Int. J. Dev. Biol. UNCORRECTED PROOF
https://doi.org/10.1387/ijdb.180270em

THE INTERNATIONAL JOURNAL OF DEVELOPMENTAL BIOLOGY
www.intjdevbiol.com

Regulatory landscape of the Hox transcriptome

ANA CASACA[#,1], GABRIEL M. HAUSWIRTH[#,2], HEIDI BILDSOE[#,2], MOISÉS MALLO[*,1] and EDWINA MCGLINN[*,2]

[1]*Instituto Gulbenkian de Ciência, Oeiras, Portugal and* [2]*EMBL Australia, Australian Regenerative Medicine Institute, Monash University, Clayton, Australia*

ABSTRACT Precise regulation of *Hox* gene activity is essential to achieve proper control of animal embryonic development and to avoid generation of a variety of malignancies. This is a multilayered process, including complex polycistronic transcription, RNA processing, microRNA repression, long noncoding RNA regulation and sequence-specific translational control, acting together to achieve robust quantitative and qualitative Hox protein output. For many such mechanisms, the *Hox* cluster gene network has turned out to serve as a paradigmatic model for their study. In this review, we discuss current knowledge of how the different layers of post-transcriptional regulation and the production of a variety of noncoding RNA species control Hox output, and how this shapes formation of developmental systems that are reproducibly patterned by complex Hox networks.

KEY WORDS: *Hox gene, post-transcriptional regulation, polycistronic, lncRNA, microRNA*

Introduction

Hox genes are key regulators of patterning processes during animal embryonic development, also playing important roles in a variety of physiological and pathological processes in the adult animal. The high conservation in function and sequence across phyla makes *Hox* genes a central component of the so called 'Evo-Devo gene toolkit'. Indeed, it is thought that changes in their spatio-temporal expression during development are at the core of the mechanisms shaping both the highly conserved anatomical structure within a given species and the morphological diversity observed across different animal phyla.

In almost all bilaterian species, *Hox* genes are organized in clusters, although the number of genes and specific cluster organization varies widely among species (Duboule, 2007). It is thought that the *Hox* clusters of extant animals originated from a common ancestral single cluster that evolved differently through the various branches of the animal phylogenetic tree (Garcia-Fernàndez, 2005). One of the best studied invertebrate example is *Drosophila*, where the ancestral cluster was split into two complexes known as the Antennapedia and Bithorax complexes (ANT-C and BX-C respectively). In vertebrates, on the other hand, the ancestral cluster underwent a series of duplications followed by gene loss and compaction to generate a variety of cluster arrangements. Mice, for instance, have 39 *Hox* genes distributed in four clusters, each containing a unique subset of 13 *Hox* gene paralogs. Teleost lineages have additional sets of *Hox* clusters resulting from an additional round of genome duplication. The best studied case is that of the zebrafish that contains seven recognizable *Hox* clusters.

Although classic studies of *Hox* gene regulation and activity were mostly centered on the protein-coding genes, it was soon recognized that transcription within the *Hox* clusters was not limited to those transcriptional units (Lipshitz *et al.*, 1987). The improvement of sequencing capabilities revealed that the portion of noncoding transcripts produced from the *Hox* clusters is far greater than initially appreciated (Mainguy *et al.*, 2007; Rinn *et al.*, 2007). These studies also showed that the noncoding transcripts, which include both microRNAs (miRNAs) and long noncoding RNA (lncRNA) transcribed in both sense and antisense directions, are produced from the *Hox* clusters of all species analyzed. Indeed, some of these noncoding elements have been conserved together with the *Hox*-coding genes throughout evolution, further supporting the common origin of the clusters and revealing evolutionary routes affecting the fate of particular *Hox* clusters. For instance,

Abbreviations used in this paper: HD, homeodomain; lncRNA, long noncoding RNA; mRNA, messenger RNA; miR, microRNA; miRNA, microRNA; LR, linker region; UTR, untranslated region; ORF, open reading frame; uORF, upstream open reading frame; IRES, internal ribosomal entry site; TSS, transcriptional start site; TIE, translational inhibitory element.

***Address correspondence to:** Moisés Mallo. Instituto Gulbenkian de Ciência, Rua da Quinta Grande 6, 2780-156 Oeiras, Portugal. Tel: +351-214464624. Fax: +351-214407970. e-mail: mallo@igc.gulbenkian.pt - web: http://www.igc.gulbenkian.pt/mmallo - https://orcid.org/0000-0000-0000-0000 or Edwina McGlinn. Australian Regenerative Medicine Institute, Monash University, Wellington Road, Clayton, 3800, Australia. Tel: +613-9902-9721. Fax: +613-9902-9727. e-mail: edwina.mcglinn@monash.edu - http://www.armi.org.au/about/our-people/edwina-mcglinn - https://orcid.org/0000-0000-0000-0000

#Note: these authors contributed equally to the work

Submitted: 27 August, 2018; *Accepted:* 13 September, 2018.

ISSN: Online 1696-3547, Print 0214-6282

the presence of one such element, a lone Hox-embedded miRNA, informs of the transient presence of the missing eighth cluster in the zebrafish genome (Woltering and Durston, 2006).

In this review we will take a closer look at the different RNA species transcribed from *Hox* clusters, focusing on what can be considered non-canonical transcripts. We will review alternative processing events that occur on *Hox* gene mRNAs and describe other types of RNA transcripts arising from the *Hox* clusters, such as miRNAs or lncRNAs. We will also try to summarize current views about the possible impact that these RNA species might have on *Hox* gene expression and function.

Canonical and non-canonical *Hox* mRNAs

While outnumbered by noncoding transcripts, *Hox* mRNAs are considered the major drivers of Hox function by specifying the amino acid sequence of protein effectors. After transcription, Hox pre-mRNAs undergo important processing events that can diversify their output and regulate quantitative protein production. For instance, alternative splicing can generate mRNA variants coding different protein products. Also, relevant information is encoded in the untranslated regions (UTRs) that include sequences regulating both stability and translational efficiency of the mRNAs. In this section, we will describe some of the regulatory mechanisms acting on *Hox* mRNAs that bear the potential to impact Hox protein function and expression profiles, thereby expanding the coding capacity of the *Hox* genome.

Alternative splicing of Hox *mRNAs originate protein isoforms*

In contrast to the complex architecture of the *Hox* clusters, individual *Hox* genes usually show a remarkably simple organization. With the exception of the presence of microexons in some species, *Hox* genes generally contain a single intron, splitting the gene in two exons (Exon I-Intron-Exon II), with the 3' exon coding for the homeodomain (HD). Despite the apparent simplicity of their structure, a series of molecular analyses have shown that *Hox* genes may undergo complex processes of alternative intron-splicing.

The *Drosophila Ubx* gene is a well-studied case of alternative splicing. The *Ubx* transcriptional unit includes, in addition to the canonical Exon I and Exon II, two additional internal microexons (O'Connor *et al.*, 1988; Kornfeld *et al.,* 1989). The presence of these extra microexons, together with two alternative splice donor sites in Exon I, can originate six different Ubx coding sequences (Fig. 1A) (O'Connor *et al.*, 1988; Kornfeld *et al.,* 1989; Reed *et al.,* 2010). Notably, the characterization of *Ubx* transcript maturation led to the discovery of a new concept of splicing, known as recursive splicing. Recursive splicing represents a mechanism whereby portions of a large intron can be spliced out sequentially, rather than in a single excision step (Hatton *et al.*, 1998; Burnette *et al.*, 1999). While other examples of recursive splicing have been identified in diverse transcriptional units of both invertebrates and vertebrates, in the *Hox* environment *Ubx* seems to be the only gene undergoing this process, despite the presence of microexons in other *Hox* genes.

Canonical RNA splicing events have been shown to generate four different coding sequences from the *Drosophila Antp* gene, through the inclusion or exclusion of microexons coding for 13 and 4 amino acids, located between the two major exons (Laughon *et al.,* 1986; Stroeher *et al.*, 1986; Bermingham and Scott, 1988). Remarkably, the alternative splicing processes identified in the *Antp* and *Ubx* genes introduce variations in the linker region of the protein (LR; the region connecting the hexapeptide motif to the HD), suggesting functional significance, as the LRs contribute actively to the functional properties of *Drosophila* Hox proteins (Gebelein *et al.*, 2002; Merabet *et al.*, 2003; Reed *et al.*, 2010; Papadopoulos *et al.*, 2011; Navas *et al.*, 2011). Interestingly, it has been reported that these alternatively spliced transcripts might be distributed according to temporally regulated and tissue specific patterns (Lopez and Hogness, 1991; Lopez *et al.,* 1996), further suggesting their functional relevance during *Drosophila* development.

Alternative splicing events have also been described for other

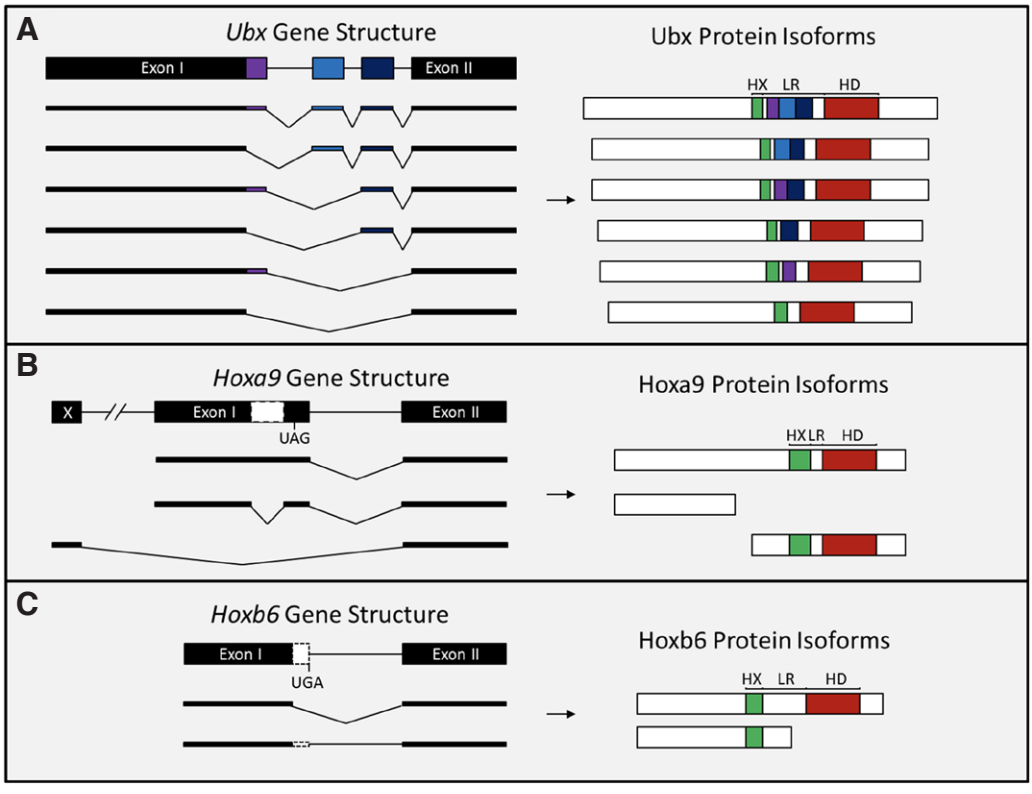


Fig. 1. Schematics of gene structure of *Ubx* (A), *Hoxa9* (B) and *Hoxb6* (C) showing the transcripts splice variants and the corresponding protein isoforms (not to scale). *The principal exons are represented in black. Microexons of* Ubx *are in blue and dark blue and the element spliced out from the recursive splicing event is in purple. Dotted-outline white boxes represent alternatively spliced elements within exons/introns. Stop codons (UAG and UGA) are indicated. HX, hexapeptide; HX', modified hexapeptide; LR, linker region; HD, homeodomain; X, exon X.*

Drosophila Hox genes, such as *Pb*, *Lab* and *Abd-B* (Cribbs *et al*., 1992; Mlodzik *et al*., 1988; Kuziora and McGinnis, 1988), although they have been studied in less detail. Altogether, these data clearly suggest that alternative splicing has the potential to play a relevant role in the modulation of Hox protein function.

