A wide spectrum of molecular mechanisms regulates Hox gene expression during animal development

Moisés Mallo[§] and Claudio R. Alonso*

§ *Instituto Gulbenkian de Ciência, Rua da Quinta Grande 6, 2780-156 Oeiras, Portugal.*

* *John Maynard Smith Building, School of Life Sciences, University of Sussex Brighton BN1 9QG, United Kingdom.*

Editorial correspondence to:

Moisés Mallo
mallo@igc.gulbenkian.pt
T: +351 214464624
F: +351 214407970

Claudio R. Alonso
c.alonso@sussex.ac.uk
T: +44 (0) 1273 876621
F: +44 (0) 1273 678433

ABSTRACT

Hox genes encode a family of transcriptional regulators that elicit distinct developmental programs along the head-to-tail axis of animals. The specific regional functions of individual Hox genes largely reflect their restricted expression patterns, whose disruption can lead to developmental defects and disease. Here we examine the spectrum of molecular mechanisms controlling Hox gene expression in model vertebrates and invertebrates and find that a diverse range of mechanisms including nuclear dynamics, RNA processing, microRNA and translational regulation all concur to control Hox gene outputs. We propose that this complex multi-tiered regulation might contribute to the robustness of Hox expression during development.

[100 words]

INTRODUCTION

Hox genes provide a paradigm for several areas in modern biology. Firstly, from a developmental perspective, they constitute a genetic system involved in the allocation of segmental identity along animal body axes. As such, they offer an opportunity to investigate how transcription factors organise networks of subordinate genes to guide the behaviour of cell populations during morphogenesis. Secondly, given their remarkable evolutionary conservation across distant animal phyla, they represent an abstract system of cardinal information able to operate within a wide spectrum of invertebrates and vertebrates, bringing about the question of how the same set of developmental genes can be involved in the generation of widely diverse developmental programmes. Third, their genomic organisation and molecular regulation is highly complex, opening up the possibility to look at the molecular mechanisms by which genomic information is extracted and subsequently converted as development progresses. This review focuses on this latter, gene regulatory dimension, aiming to provide an updated perspective on the variety of molecular mechanisms involved in Hox gene regulation (for recent reviews on the developmental and evolutionary functions of Hox genes see Duboule, 2007; Gehring et al., 2009; Lemons and McGinnis, 2006; Maeda and Karch, 2009; Mallo et al., 2010; Merabet et al., 2005; Pearson et al., 2005; Pick and Heffer, 2012). We consider that this is important given that models of how genes are molecularly regulated constrain our understanding of how they might exert their control over development in health and disease.

We start by providing some general features of the Hox genes and then use this background to look at problems of chromosomal and chromatin regulation, transcriptional control, RNA processing and microRNA (miRNA) regulation before then turning to the modulation of protein translation. The main discussion is set up using *Drosophila* and the mouse as key experimental models, bringing in information from other systems where useful and possible. Due

to space limitations, we give priority to regulatory mechanisms other than classical transcriptional regulation, given that these have been well covered elsewhere (Alexander et al., 2009; Maeda and Karch, 2009; Tschopp and Duboule, 2011). In summary, we aim to offer an account of the wide variety of molecular processes able to affect the regulation of Hox gene expression during animal development.

Key features of *Hox* genes

The Hox genes were discovered in *Drosophila* (Bridges and Morgan, 1923) where they exist in two separate gene clusters (Fig. 1): the Antennapedia and Bithorax complexes (ANT-C and BX-C, respectively) (Kaufman et al., 1980; Lewis, 1978; Lewis et al., 1980). Early genetic experiments in adult flies demonstrated that Hox genes are involved in the allocation of distinct morphological identities to each body segment: mutations affecting specific Hox genes typically lead to *homeotic* transformations, in which the morphology of a given segment is transformed into the likeness of another (Bateson, 1894). Further conceptual work on BX-C mutations by Ed Lewis showed that the gene order inferred from BX-C genetic maps was directly related to the anatomical areas influenced by the individual Hox genes: mutations in genes located at one end of the complex affected anterior larval structures while lesions on genes at the other end of the cluster affected posterior larval patterning (Lewis, 1978). This pioneering work indicated that Hox genes provide a genetic coordinate system for the allocation of developmental identities in the fly. However, understanding the mechanisms that link Hox genes to their developmental roles was only possible when the relevant genes were cloned, and their expression domains in the fly embryo, as well as their regulation, clarified. Indeed, molecular cloning of the BX-C (Bender et al., 1983) followed by expression analysis showed that Hox genes are expressed in particular sub-domains along the anteroposterior axis of the embryo (Akam, 1987; Harding et al., 1985). These

observations also revealed that the order in which individual Hox genes are expressed along the head-to-tail axis of the embryo mirrors the physical order of the Hox genes within the Hox cluster (Akam, 1987; Harding et al., 1985), a characteristic generally known as spatial collinearity, which had a strong impact in further investigations on the molecular regulation of Hox genes.

By the end of the 1980s, Hox genes had also been identified in the mouse genome. This remarkable observation suggested that they were likely to be present in most other vertebrates. Molecular analyses revealed that the basic genomic organization and expression patterns of the mouse and *Drosophila* Hox genes showed common features (Duboule and Dolle, 1989; Graham et al., 1989). In particular, it was shown that mouse Hox genes are organized in clusters and that their expression also follows the spatial collinearity principle. On the basis of these similarities, it was suggested that clustering is an intrinsic property of Hox genes that is indispensable for proper regulation of Hox gene expression. However, the cloning of Hox genes from other bilaterians, which revealed a variety of cluster structures as well as cases with a complete absence of clustering, raised some questions about the role of Hox gene clustering (Lemons and McGinnis, 2006; Duboule, 2007).

With the probable exception of lampreys (Smith et al., 2013), vertebrates possess at least four Hox clusters (teleost fishes have seven or eight clusters (Hurley et al., 2005)), possibly as a result of successive duplications of an ancestral cluster (Hurley et al., 2005) (Fig. 1). Although the configuration of vertebrate Hox clusters has been interpreted as a paradigmatic form of Hox gene organization, it might instead represent a rather exceptional case of organization and compaction (Duboule, 2007) that reflects intrinsic regulatory features of vertebrate Hox genes that are not necessarily present in other organisms (see below).

Hox transcriptional regulation in space and time

A rigorous analysis of gene interactions and Hox expression patterns during early *Drosophila* development (Akam, 1987) suggested that Hox expression domains are likely determined by at least three distinct regulatory inputs: transcriptional regulation from earlier segmentation genes (Irish et al., 1989), a cellular memory system based on the action of Polycomb (PcG) / trithorax (trxG) group proteins (Denell, 1978; Puro and Nygrén, 1975; Wedeen et al., 1986), and cross-regulatory interactions among the Hox genes themselves (Morata and Kerridge, 1982; Struhl, 1982). Regarding the latter mechanism, it was observed that the more posterior Hox genes are able to repress the expression (and suppress the function) of more anterior genes, a process termed '*posterior prevalence*' (Akam, 1987; Duboule and Morata, 1994; Hafen et al., 1984; Harding et al., 1985). Nonetheless, within single *Drosophila* segments the expression of each Hox gene is substantially modulated according to position and cellular type. Furthermore, in some developing structures, *Drosophila* Hox genes are activated in domains that are not discernibly related to the metameric specification of the animal (Akam, 1987; Akam and Martinez-Arias, 1985; White and Wilcox, 1985). These observations in the fly suggest that the control of Hox gene expression may have further dimensions to the three main components mentioned above.

In vertebrates, Hox gene expression includes an additional regulatory step. During initial stages of development, Hox genes are kept globally silent and become progressively activated during development following a temporal sequence that correlates with the gene's position within the cluster in a 3' to 5' direction (Kmita and Duboule, 2003). This property, commonly known as '*temporal collinearity*', has been observed in both the vertebrate primary body axis and developing limb buds (Kmita and Duboule, 2003). After initial activation, domains of vertebrate Hox gene expression are subsequently refined to produce the characteristic spatial distribution observed in the different tissues of older embryos (Deschamps and Wijgerde, 1993; Alexander et

al., 2009; Deschamps and van Nes, 2005). This latter and belated regulatory phase seems to rely on the activity of a second tier of control elements that produce specific Hox expression patterns, which are later maintained by the PcG and trxG systems (Alexander et al., 2009; Deschamps and van Nes, 2005). The tempo of Hox gene activation is functionally important because experimental conditions resulting in premature or delayed Hox gene activation have been shown to produce phenotypic alterations, even in cases when the final Hox expression patterns are preserved (Zakany et al., 1997; Juan and Ruddle, 2003; Gerard et al., 1997; Kondo and Duboule, 1999). This is consistent with the existence of distinct functional activities associated to early and late phases of vertebrate Hox gene expression (Carapuço et al., 2005). It has also been suggested that during early vertebrate development the general repressed state of the Hox cluster keeps the late regulatory elements in a 'silent state', and that it is only after global repression is erased that these elements become accessible to transcriptional regulators and, therefore, functional (Tschopp and Duboule, 2011) (Fig. 2).

A wealth of experimental data over the last three decades has led to the identification of many *cis*-regulatory elements that control Hox gene transcriptional patterns, thus furthering our understanding of their mechanisms of activity (Akbari et al., 2006; Alexander et al., 2009). These studies revealed that transcriptional *cis*-regulation represents a major determinant of Hox gene spatial expression, that *cis*-regulatory elements are functionally autonomous, and that the activity of *cis*-activity regulatory elements is independent of them being inserted within a Hox cluster (Whiting et al. 1991; Charite et al., 1995; Herault et al., 1998; Puschel et al., 1991). Therefore, despite the remarkable correlation between Hox expression domains along the anterior-posterior (AP) body axis and the gene's position within the cluster (Harding et al. 1985; Akam 1987; Graham, et al. 1989; Duboule and Dolle, 1989) (Fig. 1), the spatial control of Hox gene expression seems to be mostly independent of genomic arrangement. This is consistent with the

finding that in animals in which Hox genes are not clustered, Hox expression still maintains spatially collinear gene expression (considering the expression of paralogue groups) (Seo et al., 2004). Furthermore, the structure of the Hox cluster differs among species within the Drosophila genus; for example the cluster in *D. buzzattii* has two distinct splits compared to that of *D. melanogaster*, yet the changes do not seem to lead to any significant differences in Hox gene expression (Negre et al., 2005) arguing that the integrity of the gene complex is not an absolute requirement for the establishment of Hox expression patterns.

The chromatin component of Hox gene regulation

The functional autonomy of Hox *cis*-regulatory elements suggests that their activity with regards to broad transcriptional patterns of gene activation is largely independent of chromosome structural features that involve chromatin-based regulation. However, this notion does not hold true when looking at the molecular mechanisms underlying early phases of vertebrate Hox expression or those ensuring maintenance of appropriate Hox expression domains.

In vertebrates, the early global repression and subsequent collinear activation of Hox gene expression is closely associated with the physical characteristics of chromatin at the Hox clusters (Kmita and Duboule, 2003; Tschopp and Duboule, 2011). Indeed, a number of genomic manipulations in the mouse *Hoxd* locus indicate that temporal collinear gene activation is linked to its clustered structure (Tschopp and Duboule, 2011), which could explain the more compact structure of Hox loci in vertebrate genomes (Duboule, 2007).

