Hox transcription factors: an overview of multi-step regulators of gene expression

Julie Carnesecchi, Pedro Pinto, Ingrid Lohmann

Hox Transcription Factors: an overview of multi-step regulators of gene expression

Julie Carnesecchi[1*], Pedro Borges Pinto[1*] and Ingrid Lohmann[1]

* these authors contributed equally

[1]*Heidelberg University, Centre for Organismal Studies (COS) Heidelberg, Department of Developmental Biology, Heidelberg, Germany*

Key words: Hox, transcription, chromatin, promoter, enhancer

Short running title: Hox multi-step regulators of gene expression

Abbreviations

AbdA: Abdominal-A
AbdB: Abdominal-B
Antp: Antennapedia
Bip2: Bric à brac interacting protein 2
CBP: CREB-binding protein
Cdk9: Cyclin Dependent Kinase 9
CRM: Cis-Regulatory module
CTCF: CCCTC-binding factor
ChIP: Chromatin ImmunoPrecipitation
ChIP-seq: ChIP-sequencing
Dfd: Deformed
Dip1: Disconnected Interacting Protein 1
Dll: distal less
DNAseseqDNAse-seq: DNAse sequencing
Exd: Extradenticle
HDAC: histone deacetylase
Hth: Homothorax
M1BP: Motif 1 binding protein
Pbx: Pre-B cell leukemia transcription factor
Pc: Polycomb
PcG: Polycomb group
Pho: Pleiohomeotic
PIC: Polymerase Initiation Complex
Pol II: RNA-Polymerase II
PRE: Polycomb Response Element
P-TEFb: positive transcription elongation factor b
PTM: Post-Translational Modification
Ser2P: Serine 2 phosphorylated
Ser5P: Serine 5 phosphorylated
Spen: Split ends
Src: Sex combs reduced
SWI/SNF: SWItch/Sucrose Non-Fermentable
TALE: Three Amino-acid Loop Extension
TF: Transcription Factor

TFIIEβ/D: Transcription factor II Eβ / D
TRE: Trithorax Response Element
TrxG: Trithorax group
TSS: Transcription Start Site
Ubx: Ultrabithorax

Complete address of corresponding author:

Ingrid Lohmann
University of Heidelberg, Centre for Organismal Studies (COS) Heidelberg, Department of Developmental Biology, 69120 Heidelberg, Germany
Phone: +49 6221-545523
Fax: +49 6221-546424
Email: Ingrid.lohmann@cos.uni-heidelberg.de

Email of other authors:

julie.carnesecchi@cos.uni-heidelberg.de, pedro.pinto@cos.uni-heidelberg.de

Abstract

The essential class of Hox Transcription Factors (TFs) functions as key determinants in the specification of cell fates during development. They do so by triggering entire morphogenetic cascades through the activation of specific target genes. In contrast to their fundamental role in development, the molecular mechanisms employed by Hox TFs are still poorly understood.

In recent years, a new picture has emerged regarding the function of Hox proteins in gene regulation. Initial studies have primarily focused in understanding how Hox TFs recognized and bind specific enhancers to activate defined Hox targets. However, genome-wide studies on the interactions and dynamics of Hox proteins have revealed a more elaborated function of the Hox factors. It is now known that Hox proteins are involved in several steps of gene expression with potential regulatory functions in the modification of the chromatin landscape and its accessibility, recognition and activation of specific CRMs, assembly and activation of promoter transcription complexes and mRNA processing. In the coming years, the characterization of the molecular activity of the Hox TFs in these mechanisms will greatly contribute to our general understanding of Hox activity.