Alternative splicing events have also been described for vertebrate *Hox* genes, but they seem to produce isoforms different from those observed in the fly. For instance, despite the essential role that the LR plays in the rib-promoting function of the mouse Hoxb6 protein (Casaca *et al*., 2016), so far no splice variant affecting the LR size of this, or other vertebrate Hox protein, has been described. Conversely, the most typical splice variants observed for vertebrate *Hox* genes affect the production of protein products containing or lacking the HD. This is the case of both *Hoxa1* and *Hoxa9* where a portion of Exon I can undergo an additional splicing event, causing a shift in the reading frame that would produce truncated versions of the Hox proteins lacking the HD (Fig. 1B) (LaRosa and Gudas, 1988; Fujimoto *et al*., 1998). Importantly, the *Hoxa9* transcript coding for the truncated protein has been detected in mouse, chicken and human tissues at both embryonic and adult stages (Dintilhac *et al*., 2004).

A possible role for the truncated Hoxa1 variant was investigated through biochemical approaches (Fernandez and Gudas, 2009). These *in vitro* studies showed that the truncated Hoxa1 variant interacts directly with the full-length isoform, and indirectly with its cofactor Pbx1, and that this interaction could interfere with the activity of the full-length Hoxa1 during stem cell differentiation. An equivalent role for the truncated form of Hoxa9 was suggested by the observation that it can bind the Hoxa9 cofactor CBP (CREB-binding protein) as efficiently as the full-length Hoxa9 (Dintilhac *et al*., 2004), which was consistent with a dominant-negative function for the short isoform. The dominant-negative activity of the truncated Hoxa9 protein is, however, not consensual as two other studies proposed that this Hoxa9 isoform can assist the full length protein in its leukemogenic potential (He *et al*., 2012 and Stadler *et al*., 2014).

Additional alternatively spliced mRNAs from the *Hoxa9* gene have been isolated from both mouse and human samples, containing an extra 5' coding exon (Exon X) directly linked to Exon II (Fig. 1B) (Borrow *et al*., 1996; Kim *et al*., 1998). Similarly, a non-canonical *Hoxa10* mRNA was isolated from E15.5 mouse liver, containing an extra 5' exon (Exon X) linked to the canonical Exon II of the *Hoxa10* gene. Such mRNAs from both *Hoxa9* and *Hoxa10* containing this alternative exon combination have open reading frames (ORFs) coding for Hox protein variants composed of the HD linked to a small N-terminal region. Whether these proteins are indeed synthesized, as well as their possible functional roles, remain to be investigated. In this regard, we have previously shown that a mutant Hoxa10 lacking most of its N-terminal region is unable to block rib formation in the mouse axial skeleton (Guerreiro *et al*., 2012), which contrasts with the strong rib-repressing properties of the full-length protein (Carapuço *et al*., 2005).

Other studies have suggested that Hox protein variants lacking the HD might also exist, but resulting from a mechanism distinct from the ones described above. In particular, truncated proteins from HOXA1 and Hoxb6 would originate not from an extra splicing event, but rather from the unspliced variant of the mRNA, still containing the intron (Chariot *et al*., 1995; Shen *et al*., 1991). The identification of unspliced *Hoxb6* transcripts at levels that can be detected by Northern blotting or RNAse protection assays indicated that a considerable part of the original transcript might not be immediately spliced. Furthermore, these unspliced transcripts are distributed according to specific spatiotemporal expression patterns in both embryonic tissues and adult organs of mouse and human origin (Shen *et al*., 1991; Mathews *et al*., 1993). The presence of an in frame UGA stop codon in the intron just six nucleotides past the splice junction would result in the production of a truncated protein similar to the N-terminal region of Hoxb6 but lacking the homeodomain (Fig. 1C). Support for the translation of such truncated HOXB6 protein was obtained by Komuves and colleagues (2000) when studying the differentiation of human epidermis. Using two different antibodies, one recognizing both protein isoforms and the other specific for the HD-containing isoform, immunostaining experiments on skin samples indicated that the truncated HOXB6 protein was expressed in the cytoplasm of undifferentiated keratinocytes, while the full-length protein was induced upon differentiation and entered the cell nucleus (Komuves *et al.,* 2000). This supports the hypothesis that alternatively-spliced *Hox* mRNAs can generate distinct proteins.

While the above-discussed observations seem to indicate that a HOXB6 truncated isoform is indeed synthesized in animal tissues, conclusive evidence for the actual occurrence of isoforms of other Hox proteins is still missing. In particular, the studies evaluating the properties of the different Hox isoforms that we have reviewed here (such as the Hoxa9 and Hoxa1 truncated proteins), were performed in stable cell lines forced to express the truncated proteins using constructs containing the already spliced cDNA (Fernandez and Gudas, 2009; Stadler *et al*., 2014). In addition, the authors reported some technical problems in identifying Hox proteins other than the full-length (Fernandez and Gudas, 2009), further complicating the interpretation of those experiments. Therefore, additional and more conclusive experimental data concerning the protein products derived from variant mRNAs, as well as the biological contexts where they are synthesized, will be instrumental to clarify the full potential of *Hox* genes.

Polycistronism increases Hox mRNA diversity

The *Hox* cluster architecture, with tandemly organized genes, provides a suitable environment for the production of polycistronic transcripts. The first case of Hox polycistronism was reported almost 30 years ago, with the observation that the mouse *Hoxc4, Hoxc5* and *Hoxc6* genes may be transcribed as part of a single unit (Simeone *et al*., 1988). This composite transcript would then be alternatively processed to give rise to three mature mRNAs with common 5'UTRs but containing the ORFs coding for the different Hox proteins.

Over the years, analyses of transcriptional profiling have identified a considerable number of polycistronic RNAs synthesized from the mammalian *Hox* clusters (Mainguy *et al*., 2007; Brunskill and Potter, 2012). A well-studied case involves the mouse *Hoxa5* locus (Coulombe *et al*., 2010). During embryonic development, overlapping transcripts are synthesized from this region of the *HoxA* cluster, due to the use of multiple promoters and distinct splicing events. Importantly, some of the resulting transcripts are bicistronic, containing both the *Hoxa6* and *Hoxa5* coding sequences. Experiments in cultured cells showed that translation from this polycistronic transcript was restricted to the *Hoxa6* ORF, i.e. the one located closer to the 5' end of the transcript. Since multiple

translation events seem not be occurring from this specific *Hox* polycistronic transcript, it might represent an alternative regulatory strategy to control *Hox* gene transcription, eventually following a mechanism similar to those associated with lncRNAs.

Events of alternative splicing in polycistronic transcripts may also produce hybrid Hox protein-coding sequences. A particular case of such hybrid Hox transcript was described by Benson *et al.,* (1995) who detected a transcript containing the 5' exon of *Hoxa10* spliced to the HD-containing exon of the adjacent *Hoxa9* gene. Exon sharing was also described for zebrafish with the identification of an exon located far upstream of the *hoxb4a* locus used by both *hoxb4a* and *hoxb3a* genes (Hadrys *et al*., 2004) to produce distinct transcripts upon being spliced to either *hoxb4* or *hoxb3* exons (Hadrys *et al*., 2004; Hadrys *et al*., 2006). However, translation into the corresponding proteins was not directly analyzed, and thus the relevance of exon sharing between Hox genes is still to be clarified.

There is no evidence of polycistronism in *Drosophila*, probably due to the widely spaced nature of *Hox* genes within the fly clusters, contrary to the tightly clustered vertebrate *Hox* genes (Duboule, 2007). Nevertheless, bicistronic *Antp/Ubx* transcripts have been described in other arthropods, such as crustaceans (Shiga *et al*., 2006) and myriapods (Janssen and Budd, 2010), which suggests their possible importance in diverse arthropod lineages.

Untranslated regions in the control of Hox protein levels

A close analysis of published data shows that the patterns of Hox protein expression do not always correlate with the expression of the corresponding mRNAs (Nelson *et al*., 1996; Brend *et al*., 2003), suggesting that *Hox* gene regulation might also occur at the translational level. Some of this regulation relies on specific characteristics of the untranslated regions (UTRs) of their mRNAs, which have been shown to be instrumental in post-transcriptional regulatory processes. 3'UTRs, for instance, are frequent miRNAs targets, involved in the modulation of mRNA stability and protein translation (reviewed in Bartel, 2018; Section 4), suggesting that differential mRNA processing can lead to the formation of 3'UTRs carrying substantially distinct instructions.

The *Drosophila Ubx* transcripts represent an interesting example of how alternative 3'UTR RNA-processing influences gene regulation. Thomsen and collaborators (2010) observed that during *Drosophila* development, *Ubx* mRNAs are polyadenylated at two alternative positions, producing transcripts with different 3'UTRs (short and long isoforms). Remarkably, the transcripts with shorter 3'UTR are less susceptible to miRNAs regulation (Bender 2008; Thomsen *et al*., 2010) and therefore more stable. Together with the finding that short and long isoforms exhibit distinct spatiotemporal expression patterns during fly development, this clearly suggests an important contribution of 3'UTR processing in the regulation of *Ubx* gene expression. Moreover, *Antp*, *Abd-A* or *Abd-B* can also generate transcripts with various 3'UTR lengths (O'Connor *et al*., 1988; Sánchez-Herrero and Crosby, 1988; Celniker *et al*; 1989), each with a distinct spatiotemporal distribution in the embryo (Thomsen *et al*., 2010), suggesting that gene regulation by alternative 3'UTRs, might be a general strategy adopted for *Drosophila Hox* genes.

For mammalian *Hox* genes, there is no evidence of coordination between polyadenylation and 3'UTR lengthening. However, it has been reported that in general 3'UTRs tend to be longer as embryonic development progresses (Ji *et al*., 2009). Also, alternative polyadenylation sites were identified in the *Hoxa4* and *Hoxa7* 3'UTRs (Yekta *et al*., 2008). While some experimental validation is needed, it is tempting to speculate a general regulatory mechanism in *Hox* genes involving coordination of processes affecting alternative polyadenylation and miRNAs.

5'UTRs can also contain several regulatory elements that affect the translation of transcripts. In general, these consist of a variety of *cis*-regulatory sequences, including binding sites for regulatory factors, upstream ORFs (uORFs), internal ribosomal entry sites (IRES) or translational inhibitory elements (TIEs) (Xue and Barna, 2015). For this reason, variations in 5'UTRs can also function as important switches to regulate protein translation.

Multiple promoters for a given gene, which is a common theme among *Hox* genes, can generate transcripts containing the same ORF but different 5'UTR lengths and characteristics. This seems particularly relevant for *Hox* mRNAs, since their 5'UTRs often contain multiple uORFs, known to interfere with cap-dependent translation (Kozak *et al*., 2005). The various 5'UTRs may diverge in the number of uORFs, and therefore have differential impact on protein production from their main coding region.

In addition to the canonical cap-dependent translation, protein synthesis from eukaryotic mRNAs can be driven by internal ribosomal entry sites (IRES) in their 5'UTR. IRES were initially described as part of the mechanism of picornaviruses to take control of the cell's translational machinery (Pelletier and Sonenberg, 1988). Shortly after the discovery of this element, studies performed in *Drosophila* cells and transgenic flies described the occurrence of IRES in the long 5′UTRs of *Antp* and *Ubx* (Oh *et al*., 1992; Ye *et al*., 1997). Notably, the IRES activities associated with these *Hox* genes are developmentally regulated (Ye *et al*., 1997), indicating that translational control may be a strategy adopted by *Drosophila* to modulate Hox protein production. Unfortunately, the biological relevance or mechanistic details of the *Drosophila Hox* IRES has so far not been further evaluated.