Hox gene expression within the cluster correlates with the distribution of histone modifications associated with inactive (trimethylation at Lys27 of histone H3, H3K27m3) and active (trimethylation at Lys4 of histone H3, H3K4m3) chromatin (Schuettengruber et al., 2007) throughout the Hox cluster. In embryonic stem (ES) cells, which are thought to represent the

earliest stages of mammalian development and do not express Hox genes, Hox gene clusters are fully decorated by H3K27m3 and contain very low levels of H3K4m3, which is typically associated with transcriptional start sites in the configuration known as “bivalent” chromatin (Bernstein et al., 2006; Soshnikova and Duboule, 2009; Eskeland et al., 2010). When the *Hoxd* cluster was analysed in the tail tip of mouse embryos between E8.5 and E10.5 (representing different stages of temporal activation), Hox gene expression was closely followed by the disappearance of H3K27m3 and by a strong increase in H3K4m3 (Soshnikova and Duboule, 2009). A similar change in the H3K27m3 pattern was observed in *Hoxb* and *Hoxd* clusters of mouse ES cells undergoing RA-induced differentiation (Eskeland et al., 2010). This indicates that the temporal component of Hox gene activation is associated with the progressive change in chromatin marks that take the system from a silent to an active configuration (Fig. 2).

The correlation between H3K27m3 marks and inactive Hox clusters (Soshnikova and Duboule, 2009; Eskeland et al., 2010) suggests involvement of PcG in Hox gene silencing and genetic analyses in mice seem to support this hypothesis. PcG activity involves sequential action of the polycomb repressive complexes (PRC) 2 and 1. PRC2 first interacts with target DNA to introduce H3K27m3 marks into the chromatin, which are then read by PRC1 to promote silencing of the associated chromatin, perhaps by the incorporation of additional modifications to histones, most notably a ubiquitin moiety at lysine 119 of histone H2A (H2AK119) (Schuettengruber et al., 2007). ES cells with an inactive PRC2 fail to introduce the H3K27m3 mark into Hox cluster chromatin (Eskeland, et al., 2010; Shen et al., 2008). Similarly, active PRC1 is required for cluster silencing; ES cells lacking *RingB*, a E3 ligase that introduces ubiquitin into H2AK119 (de Napoles et al., 2004), fail to repress Hox gene expression, which in these cells reaches activation levels close to those obtained upon differentiation of normal ES cells (Eskeland, et al., 2010). Interestingly, the finding that Hox genes are expressed in *RingB*

mutant cells indicates that the presence of H3K27m3 alone in the Hox clusters is not sufficient to inactivate Hox gene expression because in these cells H3K27m3 distribution along the *Hoxb* or *Hoxd* clusters is largely not affected (Eskeland, et al., 2010).

It has long been established that PcG is essential for maintaining appropriate Hox gene expression domains by a repressive mechanism (Lewis, 1978). It is now recognized that PcG activity is also closely linked to modulation of specific chromatin states, which are usually not restricted to specific genes, but rather affect large chromosomal domains (Schwartz et al., 2006; Schuettengruber et al., 2007). Repression of *Drosophila* Hox gene expression is also associated with such a mechanism involving modulation of chromatin conformations. Indeed, in the fly embryo, repressed Hox genes are contained within a single PcG structural domain regardless of whether they reside within BX-C or ANT-C, indicating that the two Hox complexes belong to the same functional PcG domain (Bantignies et al., 2011; Tolhuis et al., 2011), despite being separated by about 10 Mbps.

Adding active and inactive-associated marks to Hox chromatin

The mechanisms responsible for targeting of PcG activity to specific loci seem to be different in *Drosophila* and vertebrates. In the fly, PRC2 complexes associate with *cis*-regulatory regions known as polycomb responsive elements (PREs) (Muller and Kassis, 2006; Schwartz et al., 2006). Recruitment of PcG activity to the *Drosophila* Hox genes has been extensively characterized (Akbari et al., 2006). Work in the late 1990's showed that coordinated PcG repression of BX-C and ANT-C genes is regulated by *Fab-7*, an element previously known to control PcG activity in BX-C (Mihaly et al., 1997). The finding that *Fab-7* deletion mutants also affect PcG activity on ANT-C suggests the existence of long-range interactions between the two *Drosophila* Hox complexes (Bantignies et al., 2011) (see below).

In vertebrates, the mechanisms of PcG targeting are largely unknown so very little can be said about PcG targeting of vertebrate Hox genes (Delest et al., 2012). A few studies identified some regions within vertebrate Hox clusters that could be involved in 'seeding' PcG activity (Kim et al., 2006; Woo et al., 2010). Of these, the best characterized is a phylogenetically conserved region between *HOXD11* and *HOXD12* that seems to be involved in homing PRC2 to the HOXD cluster in human stem cells (Woo et al., 2010). This region contains several of the characteristic features of *Drosophila* PREs and is able to promote repression of reporter genes when stably integrated in the genome (Woo et al., 2010). However, its role in the control of Hox gene expression is still unclear because its removal from the mouse genome had no evident deleterious effects on development or on the expression of *Hoxd* genes in mouse embryos (Beckers and Duboule, 1998; Tschopp et al., 2009). Therefore, the mechanisms that direct PcG to the Hox clusters to (i) produce global silencing during early development and (ii) secure proper spatial Hox gene expression at later developmental stages remain unclear.

The mechanisms controlling the sequential 3' to 5' loss of H3K27m3 from the Hox clusters are also far from being understood. One possibility is that this process results from regulated activity of specific demethylases. Two H3K27m3 demethylases, Kdm6b (Jmjd3) and Kdm6a (Utx) (Hong et al., 2007; Xiang et al., 2007), have been shown to interact with Hox genes and to modulate H3K27m3 levels at their promoters (Agger et al., 2007; Lan et al., 2007). Phenotypic analysis of genetically inactive and/or knocked-down *Kdm6* genes indicate that, of the two demethylases, Kdm6a emerges as a more promising Hox regulator given that experiments that remove or decrease its activity lead to Hox-like phenotypes (Lan et al., 2007; Lee et al., 2012; Satoh et al., 2010).

Hox cluster activation not only correlates with reduced H3K27m3 but also with a broad increase in H3K4m3, which seems to precede H3K27m3 disappearance in the mouse tail tip

(Soshnikova and Duboule, 2009). Therefore, the regulated addition of chromatin activation marks could be a key driver of progressive Hox cluster activation. TrxG is a leading candidate to play a role in such process because its activity adds H3K4m3 marks to chromatin (Schuettengruber et al., 2007). So far the role of trxG in Hox cluster opening has not been directly evaluated and the available data from different trxG mutants do not allow a clear assessment of its role during early phases of vertebrate Hox expression (Ayton et al., 2001; Glaser et al., 2006; Stoller et al., 2010; Yu et al., 1995; 1998). Interestingly, *Kdm6a* has been shown to establish functional interactions with the trxG member *Mll2*, which encodes for a methyl transferase that produces H3K4m3 (Issaeva, et al., 2007), suggesting that sequential activation of Hox clusters could result from a coordinated action of H3K4m3-promoting and H3K27m3-removing activities.

The mechanisms controlling the early phases of vertebrate Hox gene activation are less well known than those regulating late Hox gene expression in limbs and other domains (Alexander et al., 2009; Tschopp and Duboule, 2011). Coordinated activation of *Hoxd* genes in limb buds depends on specific global control regions (Tschopp and Duboule, 2011), although the mechanistic details of their activity are mostly unknown. Recent analyses detected physical interactions between these control regions and *Hoxd* genes (Montavon et al., 2011). However, the role of these interactions in Hox gene activation is not clear because they were also observed in tissues that do not express *Hoxd* genes. In the main body axis, Hox gene activation seems to depend on the coordinated activity of several signaling pathways, including those activated by retinoic acid, fibroblast growth factors (FGFs), Wnts and growth differentiation factor 11 (Gdf11) (Iimura and Pourquié, 2007; Alexander et al., 2009). In addition, a variety of genetic analyses suggest that Cdx transcription factors could be a core integrator of these signals in the growing posterior embryo (Young and Deschamps, 2009). However, although the impact of these factors

on Hox gene expression is clear, their relationship to the chromatin changes associated with vertebrate Hox gene activation remains to be clarified.

3D chromatin conformation and Hox gene expression

Recent data shows that PcG (and possibly trxG) function is connected with specific 3D chromatin configurations within the nucleus. Analysis of the sub-nuclear localization of PRC components revealed that PcG activity is confined to distinct nuclear regions termed polycomb (PC) bodies (Buchenau et al., 1998). In addition, extensive chromatin contacts have been detected among PcG-regulated areas, even if these are separated by considerable distances (Delest et al., 2012). *Drosophila* Hox genes provide a good example of this regulatory mechanism, showing that all inactive Hox genes nucleate at the same PC body and display extensive mutual physical interactions regardless of whether they are in the BX-C or the ANT-C complex (Bantignies et al., 2011; Tolhuis et al., 2011). These observations reinforce the notion of coordinated regulation of the *Drosophila* Hox genes by PcG, and show that this regulation involves arrangement of their associated chromatin within the same nuclear domain.

Specific 3D chromatin configurations are also likely to be involved in Hox gene regulation during early vertebrate embryo development. Analyses of the Hox clusters in undifferentiated human embryonal carcinoma and mouse ES cells using chromosome conformation capture approaches, revealed strong chromatin interactions spanning the whole cluster (Ferraiuolo et al., 2010; Noordermeer et al., 2011), indicating that inactive vertebrate Hox clusters might also be packed together within a specific nuclear compartment. A similar conclusion emerged from fluorescent *in situ* hybridization (FISH) studies showing that 3' and 5' areas of the inactive *Hoxb* and *Hoxd* clusters localize together within their respective chromosomal territory (i.e. the area of the interphase nucleus occupied by each chromosome)

(Chambeyron and Bickmore, 2004; Morey et al., 2007). Interestingly, these chromosome interactions are essentially restricted to the Hox cluster itself, with very little extension into adjacent chromatin (Noordermeer et al., 2011), suggesting the existence of a Hox cluster-specific marking mechanism that controls chromosomal configuration.

Studies using mutant ES cell lines support the involvement of PcG in Hox cluster compaction. In particular, in ES cells deficient for either the PRC2 component *Eed* or the PRC1 component *RingB*, the *Hoxb* and *Hoxd* clusters adopted an open configuration (Eskeland, et al., 2010). It was shown that undifferentiated ES cells also contain PC bodies that remodel with differentiation (Ren et al., 2008). Whether silent Hox clusters are also contained within those PC bodies is as yet unclear.

Activation of Hox gene expression both in differentiating cells and in the developing tail and limb buds is associated with major structural changes in the Hox cluster chromatin (Chambeyron and Bickmore, 2004; Chambeyron et al., 2005; Eskeland, et al., 2010; Ferraiuolo et al., 2010; Montavon et al., 2011; Morey et al., 2007; Noordermeer et al., 2011). In particular, inactive and active areas of the Hox cluster segregate into separate domains. Inactive areas keep their mutual interactions but lose contact with active regions of the same cluster (Ferraiuolo et al., 2010; Noordermeer et al., 2011; Montavon et al., 2011). Interestingly, interactions were still observed throughout H3K4m3-rich active areas of the Hox clusters (Noordermeer et al., 2011; Montavon et al., 2011). A similar type of interaction between the H3K4m3-containing active areas of the *HOXA* cluster was also observed in primary human fibroblasts and in a macrophage cell line (Guenther et al., 2005; Wang et al., 2011), indicating that such interactions are a general characteristic of active Hox clusters. Therefore, active Hox chromatin might be precisely organized within a defined nuclear domain that is physically and functionally different to the nuclear domain containing the inactive Hox cluster.

This view is consistent with FISH analyses showing that gene activation within the *Hoxb* and *Hoxd* clusters is associated with the relocation of their active domain into a position within the nucleus that is physically separated from that occupied by the silent part of the cluster (Chambeyron and Bickmore, 2004; Chambeyron et al., 2005; Morey et al., 2007). It was recently reported that chromatin in the interphase nucleus is organized into distinct globular compartments, where active and inactive chromatin are segregated from each other (Lieberman-Aiden et al., 2009). The 3D structural characteristics of Hox clusters at different activation stages are consistent with a model in which active and inactive Hox chromatin resides separately within such globular compartments and the activation process requires a transit between them. Because the activation sequence of vertebrate Hox genes follows the gene's position in the cluster, it can be hypothesized that the movement of Hox cluster DNA between compartments is produced by linear sliding of the chromatin fiber in a 3'-to-5' direction. The characteristic knot-free organization of fibres within the globular compartments (Lieberman-Aiden et al., 2009) would facilitate the chromatin linear unwinding and rewinding required for this type of mechanism. Such a model also fits with the alterations observed in the temporal activation of mouse *Hoxd* genes when their distance to the 3' end of the cluster was modified either by deletions or duplications in the cluster (Tschopp et al., 2009).