Introduction: From morphogenesis to molecular function

The Hox transcription factors (TFs) are master regulators of animal body plan (Pearson *et al.*, 2005, Hombria *et al*., 2016). At the cellular level, Hox TFs regulate several biological processes such as cell death, proliferation, differentiation, migration, size and shape (Domsch *et al.*, 2015; Rezsohazy *et al.*, 2015; Sánchez-Herrero, 2013). It is known that Hox TFs are able to regulate these processes through the activation of genetic networks via spatial and temporal activation of specific target genes. These include secondary TFs as well as genes that directly regulate cell behaviour required for tissue morphogenesis, the so-called realizator genes (Sánchez-Herrero, 2013). The function of Hox TFs in the control of gene networks has been extensively covered in previous publications (Hombría and Lovegrove, 2003; Sánchez-Herrero, 2013, Hombria *et al*., 2016). In general terms, the deployment of these genetic networks is usually seen as the result of the recognition and binding of Hox TFs to *cis*-regulatory modules (CRMs) regulating Hox target gene expression. In other words, Hox TFs are placed on top of morphogenetic cascades, with Hox TFs regulating events by activating or repressing downstream targets. Thus, strong emphasis has been placed not only on the identification of Hox targets but also in understanding how Hox TFs perform such diverse and highly specific functions at the molecular level. Considerable effort has been and still is devoted to understand the molecular mechanism by which different Hox proteins recognize and bind specific enhancers, in particular in the context of cofactor interaction such as the well described TALE family of TFs (Mann *et al.*, 2009; Merabet and Hudry, 2011; Merabet and Lohmann, 2015; Slattery, Riley, *et al.*, 2011). However, recently, genome-wide DNA interactions studies revealed that Hox TFs prominently interact not only with enhancers but also with other chromatin regions, including promoters, intragenic and intergenic regions (Beh *et al.*, 2016; Choo *et al.*, 2011;

Shlyueva *et al.*, 2016; Slattery *et al.*, 2011; Sorge *et al.*, 2012; Zouaz *et al.*, 2017). These binding profiles suggest that Hox proteins act at different steps of gene expression in order to promote different functional outputs. In addition, protein interaction screens showed that Hox TFs interact with a variety of transcriptional regulators such as chromatin-, RNA Polymerase II- (Pol II) and mRNA- associated proteins (reviewed in Rezsohazy *et al.*, 2015). This strongly suggests that, in contrast with the classical view that Hox TFs function primarily on enhancers, they regulate a broader range of transcriptional events as well as intervening at different levels of gene expression.

In this review we aim to present a general overview of the most recent data demonstrating the involvement of Hox TFs in key steps of transcriptional regulation. The picture that emerges from these data reveals intricate molecular mechanisms employed by the Hox proteins in the regulation of gene expression, allowing them to activate or repress target genes in a precise and highly specific spatial-temporal manner.

A. Hox & Chromatin Landscape

1. Hox & Histone Marks

Transcription depends on chromatin accessibility thus, the first level of gene regulation is the modification of the chromatin landscape. Heterochromatin or condensed chromatin prevents gene expression by abolishing the recruitment of TFs to their corresponding CRMs as well as by interfering with chromatin loop dynamics thereby affecting CRM-promoter interactions (Huisinga *et al*., 2006). In contrast, euchromatic environments are not necessarily correlated with transcriptional activation (Huisinga *et al*., 2006). Active and repressive states of chromatin are regulated through changes in nucleosome dynamics, in which post-translational modifications (PTM) of histones, the core component of nucleosome, play a major role (Kouzarides, 2007). These modifications such as acetylation, methylation, ubiquitylation and phosphorylation occur mainly on the free and unstructured N-terminal tail of histones. The detailled characterization of these histone “marks” highlighted the existence of a “histone code” that orchestrates gene expression by promoting permissive or repressive chromatin states (Kouzarides, 2007). The histone PTMs are deposited or removed by modifiers and thus, the regulation of these enzymes has been the focus of many studies (Kouzarides, 2007). It has been shown that TFs are able to orchestrate the recruitment or the activity of these modifiers in order to promote their precise transcriptional program (Benveniste *et al.*, 2014; Carnesecchi *et al.*, 2017; Magnani *et al*., 2011).