In the mouse, however, recent data showed that IRES might play important roles in *Hox* mRNA translation. Work from the Barna group led not only to the discovery of IRES in the 5'UTRs of *HoxA* mRNAs, but also showed that these IRES could enable translation of *Hox* transcripts in areas where cap-dependent translation was locally inhibited (Xue *et al*., 2015). It is important to mention that a previous study had proposed a regulatory function for the ribosome in the differential translational control of *Hox* genes. In particular, it was shown that deficiency in ribosomal protein L38 (RPL38) led to reduced translation of a subset of *Hox* mRNAs, while global protein synthesis was not perturbed (Kondrashov *et al*., 2011). These observations provided the basis to explain the patterning defects of mouse mutants for the Rpl38 gene, which resembled those associated with mutations in *Hox* genes. In addition, they suggested differential requirement of various *Hox* transcripts for RPL38 integrity within the ribosome (Kondrashov *et al*., 2011). Interestingly, Xue *et al.,* (2015) have demonstrated that Rpl38-dependent translation of *Hox* mRNAs relies on the presence of IRES sequences within their 5'UTRs.

Another important finding from this work was that *Hox* mRNAs have the ability to inhibit cap-dependent translation in order to favor IRES-mediated translation. This property seems to depend on specific inhibitory sequences known as translational inhibitory elements (TIEs), that have been identified in the 5'UTRs of IRES-

containing *Hox* mRNAs. However, the mechanism underlying TIE activity is still unclear. One possibility is suggested by a report showing that 4EHP (mRNA cap-binding eIF4E-related protein) contributes to the repression of *Hoxb4* mRNA translation by binding to the RNA-binding-protein Prep1 (Villaescusa *et al*., 2009). Prep1 would act as a bridge linking the 3'UTR of the *Hoxb4* mRNA to 4EHP that in turn would bind to the 5' cap structure. Recently, a more general role of 4EHP as translational repressor was reported by showing that it makes part of the miRNA-dependent translational repressor machinery complex promoting the formation of a closed-loop structure that approximates the mRNA 3'UTR to the 5' cap, blocking translation (Chapat *et al*., 2017). Whether *Hox* TIEs participate in this or a similar mechanism to inhibit cap-dependent mRNA translation, still needs experimental evaluation.

Phua *et al.,*2011) described the presence of an uncapped *Hox* mRNA species. These researchers showed that when the *Hoxd4* mRNAs is first produced, it can either be capped and spliced, or spliced and then cleaved by the Drosha RNAse, in coordination with the release of a miRNA from the 5' region of the transcript. Quite interestingly, the shorter 5'UTR is not capped, which could hinder the cap-dependent translation of this mRNA. Moreover, contrary to other *Hox* mRNAs, the 5'UTR of this uncapped transcript does not contain IRES activity. Further experimental analysis may help to elucidate whether Drosha processing represents an additional strategy to regulate *Hox* gene expression.

Hox-associated long noncoding RNA

In the last decade, it has become evident that *Hox* clusters are heavily enriched in long noncoding RNAs (lncRNAs), that is, transcripts of >200 nucleotides with no discernable protein-coding potential (Rinn *et al*., 2007; Sessa *et al*., 2007; Mainguy *et al*., 2007; Sasaki *et al*., 2007; De Kumar *et al*., 2015). Some of the first lncRNAs were described within the *bithoraxoid* (*bxd*) region of *Drosophila* BX-C (Lipshitz *et al*., 1987; Sanchez-Herrero and Akam, 1989; Cumberledge *et al*., 1990). Subsequently, sequencing and *in situ* hybridization screening have demonstrated pervasive transcription of spatially restricted lncRNAs from both strands of the BX-C cluster (Bae *et al*., 2002, Petruk *et al*., 2006; Pease *et al*., 2013) and emerging evidence of lncRNAs arising from the ANT-C cluster (Pettini and Ronshaugen, 2016). Similarly in vertebrates such as humans and mouse, spliced unannotated transcripts are broadly transcribed from all four *Hox* clusters (Mainguy *et al*., 2007, reviewed De Kumar and Krumlauf, 2016), with many of these exhibiting evolutionary conservation suggestive of function. Considerable effort has been made to understand whether this extensive noncoding transcriptional output represents noise associated with clustered loci harbouring strong enhancers, or rather, denotes a mechanism that adds to the regulatory toolbox controlling Hox output. Accumulating evidence, at least at the molecular level, supports the latter hypothesis. In Fig. 2 we summarize the

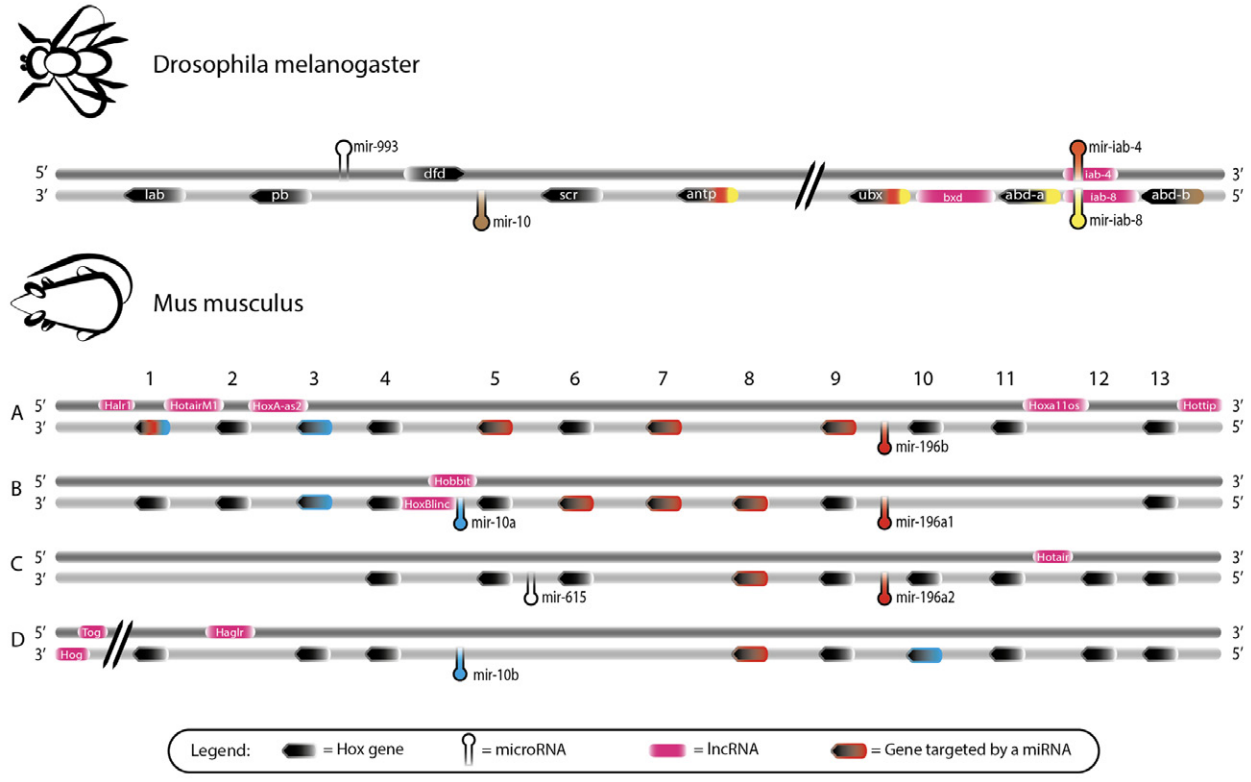


Fig. 2. Schematic illustration of the *Drosophila* and mouse *Hox* clusters showing the locations of miRNAs and lncRNAs relative to *Hox* genes (not to scale). *The* Drosophila Hox *genes are identified by the naming inside the features (black arrows). Mouse* Hox *genes are identified by their respective numbering and lettering along the axes of the* Hox *clusters. Depicted are both sense (dark grey) and antisense (light grey) DNA strands. miRNAs and their respective targets are color-coded:* Drosophila miR-10 *brown,* miR-iab-4 *red,* miR-iab-8 *yellow; Mouse* miR-10 *paralogs blue,* miR-196 *paralogs red. miRNA targets include those currently predicted in Targetscan 7.1 and those with published experimental support. Regions where lncRNAs originate are colored in pink.*

lncRNAs located within or flanking the different *Hox* loci that have been shown to regulate *Hox* expression. Various modes of action are associated with these lncRNAs, including consequences relating to the process of transcription of the gene/loci itself, functionality of the noncoding RNA transcript and/or enhancer function. In this section, we will use key examples to demonstrate the breadth of mechanisms utilized by Hox-associated lncRNAs.

LncRNA: modes of action

Hox-associated lncRNAs have been shown to act in *cis* or *trans*, implying that they affect *Hox* genes from the same or different clusters, respectively. Clear *trans*-acting examples include the minority of Hox-associated lncRNAs that produce miRNAs (the relevance of which will be further explored in the next section). It is important to note however, that many Hox-embedded miRNAs are produced from long transcripts, such as *Drosophila iab-8*, which spans at least 120Kb (Bae *et al*., 2002; Bender, 2008). From a cellular perspective, the generation of such long transcripts to merely obtain short miRNA molecules seems an inefficient strategy to adopt. It is therefore possible that miRNA-harbouring lncRNAs may have additional regulatory roles yet to be uncovered. A more controversial *trans*-acting example is that of *HOTAIR*, a lncRNA generated from the intergenic region between *HOXC11* and *HOXC12*, first identified in humans (Rinn *et al*., 2007) and shown to be conserved in mice (Schorderet and Duboule, 2011) (Fig. 2). Numerous studies demonstrate that *HOTAIR*/*Hotair* binds to the Polycomb repressive complex 2 (PRC2) in cells from distal anatomical regions (i.e. foot fibroblasts or tail tip fibroblasts) and also indicate that this lncRNA functions in *trans* to repress expression of *HoxD* genes (Rinn *et al*., 2007; Khalil *et al*., 2009; Tsai *et al*., 2010). Supporting this, *Hotair* inactivation in mice by generating a 4 kb genomic deletion led to derepression of *HoxD* gene expression, minor limb malformations and homeotic transformation of the last lumbar vertebral element (Li *et al*., 2013), the latter being consistent with the expected phenotype resulting from mild precocious posterior *Hox* gene activation (Boulet and Capecchi, 2002). The conclusions from this study have recently been challenged. Using the same *Hotair* deletion mouse mutant on an outbred background, and performing genomic analyses on embryonic tissue instead of tail tip fibroblasts, the authors did not observe the same molecular or morphological outcomes (Amandio *et al*., 2016), indicating that *Hotair* does not regulate lumbo-sacral patterning, nor *HoxD* suppression in *trans.* Therefore, *Hotair* appears to have, at best, a relatively minor role in early embryonic development, and the function of *Hotair* as a *trans* repressor of *HoxD* expression remains debated.