Although 3D organization might follow similar general principles for all four mammalian Hox clusters, it is possible that there are also cluster- and tissue-specific features. For instance, global interactions within the inactive *Hoxb* cluster seem to differ slightly from those observed in the other clusters (Noordermeer et al., 2011). In addition, while active regions of the *Hoxb* and *Hoxd* clusters move to a position outside their chromosomal territory within the nuclei of ES and tail bud cells (Chambeyron and Bickmore, 2004; Chambeyron et al., 2005; Morey et al., 2007), in

the limb bud cell nuclei active *Hoxd* genes separate from the inactive areas of the cluster but stay within their chromosomal territory (Morey et al., 2007).

In conclusion, several aspects of Hox gene expression are strongly connected to higher order structural features of chromatin, suggesting that a better understanding of the structural and functional characteristics of nuclear compartments and the dynamics of chromatin transit among them is likely to impact current models of Hox gene regulation.

Long non-coding RNAs in the regulation of Hox gene expression

Recent studies suggest that long non-coding RNAs (lncRNAs) could interact with transcription factors and chromatin modifiers to modulate gene function during development (Lee, 2012). In this context, one pervasive role of lncRNAs is to act as a device that targets PcG and trxG activity to specific functional targets.

Drosophila Hox genes have been reported to be among the targets of lncRNA-mediated gene regulation. A well-studied case is that of the lncRNA *bithoraxoid* (*bxd*) in the control of *Ubx* transcription. *bxd* is expressed from the genomic region upstream of *Ubx* and is transcribed in the same direction as *Ubx*. The exact functional relationship between *bxd* and *Ubx* remains unclear, however, as two groups have reported conflicting reports on this issue (Petruk et al., 2006; Sanchez-Elsner et al., 2006). Nonetheless, in both cases *bxd* seems to recruit trxG proteins to the *cis* elements regulating expression of these RNAs. In one of the cases, it was proposed that *bxd*-mediated trxG recruitment promotes its own transcription, eventually resulting in the silencing of *Ubx* (Petruk et al., 2006). In the other study, it was argued that *bxd* brings trxG proteins to the *Ubx* promoter, thus promoting *Ubx* transcription (Sanchez-Elsner et al., 2006).

The lncRNA *iab-8* has also been implicated in the control of Hox gene expression in *Drosophila*. Posterior expression of *abdA* in the *Drosophila* embryonic central nervous system is

repressed by the lncRNA *iab-8*, which is transcribed from the genomic region between *AbdB* and *abdA* and seems to repress *abdA* transcription in *cis* by a transcriptional interference mechanism (Gummalla et al., 2012). These observations led to the suggestion that *iab-8* might play a role in the process of posterior prevalence (Gummalla et al., 2012).

lncRNAs have also been implicated in the control of vertebrate Hox gene expression. One of them, *HOTTIP*, is transcribed from a region 5' of *HOXA13*, in the opposite direction to the *HOXA* genes (Wang et al., 2011). *HOTTIP* activity in human cells seems to be linked to interactions with the trxG protein WDR5, which recruits other trxG members to introduce the H3K4m3 mark into the 5' area of the *Hoxa* cluster to keep it transcriptionally active (Wang et al., 2011). Experiments performed in chicken embryos also suggest that *HOTTIP* is required to induce and/or maintain expression of the 5' *Hoxa* cluster genes in the distal part of developing limb buds (Wang et al., 2011). A regulatory role has also been described for another non-coding antisense transcript spanning the *Hoxb5* and *Hoxb6* loci (*Hoxb5/b6as*) (Dinger et al., 2008). *Hoxb5/b6as* has also been reported to interact with the trxG regulatory network in differentiating mouse embryoid bodies, most notably with Mll1 and with H3K4m3, either to activate or maintain the transcription of *Hoxb5* and *Hoxb6*.

HOTAIR is another lncRNA that has been proposed to regulate vertebrate Hox genes. In human cells, *HOTAIR* is transcribed from the region between *HOXC11* and *HOXC12* (Rinn et al., 2007). Intriguingly, unlike most other characterized lncRNAs, *HOTAIR* does not seem to regulate the expression of adjacent transcripts, but it rather appears to block expression of genes within the posterior *HOXD* cluster. HOTAIR seems to work by co-ordinately targeting PRC2 and the histone demethylase LSD1 (Shi et al., 2004) to specific areas of the *HOXD* complex (Rinn et al., 2007; Tsai et al., 2010). However, HOTAIR sequence conservation between human and mouse is surprisingly low, and its removal from the mouse genome has no detectable negative effects on

embryonic development or on the distribution of the H3K27m3 mark throughout the *Hoxd* cluster (Schorderet and Duboule, 2011), suggesting that *HOTAIR* represents a novel mechanism that emerged in the human lineage and/or its roles are compensated by other factors.

In summary, experiments so far indicate that, in *Drosophila*, lncRNAs do play a role in the regulation of Hox expression and developmental functions in embryos (Gummalla et al., 2012; Petruk et al., 2006; Sanchez-Elsner et al., 2006). In vertebrates, although growing evidence supports the notion that lncRNAs control Hox gene expression, their participation in physiological processes is still debated.

Hox regulation via RNA processing

Although transcriptional and chromatin-mediated regulation play a key role in gene expression, a range of co-transcriptional and post-transcriptional processes are also able to affect mRNA quality and/or quantity and thus represent an additional layer of gene control in eukaryotic organisms (Licatalosi and Darnell, 2010). For instance, differential RNA processing via alternative splicing, alternative polyadenylation, and/or alternative promoter usage can lead to the formation of mRNA transcripts that carry substantially different information and/or structures (Licatalosi and Darnell, 2010; Nilsen and Graveley, 2010; Proudfoot, 2011). In line with this general regulatory potential, RNA processing does play an important role in the regulation of Hox gene expression.

One of the best examples of the impact of alternative splicing on Hox function concerns the *Drosophila Ubx* gene (Lewis, 1978; Sanchez-Herrero et al., 1985). *Drosophila Ubx* produces six distinct mRNA isoforms (Kornfeld et al., 1989; O'Connor et al., 1988) via a recursive splicing mechanism (Hatton et al., 1998) that releases a family of transcripts bearing common 5’ and 3’ exons but carrying distinctive small exonic sequences (microexons m1 and m2) and/or a 9-aa

extension to the 3'exon (the b element) (Kornfeld et al., 1989; O'Connor et al., 1988). Notably, the same spectrum of *Ubx* isoforms found in *D. melanogaster* is present in distantly related *Drosophila* species, implying that the distinctive pattern of *Ubx* splicing has been maintained for over 60 million years of independent evolution (Bomze and Lopez, 1994). Furthermore, each alternatively spliced *Ubx* mRNA isoform displays a characteristic developmental expression pattern in *D. melanogaster* and such patterns are also evolutionary conserved (Bomze and Lopez, 1994). This strongly suggests functional implications. Yet, despite initial reports in cell culture, which highlighted specific functions for *Ubx* isoforms (Gavis and Hogness, 1991; Krasnow et al., 1989), early genetic experiments in flies suggested that all isoforms were largely equivalent in their developmental roles (Busturia et al., 1990). Nonetheless, further studies focused on specific developmental, cellular and molecular roles under Ubx control have consistently and conclusively demonstrated that alternatively spliced Ubx isoforms perform distinct functions during the development of the embryo and adult fly. For example, heat-shock induction of *Ubx* alternatively spliced isoforms leads to isoform-specific transformations in the peripheral nervous system (PNS). In particular, ectopic expression of the *Ubx* splicing isoform *Ia* but not *Ubx-IVa* transforms the PNS of thoracic parasegments into the PNS of an abdominal-like segment (Mann and Hogness, 1990). Further genetic work on the PNS showed that indeed *Ubx-IVa* cannot substitute functionally for other isoforms to promote normal development of the PNS (Subramaniam et al., 1994). Recent studies using the *Drosophila* UAS/Gal4 system further demonstrated the impact of *Ubx* alternative splicing on the activation of Ubx target genes during embryogenesis (Reed et al., 2010) and in the ability of *Ubx* to control the morphology and underlying gene networks of adult appendage development (de Navas et al., 2011).

Notably, the molecular characterization of five of the remaining seven *Drosophila Hox* genes, including *proboscipedia* (*pb*), *Antennapedia* (*Antp*), *labial* (*lab*), *abdominal A* (*Abd-A*) and

Abdominal-B (*Abd-B*) revealed that they also undergo differential RNA processing either in the form of alternative splicing, alternative promoter usage and/or alternative polyadenylation (Celniker et al., 1989; Celniker et al., 1990; Kuziora and McGinnis, 1988; Laughon et al., 1986; Mlodzik et al., 1988; O'Connor et al., 1988; Sanchez-Herrero and Crosby, 1988; Schneuwly et al., 1986; Scott et al., 1983; Stroeher et al., 1986). Furthermore, inspection of recent RNA-sequencing data suggests that *Sex combs reduced* (*Scr*) may also undergo a form of RNA processing via alternative promoter usage (Patraquim and Alonso, unpublished). Altogether these observations indicate that RNA processing affects the expression of most –perhaps even all– *Drosophila* Hox genes.

In contrast, much less is known about the extent to which RNA processing affects the expression of mammalian Hox genes. A number of studies provide strong evidence that several rodent Hox genes undergo RNA processing, mostly via alternative splicing. For example, *Hoxa1* was shown to produce two alternatively spliced transcripts that encode two different protein isoforms (LaRosa and Gudas, 1988). Similarly, several transcripts were described for *Hoxa9*, originating from both alternative splicing and differential promoter usage (Fujimoto et al., 1998; Rubin and Nguyen-Huu, 1990). The different mRNA species of *Hoxa9* encode different versions of the Hoxa9 protein, two of them containing the same homeodomain but differing slightly upstream of this motif (Rubin and Nguyen-Huu, 1990), and another encoding a homeodomain-less protein (Fujimoto et al., 1998). Transcripts encoding different protein versions can be detected in embryonic tissue (LaRosa and Gudas, 1988; Fujimoto et al., 1998), suggesting that they could be functionally relevant, but a direct evaluation of this possibility remains to be performed. Interestingly, alternatively spliced mRNAs are also produced from the *Hoxa10* locus, and encode a predicted protein that is essentially the homeodomain without any significant N-terminal attached to it (Benson et al., 1995). The role of such a molecule is still unclear, as

functional analysis in transgenic mice indicate that it is not able to activate the typical Hoxa10 patterning program in the axial skeleton or interfere with it (Guerreiro et al., 2012). *Hoxb3* and *Hoxa5* also produce different mRNA isoforms through a combination of alternative splicing and alternative promoter usage; interestingly the resulting isoforms are expressed in different domains during embryonic development (Chan et al., 2010; Coulombe et al., 2010; Sham et al., 1992), once again suggesting that the resulting isoforms might carry out specific functions. Finally, a recent bioinformatic study developed by Pedro Patraquim and Claudio Alonso focused on mouse Hox genes revealed that approximately a third of them produce a spectrum of mRNA isoforms generated by different types of RNA processing, many of which are evolutionary conserved between mouse and human. These findings strongly suggest that RNA processing represents a widespread and somewhat overlooked form of gene regulation that can affect mammalian Hox genes.