Because they initiate complete developmental programs, the Hox TFs have long been considered as potential regulators of histone modifier complexes, such as the Trithorax-group (TrxG) and Polycomb-group (PcG) proteins. These complexes promote permissive (via histone acetylation, H3K4 methylation, chromatin remodelling) as well as repressive (mainly via H3K27me3) chromatin environments, respectively. In spite of having opposing functions, TrxG Response Elements (TREs) and PcG Response Elements (PREs) are found frequently co-

localised on the chromatin, which is suggestive of a dynamic transcriptional response. This is illustrated by the dual function of the TrxG TF GAGA (Trl/GAF) and the PcG factor Zeste, both promoting gene repression and activation (Kingston and Tamkun, 2014).

PcG and TrxG have been primarily shown to regulate *Hox* expression during *Drosophila* embryogenesis (reviewed in Schuettengruber *et al.*, 2017). However, a recent study has reported protein-protein interactions between Hox TFs and PcG proteins (Cao *et al.,* 2014). Affinity purification of PcG proteins EED and Ring1B followed by mass spectrometry analysis has identified HOXB13 as a possible interactor of PcG. Although the functional significance of these interactions is unknown and further studies are required to support these findings. Nevertheless, despite the lack of evidence showing direct interaction between PcG and Hox TFs, it is clear that they can act in concert in order to engage downstream transcriptional programs. A recent example is the interplay between JNK, Polycomb (Pc), Abdominal-A (AbdA) and Abdominal-B (AbdB) in the remodelling of segment boundary during dorsal closure in *Drosophila* embryos (Roumengous *et al.*, 2017). Additionally, Garaulet *et al*. provided evidences that the Hox TF Ultrabithorax (Ubx) repressed its own transcription in a PcG-dependent manner (Garaulet *et al.*, 2008). Using different *Ubx-GAL4* lines with or without PRE sites close to the Ubx regulatory region, they identified a permanent repression of endogenous Ubx transcription in the larval haltere disc when Ubx was expressed at high level.

More recently, Shlyueva *et al* established a Ubx DNA-binding profile (ChIP-seq) in *Drosophila* embryos (Shlyueva *et al.*, 2016) and compared it to the genomic interactions of two PcG proteins, Pc and Pleiohomeotic (Pho) (Kwong *et al.*, 2008). They observed a co-occurrence of Pc and Pho at Ubx sites. However, it remained unclear whether Ubx was required for Pc and/or Pho recruitment or if it regulated Pc methyltransferase activity. Importantly, due to the expression of Ubx and Pc/Pho in different cell types, it is impossible to discriminate between a mutually exclusive binding of PcG and Ubx and a co-occurrence of the proteins at the same genomic location when obtaining profiles from whole embryos. To circumvent this problem, Zouaz *et al* adopted a different strategy for deciphering a molecular network between Hox TFs and the PcG complex (Zouaz *et al.*, 2017). Using S2 *Drosophila* cells, they transiently expressed AbdA and performed ChIP-seq analysis of AbdA, dRing, Enhancer of zeste E(z), Pc, three components of the PcG complex, as well of the active H3K4me3 and repressive H3K27me3 chromatin marks. Their analysis revealed that AbdA binding to proximal promoters (near the transcription start site or TSS) induced decrease of dRing binding on a defined set of genes (containing paused Pol II). However H3K27me3, the major mark set by PcG complex, was not detected on these genes. Thus, this study highlighted an important transcriptional interplay between Hox TFs and PcG, but left the question open of how Hox TFs influence H3K27me3 as well as other repressive histone marks.

While there is some evidence that Hox TFs functionally interact with PcG proteins, little is known concerning the functional interplay between TrxG and Hox TFs. Similarly to PcG proteins, *in vivo* co-occurrence of TrxG and Hox TFs binding strongly suggests a molecular

interaction between them. For example, ChIP on ChIP experiments in the *Drosophila* haltere disc and ChIP-seq experiment in S2 *Drosophila* cells identified an enrichment of GAGA motif in the vicinity of Ubx (and AbdA) binding sites (Agrawal *et al.*, 2011; Zouaz *et al.*, 2017). However, so far, a direct interaction between Hox TFs and TrxG has not been shown.