Many *cis*-acting Hox-associated lncRNAs have been reported in both vertebrates and invertebrates. The *Drosophila* sense-transcribed *bxd* lncRNA has been shown to repress *Ubx* expression, likely by transcriptional interference (Petruk *et al*., 2006; Pease *et al*., 2013). However, loss of *bxd* expression by promoter inversion did not result in overt phenotypic alterations (Pease *et al*., 2013), a rather surprising finding given the complex array of alternatively spliced lncRNAs produced from the locus. In the mouse, four antisense transcripts have been identified initiating either upstream or from within the first exon of *Hoxa11* (*Hoxa11as-a* and *Hoxa11as-b,* respectively) (Hsieh-Lie *et al*., 1995; Potter and Branford, 1998; Kherdjemil *et al*., 2016). Loss of *Hoxa11as-b* expression led to an expansion of *Hoxa11* sense transcription into the distal part (autopod) of the developing mouse limb (Kherdjemil *et al*., 2016), a region normally devoid of *Hoxa11* expression (Hsieh-Lie *et al*., 1995). Intriguingly, proximal restriction of *Hoxa11* is not observed in the fin buds of teleost and other fish (Sakamoto *et al*., 2009; Davis *et al*., 2007, Metscher *et al*., 2005), which correlates with the reported absence of an enhancer shown to be important for *Hoxa11as-b* expression (Kherdjemil *et al*., 2016). Combined with the observation that ectopic distal *Hoxa11* expression results in polydactyly, these data suggest that the appearance of *Hoxa11as* provided a mechanism to restrict *Hoxa11* expression from the distal limb, contributing to the emergence of the wrist and ankle joints as well as pentadactyly in tetrapods. Mechanistically, *Hoxa11as* does not appear to act via sense-antisense pairing of RNA transcripts nor via promoter interference (Kherdjemil *et al*., 2016; Chau *et al*., 2002) and thus exactly how this lncRNA exerts its inhibitory effect is yet to be determined.

The role of lncRNAs in chromatin remodelling

The role of lncRNAs in eliciting chromatin alterations has been well established in both humans and mice (Luo *et al.,* 2016; Quinn and Chang, 2016; Nakagawa *et al*., 2014 (references herein); Khalil *et al*., 2009). This can be achieved through the recruitment of histone modifying complexes by the lncRNAs, or by setting up topological domains that facilitate or maintain correct interactions between enhancers and genes. (Wang *et al*., 2011; Ørom *et al*., 2010). Within the vertebrate *Hox* clusters, *Hottip* is a good example of a two-tiered mode of action, as both the lncRNA gene locus and the noncoding RNA transcript play a role in chromatin remodelling and recruitment of histone modifiers to facilitate the transcription of neighbouring *Hox* genes. Located upstream of *Hoxa13* in both mice (Fig 2; Sasaki *et al.,* 2007; Wang *et al.,* 2011) and humans (Rinn *et al*., 2007; Wang *et al*., 2011), *Hottip* recruits the MLL complex via binding to Wdr5, resulting in the deposition of H3K4me4 marks on 5' *HoxA* genes and thus keeping them transcriptionally active (Wang *et al*., 2011; Pradeepa *et al*., 2017). Additionally, chromatin conformation capture technology identified physical interactions between the *Hottip* locus and 5' *HoxA* genes by chromatin looping (Wang *et al*., 2011). The importance of these interactions is underpinned by the observation that ectopic expression of *Hottip* RNA does not activate expression of 5' *HoxA* genes, nor rescues their expression in cells in which the endogenous *Hottip* had been deleted (Wang *et al*., 2011). A similar dual mechanism was observed for *HoxBlinc*, a lncRNA regulating the activation of nearby 3' *HoxB* genes during hematopoietic and cardiac cell lineage specification (Deng *et al*., 2016). Whether chromosomal looping and transcription of the lncRNA proceed following a specific temporal sequence (i.e. one of them preceding the other), or if both occur concomitantly is currently not known.

An enhancer-like function has been demonstrated for a particular transcriptional start site (TSS) shared by *Hog* and *Tog*, two divergent lncRNAs flanking the *HoxD* cluster (Delpretti *et al*., 2013). These lncRNAs are exclusively transcribed in the developing cecum and are required for correct activation of specific *HoxD* genes that pattern this region (Delpretti *et al*., 2013). This includes proper temporal activation of *Hoxd4* to *Hoxd11* genes while keeping *Hoxd12* and *Hoxd13* silent. At least eight enhancers have been identified in the gene desert spanning the telomeric region including the TSS of *Hog* and *Tog*, which physically interact with the *HoxD* genes, in particular *Hoxd4* and *d11* (Delpretti *et al*., 2013). These interactions were lost upon inversion of the *HoxD* cluster and led

to decreased *Hox* gene expression, indicating that changing the distance between *HoxD* genes and the region around *Hog* and *Tog* can modify chromatin topology, thereby affecting the activation of the *Hox* genes. This is underpinned by the ectopic activation of *Hoxd12* and *d13* in the inversion mutant, suggesting that the changed chromatin microarchitecture now allows *Hoxd12* and *d13* to interact with activating enhancers and induce their expression (Delpretti *et al*., 2013). Whether *Hog*/*Tog* transcripts are purely a by-product of this physical interaction or are actively involved in setting up the chromatin interaction and looping remains to been addressed.

Competing endogenous RNA

Finally, emerging data indicates that some lncRNAs might function as miRNA sponges, i.e. competing endogenous RNAs (ceRNAs). Indeed, the Hox-associated lncRNAs *HOTAIR* and *HOTTIP* have been implicated as sponges for several miRNAs in various cancers (Luan *et al*., 2017, Ma *et al*., 2017 Sun *et al*., 2017 and Liu *et al.,* 2014), but their function as sponges to regulate *Hox* gene expression during embryonic development remains to be explored.

Together, accumulating data have shown that lncRNAs are multifaceted molecules working within different layers of *Hox* gene regulation. The evolutionary conservation of many *Hox*-associated lncRNAs from human to mouse (De Kumar and Krumlauf, 2016), suggests functional importance. However, while robust changes in *Hox* expression have clearly been associated with lncRNAs, much work is still required to evaluate their relative functional relevance in developmental outputs reliant on Hox networks.

Hox-embedded miRNAs shaping Hox cluster output

MicroRNAs are found in all annotated bilaterian *Hox* clusters. These short noncoding RNAs act to repress gene expression post-transcriptionally, by deploying the RNA-induced silencing complex to target transcripts in a sequence-dependent manner (reviewed in Bartel, 2018). *Hox*-embedded miRNAs have a plethora of target transcripts, however, it is clear that across species, *Hox* genes themselves are present and often enriched within target gene lists (Yekta *et al*., 2008). Moreover, there is now ample evidence that the function of Hox-embedded miRNAs is essential for a wide range of developmental processes and aberrant in numerous pathological states. In this section, we review the evolution, expression and functional assessment of *Hox*-embedded miRNAs.

Evolutionary acquisition of Hox-embedded miRNAs

MicroRNAs have continually been acquired within *Hox*-clusters over the course of evolution (reviewed in Heimberg and McGlinn, 2012). *Drosophila Hox* clusters house four miRNAs; *miR-10*, *miR-993*, *miR-iab-4* and *miR-iab-8* (Fig. 2). *miR-10* is common to almost all bilaterian animals, likely arising soon after the cnidarian-bilaterian split and expanding throughout the lineages. This expansion includes the protostome-specific *miR-993*, a *miR-10* family member. In most arthropods, *miR-iab-4* and *miR-iab-8* are located between *abd-A* and *Abd-B* (Aravin *et al*., 2003; Miura *et al*., 2011; Ronshaugen *et al*., 2005). These miRNAs are transcribed bidirectionally, *miR-iab-8* representing the antisense transcript of *miR-iab-4*, with all four mature miRNAs generated from this locus being highly conserved (Miura *et al*., 2011; Ronshaugen *et al*., 2005). Murine *Hox* clusters house six miRNAs; *miR-10a*, *miR-10b*, *miR-196a1*, *miR-196a2*, *miR-196b* and *miR-615* (Fig. 2). The genomic position of *miR-10,* between *Hox4* and *Hox5* orthologues, has been conserved from fly to mouse. *miR-196* arose at the base of chordates and urochordates, while *miR-615* is conserved across eutherian mammals. As a general statement, once miRNAs have been acquired within *Hox* clusters, they are often fixed in the genome, consistent with what has been observed for miRNAs more broadly (Sempere *et al*., 2006; Wheeler *et al*., 2009). Moreover, when comparing Hox-embedded miRNA acquisition across a range of vertebrate and invertebrate species (reviewed in Heimberg and McGlinn, 2012), it becomes apparent that the positioning of miRNAs relative to the *Hox* genes may not be random (Yekta *et al*., 2008). For example, vertebrate *miR-196* occupies a similar, though not syntenic, position to *miR-iab4/8* in arthropods. Perhaps more strikingly, *miR-10* has been duplicated in the basal chordate *Amphioxus*, with two additional *miR-10* copies located at similar positions occupied by *miR-615* and *miR-196* in vertebrates (Campo-Paysaa *et al*., 2011). A similar duplication and positioning of *miR-10* in the evolutionarily distant *Capitella telata* (Wheeler *et al*., 2009) supports the view that the genomic position of Hox-embedded miRNAs is important.

Regulated expression of Hox-embedded miRNAs

In the vertebrate *Hox* clusters, ample evidence suggests co-transcription of Hox-embedded miRNAs with neighboring *Hox* genes. Transcripts containing *Hoxb4* and *miR-10a* or *Hoxd4* and *miR-10b* have been identified in mouse cell lines and in human datasets (Phua *et al*., 2011; Mainguy *et al*., 2007), while in zebrafish, *miR-10c* was identified as a polycistronic transcript together with *HoxB3a*, originating close to the *HoxB5a* transcriptional start site (Woltering and Durston, 2008). Additionally, *miR-10a* expression was also shown to be under the control of a proximal promoter during smooth muscle differentiation in mouse (Huang *et al*., 2010), although its relevance in other tissues remains unexplored. Notably, the spatio-temporal expression of *miR-10a* recapitulates the pattern of *Hoxb4* during mouse development, in agreement with co-transcriptional expression (Mansfield *et al*., 2004; For a detailed description of all Hox-embedded miRNA expression patterns, see Mansfield and McGlinn, 2012). For *miR-196*, there is evidence indicating both polycistronic transcription with neighboring *Hox* genes (Mainguy *et al*., 2007) as well as individual transcription starting from an autonomous promoter (Fantini *et al*., 2015). Which of the two mechanisms represents the major form of *miR-196* transcription remains unexplored. Certainly however, reporter gene expression driven from individual *miR-196a* paralog loci is consistent with their collinear activation relative to adjacent *Hox9/10* paralogs (Wong *et al*., 2015), with clear differences in the timing of *miR-196a1* and *miR-196a2* activation reflecting differences in individual cluster dynamics. *miR-615*, genomically positioned within the *Hoxc5* intron, is assumed to be co-expressed with *Hoxc5* in a single transcript. Nonetheless, integrated sequencing data indicates the existence of a *miR-615* autonomous promoter located within the coding sequence of *Hoxc5* (Quah and Holland, 2015). Evidence exists for both concordant and discordant expression between the host gene and the miRNA, supporting both intron processing and independent transcription in the generation of *miR-615*. The spatial expression pattern of *miR-615* during mouse development is yet to be assessed, although whole-body RNA-seq data indicates that *miR-615* expression increases throughout mouse embryonic development and is sustained in newborns (Quah and Holland, 2015).

Unlike vertebrates, *Drosophila* Hox miRNAs do not seem to be co-transcribed with *Hox* genes but rather, they are part of long noncoding transcripts. *miR-10*, for instance, originates from a 7.5 kb polyadenylated transcript (Lemons *et al.*, 2012). Alternative splicing of the primary miRNA transcript has been observed, producing identical mature miRNAs, with an underlying regulatory function yet unknown (Qian *et al.*, 2011). *miR-iab-4/8* stem from within long sense and antisense transcripts (*iab-4* and *iab-8*; Bender, 2008; Tyler *et al.*, 2008, Lemons *et al.*, 2012). Different splice isoforms of *iab-8* have been detected, exclusively expressed in embryos and adult male flies, eg. male-specific abdominal (Graveley *et al.*, 2011). For both *miR-iab-4* and *miR-iab-8*, the 5p arm represents the dominant mature miRNA (Aravin *et al.*, 2003; Bender, 2008; Ruby *et al.*, 2007; Stark *et al.*, 2008; Tyler *et al.*, 2008). The posterior expression patterns of *mir-iab-4* and *miR-iab-8* during development are in register with their genomic positioning, though interestingly, the two miRNAs exhibit no spatial overlap (Bender, 2008; Ronshaugen *et al.*, 2005; Stark *et al.*, 2008; Tyler *et al.*, 2008; Garaulet *et al.*, 2014). This mutual exclusion established at the boundary of abdominal segment precursors of A7 and A8 may result from cross-repressive mechanisms such as transcriptional interference or mature miRNA annealing (Stark *et al.*, 2008).