Hox regulation via miRNAs

Studies in the mouse and *Drosophila* show that regulation via miRNAs represents another tier in the complex molecular regulatory circuit controlling Hox gene expression. miRNAs are short (22-mer) non-coding RNAs that repress gene expression by binding to complementary target sequences in 3'UTRs of mRNAs leading to mRNA degradation and/or translational repression (Bartel, 2009; Lee and Shin, 2012). Several miRNAs have been shown to regulate Hox gene expression. Notably, most examples available to date show that Hox-regulatory miRNAs are encoded within the Hox clusters. This genomic arrangement may provide an effective mechanism to guide the production of miRNAs in the "same cells and at the same time" as their Hox mRNA targets (Alonso, 2012). In vertebrates, these miRNAs include miR-196, miR-10 and miR-615 (Fig. 1).

The mouse has three miR-196 gene paralogues - miR-196b, miR-196a-1, miR-196a-2 - that are situated within Hox clusters a, b and c, respectively. Intriguingly, the mouse miR-196 genes are always situated between the Hox9 and Hox10 genes, a configuration that is also observed for the miR-196 genes present in the frog (*Xenopus laevis*), zebrafish (*Danio rerio*) and human genomes. Assuming that sufficient evolutionary time has elapsed to probe different chromosomal orders, this absence of change in miRNA location could reflect that the chromosomal position of this miRNA is essential for cell survival and/or normal development. In the case of miR-10, the mouse has two genes encoding this miRNA, miR-10a and miR-10b, that are located in Hox clusters a and d, respectively. Unlike miR-196, which is specific to chordates (Mansfield and McGlinn 2012), miR-10 represents a much more ancient miRNA system (see below). Remarkably, miR-10 genes are always found in-between the same Hox genes (i.e. paralogue groups (PGs) 4 and 5). Finally, miR-615 is confined to Eutherians (non-monotreme, non-marsupial mammals) where it is produced from an intron in *Hoxc5*.

miR-196 was the first miRNA to be experimentally shown to regulate the expression of a Hox gene (*Hoxb8*) (Yekta et al., 2004). Downregulation of miR-196 in the chick (McGlinn et al., 2009) and zebrafish (He et al., 2011) leads to Hox protein de-repression and the generation of homeotic transformations, highlighting the importance of this miRNA. Similarly, downregulation of miR-10 in zebrafish led to over-expression of *Hoxb1a* and *Hoxb3a* (Woltering and Durston, 2008), further confirming the notion that Hox complex miRNAs are important factors that modulate Hox protein expression during vertebrate development. Although miR-615 has been implicated in several pathologies affecting human liver function (El Tayebi et al., 2012), its roles in Hox gene expression have not yet been investigated.

In *Drosophila,* at least four miRNA genes are located within the Hox clusters: miR-iab4, miR-iab8, miR-10 and miR-993. Of these, miR-iab4 and miR-iab8 were shown to target several

Hox genes including *Ubx*, *Antp*, *Abd-a* and *Abd-b* (Bender, 2008; Ronshaugen et al., 2005; Stark et al., 2008; Thomsen et al., 2010; Tyler et al., 2008). Curiously, miR-iab4 and miR-iab8 are both encoded by the opposite DNA strands of the same locus, which is located between the BX-C genes *Abd-A* and *Abd-B* (Bender, 2008; Gummalla et al., 2012; Stark et al., 2008; Thomsen et al., 2010; Tyler et al., 2008). Despite the fact that the *Drosophila* miR-iab4/8 system and the mouse miR-196 are located in equivalent positions relative to orthologue Hox genes in each species, the murine and fly miRNAs show no sequence similarity with each other arguing that these miRNAs have evolved independently in the insect and mammalian lineages (Lemons et al., 2012). This observation is remarkable in that it might imply the requirement of a Hox-controlling miRNA system to be present within these particular genomic locations for Hox expression to be adequately regulated.

Ectopic over-expression of miR-iab4 in the *Drosophila* haltere disc leads to a mild, but clear, haltere-to-wing transformation closely reminiscent of a *Ubx*-like phenotype and also produces a reduction in Ubx protein levels indicating that this miRNA can indeed regulate *Ubx* gene expression during appendage development (Ronshaugen et al., 2005). However, genetic removal of the miR-iab4/8 locus does not lead to any evident homeotic transformation in the haltere or elsewhere in the developing fly (Bender, 2008; Thomsen et al., 2010). These experiments raise doubts about the biological significance of Hox regulation by miR-iab4/8. However, expression of Ubx protein within the embryonic central nervous system (CNS) is markedly increased in miR-iab4/8 mutants suggesting that miR-iab4/8 might exert a biologically relevant regulatory function within this tissue (Bender, 2008; Thomsen et al., 2010).

A possible explanation for the absence of biological interactions between miR-iab4/8 and *Ubx* during early development emerged from the observation that several *Drosophila* Hox genes undergo developmentally-regulated alternative polyadenylation so that longer *Ubx* 3'UTR

isoforms (i.e. those containing more miRNA target sites) are only expressed within the CNS (Thomsen et al., 2010), making it likely that such longer forms are specifically subjected to miR-iab4/8 repression while shorter isoforms –those expressed at earlier developmental stages– might be immune to miR-iab4/8 downregulation or segregated to distinct cellular expression domains (Thomsen et al., 2010). Notably, alternative polyadenylation of Hox transcripts is not an exclusive feature of *D. melanogaster*. A computational study of different *Drosophila* species (Patraquim et al. 2011) revealed that alternative Hox transcript polyadenylation patterns are maintained across *Drosophila* lineages that have evolved independently from each other for c. 60 million years, implying that this is an ancestral feature of the *Drosophila* group with functional relevance in modern fruitflies. Notably, the interplay between Hox RNA processing via alternative polyadenylation and miRNA regulation (Fig. 3) is not confined to the *Drosophila* Hox genes.

The evolution of the miR-10 gene is markedly different from that of the miR-iab4/8 locus. A very high level of sequence similarity between fly and mouse miR-10 genes strongly suggests that this miRNA system was a feature of the ancient common ancestor between insects and mammals (Lemons and McGinnis, 2006). This remarkable evolutionary conservation strongly suggests an important functional role for miR-10 across the bilaterians. These considerations are, however, at odds with recent work in *Drosophila* embryos (Lemons et al., 2012), which showed that the two functional miRNAs derived from miR-10 (i.e. miR-10-5p and miR-10-3p), which are predicted to target *sex combs reduced* (*scr*) and *Abdominal-b* (*Abd-B*) respectively, have no major effects on the expression of these *Hox* genes. Indeed, this study also failed to detect the interaction of miR-iab4-5p with one of its *Hox* targets, *Antp*, at least during embryogenesis (Lemons et al., 2012). This suggests that whatever the biological roles of miR-10 are during *Hox* expression, they might only be detectable in specific developmental contexts or come about as a

result of combinatorial interactions with other RNA regulators or when coupled with selective RNA processing patterns of their targets.

Thus, it is still early days with regards to understanding the underlying mechanisms and biological roles of *Hox* regulation by miRNAs. Although our discussion has been focused on Hox-embedded miRNAs, many other miRNA systems encoded outside the Hox clusters could potentially play key roles in the regulation of Hox gene expression.

Translational control

Examples of translational control of Hox protein production exist but are, to date, scarce. Yet this might not necessarily reflect the lack of relevance of this type of regulation but instead the absence of a systematic search for such translational regulatory effects on Hox gene expression.

The *Drosophila* genes *Antp* and *Ubx* for example, have been suggested to be under translational control during embryonic development (Oh et al. 1992; Ye et al., 1997). In particular, it has been shown that a subgroup of transcripts produced from the *Antp* and *Ubx* loci contain functional internal ribosomal entry sites (IRES) that allow their translation using a "cap"-independent mechanism (Oh et al. 1992; Ye et al., 1997). The IRES elements of both gene transcripts were reported to be functional in the *Drosophila* embryo and their activity to be developmentally regulated (Ye et al., 1997), providing patterns of activity that match the known expression domains of the endogenous Antp and Ubx proteins (Ye et al., 1997). This suggests that IRES-mediated translational regulation could indeed be involved in controlling the spatial/temporal production of the Antp and Ubx proteins during *Drosophila* development. Interestingly, IRES-mediated regulation of Antp must be coordinated with other regulatory processes at the transcriptional and splicing levels because production of the IRES-containing

mRNA requires transcription from the alternative P2 promoter and a specific splicing scheme (Oh et al. 1992).

Translational regulation of Hox expression was also documented for mouse Hoxb4 a decade ago (Brend et al. 2003). Within the murine developing spinal cord *Hoxb4* transcripts are localized in broader domains than those of the corresponding protein. Interestingly, this regulation is tissue specific because in the paraxial mesoderm both *Hoxb4* mRNA and protein were detected in similar areas. However, the mechanistic details and functional significance of this form of *Hoxb4* regulation are still unknown.

More recently, cloning of the mutant allele of three mouse mutant strains, known as *Tail short* (*Ts*), *Tail-short Shionogi* (*Tss*) and *Rabo torcido* (*Rbt*), led to the identification of a new type of translational regulation that might impact the production of several Hox proteins (Kondrashov et al., 2011). The axial skeletons of these mutants presented a variety of homeotic transformations reminiscent of those associated with inactivating mutations of several Hox genes. All three strains are heterozygous for an inactivating mutation in the ribosomal protein RPL38, an integral component of the 60S ribosomal subunit. The RPL38 deficiency resulted in inefficient translation of a specific set of transcripts, without apparent negative effects on global translation. Transcripts for eight different Hox genes were among those that failed to be translated in *Ts* mice. The skeletal alterations observed in these mutants fitted well with a combination of the phenotypes observed in mutant mice for the Hox genes affected by the RPL38 deficiency, supporting the involvement of alterations in Hox protein levels in the *Ts* mutant phenotype. Biochemical analyses indicated that RPL38 is required during translational initiation to build a mature 80S ribosomal particle (Kondrashov et al., 2011). It is thus possible that *Ts* mutant tissues contain suboptimal amounts of RPL38 creating a partial deficiency of functional (RPL38-containing) 60S ribosomal particles that are able to engage in active translation. Accordingly, the

different translational efficiencies of different transcripts in the tissues of *Ts* mice might reflect their intrinsic ability to compete for limiting amounts of functional 60S subunits.

It is not clear if the molecular phenotype of *Ts* mice represents a hypomorphic phenotype with no relevance to physiological regulatory processes or if it has uncovered a new *bona fide* regulatory mechanism. If the latter case is true, it suggests that regulated production of the RNA and protein components of the ribosome could be relevant for the control of specific developmental processes. Interestingly, analysis of the levels and distribution of a large set of ribosomal proteins in the mouse embryo revealed an unexpectedly high degree of tissue specific expression (Kondrashov et al., 2011). Particularly interesting are ribosomal proteins, like RPS25, that have been suggested to interact with IRES (Nishiyama et al., 2007; Landry et al., 2009) and could then be involved in regulatory processes similar to those suggested above for Ubx or Antp in *Drosophila*.

An important question that remains to be addressed is whether translational control mechanisms similar to those observed in *Ts* mice are vertebrate-specific (or even mouse-specific) or if they represent an evolutionary conserved control mechanism.

CONCLUDING REMARKS

Our discussion here reveals a remarkably wide spectrum of regulatory mechanisms involved in the control of Hox expression, many of which have been evolutionary conserved between insects and mammals (Table 1). Although each Hox-regulatory mechanism is likely to have its own intrinsic properties it is important to note that many of the mechanisms presented seem to be linked to one another either by the regulatory input imparted by common factors or other forms of functional coupling.