The possible interaction between Hox TFs and chromatin modifiers is not restricted to *Drosophila*, as similar observations have been described in vertebrates (Hassan *et al.*, 2007). Hassan *et al* reported the binding and regulation of the osteoblast-differentiation master gene Runx2 by HoxA10 (AbdB vertebrate homolog). Knock-down of HoxA10 induced hypo-acetylation of histone H4 and a decrease of H3K4me3 at the Runx2 promoter. This is in line with the observation reported by Zouaz *et al*, showing that AbdA binding on promoters correlated with H3K4me3 deposition. Although the methyltransferase activity has not been identified, Hassan *et al* (2007) identified the p300/CREB-binding protein (CBP) as the acetyltransferase acting at HoxA10 target genes. The histone (and non-histone) acetyltransferase CBP is a major transcriptional co-regulator, conserved in mammals (p300/CBP) and *Drosophila* (Nejire). Several genetic and physical interactions between CBP and Hox TFs have been described in *Drosophila* (Florence and McGinnis, 1998; Sorge *et al.*, 2012), as well as in mammal systems (Chariot *et al.*, 1999; Shen *et al.*, 2001). For example, Bei *et al* showed that CBP physically interacted with HoxA10, thereby promoting integrin beta 3 expression via H2A acetylation (Bei *et al.*, 2007). These data is line with the observations of Huang *et al.,* 2012 where it was shown that binding of HoxA9 at enhancers correlates with a decrease in P300 and CBP and a decrease in H3 and H4 acetylation. Additionally, Hox TFs may also act as transcriptional repressors by regulating histone acetylation (Huang *et al*, 2012). Moreover, Gordon *et al* proposed a model where the TALE TF Pbx1 was ejected from osteoblast-proximal gene promoters allowing the recruitments of Runx2 and p300/CBP by HoxA10 upon induction of osteoblast differentiation (Gordon *et al.*, 2010). HoxA10 was also shown to act as a transcriptional repressor via histone deacetylation by HDAC2 (Lu *et al.*, 2003). A similar interplay between Hox/TALE, p300/CBP and HDAC has also been described in other mammalian systems (Saleh *et al.*, 2000) and in zebrafish (Choe *et al.*, 2009). In sum, all these data strongly suggest the existence of an interplay between Hox factors, CBP and HDAC that regulates gene expression via the dynamic deposition or removal of histone acetylation.

Strikingly, Petruk *et al* identified an interaction between TrxG and CBP in *Drosophila* (Petruk, 2001). Subsequent work by the laboratory of Peter Harte dissected the cross-talk between TrxG, PcG and CBP, which showed that CBP and the TrxG methyltransferase Trx mediated the deposition of H3K27ac marks and that CBP and Pc interacted in *Drosophila* embryos and mammalian system (Tie *et al.*, 2009; Tie *et al.*, 2014; Tie *et al.*, 2016). In sum, evidences strongly suggest that the assembly of multi-protein complexes between Hox proteins and chromatin modifiers is one of the mechanisms employed by Hox TFs to regulate gene expression via the deposition of histone marks. They may induce a repressive environment via

histone de-acetylation and H3K27me3 by HDAC and PcG respectively, or a permissive state via H3K27ac and H3K4me3 deposition by CBP and TrxG interplay.