Post-transcriptional processing of primary miRNA transcripts can be regulated at multiple steps, resulting in diversification of mature miRNA transcripts within and between species. This is particularly evident for the *miR-10* family where both the 3p- and 5p-arms are processed to maturity and expressed in different ratios across species, an event known as arm switching (Landgraf *et al.*, 2007; Lim *et al.*, 2003; Ruby *et al.*, 2007; Stark *et al.*, 2007; Wheeler *et al.*, 2009; Griffiths-Jones *et al.*, 2011). Seed shifting and RNA editing are additional processing mechanisms that generate variety in mature *miR-10* species born from a single transcript (Marco *et al.*, 2010; Wheeler *et al.*, 2009).

Functional assessment of Hox-embedded miRNAs

The genomic positioning of Hox-embedded miRNAs becomes particularly relevant when considering their predicted Hox targets (Yekta *et al.*, 2008; Woltering and Durston, 2008). In vertebrates, the predicted Hox-targets are either biased (miR-10) or exclusively (miR-196) positioned more 3' (proximal) than the miRNA (Fig. 2). In *Drosophila*, the same can be observed for miR-iab-4/8, though miR-10 predicted Hox targets are biased more 5' (distal) relative to the miRNA (Fig. 2). Given the intrinsic collinear expression pattern of Hox-clusters genes along the anterior-to-posterior (A-P) axis, this suggests a complex spatio-temporal interplay between the miRNAs and their Hox targets with examples of both overlapping and non-overlapping miRNA-target expression patterns observed, and a function in A-P patterning predicted.

Despite predictions, there is little evidence for a strong developmental function for miR-10 across species. Predicted miR-10 target binding sites in Ubx and Abd-B are evolutionarily conserved across Drosophilids suggesting functionality, however gain-of-function approaches did not demonstrate significant targeting *in vivo* (Lemons *et al.*, 2012). Moreover, miR-10 locus deletion in *Drosophila* does not change the levels or spatial pattern of Abd-B protein, though phenotypic characterisation was not presented for this line (Lemons *et al.*, 2012). In vertebrates such as zebrafish, miR-10 synergises with HoxB4a in the molecular repression of hoxB1a and hoxB3a *in vivo* (Woltering and Durston, 2008), demonstrating a coordinated, co-transcribed, system that represses an anterior Hox program. However, miR-10 gain-of-function resulted in relatively minor neural patterning defects and miR-10 loss-of-function phenotypes were not apparent (Woltering and Durston, 2008). In mouse, genetic deletion of miR-10a has been characterised in the context of disease, with enhanced susceptibility to intestinal neoplasia demonstrated (Stadthagen *et al.*, 2013). The generation of miR10a/b compound mouse knockouts awaits.

In contrast, *Drosophila* exhibits a stage-specific, essential, requirement for miR-iab4/8 in regulating Hox output, both at the molecular and phenotypic level. At embryonic stages, efficient posterior repression of Ubx and Abd-A requires minimal contribution from miR-iab4/8, consistent with the lack of overt A-P patterning defects observed in mutant flies (Bender, 2008). Here, the action of posterior Hox gene Abd-B dominates, at least in the repression of Ubx. In the larval central nervous system however, the dominant mode of repression transitions to a miRNA-dependent one (Garaulet *et al.*, 2014), with significant phenotypic defects arising. These include sterility (Bender, 2008) and self-righting behavioural defects (Picao-Osorio *et al.*, 2015), downstream of altered Hox repression in separate motor neuron populations (Garaulet *et al.*, 2014; Picao-Osorio *et al.*, 2015). A likely explanation for this stage-specific repression of Hox targets by miR-iab4/8 is the elegant avoidance strategy described in Section 2, whereby dynamically regulated 3'UTR processing initially precludes many miRNA binding sites from the Hox transcript, with subsequent incorporation as development proceeds (Thomsen *et al.*, 2010). However, overexpression of exogenous miR-iab-4 was capable of transforming halteres to wings (Ronshaugen *et al.*, 2005), a classic early Ubx loss of function phenotype, reinforcing that miRNA-target cellular coexpression and/or levels of the miRNA are also critical parameters. Together, these data demonstrate a critical, albeit later than anticipated, role for miR-iab4/8 in controlling functional Hox output. It is interesting to note that *Drosophila* is a long germ insect where segments arise simultaneously, in contrast to insects such as the red flour beetle Tribolium castaneum, where segmentation happens in a posterior growth zone. As such, it would be interesting to assess whether Hox-embedded miRNAs exhibit a more prominent early A-P patterning role in such species.

Analysis of vertebrate *miR-196* has provided the clearest evidence to date for a classic homeotic functioning Hox-embedded miRNA. The *miR-196* paralogs form a genomic posterior boundary for Hox genes known to pattern the thoracic region across vertebrates (Yekta *et al.*, 2004), and importantly, the majority of predicted Hox targets lie within this thoracic-forming region. *Hoxb8* has been extensively characterised as a miR-196 target, being a unique example in animals of miRNA-directed endonucleolytic target cleavage (Mansfield *et al.,* 2004; Yekta *et al.,* 2004). Consistent with predictions that miR-196 may establish or reinforce posterior *Hox* boundaries, over-expression and knockdown studies have shown important roles for miR-196 in clearing unwanted *Hoxb8* in the developing neural tube and hindlimb (Asli and Kessel, 2010; Hornstein *et al.*, 2005). Additionally, evidence exists for a function of miR-196 within the endogenous *Hox* target expression domain (McGlinn *et al.*, 2009). Our lab has generated the first complete allelic deletion series for the three *miR-196* paralogs in mouse (Wong *et al.*, 2015). This allowed us to perform an unbiased molecular characterisation, revealing that 7 of the 10 predicted miR-196 Hox targets are statistically upregulated following loss

of miR-196 function at a time-point relevant to the generation of thoracic and lumbar vertebrae (Note: some target predictions have changed in the latest version of Targetscan (Agarwal *et al.*, 2015)). Compound knockout of *miR-196* paralogs demonstrated a dose-dependent, non-redundant, role for miR-196 in restricting the number of thoracic elements (Wong *et al.*, 2015). Surprisingly, we also observe a role in restricting the total number of vertebral elements, a phenotype difficult to reconcile with previous Hox literature, though a phenotype consistent with morpholino knockdown of miR-196 in zebrafish (He *et al.*, 2011a/b). Whether this latter phenotype arises due to the observed global posterior shift in Hox code (Wong *et al.*, 2015) and thus delayed activation of "terminator" Hox13 paralogs, or whether miR-196 acts on additional signalling pathways in this context remains to be elucidated. The recent observation that 5' posterior Hox genes directly activate *miR-196b* transcription *in vivo* (Fantini *et al.*, 2015) highlights miRNA activity as a novel mechanism whereby a posterior Hox protein represses the activity of a more anterior Hox program (ie. posterior prevalence).

miR-615 has no current obvious link in regulating Hox output during development, though interestingly, both miR-615 and the product of its host gene *Hoxc5* were found to repress the telomerase subunit hTERT during cellular differentiation (Yan *et al.*, 2018). This provided another example of a coordinated, co-transcribed, system whereby a Hox protein and a *Hox*-embedded miRNA repress gene function at different levels.

Concluding remarks

Over the years, *Hox* research has significantly focused on the transcriptional regulation *of Hox* genes, more specifically their conventional Hox coding transcripts, and how the resulting proteins control developmental and pathogenic processes. However, emerging data shows that *Hox* clusters exhibit remarkable transcriptional complexity, in such a way that one genomic sequence can be incorporated into distinct types of transcripts, either coding or noncoding, sense or antisense. This, together with multiple RNA processing mechanisms, results in a highly diverse *Hox* transcriptome, including alternatively spliced mRNAs, multifunctional lncRNAs and miRNAs targeting multiple genes. Clearly, the production of these RNA species from the *Hox* clusters is not a byproduct of the complex transcriptional activity occurring within a rather unconventional chromosomal region packed with genes that should be produced according to very precise spatial and temporal patterns. Just the opposite, the association of distinct transcriptomic profiles with specific biological processes, together with the functional characteristics of these various transcripts indicate that they must be an integral component of the mechanisms guiding proper production of Hox proteins (or variants of those), most likely the main ultimate effectors of *Hox* function. The study of those transcripts, and particularly their functional relevance, has been complicated by the scarcity of methods allowing efficient interference with the normal activity of those gene products in relevant biological contexts. The emergence of new gene editing technologies, together with the fast improvement of high throughput methods and the refinement of bioinformatic predictive and analytic capabilities, will help in evaluating the real dimension and impact this complex transcriptome has on the activities regulated from the *Hox* clusters.

Acknowledgements

A.C. is supported by postdoctoral fellowship SFRH/BPD/89500/2012 from the Fundação para a Ciência e a Tecnologia. G.M.H. is supported by an Australian Government Research Training Program (RTP) Scholarship. M.M. is supported by the Fundação para a Ciência e a Tecnologia, grant PTDC/BEX-BID/0899/2014 and by Santa Casa da Misericordia de Lisboa, grant SCML-MC-60-2014. E.M. is supported by National Health and Medical Research Council Project Grant APP1051792 and Australian Research Council Discovery Project DP180102157. The Australian Regenerative Medicine Institute is supported by grants from the State Government of Victoria and the Australian Government.

References

AGARWAL V, BELL GW, NAM J-W, BARTEL DP (2015). Predicting effective microRNA target sites in mammalian mRNAs. *eLife* 4.

AMÂNDIO AR, NECSULEA A, JOYE E, MASCREZ B, DUBOULE D (2016). Hotair Is Dispensible for Mouse Development. *PLOS Genetics* 12.

ARAVIN AA, LAGOS-QUINTANA M, YALCIN A, ZAVOLAN M, MARKS D, SNYDER B, GAASTERLAND T, MEYER J, TUSCHL T (2003). The Small RNA Profile during *Drosophila melanogaster* Development. *Dev Cell* 5: 337–350.

ASLI NS, KESSEL M (2010). Spatiotemporally restricted regulation of generic motor neuron programs by miR-196-mediated repression of Hoxb8. *Dev Biol* 344: 857–868.

BAE E, CALHOUN VC, LEVINE M, LEWIS EB, DREWELL RA (2002). Characterization of the intergenic RNA profile at abdominal-A and Abdominal-B in the *Drosophila* bithorax complex. *Proc Natl Acad Sci USA* 99: 16847–16852.

BARTEL DP (2018). Metazoan MicroRNAs. *Cell* 173: 20–51.

BENDER W (2008). MicroRNAs in the *Drosophila* bithorax complex. *Gen Dev* 22: 14–19.

BENSON GV, NGUYEN TH, MAAS RL (1995). The expression pattern of the murine Hoxa-10 gene and the sequence recognition of its homeodomain reveal specific properties of Abdominal B-like genes. *Mol Cell Biol* 15: 1591–1601.

BERMINGHAM JR, SCOTT MP (1988). Developmentally regulated alternative splicing of transcripts from the *Drosophila* homeotic gene Antennapedia can produce four different proteins. *EMBO J* 7: 3211–22.