Examples in *Drosophila* show, for instance, that the process of alternative promoter usage affecting the expression of mRNAs (Oh et al. 1992) dictate the type of 5'UTRs included in the mRNA, which, in turn, could affect protein translation. Definition of Hox mRNA 'ends' via the coupled reactions of cleavage, polyadenylation and transcriptional termination determines the nature of 3'UTR information on board Hox transcripts, which will have direct impact on the set of regulatory interactions between Hox mRNAs and RNA regulators such as miRNAs and RNA binding proteins involved in the control of mRNA degradation and translation efficiency (Thomsen et al., 2010; Alonso, 2012). Further examples of gene expression coupling include the effects of transcriptional elongation rates on the spectrum of alternatively spliced forms of *Ubx* (de la Mata, et al. 2003) and the consequences of nonsense-mutations on *Ubx* mRNA stability and alternative splicing (Alonso and Akam, 2003).

In vertebrates, Hox cluster miRNAs seem to follow a temporal expression pattern highly similar to that of their targets, leading to the suggestion that miRNAs could be an integral part of the mechanisms behind a long known, but still poorly understood, property of Hox genes: posterior prevalence (Yekta et al., 2008). miRNAs synthesized within the Hox clusters have also been suggested to play a role in posterior prevalence in *Drosophila* (Gummalla et al., 2012), thus indicating that similar mechanisms could be operating in vertebrates and flies. Interestingly, it has been shown that the role of miRNAs in posterior prevalence can be reinforced by a lncRNA transcribed from the same genomic area (Gummalla et al., 2012). Therefore, the coordinated evolution of the different regulatory processes might be a core feature underlying the Hox expression program.

In addition, we have discussed the impact of 3D chromatin conformation on several aspects of Hox gene regulation. Although it is still too early to draw a clear picture on this process, it is plausible that early stages of vertebrate Hox gene activation involve a sequential

relocation of Hox chromatin from transcriptionally silent nuclear domains into nuclear areas engaged in active transcription. Also, although it now seems clear that both Hox chromatin movements within the nucleus and Hox gene silencing/activation are connected to PcG and trxG activities, it will be important to understand how these processes are mechanistically linked together. In this context, it is conceivable that recently discovered regulatory players, including non-coding RNAs, like lncRNAs or enhancer-associated transcripts, might influence chromatin structure and cluster-wide control mechanisms.

Yet these functional links across the many processes involved in gene expression are not a unique feature of the Hox genes. Indeed, more than a decade ago Tom Maniatis and Robin Reed suggested the existence of a pervasive level of functional coupling across all the molecular machines involved in gene expression control (Maniatis and Reed, 2002). However most work addressing the mechanisms and biological implications of gene expression coupling has so far been conducted in mammalian cultured cells or yeast leaving the question of the extent to which these interconnections play a relevant role within the physiological context of development largely unexplored. In this context, Hox genes might provide an excellent system to investigate both the mechanisms and biological roles of the many interconnections across different gene regulatory levels that are traditionally studied in isolation. One possibility in regards to the biological roles of a highly interconnected network of Hox-regulatory interactions is that this has emerged to increase robustness of the Hox expression program, ensuring that correct spatial and temporal patterns of Hox expression are achieved despite intrinsic molecular variation and extrinsic fluctuations in embryonic environment. More generally, the existence of this highly interconnected form of control might suggests that animal embryos must carefully control Hox gene expression in order to complete development successfully. This interlocked regulatory network arrangement might also provide an explanation for why core features of Hox expression,

especially those related to expression domains, have been maintained largely unchanged during the extended periods of bilaterian evolution.

ACKNOWLEDGEMENTS

CRA wishes to thank Michael Akam, as well as present and past members of the Alonso Lab for inspirational discussions about the workings of the Hox genes. Work in CRA's lab was funded by The Wellcome Trust. MM's laboratory was supported by grants PTDC/BIA-BCM/110638/2009 and PTDC/SAU-BID/110640/2009.

REFERENCES

Agger, K., Cloos, P. A., Christensen, J., Pasini, D., Rose, S., et al., (2007). UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. *Nature* **449,** 731–734.

Akam, M. (1987). The molecular basis for metameric pattern in the Drosophila embryo. *Development* **101**, 1-22.

Akam, M. E. and Martinez-Arias, A. (1985). The distribution of Ultrabithorax transcripts in Drosophila embryos. *EMBO J.* **4**, 1689-700.

Akbari, O. S., Bousum, A., Bae, E. and Drewell, R. A. (2006). Unraveling cis-regulatory mechanisms at the abdominal-A and Abdominal-B genes in the Drosophila bithorax complex. *Dev. Biol.* **293,** 294-304.

Alexander, T., Nolte, C. and Krumlauf, R. (2009). Hox genes and segmentation of the hindbrain and axial skeleton. *Annu. Rev. Cell. Dev. Biol.* **25,** 431-456.

Alonso, C. R. (2012). A complex 'mRNA degradation code' controls gene expression during animal development. *Trends Genet.* **28**, 78-88.

Alonso, C. R. and Akam, M. (2003) A *Hox* gene mutation that triggers Nonsense-mediated RNA decay and affects alternative splicing during *Drosophila* development, *Nucleic Acids Res.* **31,** 3873-3880.

Ayton, P., Sneddon, S. F., Palmer, D. B., Rosewell, I. R., Owen, M. J., et al., (2001). Truncation of the Mll gene in exon 5 by gene targeting leads to early preimplantation lethality of homozygous embryos. *Genesis* **30,** 201-212.

Bantignies, F., Roure, V., Comet, I., Leblanc, B., Schuettengruber, B., et al. (2011). Polycomb-dependent regulatory contacts between distant Hox loci in Drosophila. *Cell* **144,** 214-226.

Bartel, D. P. (2009). MicroRNAs: target recognition and regulatory functions. *Cell* **136**, 215-33.

Bateson, W. (1894). Materials for the Study of Variation. New York: Macmillan and Co.

Beckers, J., and Duboule, D. (1998). Genetic analysis of a conserved sequence in the HoxD complex: regulatory redundancy or limitations of the transgenic approach? *Dev. Dyn.* **213,** 1–11.

Bender, W. (2008). MicroRNAs in the Drosophila bithorax complex. *Genes Dev.* **22**, 14-9.

Bender, W., Akam, M., Karch, F., Beachy, P. A., Peifer, M., Spierer, P., Lewis, E. B. and Hogness, D. S. (1983). Molecular Genetics of the Bithorax Complex in Drosophila melanogaster. *Science* **221**, 23-9.

Benson, G. V., Nguyen, T. H. and Maas, R. L. (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.

Bernstein, B. E., Mikkelsen, T. S., Xie, X., Kamal, M., Huebert, D. J., et al., (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125,** 315-326.

Bomze, H. M. and Lopez, A. J. (1994). Evolutionary conservation of the structure and expression of alternatively spliced Ultrabithorax isoforms from Drosophila. *Genetics* **136**, 965-77.

Brend, T., Gilthorpe, J., Summerbell, D. and Rigby, P. W. (2003). Multiple levels of transcriptional and post-transcriptional regulation are required to define the domain of Hoxb4 expression. *Development* **130,** 2717-2728.

Bridges, C. B. and Morgan, T. H. (1923). The third-chromosome group of mutant characters of Drosophila melanogaster. *Carnegie Inst. Washington Publ.* **327**, 1–251.

Buchenau, P., Hodgson, J., Strutt, H. and Arndt-Jovin, D.J. (1998). The distribution of polycomb-group proteins during cell division and development in *Drosophila* embryos: impact on models for silencing. *J. Cell Biol.* **141**, 469–481.

Busturia, A., Vernos, I., Macias, A., Casanova, J. and Morata, G. (1990). Different forms of Ultrabithorax proteins generated by alternative splicing are functionally equivalent. *EMBO J.* **9**, 3551-5.

Carapuço, M., Nóvoa, A. Bobola, N. and Mallo, M. (2005). *Hox* genes specify vertebral types in the presomitic mesoderm. *Genes Dev.* **19,** 2116-2121.

Carroll, S.B. (1995). Homeotic genes and the evolution of arthropods and chordates. *Nature* **376**, 479-85.

Celniker, S. E., Keelan, D. J. and Lewis, E. B. (1989). The molecular genetics of the bithorax complex of Drosophila: characterization of the products of the Abdominal-B domain. *Genes Dev.* **3**, 1424-36.

Celniker, S. E., Sharma, S., Keelan, D. J. and Lewis, E. B. (1990). The molecular genetics of the bithorax complex of Drosophila: cis-regulation in the Abdominal-B domain. *EMBO J.* **9**, 4277-86.

Chambeyron, S. and Bickmore, W. A. (2004). Chromatin decondensation and nuclear reorganization of the HoxB locus upon induction of transcription. *Genes Dev.* **18,** 1119–1130.

Chambeyron, S., Da Silva, N.R., Lawson, K.A., and Bickmore, W.A. (2005). Nuclear re-organisation of the Hoxb complex during mouse embryonic development. *Development* **132,** 2215–2223.

Chan, K. T., Qi, J. and Sham, M. H. (2010). Multiple coding and non-coding RNAs in the Hoxb3 locus and their spatial expression patterns during mouse embryogenesis. *Biochem. Biophys. Res. Commun.* **398**, 153-9.

Charite, J., de Graaff, W., Vogels, R., Meijlink, F. and Deschamps, J. (1995). Regulation of the Hoxb-8 gene: synergism between multimerized cis-acting elements increases responsiveness to positional information. *Dev. Biol.* **171,** 294-305.

Coulombe, Y., Lemieux, M., Moreau, J., Aubin, J., Joksimovic, M., et al. (2010). Multiple promoters and alternative splicing: Hoxa5 transcriptional complexity in the mouse embryo. *PLoS One* **5**, e10600.

De la Mata, M., Alonso, C. R., Kadener, S., Fededa J. P., Blaustein M., Pelisch, F., Cramer, P., Bentley, D. and Kornblihtt, A. R. (2003) A slow RNA polymerase II affects alternative splicing in vivo, *Mol. Cell* **12,** 525-532.

Delest, A, Sexton, T. and Cavalli, G. (2012). Polycomb: a paradigm for genome organization from one to three dimensions. *Curr. Op. Cell Biol.* **24,** 405-414.

de Navas, L. F., Reed, H., Akam, M., Barrio, R., Alonso, C. R. and 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**, 107-16.

de Napoles, Mermoud, J. E., Wakao, R., Tang, Y. A., Endoh, M., et al. (2004). Polycomb group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation. *Dev. Cell* **7,** 663–676.

Denell, R. E. (1978). Homoeosis in Drosophila. II. a Genetic Analysis of Polycomb. *Genetics* **90**, 277-89.

Deschamps, J. and van Nes, J. (2005). Developmental regulation of the Hox genes during axial morphogenesis in the mouse. *Development* **132,** 2931-2942.

Deschamps, J. and Wijgerde, M. (1993). Two phases in the establishment of Hox expression domains. *Dev. Biol.* **156,** 473-480.

Dinger, M. E., Amaral, P. P., Mercer, T. R., Pang, K. C., Bruce, S. J., et al. (2008). Long noncoding RNAs in mouse embryonic stem cell pluripotency and differentiation. *Genome Res.* **18,** 1433-1445.

Duboule, D. (2007). The rise and fall of Hox gene clusters. *Development* **134**, 2549-60.

Duboule, D. and Dolle, P. (1989). The structural and functional organization of the murine HOX gene family resembles that of Drosophila homeotic genes. *EMBO J.* **8**, 1497-505.

Duboule, D. and Morata, G. (1994). Colinearity and functional hierarchy among genes of the homeotic complexes. *Trends Genet.* **10**, 358-64.

El Tayebi, H. M., Hosny, K. A., Esmat, G., Breuhahn, K. and Abdelaziz, A. I. (2012). miR-615-5p is restrictedly expressed in cirrhotic and cancerous liver tissues and its overexpression alleviates the tumorigenic effects in hepatocellular carcinoma. *FEBS Lett.* **586**, 3309-16.

Eskeland, R., Leeb, M., Grimes, G. R., Kress, C., Boyle, S., et al. (2010). Ring1B compacts chromatin structure and represses gene expression independent of histone ubiquitination. *Mol. Cell* **38,** 452–464.

