M. (2016) Neofunctionalization of a duplicate *dachshund* gene underlies the evolution of a novel leg segment in arachnids. *Molecular Biology and Evolution*, 33, 109–121.

Urbach, R. and Technau, G.M. (2003) Early steps in building the insect brain: neuroblast formation and segmental patterning in the developing brain of different insect species. *Arthropod Structure & Development*, 32, 103–123.

Vachon, G., Cohen, B., Pfeifle, C., Mcguffin, M.E., Botas, J. and Cohen, S.M. (1992) Homeotic genes of the bithorax complex repress limb development in the abdomen of the *Drosophila* embryo through the target gene *Distal-less*. *Cell*, 71, 437–450.

Weatherbee, S.D., Nijhout, H.F., Grunert, L.W., Halder, G., Galant, R., Selegue, J. and Carroll, S. (1999) *Ultrabithorax* function in butterfly wings and the evolution of insect wing patterns. *Current Biology*, 9, 109–115.

Weigmann, K. and Cohen, S.M. (1999) Lineage-tracing cells born in different domains along the PD axis of the developing *Drosophila* leg. *Development*, 126, 3823–3830.

Weihe, U., Dorfman, R., Wernet, M.F., Cohen, S.M. and Milan, M. (2004) Proximodistal subdivision of *Drosophila* legs and wings: the *elbow-no ocelli* gene complex. *Development*, 131, 767–774.

Wheeler, W.C., Whiting, M., Wheeler, Q.D. and Carpenter, J.M. (2001) The phylogeny of the extant hexapod orders. *Cladistics–the International Journal of the Willi Hennig Society*, 17, 113–169.

White, R.A.H., Aspland, S.E., Brookman, J.J., Clayton, L. and Sproat, G. (2000) The design and analysis of a homeotic response element. *Mechanisms of Development*, 91, 217–226.

Williams, T.A. (1998) *Distal-less* expression in crustaceans and the patterning of branched limbs. *Development Genes and Evolution*, 207, 427–434.

Williams, T.A. and Nagy, L.M. (1996) Comparative limb development in insects and crustaceans. *Seminars in Cell & Developmental Biology*, 7, 615–628.

Williams, T.A., Nulsen, C. and Nagy, L.M. (2002) A complex role for *Distal-less* in crustacean appendage development. *Developmental Biology*, 241, 302–312.

Wu, J. and Cohen, S.M. (1999) Proximodistal axis formation in the *Drosophila* leg: subdivision into proximal and distal domains by *Homothorax* and *Distal-less*. *Development*, 126, 109–117.

Zerucha, T. and Ekker, M. (2000) *Distal-less*-related homeobox genes of vertebrates: Evolution, function, and regulation. *Biochimie et Biologie Cellulaire*, 78, 593–601.

Zheng, Z., Khoo, A., Fambrough, D., Garza, L. and Booker, R. (1999) Homeotic gene expression in the wild-type and a homeotic mutant of the moth *Manduca sexta*. *Development Genes and Evolution*, 209, 460–472.

Accepted February 4, 2016