BORROW J, SHEARMAN AM, STANTON VP, BECHER R, COLLINS T, WILLIAMS AJ, DUBÉ I, KATZ F, KWONG YL, MORRIS C, OHYASHIKI K, TOYAMA K, ROWLEY J, HOUSMAN DE (1996). The t(7;11)(p15;p15) translocation in acute myeloid leukaemia fuses the genes for nucleoporin NUP96 and class I homeoprotein HOXA9. *Nat Genet* 12: 159–167.

BOULET AM, CAPECCHI MR (2002). Duplication of the Hoxd11 Gene Causes Alterations in the Axial and Appendicular Skeleton of the Mouse. *Dev Biol* 249: 96–107.

BREND T, GILTHORPE J, SUMMERBELL D, RIGBY PWJ (2003). Multiple levels of transcriptional and post-transcriptional regulation are required to define the domain of Hoxb4 expression. *Development* 130: 2717–2728.

BRUNSKILL EW, POTTER SS (2012). RNA-Seq defines novel genes, RNA processing patterns and enhancer maps for the early stages of nephrogenesis: Hox supergenes. *Dev Biol* 368: 4–17.

BURNETTE JM, HATTON AR, LOPEZ AJ (1999). Trans-acting factors required for inclusion of regulated exons in the Ultrabithorax mRNAs of *Drosophila melanogaster*. *Genetics* 151: 1517–1529.

CAMPO-PAYSAA F, SEMON M, CAMERON RA, PETERSON KJ, SCHUBERT M (2011). microRNA complements in deuterostomes: origin and evolution of microRNAs. *Evol Dev* 13: 15-27.

CARAPUÇO M, VINAGRE T, BOBOLA N, MALLO M (2005). Hox genes specify vertebral types in the presomitic mesoderm. *Genes Dev* 19: 2116–2121.

CASACA A, NOVOA A, MALLO M (2016). Hoxb6 can interfere with somitogenesis in the posterior embryo through a mechanism independent of its rib-promoting activity. *Development* 143: 437–448.

CELNIKER SE, KEELAN DJ, LEWIS EB (1989). The molecular genetics of the bithorax complex of *Drosophila*: characterization of the products of the Abdominal-B domain. *Genes Dev* 3: 1424–1436.

CHAPAT C, JAFARNEJAD SM, MATTA-CAMACHO E, HESKETH GG, GELBART IA, ATTIG J, GKOGKAS CG, ALAIN T, STERN-GINOSSAR N, FABIAN MR, GINGRAS A-C, DUCHAINE TF, SONENBERG N (2017). Cap-binding protein 4EHP effects

translation silencing by microRNAs. *Proc Natl Acad Sci USA* 114: 5425–5430.

CHARIOT A, MOREAU L, SENTERRE G, SOBEL M, CASTRONOVO V (1995). Retinoic Acid Induces Three Newly Cloned HOXA1 Transcripts in MCF7 Breast Cancer Cells. *Biochem Biophys Res Commun* 215: 713–720.

CHAU YM, PANDO S, TAYLOR HS (2002). HOXA11 Silencing and Endogenous HOXA11 Antisense Ribonucleic Acid in the Uterine Endometrium. *J Clin Endocrin Metab* 87: 2674–2680.

COULOMBE Y, LEMIEUX M, MOREAU J, AUBIN J, JOKSIMOVIC M, BÉRUBÉ-SIMARD FA, TABARIÈS S, BOUCHERAT O, GUILLOU F, LAROCHELLE C, TUGGLE CK, JEANNOTTE L (2010). Multiple Promoters and Alternative Splicing: Hoxa5 Transcriptional Complexity in the Mouse Embryo. *PLoS ONE* 5.

CRIBBS DL, PULTZ M a, JOHNSON D, MAZZULLA M, KAUFMAN TC (1992). Structural complexity and evolutionary conservation of the *Drosophila* homeotic gene proboscipedia. *EMBO J* 11: 1437–49.

CUMBERLEDGE S, ZARATZIAN A, SAKONJU S (1990).Characterization of two RNAs transcribed from the cis-regulatory region of the abd-A domain within the *Drosophila* bithorax complex. *Proc Natl Acad Sci USA* 87:3259-63

DAVIS MC, DAHN RD, SHUBIN NH (2007). An autopodial-like pattern of Hox expression of a basal actinopterygian fish. *Nature* 447:473-6.

DE KUMAR B, KRUMLAUF R (2016). HOXs and lincRNAs: Two sides of the same coin. *Sci Adv* 2.

DE KUMAR B, PARRISH ME, SLAUGHTER BD, UNRUH JR, GOGOL M, SEIDEL C, PAULSON A, LI H, GAUDENZ K, PEAK A, MCDOWELL W, FLEHARTY B, AHN Y, LIN C, SMITH E, SHILATIFARD A, KRUMLAUF R (2015). Analysis of dynamic changes in retinoid-induced transcription and epigenetic profiles of murine Hox clusters in ES cells. *Genome Res* 25: 1229–1243.

DELPRETTI S, MONTAVON T, LELEU M, JOYE E, TZIKA A, MILINKOVITCH M, DUBOULE D (2013). Multiple Enhancers Regulate Hoxd Genes and the Hotdog LncRNA during Cecum Budding. *Cell Rep* 5: 137–150.

DENG C, LI Y, ZHOU L, CHO J, PATEL B, TERADA N, LI Y, BUNGERT J, QIU Y, HUANG S (2016). HoxBlinc RNA Recruits Set1/MLL Complexes to Activate Hox Gene Expression Patterns and Mesoderm Lineage Development. *Cell Rep* 14: 103–114.

DINTILHAC A, BIHAN R, GUERRIER D, DESCHAMPS S, PELLERIN I (2004). A conserved non-homeodomain Hoxa9 isoform interacting with CBP is co-expressed with the 'typical' Hoxa9 protein during embryogenesis. *Gene Expr Patterns* 4: 215–222.

DUBOULE D (2007). The rise and fall of Hox gene clusters. *Development* 134: 2549–2560.

FANTINI S, SALSI V, VITOBELLO A, RIJLI FM, ZAPPAVIGNA V (2015). MicroRNA-196b is transcribed from an autonomous promoter and is directly regulated by Cdx2 and by posterior Hox proteins during embryogenesis. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* 1849: 1066–1080.

FERNANDEZ CC, GUDAS LJ (2009). The truncated Hoxa1 protein interacts with Hoxa1 and Pbx1 in stem cells. *J Cell Biochem* 106: 427–443.

FUJIMOTO S, ARAKI K, CHISAKA O, ARAKI M, TAKAGI K, YAMAMURA KI (1998). Analysis of the murine Hoxa-9 cDNA: an alternatively spliced transcript encodes a truncated protein lacking the homeodomain. *Gene* 209: 77–85.

GARAULET DL, CASTELLANOS MC, BEJARANO F, SANFILIPPO P, TYLER DM, ALLAN DW, SÁNCHEZ-HERRERO E, LAI EC (2014). Homeotic Function of *Drosophila* Bithorax-Complex miRNAs Mediates Fertility by Restricting Multiple Hox Genes and TALE Cofactors in the CNS. *Dev Cell* 29: 635–648.

GARCIA-FERNÀNDEZ J (2005). Hox, ParaHox, ProtoHox: facts and guesses. *Heredity* 94: 145–152.

GEBELEIN B, CULI J, RYOO HD, ZHANG W, MANN RS (2002). Specificity of Distalless Repression and Limb Primordia Development by Abdominal Hox Proteins. *Dev Cell* 3: 487–498.

GRAVELEY BR, BROOKS AN, CARLSON JW, DUFF MO, LANDOLIN JM, YANG L, ARTIERI CG, BAREN MJV, BOLEY N, BOOTH BW, *et al.,* (2011). The developmental transcriptome of *Drosophila melanogaster*. *Nature* 471: 473–479.

GRIFFITHS-JONES S, HUI JHL, MARCO A, RONSHAUGEN M (2011). MicroRNA evolution by arm switching. *EMBO Rep* 12: 172–177.

GUERREIRO I, CASACA A, NUNES A, MONTEIRO S, NOVOA A, FERREIRA RB, BOM J, MALLO M (2012). Regulatory role for a conserved motif adjacent to the homeodomain of Hox10 proteins. *Development* 139: 2703–2710.

HADRYS T, PRINCE V, HUNTER M, BAKER R, RINKWITZ S (2004). Comparative genomic analysis of vertebrate Hox3 and Hox4 genes. *J Exp Zool* 302B: 147–164.

HADRYS T, PUNNAMOOTTIL B, PIEPER M, KIKUTA H, PEZERON G, BECKER TS, PRINCE V, BAKER R, RINKWITZ S (2006). Conserved co-regulation and promoter sharing of hoxb3a and hoxb4a in zebrafish. *Dev Biol* 297: 26–43.

HATTON AR, SUBRAMANIAM V, LOPEZ A (1998). Generation of Alternative Ultrabithorax Isoforms and Stepwise Removal of a Large Intron by Resplicing at Exon–Exon Junctions. *Mol Cell* 2: 787–796.

HE M, CHEN P, ARNOVITZ S, LI Y, HUANG H, NEILLY MB, WEI M, ROWLEY JD, CHEN J, LI Z (2012). Two isoforms of HOXA9 function differently but work synergistically in human MLL-rearranged leukemia. *Blood Cells, Mol Dis* 49: 102–106.

HE X, YAN Y-L, DELAURIER A, POSTLETHWAIT JH (2011a). Observation of miRNA Gene Expression in Zebrafish Embryos by *In situ* Hybridization to MicroRNA Primary Transcripts. *Zebrafish* 8: 1–8.

HE X, YAN Y-L, EBERHART JK, HERPIN A, WAGNER TU, SCHARTL M, POSTLETHWAIT JH (2011b). miR-196 regulates axial patterning and pectoral appendage initiation. *Dev Biol* 357: 463–477.

HEIMBERG A, MCGLINN E (2012). Building a Robust A-P Axis. *Curr Genomics* 13: 278–288.

HORNSTEIN E, MANSFIELD JH, YEKTA S, HU JK-H, HARFE BD, MCMANUS MT, BASKERVILLE S, BARTEL DP, TABIN CJ (2005). The microRNA miR-196 acts upstream of Hoxb8 and Shh in limb development. *Nature* 438: 671–674.

HSIEH-LIE HM, WITTE DP, WEINSTEIN M, BRANFORD W, LI H, SMALL K, POTTER SS (1995). *Hoxa11* structure, extensive antisense transcriptiona, and function in male and female fertility. *Development* 121:1373-1385.

HUANG H, XIE C, SUN X, RITCHIE RP, ZHANG J, CHEN YE (2010). miR-10a Contributes to Retinoid Acid-induced Smooth Muscle Cell Differentiation. *J. Biol. Chem.* 285: 9383–9389.

JANSSEN R, BUDD GE (2010). Gene expression suggests conserved aspects of Hox gene regulation in arthropods and provides additional support for monophyletic Myriapoda. *EvoDevo* 1: 1-11.

JI Z, LEE JY, PAN Z, JIANG B, TIAN B (2009). Progressive lengthening of 3' untranslated regions of mRNAs by alternative polyadenylation during mouse embryonic development. *Proc Natl Acad Sci USA* 106: 7028–7033.

KHALIL AM, GUTTMAN M, HUARTE M, GARBER M, RAJ A, MORALES DR, THOMAS K, PRESSER A, BERNSTEIN BE, OUDENAARDEN AV, REGEV A, LANDER ES, RINN JL (2009). Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proc Natl Acad Sci USA* 106: 11667–11672.

KHERDJEMIL Y, LALONDE RL, SHETH R, DUMOUCHEL A, MARTINO GD, PINEAULT KM, WELLIK DM, STADLER HS, AKIMENKO M-A, KMITA M (2016). Evolution of Hoxa11 regulation in vertebrates is linked to the pentadactyl state. *Nature* 539: 89–92.