Ferraiuolo, M. A., Rousseau, M., Miyamoto, C., Shenker, S., Wang, X. Q., et al. (2010). The three-dimensional architecture of Hox cluster silencing. *Nucl. Acid. Res.* **38,** 7472-7484.

Fujimoto, S., Araki, K., Chisaka, O., Araki, M., Takagi, K. and Yamamura, K. (1998). Analysis of the murine Hoxa-9 cDNA: an alternatively spliced transcript encodes a truncated protein lacking the homeodomain. *Gene* **209**, 77-85.

Gavis, E. R. and Hogness, D. S. (1991). Phosphorylation, expression and function of the Ultrabithorax protein family in Drosophila melanogaster. *Development* **112**, 1077-93.

Gehring, W. J., Kloter, U. and Suga, H. (2009). Evolution of the Hox gene complex from an evolutionary ground state. *Curr. Top. Dev. Biol.* **88**, 35-61.

Gérard, M., Zákány, J. and Duboule, D. (1997). Interspecies exchange of a Hoxd enhancer in vivo induces premature transcription and anterior shift of the sacrum. *Dev. Biol.* **190,** 32-40.

Glaser, S., Schaft, J., Lubitz, S., Vintersten, K., van der Hoeven, F., et al. (2006). Multiple epigenetic maintenance factors implicated by the loss of Mll2 in mouse development. *Development* **133,** 1423-1432.

Graham, A., Papalopulu, N. and Krumlauf, R. (1989). The murine and Drosophila homeobox gene complexes have common features of organization and expression. *Cell* **57**, 367-78.

Guenther, M. G., Jenner, R. G., Chevalier, B., Nakamura, T., Croce, C. M., et al., (2005). Global and Hox-specific roles for the MLL1 methyltransferase. *Proc. Natl. Acad. Sci. U S A.* 102, 8603-8608.

Guerreiro, I., Casaca, A., Nunes, A., Monteiro, S., Nóvoa, A., et al., (2012). Regulatory role for a conserved motif adjacent to the homeodomain of Hox10 proteins. *Development* **139,** 2703-2710.

Gummalla, M., Maeda, R. K., Castro Alvarez, J. J., Gyurkovics, H., Singari, S., Edwards, K. A., Karch, F. and Bender, W. (2012). abd-A regulation by the iab-8 noncoding RNA. *PLoS Genet.* **8**, e1002720.

Hafen, E., Levine, M. and Gehring, W. J. (1984). Regulation of Antennapedia transcript distribution by the bithorax complex in Drosophila. *Nature* **307**, 287-9.

Harding, K., Wedeen, C., McGinnis, W. and Levine, M. (1985). Spatially regulated expression of homeotic genes in Drosophila. *Science* **229**, 1236-42.

Hatton, A. R., Subramaniam, V. and Lopez, A. J. (1998). Generation of alternative Ultrabithorax isoforms and stepwise removal of a large intron by resplicing at exon-exon junctions. *Mol. Cell* **2**, 787-96.

He, X., Yan, Y. L., Eberhart, J. K., Herpin, A., Wagner, T. U., Schartl, M. and Postlethwait, J. H. (2011). miR-196 regulates axial patterning and pectoral appendage initiation. *Dev. Biol.* **357**, 463-77.

Herault, Y., Beckers, J., Kondo, T., Fraudeau, N. and Duboule, D. (1998). Genetic analysis of a Hoxd-12 regulatory element reveals global versus local modes of controls in the HoxD complex. *Development* **125,** 1669–1677.

Hong, S., Cho, Y.W., Yu, L.R., Yu, H., Veenstra, T.D., and Ge, K. (2007). Identification of JmjC domain-containing UTX and JMJD3 as histone H3 lysine 27 demethylases. *Proc. Natl. Acad. Sci. USA* **104,** 18439–18444.

Hurley, I., Hale, M. E. and Prince, V. E. (2005). Duplication events and the evolution of segmental identity. *Evol. Dev.* **7,** 556-567.

Iimura, T. and Pourquié, O. (2007). *Hox* genes in time and space during vertebrate body formation. *Develop. Growth Differ.* **49**, 265–275.

Irish, V. F., Martinez-Arias, A. and Akam, M. (1989). Spatial regulation of the Antennapedia and Ultrabithorax homeotic genes during Drosophila early development. *EMBO J.* **8**, 1527-37.

Issaeva, I., Zonis, Y., Rozovskaia, T., Orlovsky, K., Croce, C. M. ,et al. (2007). Knockdown of ALR (MLL2) reveals ALR target genes and leads to alterations in cell adhesion and growth. *Mol. Cell. Biol.* **27,** 1889–1903.

Juan, A. H. and Ruddle, F. H. (2003). Enhancer timing of Hox gene expression: deletion of the endogenous Hoxc8 early enhancer. *Development* **130,** 4823-4834.

Kaufman, T. C., Lewis, R. and Wakimoto, B. (1980). Cytogenetic Analysis of Chromosome 3 in Drosophila melanogaster: The Homoeotic Gene Complex in Polytene Chromosome Interval 84a-B. *Genetics* **94**, 115-33.

Kim, S. Y., Paylor, S. W., Magnuson, T and Schumacher, A. (2006). Juxtaposed Polycomb complexes co-regulate vertebral identity. *Development* **133,** 4957-4968

Kmita, M., and Duboule, D. (2003). Organizing axes in time and space; 25 years of colinear tinkering. *Science* **301,** 331–333.

Kondo, T., and Duboule, D. (1999). Breaking colinearity in the mouse HoxD complex. *Cell* **97,** 407–417.

Kondrashov, N., Pusic, A., Stumpf, C. R., Shimizu, K., Hsieh, A. C., et al. (2011). Ribosome-mediated specificity in Hox mRNA translation and vertebrate tissue patterning. *Cell* **145,** 383-397.

Kornfeld, K., Saint, R. B., Beachy, P. A., Harte, P. J., Peattie, D. A. and Hogness, D. S. (1989). Structure and expression of a family of Ultrabithorax mRNAs generated by alternative splicing and polyadenylation in Drosophila. *Genes Dev.* **3**, 243-58.

Krasnow, M. A., Saffman, E. E., Kornfeld, K. and Hogness, D. S. (1989). Transcriptional activation and repression by Ultrabithorax proteins in cultured Drosophila cells. *Cell* **57**, 1031-43.

Kuziora, M. A. and McGinnis, W. (1988). Different transcripts of the Drosophila Abd-B gene correlate with distinct genetic sub-functions. *EMBO J* **7**, 3233-44.

Lan, F., Bayliss, P. E., Rinn, J. L., Whetstine, J. R., Wang, J. K., et al. (2007). A histone H3 lysine 27 demethylase regulates animal posterior development. *Nature* **449,** 689–694.

Landry, D. M., Hertz, M. I., and Thompson, S. R. (2009). RPS25 is essential for translation initiation by the Dicistroviridae and hepatitis C viral IRESs. *Genes Dev.* **23,** 2753–2764.

LaRosa, G. J. and Gudas, L. J. (1998). 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, A. M., Bermingham, J. R., Jr., Laymon, R. A. and Scott, M. P. (1986). Structure of transcripts from the homeotic Antennapedia gene of Drosophila melanogaster: two promoters control the major protein-coding region. *Mol. Cell. Biol.* **6**, 4676-89.

Lee, J. T. (2012). Epigenetic regulation by long noncoding RNAs. *Science* **338,** 1435-1439.

Lee, D. and Shin, C. (2012). MicroRNA-target interactions: new insights from genome-wide approaches. *Ann. N Y Acad. Sci.* **1271**, 118-28.

Lee, S., Lee, J.W. and Lee, S.K. (2012). UTX, a histone H3-lysine 27 demethylase, acts as a critical switch to activate the cardiac developmental program. *Dev. Cell* **22,** 25-37.

Lemons, D. and McGinnis, W. (2006). Genomic evolution of Hox gene clusters. *Science* **313**, 1918-22.

Lemons, D., Pare, A. and 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**, e31365.

Lewis, E. B. (1978). A gene complex controlling segmentation in Drosophila. *Nature* **276**, 565-70.

Lewis, R. A., Wakimoto, B. T., Denell, R. E. and Kaufman, T. C. (1980). Genetic Analysis of the Antennapedia Gene Complex (Ant-C) and Adjacent Chromosomal Regions of DROSOPHILA MELANOGASTER. II. Polytene Chromosome Segments 84A-84B1, 2. *Genetics* **95**, 383-97.

Licatalosi, D. D. and Darnell, R. B. (2010). RNA processing and its regulation: global insights into biological networks. *Nat. Rev. Genet.* **11**, 75-87.

Lieberman-Aiden, E., van Berkum, N. L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., et al. (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326,** 289-293.

Maeda, R. K. and Karch, F. (2009). The bithorax complex of Drosophila an exceptional Hox cluster. *Curr. Top. Dev, Biol.* **88**, 1-33.

Mallo, M., Wellik, D. M. and Deschamps, J. (2010). *Hox Genes* and Regional Patterning of the Vertebrate Body Plan. *Dev. Biol.* **344,** 7-15.

Maniatis, T. and Reed, R. (2002). An extensive network of coupling among gene expression machines. *Nature* **416,** 499-506.

Mann, R. S. and Hogness, D. S. (1990). Functional dissection of Ultrabithorax proteins in D. melanogaster. *Cell* **60**, 597-610.

Mansfield, J. H. and McGlinn, E. (2012). Evolution, expression, and developmental function of Hox-embedded miRNAs. *Curr. Top. Dev. Biol.* **99,** 31-57.

McGlinn, E., Yekta, S., Mansfield, J. H., Soutschek, J., Bartel, D. P. and Tabin, C. J. (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. U S A* **106**, 18610-5.

Merabet, S., Pradel, J. and Graba, Y. (2005). Getting a molecular grasp on Hox contextual activity. *Trends Genet.* **21**, 477-80.

Mihaly, J., Hogga, I., Gausz, J., Gyurkovics, H. and Karch, F. (1997). In situ dissection of the Fab-7 region of the bithorax complex into a chromatin domain boundary and a Polycomb-response element. *Development* **124,** 1809-1820.

Mlodzik, M., Fjose, A. and Gehring, W. J. (1988). Molecular structure and spatial expression of a homeobox gene from the labial region of the Antennapedia-complex. *EMBO J.* **7**, 2569-78.

Montavon, T., Soshnikova, N., Mascrez, B., Joye, E., Thevenet, L., et al. (2011). A regulatory archipelago controls Hox genes transcription in digits. *Cell* **147,** 1132-1145.

Morata, G. and Kerridge, S. (1982). The role of position in determining homoeotic gene function in Drosophila. *Nature* **300**, 191-2.

Morey, C., Da Silva, N. R., Perry, P. and Bickmore, W. A. (2007). Nuclear reorganisation and chromatin decondensation are conserved, but distinct, mechanisms linked to Hox gene activation. *Development* **134,** 909-919.

Muller, J., and Kassis, J. A. (2006). Polycomb response elements and targeting of Polycomb group proteins in Drosophila. *Curr. Opin. Genet. Dev.* **16,** 476–484.

Negre, B., Casillas, S., Suzanne, M., Sánchez-Herrero, E., Akam, M., et al. (2005). Conservation of regulatory sequences and gene expression patterns in the disintegrating Drosophila Hox gene complex. *Genome Res.* **15,** 692-700.

Nilsen, T. W. and Graveley, B. R. (2010). Expansion of the eukaryotic proteome by alternative splicing. *Nature* **463**, 457-63.

Nishiyama, T., Yamamoto, H., Uchiumi, T., and Nakashima, N. (2007). Eukaryotic ribosomal protein RPS25 interacts with the conserved loop region in a dicistroviral intergenic internal ribosome entry site. *Nucleic Acids Res.* **35,** 1514–1521.

Noordermeer, D., Leleu, M., Splinter, E., Rougemont, J., De Laat, W. and Duboule, D. (2011). The dynamic architecture of Hox gene clusters. *Science* **334,** 222-225.

O'Connor, M. B., Binari, R., Perkins, L. A. and Bender, W. (1988). Alternative RNA products from the Ultrabithorax domain of the bithorax complex. *EMBO J.* **7**, 435-45.

Oh, S. K., Scott, M. P., and Sarnow, P. (1992). Homeotic gene Antennapedia mRNA contains 5'-noncoding sequences that confer translational initiation by internal ribosome binding. *Genes Dev.* **6,** 1643–1653.

Patraquim, P., Warnefors, M. and Alonso, C. R. (2011). Evolution of Hox post-transcriptional regulation by alternative polyadenylation and microRNA modulation within 12 Drosophila genomes. *Mol. Biol. Evol.* **28**, 2453-60.

Pearson, J. C., Lemons, D. and McGinnis, W. (2005). Modulating Hox gene functions during animal body patterning. *Nat. Rev. Genet.* **6**, 893-904.

Petruk, S., Sedkov, Y., Riley, K. M., Hodgson, J., Schweisguth, F., et al. (2006). Transcription of bxd noncoding RNAs promoted by trithorax represses Ubx in cis by transcriptional interference *Cell* **127**, 1209–1221.

Pick, L. and Heffer, A. (2012). Hox gene evolution: multiple mechanisms contributing to evolutionary novelties. *Ann. N Y Acad. Sci.* **1256**, 15-32.

Proudfoot, N. J. (2011). Ending the message: poly(A) signals then and now. *Genes Dev.* **25**, 1770-82.

Puro, J. and Nygrén, T. (1975). Mode of action of a homoeotic gene in Drosophila melanogaster. Localization and dosage effects of Polycomb. *Hereditas* **81**, 237-48.

Puschel, A. W., Balling, R. and Gruss, P. (1991). Separate elements cause lineage restriction and specify boundaries of Hox-1.1 expression. *Development* **112,** 279–287.

Reed, H. C., Hoare, T., Thomsen, S., Weaver, T. A., White, R. A., Akam, M. and Alonso, C. R. (2010). Alternative splicing modulates Ubx protein function in Drosophila melanogaster. *Genetics* **184**, 745-58.

Ren, X., Vincenz, C. and Kerppola, T.K. (2008). Changes in the distributions and dynamics of polycomb repressive complexes during embryonic stem cell differentiation. *Mol. Cell. Biol.* **28,** 2884-2895.

Rinn, J. L., Kertesz, M., Wang, J. K., Squazzo, S. L., Xu, X., et al. (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. and Lai, E. C. (2005). The Drosophila microRNA iab-4 causes a dominant homeotic transformation of halteres to wings. *Genes Dev.* **19**, 2947-52.

Rubin, M. R. and Nguyen-Huu, M. C. (1990). Alternatively spliced Hox-1.7 transcripts encode different protein products. *DNA Seq* **1**, 115-24.

Sanchez-Elsner, T., Gou, D., Kremmer, E. and Sauer. F. (2006). Noncoding RNAs of trithorax response elements recruit Drosophila Ash1 to Ultrabithorax. *Science* **311,** 1118-1123.

Sanchez-Herrero, E. and Crosby, M. A. (1988). The Abdominal-B gene of Drosophila melanogaster: overlapping transcripts exhibit two different spatial distributions. *EMBO J.* **7**, 2163-73.

Sanchez-Herrero, E., Vernos, I., Marco, R. and Morata, G. (1985). Genetic organization of Drosophila bithorax complex. *Nature* **313**, 108-13.

Satoh, T., Takeuchi, O., Vandenbon, A., Yasuda, K., Tanaka, Y. et al. (2010). The Jmjd3-Irf4 axis regulates M2 macrophage polarization and host responses against helminth infection. *Nature Immunol.* **11,** 936-944.

Schneuwly, S., Kuroiwa, A., Baumgartner, P. and Gehring, W. J. (1986). Structural organization and sequence of the homeotic gene Antennapedia of Drosophila melanogaster. *EMBO J.* **5**, 733-9.

Schorderet, P. and Duboule, D. (2011). Structural and Functional Differences in the Long Non-Coding RNA Hotair in Mouse and Human. *PLoS Genet.* **7,** e1002071.

Schuettengruber, B., Chourrout, D., Vervoort, M., Leblanc, B., and Cavalli, G. (2007). Genome regulation by polycomb and trithorax proteins. *Cell* **128,** 735–745.

Schwartz, Y. B., Kahn, T. G., Nix. D. A., Li, X. Y., Bourgon, R. et al. (2006). Genome-wide analysis of Polycomb targets in Drosophila melanogaster. *Nat. Genet.* **38,** 700–705.

Scott, M. P., Weiner, A. J., Hazelrigg, T. I., Polisky, B. A., Pirrotta, V., Scalenghe, F. and Kaufman, T. C. (1983). The molecular organization of the Antennapedia locus of Drosophila. *Cell* **35**, 763-76.

Seo, H. C., Edvardsen, R. B., Maeland, A. D., Bjordal, M., Jensen, M. F., et al., (2004). Hox cluster disintegration with persistent anteroposterior order of expression in Oikopleura dioica. *Nature* **431,** 67–71.

Sham, M. H., Hunt, P., Nonchev, S., Papalopulu, N., Graham, A., Boncinelli, E. and Krumlauf, R. (1992). Analysis of the murine Hox-2.7 gene: conserved alternative transcripts with differential distributions in the nervous system and the potential for shared regulatory regions. *EMBO J.* **11**, 1825-36.

Shen, X., Liu, Y., Hsu, Y.J., Fujiwara, Y., Kim, J., et al. (2008). EZH1 mediates methylation on histone H3 lysine 27 and complements EZH2 in maintaining stem cell identity and executing pluripotency. *Mol. Cell* **32,** 491-502.

Shi, Y, Lan, F., Matson, C., Mulligan, P., Whetstine, J. R., et al. (2004). Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* **119,** 941-953.

Smith, J. J., Kuraku, S., Holt, C., Sauka-Spengler, T., Jiang, N., et al., (2013). Sequencing of the sea lamprey (Petromyzon marinus) genome provides insights into vertebrate evolution. *Nat. Genet.* **45,** 415-421.

Soshnikova, N., and Duboule, D. (2009). Epigenetic Temporal Control of Mouse Hox Genes in Vivo. *Science* **324,** 1320–1323.

Stark, A., Bushati, N., Jan, C. H., Kheradpour, P., Hodges, E., et al. (2008). A single Hox locus in Drosophila produces functional microRNAs from opposite DNA strands. *Genes Dev.* **22**, 8-13.

Stoller, J. Z. Huang, L., Tan, C. C., Huang, F., Zhou, D. D., et al. (2010). Ash2l interacts with Tbx1 and is required during early embryogenesis. *Exp. Biol. Med.* **235,** 569-576.

Stroeher, V. L., Jorgensen, E. M. and Garber, R. L. (1986). Multiple transcripts from the Antennapedia gene of Drosophila melanogaster. *Mol. Cell. Biol.* **6**, 4667-75.

Struhl, G. (1982). Genes controlling segmental specification in the Drosophila thorax. *Proc. Natl. Acad. Sci. U S A* **79**, 7380-4.

Subramaniam, V., Bomze, H. M. and Lopez, A. J. (1994). Functional differences between Ultrabithorax protein isoforms in Drosophila melanogaster: evidence from elimination, substitution and ectopic expression of specific isoforms. *Genetics* **136**, 979-91.

Thomsen, S., Azzam, G., Kaschula, R., Williams, L. S. and Alonso, C. R. (2010). Developmental RNA processing of 3'UTRs in Hox mRNAs as a context-dependent mechanism modulating visibility to microRNAs. *Development* **137**, 2951-60.

Tolhuis, B., Blom, M., Kerkhoven, R. M., Pagie, L., Teunissen, H., et al. (2011). Interactions among Polycomb domains are guided by chromosome architecture. *PLoS Genet.* **7,** e1001343.

Tsai, M. C., Manor, O., Wan, Y., Mosammaparast, N., Wang, J. K., et al. (2010). Long noncoding RNA as modular scaffold of histone modification complexes. *Science* **329,** 689–693.

Tschopp, P. and Duboule, D. (2011). A genetic approach to the transcriptional regulation of Hox gene clusters. *Annu. Rev. Genet.* **45**, 145-66.

Tschopp, P., Tarchini, B., Spitz, F., Zakany, J. and Duboule, D. (2009). Uncoupling time and space in the collinear regulation of Hox genes. *PLoS Genet.* **5,** e1000398.

Tyler, D. M., Okamura, K., Chung, W. J., Hagen, J. W., Berezikov, E., Hannon, G. J. and Lai, E. C. (2008). Functionally distinct regulatory RNAs generated by bidirectional transcription and processing of microRNA loci. *Genes Dev.* **22**, 26-36.

Wang, K.C., Yang, Y. W., Liu, B., Sanyal, A., Corces-Zimmerman, R., Chen, Y., et al. (2011). A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature* **472,** 120–124.

Wedeen, C., Harding, K. and Levine, M. (1986). Spatial regulation of Antennapedia and bithorax gene expression by the Polycomb locus in Drosophila. *Cell* **44**, 739-48.

White, R. A. and Wilcox, M. (1985). Distribution of Ultrabithorax proteins in Drosophila. *EMBO J.* **4**, 2035-43.

Whiting, J., Marshall, H., Cook, M., Krumlauf, R., Rigby, P. W., et al. (1991). Multiple spatially specific enhancers are required to reconstruct the pattern of Hox-2.6 gene expression. *Genes Dev.* **5,** 2048–2059

Woltering, J. M. and Durston, A. J. (2008). MiR-10 represses HoxB1a and HoxB3a in zebrafish. *PLoS One* **3**, e1396.

Woo, C. J., Kharchenko, P. V., Daheron, L., Park, P. J. and Kingston, R. E. (2010). A region of the human HOXD cluster that confers polycomb-group responsiveness. *Cell* **140,** 99-110.

Xiang, Y., Zhu, Z., Han, G., Lin, H., Xu, L. and Chen, C. D. (2007). JMJD3 is a histone H3K27 demethylase. *Cell Res.* **17,** 850-857.

Ye, X., Fong, P., Iizuka, N., Choate, D., and Cavener, D. R. (1997). Ultrabithorax and Antennapedia 5' untranslated regions promote developmentally regulated internal translation initiation. *Mol. Cell. Biol.* **17,** 1714–1721.

Yekta, S., Shih, I. H. and Bartel, D. P. (2004). MicroRNA-directed cleavage of HOXB8 mRNA. *Science* **304**, 594-6.

Yekta, S., Tabin, C. J. and Bartel, D. P. (2008). MicroRNAs in the Hox network: an apparent link to posterior prevalence. *Nat. Rev. Genet.* **9,** 789-796.

Young, T. and Deschamps, J. (2009). Hox, Cdx, and anteroposterior patterning in the mouse

embryo. *Curr. Top. Dev. Biol.* **88,** 235-255.

Yu, B. D., Hess, J. L., Horning, S. E., Brown, G. A. and Korsmeyer, S. J. (1995). Altered Hox expression and segmental identity in Mll-mutant mice. Mll1 mutant. *Nature* **378,** 505-508.

Zakany, J., Gérard, M., Favier, B. and Duboule, D. (1997). Deletion of a HoxD enhancer induces transcriptional heterochrony leading to transposition of the sacrum. *EMBO J.* **16,** 4393-4402.