KIM MH, CHANG HH, SHIN C, CHO M, PARK D, PARK HW (1998). Genomic structure and sequence analysis of human HOXA-9. *DNA Cell Biol* 17: 407–414.

KONDRASHOV N, PUSIC A, STUMPF CR, SHIMIZU K, HSIEH AC, XUE S, ISHIJIMA J, SHIROISHI T, BARNA M (2011). Ribosome-Mediated Specificity in Hox mRNA Translation and Vertebrate Tissue Patterning. *Cell* 145: 383–397.

KORNFELD K, SAINT RB, BEACHY PA, HARTE PJ, PEATTIE DA, HOGNESS DS (1989). Structure and expression of a family of Ultrabithorax mRNAs generated by alternative splicing and polyadenylation in *Drosophila*. *Genes Dev* 3: 243–258.

KÖMÜVES LG, SHEN WF, KWONG A, STELNICKI E, ROZENFELD S, ODA Y, BLINK A, KRISHNAN K, LAU B, MAURO T, LARGMAN C (2000). Changes in HOXB6 homeodomain protein structure and localization during human epidermal development and differentiation. *Dev Dyn* 218: 636–647.

KUZIORA MA, MCGINNIS W (1988). Different transcripts of the *Drosophila* Abd-B gene correlate with distinct genetic sub-functions. *EMBO J* 7: 3233–3244.

LANDGRAF P, RUSU M, SHERIDAN R, SEWER A, IOVINO N, ARAVIN A, PFEFFER S, RICE A, KAMPHORST AO, LANDTHALER M, add. (2007). A Mammalian microRNA Expression Atlas Based on Small RNA Library Sequencing. *Cell* 129: 1401–1414.

LAROSA GJ, GUDAS LJ (1988). Early retinoic acid-induced F9 teratocarcinoma stem cell gene ERA-1: alternate splicing creates transcripts for a homeobox-containing protein and one lacking the homeobox. *Mol Cell Biol* 8: 3906–3917.

LAUGHON A, BOULET AM, BERMINGHAM JR, LAYMON RA, SCOTT MP (1986). Structure of transcripts from the homeotic Antennapedia gene of *Drosophila melanogaster*: two promoters control the major protein-coding region. *Mol Cell*

Biol 6: 4676–4689.

LEMONS D, PARÉ A, MCGINNIS W (2012). Three *Drosophila* Hox Complex microRNAs Do Not Have Major Effects on Expression of Evolutionarily Conserved Hox Gene Targets during Embryogenesis. *PLoS ONE* 7.

LI L, LIU B, WAPINSKI OL, TSAI M-C, QU K, ZHANG J, CARLSON JC, LIN M, FANG F, GUPTA RA, HELMS JA, CHANG HY (2013). Targeted Disruption of Hotair Leads to Homeotic Transformation and Gene Derepression. *Cell Rep* 5: 3–12.

LIM LP, GLASNER ME, YEKTA S, BURGE CB, BARTEL DP (2003). Vertebrate MicroRNA Genes. *Science* 299: 1540–1540.

LIPSHITZ HD, PEATTIE DA, HOGNESS DS (1987). Novel transcripts from the Ultrabithorax domain of the bithorax complex. *Gen Dev* 1: 307–322.

LIU X-H, SUN M, NIE F-Q, GE Y-B, ZHANG E-B, YIN D-D, KONG R, XIA R, LU K-H, LI J-H, DE W, WANG K-M, WANG Z-X (2014). Lnc RNA HOTAIR function as a competing endogenous RNA to regulate HER2 expression by sponging miR-331-3p in gastric cancer. *Mol Cancer* 13: 92.

LOPEZ AJ, HOGNESS DS (1991). Immunochemical dissection of the Ultrabithorax homeoprotein family in *Drosophila melanogaster*. *Proc Natl Acad Sci USA* 88: 9924–9928.

LOPEZ AJ, ARTERO RD, PEREZ-ALONSO M (1996). Stage, tissue, and cell specific distribution of alternative Ultrabithorax mRNAs and protein isoforms in the *Drosophila* embryo. *Roux's Arch Dev Biol* 205: 450–459.

LUAN W, LI R, LIU L, NI X, SHI Y, XIA Y, WANG J, LU F, XU B (2017). Long noncoding RNA HOTAIR acts as a competing endogenous RNA to promote malignant melanoma progression by sponging miR-152-3p. *Oncotarget* 8.

LUO S, LU JY, LIU L, YIN Y, CHEN C, HAN X, WU B, XU R, LIU W, YAN P, SHAO W, LU Z, LI H, NA J, TANG F, WANG J, ZHANG YE, SHEN X (2016). Divergent lncRNAs Regulate Gene Expression and Lineage Differentiation in Pluripotent Cells. *Cell Stem Cell* 18: 637–652.

MA J, FAN Y, FENG T, CHEN F, XU Z, LI S, LIN Q, HE X, SHI W, LIU Y, CAO X, ZHU B, LIU Z (2017). HOTAIR regulates HK2 expression by binding endogenous miR-125 and miR-143 in oesophageal squamous cell carcinoma progression. *Oncotarget* 8.

MAINGUY G, KOSTER J, WOLTERING J, JANSEN H, DURSTON A (2007). Extensive Polycistronism and Antisense Transcription in the Mammalian Hox Clusters. *PLoS ONE* 2.

MANSFIELD JH, MCGLINN E (2012). Evolution, Expression, and Developmental Function of Hox-Embedded miRNAs. MicroRNAs in Development *Curr. Top. Dev. Biol.*: 31–57.

MANSFIELD JH, HARFE BD, NISSEN R, OBENAUER J, SRINEEL J, CHAUDHURI A, FARZAN-KASHANI R, ZUKER M, PASQUINELLI AE, RUVKUN G, SHARP PA, TABIN CJ, MCMANUS MT (2004). MicroRNA-responsive 'sensor' transgenes uncover Hox-like and other developmentally regulated patterns of vertebrate microRNA expression. *Nature Genetics* 36: 1079–1083.

MARCO A, HUI JHL, RONSHAUGEN M, GRIFFITHS-JONES S (2010). Functional shifts in insect microRNA evolution. *Genome Biol. Evol.* 2: 686-696.

MATHEWS CH, DETMER K, LAWRENCE HJ, LARGMAN C (1993). Expression of the Hox 2.2 homeobox gene in murine embryonic epidermis. *Differentiation* 52: 177–184.

MCGLINN E, YEKTA S, MANSFIELD JH, SOUTSCHEK J, BARTEL DP, TABIN CJ (2009). *in ovo* application of antagomiRs indicates a role for miR-196 in patterning the chick axial skeleton through Hox gene regulation. *Proc Natl Acad Sci USA* 106: 18610–18615.

MERABET S, KAMBRIS Z, CAPOVILLA M, BÉRENGER H, PRADEL J, GRABA Y (2003). The Hexapeptide and Linker Regions of the AbdA Hox Protein Regulate Its Activating and Repressive Functions. *Dev Cell* 4: 761–768.

METSCHER BD, TAKAHASHI K, CROW K, AMEMIYA C, NONAKA DF, WAGNER GP (2005). Expression of Hoxa-11 and Hoxa-13 in the pectoral fin of a basal ray-finned fish, Polyodon spathula: implications for the origin of tetrapod limbs. *Evol Dev* 7: 186–195.

MIURA S, NOZAWA M, NEI M (2010). Evolutionary Changes of the Target Sites of Two MicroRNAs Encoded in the Hox Gene Cluster of *Drosophila* and Other Insect Species. *Genome Biology and Evolution* 3: 129–139.

MLODZIK M, FJOSE A, GEHRING WJ (1988). Molecular structure and spatial expression of a homeobox gene from the labial region of the Antennapedia-complex. *EMBO J* 7: 2569–78.

NAKAGAWA S, KAGEYAMA Y (2014). Nuclear lncRNAs as epigenetic regulators—Beyond skepticism. *Biochim Biophys Acta (BBA) - Gene Regulatory Mechanisms* 1839: 215–222.

NAVAS LFD, REED H, AKAM M, BARRIO R, ALONSO CR, SANCHEZ-HERRERO E (2011). Integration of RNA processing and expression level control modulates the function of the *Drosophila* Hox gene Ultrabithorax during adult development. *Development* 138: 107–116.

NELSON CE, MORGAN BA, BURKE AC, LAUFER E, DIMAMBRO E, MURTAUGH LC, GONZALES E, TESSAROLLO L, PARADA LF, TABIN C (1996). Analysis of Hox gene expression in the chick limb bud. *Development* 122: 1449–1466.

O'CONNOR MB, BINARI R, PERKINS LA, BENDER W, GEHRING WJ (1988). Alternative RNA products from the Ultrabithorax domain of the bithorax complex. *EMBO J* 7: 435–445.

OH SK, SCOTT MP, SARNOW P (1992). Homeotic gene Antennapedia mRNA contains 5'-noncoding sequences that confer translational initiation by internal ribosome binding. *Genes Dev* 6: 1643–1653.

PAPADOPOULOS DK, RESENDEZ-PEREZ D, CARDENAS-CHAVEZ DL, VILLANUEVA-SEGURA K, CANALES-DEL-CASTILLO R, FELIX DA, FUNFSCHILLING R, GEHRING WJ (2011). Functional synthetic Antennapedia genes and the dual roles of YPWM motif and linker size in transcriptional activation and repression. *Proc Natl Acad Sci USA* 108: 11959–11964.

PEASE B, BORGES AC, BENDER W (2013). Noncoding RNAs of the Ultrabithorax Domain of the *Drosophila* Bithorax Complex. *Genetics* 195: 1253–1264.

PETRUK S, SEDKOV Y, RILEY KM, HODGSON J, SCHWEISGUTH F, HIROSE S, JAYNES JB, BROCK HW, MAZO A (2006). Transcription of bxd Noncoding RNAs Promoted by Trithorax Represses Ubx in cis by Transcriptional Interference. *Cell* 127: 1209–1221.

PETTINI T, RONSHAUGEN M (2016). Transvection and pairing of a *Drosophila* Hox long noncoding RNA in the regulation of *Sex combs reduced*. bioRxiv.

PHUA SLC, SIVAKAMASUNDARI V, SHAO Y, CAI X, ZHANG L-F, LUFKIN T, FEATHERSTONE M (2011). Nuclear Accumulation of an Uncapped RNA Produced by Drosha Cleavage of a Transcript Encoding miR-10b and HOXD4. *PLoS ONE* 6.

PICAO-OSORIO J, JOHNSTON J, LANDGRAF M, BERNI J, ALONSO CR (2015) MicroRNA-encoded behaviour in *Drosophila*. *Science* 350: 815-820

POTTER SS, BRANFORD WW (1998). Evolutionary conservation and tissue-specific processing of Hoxa 11 antisense transcripts. *Mamm Genome* 9: 799–806.

PRADEEPA MM, MCKENNA F, TAYLOR GCA, BENGANI H, GRIMES GR, WOOD AJ, BHATIA S, BICKMORE WA (2017). Psip1/p52 regulates posterior Hoxa genes through activation of lncRNA Hottip. *PLOS Genetics* 13.

QIAN J, ZHANG Z, LIANG J, GE Q, DUAN X, MA F, LI F (2011). The full-length transcripts and promoter analysis of intergenic microRNAs in *Drosophila melanogaster*. *Genomics* 97: 294–303.

QUAH S, HOLLAND PWH (2015). The Hox cluster microRNA miR-615: a case study of intronic microRNA evolution. *EvoDevo* 6: 1-12.

QUINN JJ, CHANG HY (2016). Unique features of long non-coding RNA biogenesis and function. *Nat Rev Genet* 17: 47–62.