FIGURE LEGENDS

Figure 1. Organisation of Hox clusters in *Drosophila* and mouse. The *Drosophila* Hox genes (top) are grouped into two genomic clusters, the Antennapedia (ANT-C) and Bithorax clusters (BX-C). Expression domains of the individual Hox genes within the ANT-C and BX-C along the antero-posterior (AP) axis of the fruitfly embryo match the array of the genes along the chromosome, displaying a property termed collinearity. Experiments in the late 1980s revealed that mice also possess a set of Hox genes (bottom) similar to those found in *Drosophila* and that their organisation along the chromosome as well as their order of expression along the AP axis also displayed collinearity. An important difference between these two systems is that the mouse has four clusters (Hoxa, Hoxb, Hoxc and Hoxd) instead of one, as a result of two rounds of gene duplication. The Hox complement of the mouse also reveals that some individual Hox genes have been duplicated while others have been lost in each cluster. Based on sequence and genomic comparisons across a wide range of phyla, the structure of the cluster ancestral to insects and mammals can be inferred (middle). Several miRNAs (miRNAs) are also encoded within the *Drosophila* and mouse Hox clusters, and many of these miRNAs have been shown to target Hox genes. Furthermore, the relative position of many of these miRNA genes in reference to nearby Hox genes appears to be invariant despite substantial periods of evolution suggesting that the position of miRNA genes might be functionally constrained. In the mouse, paralogue groups (PGs) 1-13 are indicated. This diagram has been adapted from Carroll, 1995.

Figure 2. Representation of different aspects of Hox gene regulation associated with chromatin/nuclear characteristics. In the undifferentiated state (left), the Hox clusters are associated with nucleosomes that contain inactivating (H3K27m3) marks and remain in a

compacted configuration within a silent nuclear domain (yellow). In this chromosomal configuration, the *cis*-regulatory elements involved in spatial control of Hox gene expression are kept inactive. Activation of the Hox clusters is progressive, following a 3' to 5' sequence (represented in the middle and right images). This activation is associated with a change in nucleosome modifications, which now include activating (H3K4m3) marks. During activation, the relevant areas of the cluster enter a nuclear territory of active transcription (orange). This relocation of Hox chromatin leads to exposure of *cis*-regulatory elements that can now be regulated by relevant transcription factors leading to Hox gene expression on Noordermeer et al., (2011).

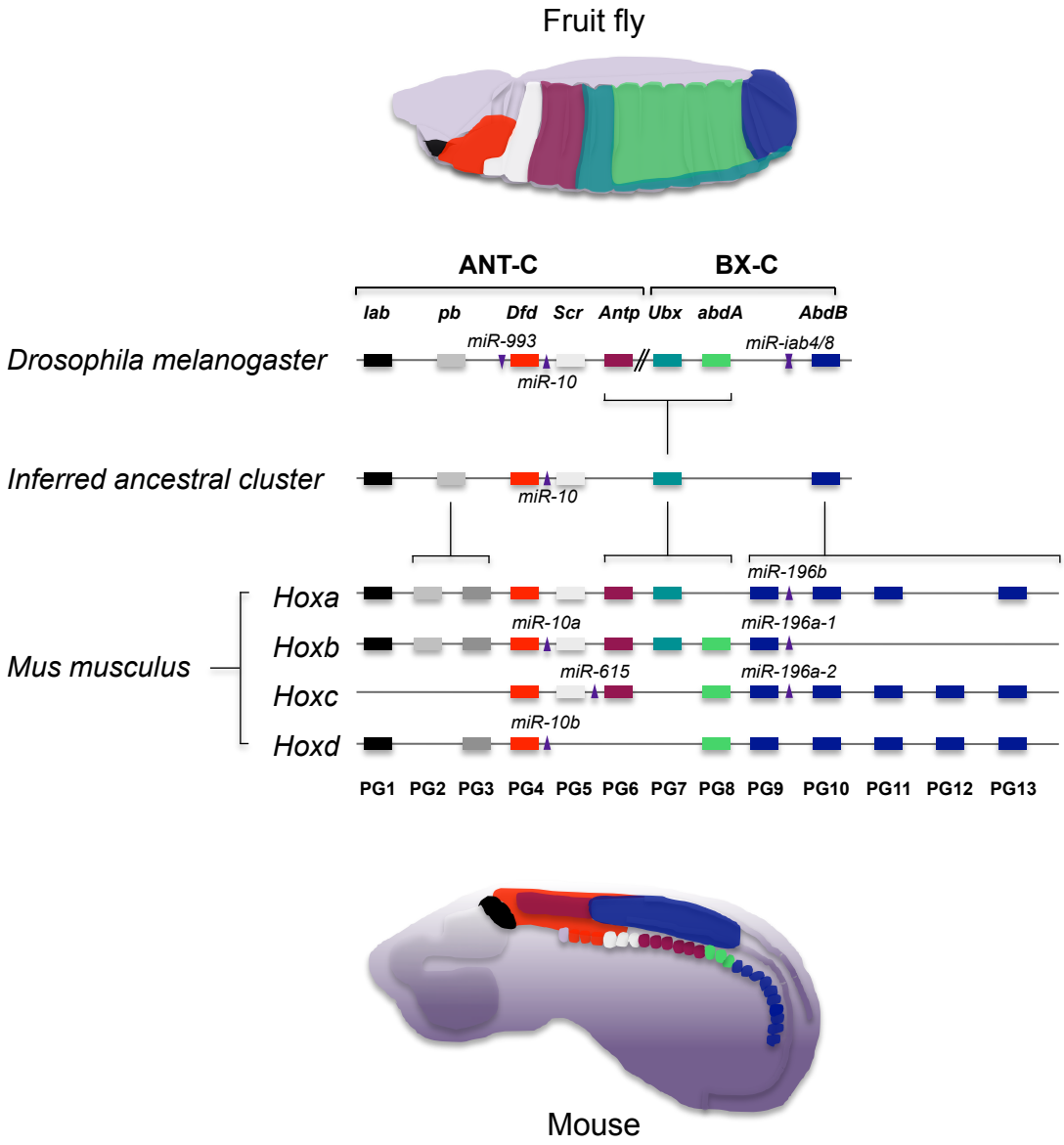
Figure 3. The effects of RNA processing on Hox gene expression. Variations in Hox RNA processing lead to the production of Hox mRNAs of different sizes that bear different sequences and secondary structure. The use of alternative promoters (left) leads to the production of mRNAs with distinct 5'UTRs, which in turn lead to differential translational efficiency. Co-transcriptional alternative splicing reactions (middle) affect the composition of many Hox mRNAs, thus giving rise to different protein isoforms that can perform specific functions during development. The use of alternative polyadenylation signals (right) determines the release of mRNAs bearing different 3'UTR tails that contain distinct sets of miRNA target sequences, thus leading to differential regulation by miRNAs.

Table 1. Summary of the main molecular mechanisms controlling Hox gene expression in Drosophila and mice.

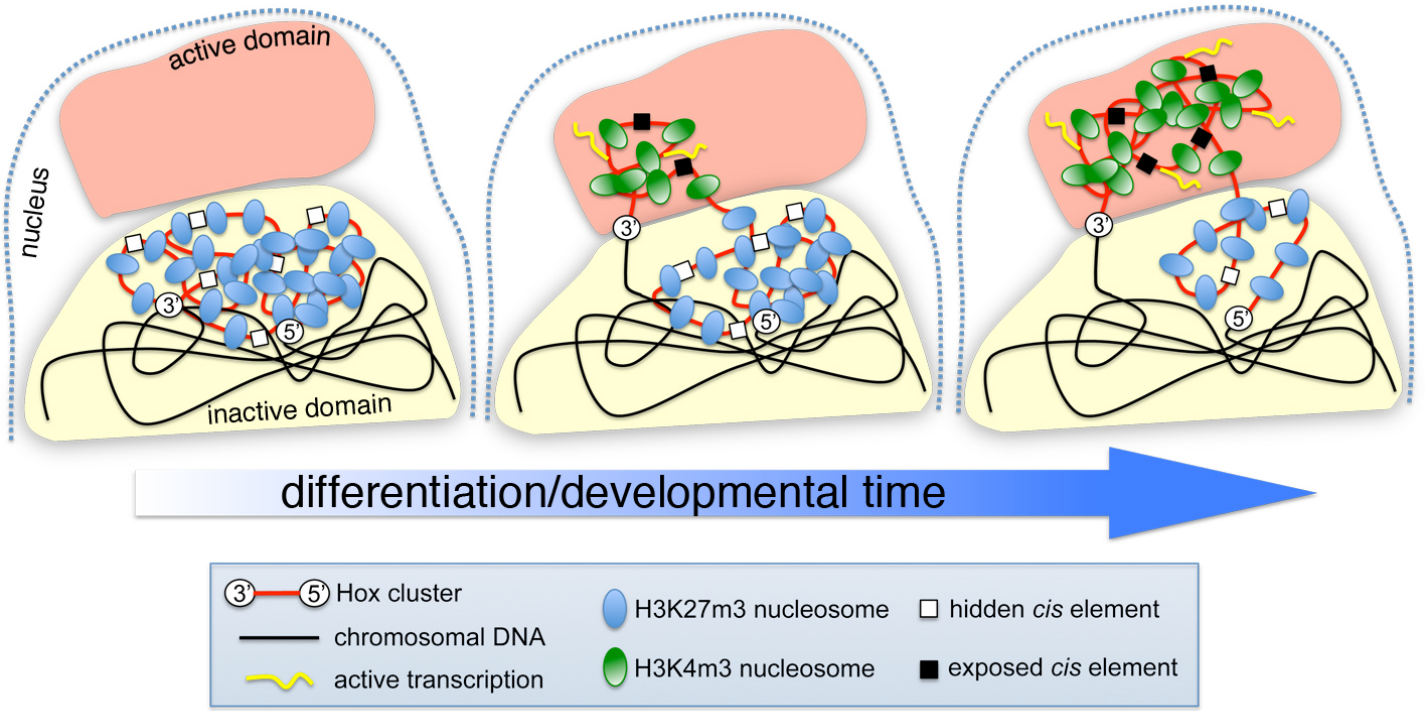
Regulatory Mechanism	*Drosophila melanogaster* (References)#	*Mus musculus* (References)#
Nuclear dynamics	YES (Bantignies et al., 2011; Tolhuis et al., 2011)	YES (Chambeyron and Bickmore, 2004; Ferraiuolo et al., 2010; Noordermeer et al., 2011)
Chromatin remodelling (Early stage)	NO	YES (Soshnikova and Duboule, 2009; Eskeland, et al., 2010)
Chromatin remodelling (Consolidation stage)	YES (Bantignies et al., 2011; Tolhuis et al., 2011)	YES (Noordermeer et al., 2011; Yu et al., 1995; 1998)
Transcriptional regulation	YES (Reviewed in Alexander et al., 2009; Maeda and Karch, 2009; Tschopp and Duboule, 2011)	YES (Reviewed in Alexander et al., 2009; Maeda and Karch, 2009; Tschopp and Duboule, 2011)
lncRNA regulation	YES (Petruk et al., 2006; Sanchez-Elsner et al., 2006)	YES (Rinn et al., 2007; Dinger et al., 2008; Wang et al., 2011)
RNA processing	YES (Kornfeld et al., 1989; O'Connor et al., 1988; Hatton et al., 1998; Reed et al., 2010; Thomsen et al., 2010)	YES (LaRosa and Gudas, 1988; Benson et al., 1995; Fujimoto et al., 1998)
miRNA regulation	YES (Ronshaugen et al., 2005; Tyler et al., 2008; Stark et al., 2008; Bender, 2008; Thomsen et al., 2010)	YES (Yekta et al., 2004; McGlinn et al., 2009; He et al., 2011)
Translational regulation	Likely (Oh et al. 1992; Ye et al., 1997)	YES (Brend et al. 2003; Kondrashov et al., 2011)

Please note that due to space limitations only a few key references are listed here. For further details see main text.

Mallo and Alonso
Figure 1



Mallo and Alonso
Figure 2



Mallo and Alonso
Figure 3