REED HC, HOARE T, THOMSEN S, WEAVER TA, WHITE RAH, AKAM M, ALONSO CR (2010). Alternative Splicing Modulates Ubx Protein Function in *Drosophila melanogaster*. *Genetics* 184: 745–758.

RINN JL, KERTESZ M, WANG JK, SQUAZZO SL, XU X, BRUGMANN SA, GOODNOUGH LH, HELMS JA, FARNHAM PJ, SEGAL E, CHANG HY (2007). Functional Demarcation of Active and Silent Chromatin Domains in Human HOX Loci by Noncoding RNAs. *Cell* 129: 1311–1323.

RONSHAUGEN M, BIEMAR F, PIEL J, LEVINE M, LAI EC (2005). The *Drosophila* microRNA iab-4 causes a dominant homeotic transformation of halteres to wings. *Gen Dev* 19: 2947–2952.

RUBY JG, JAN CH, BARTEL DP (2007). Intronic microRNA precursors that bypass Drosha processing. *Nature* 448: 83–86.

SAKAMOTO K, ONIMARU K, MUNAKATA K, SUDA N, TAMURA M, OCHI H, TANAKA M (2009). Heterochronic Shift in Hox-Mediated Activation of Sonic hedgehog Leads to Morphological Changes during Fin Development. *PLoS ONE* 4.

SÁNCHEZ-HERRERO E, CROSBY MA (1988). The Abdominal-B gene of *Drosophila melanogaster*: overlapping transcripts exhibit two different spatial distributions. *EMBO J* 7: 2163–73.

SASAKI YT, SANO M, KIN T, ASAI K, HIROSE T (2007). Coordinated expression of ncRNAs and HOX mRNAs in the human HOXA locus. *Biochem Biophys Res Commun* 357: 724–730.

SCHORDERET P, DUBOULE D (2011). Structural and Functional Differences in the Long Non-Coding RNA Hotair in Mouse and Human. *PLoS Genetics* 7.

SEMPERE LF, COLE CN, MCPEEK MA, PETERSON KJ (2006). The phylogenetic distribution of metazoan microRNAs: insights into evolutionary complexity and constraint. *JEZ-B Molec. Dev. Evol.* 206B: 575-588.

SESSA L, BREILING A, LAVORGNA G, SILVESTRI L, CASARI G, ORLANDO V (2007). Noncoding RNA synthesis and loss of Polycomb group repression accompanies the colinear activation of the human HOXA cluster. *RNA* 13: 223–239.

SHEN W-F, DETMER K, SIMONITCH-EASON TA, LAWRENCE H, LARGMAN C (1991). Alternative splicing of the HOX 2.2 homeobox gene in human hematopoietic cells and murine embryonic and adult tissues. *Nucleic Acids Res* 19: 539–545.

SHIGA Y, SAGAWA K, TAKAI R, SAKAGUCHI H, YAMAGATA H, HAYASHI S (2006). Transcriptional readthrough of Hox genes Ubx and Antp and their divergent post-transcriptional control during crustacean evolution. *Evol Dev* 8: 407–414.

SIMEONE A, PANNESE M, ACAMPORA D, D'ESPOSITO M, BONCINELLI E (1988). At least three human homeoboxes on chromosome 12 belong to the same transcription unit. *Nucleic Acids Res* 16: 5379–5390.

STADLER CR, VEGI N, MULAW MA, EDMAIER KE, RAWAT VPS, DOLNIK A, BULLINGER L, HEILMEIER B, QUINTANILLA-FEND L, SPIEKERMANN K, HIDDEMANN W, DÖHNER K, DÖHNER H, FEURING-BUSKE M, BUSKE C (2014). The leukemogenicity of Hoxa9 depends on alternative splicing. *Leukemia* 28: 1838–1843.

STADTHAGEN G, TEHLER D, HOYLAND-KROGHSBO NM, WEN J, KROGH A, JENSEN KT (2013) Loss of miR-10a Activates *Lpo* and Collaborates with Activated Wnt Signaling in Inducing Intestinal Neoplasia in Female Mice. *PLoS Genet* 9: e1003913.

STARK A, KHERADPOUR P, PARTS L, BRENNECKE J, HODGES E, HANNON GJ, KELLIS M (2007). Systematic discovery and characterization of fly microRNAs using 12 *Drosophila* genomes. *Genome Res* 17: 1865–1879.

STARK A, BUSHATI N, JAN CH, KHERADPOUR P, HODGES E, BRENNECKE J, BARTEL DP, COHEN SM, KELLIS M (2008). A single Hox locus in *Drosophila* produces functional microRNAs from opposite DNA strands. *Gen Dev* 22: 8–13.

STROEHER VL, JORGENSEN EM, GARBER RL (1986). Multiple transcripts from the Antennapedia gene of *Drosophila melanogaster*. *Mol Cell Biol* 6: 4667–4675.

SUN Y, ZHOU Y, BAI Y, WANG Q, BAO J, LUO Y, GUO Y, GUO L (2017). A long non-coding RNA HOTTIP expression is associated with disease progression and predicts outcome in small cell lung cancer patients. *Mol Cancer* 16.

THOMSEN S, AZZAM G, KASCHULA R, WILLIAMS LS, ALONSO CR (2010). Developmental RNA processing of 3'UTRs in Hox mRNAs as a context-dependent mechanism modulating visibility to microRNAs. *Development* 137: 2951–2960.

TSAI M-C, MANOR O, WAN Y, MOSAMMAPARAST N, WANG JK, LAN F, SHI Y, SEGAL E, CHANG HY (2010). Long Noncoding RNA as Modular Scaffold of Histone Modification Complexes. *Science* 329: 689–693.

TYLER DM, OKAMURA K, CHUNG W-J, HAGEN JW, BEREZIKOV E, HANNON GJ, LAI EC (2008). Functionally distinct regulatory RNAs generated by bidirectional transcription and processing of microRNA loci. *Gen Dev* 22: 26–36.

VILLAESCUSA JC, BURATTI C, PENKOV D, MATHIASEN L, PLANAGUMÀ J, FERRETTI E, BLASI F (2009). Cytoplasmic Prep1 Interacts with 4EHP Inhibiting Hoxb4 Translation. *PLoS ONE* 4.

WANG KC, YANG YW, LIU B, SANYAL A, CORCES-ZIMMERMAN R, CHEN Y, LAJOIE BR, PROTACIO A, FLYNN RA, GUPTA RA, WYSOCKA J, LEI M, DEKKER J, HELMS JA, CHANG HY (2011). A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature* 472: 120–124.

WHEELER BM, HEIMBERG AM, MOY VN, SPERLING EA, HOLSTEIN TW, HEBER S, PETERSON KJ (2009). The deep evolution of metazoan microRNAs. *Evol Dev* 11: 50–68.

WOLTERING JM, DURSTON AJ (2008). MiR-10 Represses HoxB1a and HoxB3a in Zebrafish. *PLoS ONE* 3.

WOLTERING JM, DURSTON AJ (2006). The zebrafish hoxDb cluster has been reduced to a single microRNA. *Nature Genet.* 38: 601–602.

WONG SFL, AGARWAL V, MANSFIELD JH, DENANS N, SCHWARTZ MG, PROSSER HM, POURQUIÉ O, BARTEL DP, TABIN CJ, MCGLINN E (2015). Independent regulation of vertebral number and vertebral identity by microRNA-196 paralogs. *Proc Natl Acad Sci USA* 112: E4884–E4893.

XUE S, BARNA M (2015). Cis-regulatory RNA elements that regulate specialized ribosome activity. *RNA Biol* 12: 1083–1087.

XUE S, TIAN S, FUJII K, KLADWANG W, DAS R, BARNA M (2015). RNA regulons in Hox 5′ UTRs confer ribosome specificity to gene regulation. *Nature* 517: 33–38.

YAN T, OOI WF, QAMRA A, CHEUNG A, MA D, SUNDARAM GM, XU C, XING M, POON L, WANG J, *et al.,* (2018). HoxC5 and miR-615-3p target newly evolved genomic regions to repress hTERT and inhibit tumorigenesis. *Nat Commun* 9: 100.

YE X, FONG P, IIZUKA N, CHOATE D, CAVENER DR (1997). Ultrabithorax and Antennapedia 5' untranslated regions promote developmentally regulated internal translation initiation. *Mol Cell Biol* 17: 1714–1721.

YEKTA S, SHIH I, BARTEL DP (2004). MicroRNA-Directed Cleavage of HOXB8 mRNA. *Science* 304: 594–596.

YEKTA S, TABIN CJ, BARTEL DP (2008). MicroRNAs in the Hox network: an apparent link to posterior prevalence. *Nat Rev Genet* 9: 789–796.

ØROM UA, DERRIEN T, BERINGER M, GUMIREDDY K, GARDINI A, BUSSOTTI G, LAI F, ZYTNICKI M, NOTREDAME C, HUANG Q, GUIGO R, SHIEKHATTAR R (2010). Long Noncoding RNAs with Enhancer-like Function in Human Cells. *Cell* 143: 46–58.

Further Related Reading, published previously in the *Int. J. Dev. Biol.*

Evolution of the activation domain in a Hox transcription factor
Ying Liu, Annie Huang, Rebecca M. Booth, Gabriela Geraldo Mendes, Zabeena Merchant, Kathleen S. Matthews and Sarah E. Bondos
Int. J. Dev. Biol. (2018) 62: XXXX-XXXX (this issue)
https://doi.org/10.1387/ijdb.180151sb

The road to the vertebral formula
Moisés Mallo, Tânia Vinagre and Marta Carapuço
Int. J. Dev. Biol. (2009) 53: 1469-1481
https://doi.org/10.1387/ijdb.072276mm

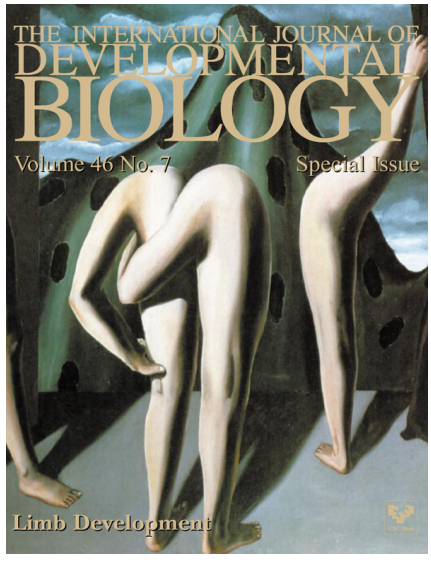
The significance of Hox gene collinearity
Stephen J. Gaunt
Int. J. Dev. Biol. (2015) 59: 159-170
https://doi.org/10.1387/ijdb.150223sg

Synergistic action in P19 pluripotential cells of retinoic acid and Wnt3a on Cdx1 enhancer elements
Stephen J. Gaunt and Yu-Lee Paul
Int. J. Dev. Biol. (2014) 58: 307-314
https://doi.org/10.1387/ijdb.140003sg

The Parahox gene Pdx1 is required to maintain positional identity in the adult foregut
Andrew M. Holland, Sonia Garcia, Gaetano Naselli, Raymond J. MacDonald and Leonard C. Harrison
Int. J. Dev. Biol. (2013) 57: 391-398
https://doi.org/10.1387/ijdb.120048ah

Hox collinearity - a new perspective
Antony J. Durston, Hans J. Jansen, Paul In der Rieden and Michiel H.W. Hooiveld
Int. J. Dev. Biol. (2011) 55: 899-908
https://doi.org/10.1387/ijdb.113358ad

Origins of Cdx1 regulatory elements suggest roles in vertebrate evolution
Stephen J. Gaunt and Yu-Lee Paul
Int. J. Dev. Biol. (2011) 55: 93-98
https://doi.org/10.1387/ijdb.103252sg



5 yr ISI Impact Factor (2016) = 2.421