2. *Hox & Chromatin Conformation*

Activation of transcription requires CRMs to be accessible for TFs, which direct and coordinate transcriptional programs. Chromatin has to be decondensed and DNA made available for protein recognition and interaction. In this context, nucleosome positioning plays a key role in the regulation of gene expression (Workman, 2006). Nucleosomes are displaced by combinatorial activity of histone modifiers (acetylation), histone chaperones, histone variants (H2Av) and the ATP-dependent nucleosome remodelling machinery. The latter are divided in four families, SWI/SNF, INO80/SWR1, ISWI and CHD, which use ATP to promote nucleosome remodelling via an ATPase domain-containing helicase. The conserved SWI/SNF member, Brahma remodelling complex, belongs to the TrxG family. It is recruited by TFs thereby regulating precise programs such as muscle development, leg and wing morphogenesis in *Drosophila* by the assembly of ubiquitous and cell-type specific subunits (reviewed in Meier and Brehm, 2014). Study of TF-remodeller activity has been greatly improved by the development of next-generation sequencing detection of nucleosome occupancy such as FAIRE-seq and ATAC-seq. Subsequent comparison of nucleosome positioning and TF binding profiles led to a better understanding of the molecular function of TFs. For example, Beh *et al* performed a genome wide chromatin interaction study for the Hox TFs Ubx, AbdA and AbdB in *Drosophila* Kc67 cells (Beh *et al.*, 2016). The binding profiles of these TFs were shown to be very similar. However, some discrepancies in their binding were observed. For example 25% of the AbdB binding regions did not overlap with DNase-print accessible regions. Thus the authors suggested that AbdB was able to interact with closed chromatin and to induce DNA accessibility through chromatin remodelling. Similarly, they showed that 17% of closed chromatin regions were bound by Ubx, however, only when it was co-expressed with its cofactor Homothorax (Hth), a TALE/Meis class TF. This suggested that Ubx might extend its repertoire of target genes via chromatin remodelling and cooperation with Exd (Extradenticle)/Hth. These results are supported by the identification of proteins associated with chromatin remodelling in several Hox-protein interaction screens (reviewed in Rezsohazy *et al.*, 2015). Beh *et al* also proposed that differences of sequence and number of AbdB binding sites could affect nucleosome stability via the reduction of DNA-histone affinity. This assumption is strengthened by several observations. First, the core of Hox binding site is an “AT” rich sequence. Second, AT sites can destabilize histone-DNA affinity and have also been described as recognition regions for ATP-dependent remodelling complexes (Lorch *et al.*, 2014). Thus, it seems plausible that a synergistic effect between “passive mechanism” on DNA-histone affinity and “active mechanism” of recruitment of nucleosome remodellers by Hox TFs facilitates chromatin remodelling and subsequent gene expression.

It has been suggested that Hox TFs could affect chromatin conformation via deposition of histone variants (Agelopoulos *et al.*, 2012). Agelopoulos *et al* established an elegant tool called cgChIP (gene-specific ChIP, lacI/lacO) to monitor chromatin loops on defined target genes in specific cell types. They showed that Ubx and AbdA repress *Distal-less* (*Dll*) transcription by impairing the enhancer-promoter communication. Based on these results, they proposed two different mechanisms by which Hox TFs might exert their repressive function. First, they could inhibit the recruitment of the TrxG member GAGA, a major facilitator of chromatin "loops" that allows the communication between enhancer and promoter thereby inhibiting transcriptional activation. Secondly, binding of Hox TFs to the *Dll* regulatory region was associated with an enrichment of the histone 2A variant (H2Av) that could impair chromatin-chromatin interaction.

The chromatin is not only regulated at the gene level but also at the level of higher-order structures. From chromatin loop between enhancer and promoter to chromosome territories in the nucleus, TFs act on a variety of levels to orchestrate their highly precise yet diverse transcriptional outputs (de Graaf and van Steensel, 2013). In vertebrates, the interplay between the TF CTCF and HoxD9-13 underlined the role of Hox TFs in controlling chromatin structure (Jerković *et al.*, 2017). The insulator CTCF has been described as a repressor of gene expression by blocking enhancer-promoter interactions (reviewed in Roy *et al.*, 2018). It also regulates boundary regions between active and heterochromatin chromosomal domains. Binding profiles for 9 Hox TFs revealed two Hox subgroups, in which one group was enriched in CTCF-binding motif (and interact with CTCF) but poorly enriched in Hox motif, suggestive of an indirect recruitment of Hox proteins by CTCF. Thus, Hox TFs could regulate enhancer-promoter communication via CTCF binding and act on higher-order structure by the regulation of chromosomal interaction (see Beccari *et al.*, 2016 for Hox and topologically associating domain in wrist development).

In summary, there are strong evidences pointing to Hox TFs as regulators of the chromatin landscape. Through defined protein-interactions, they seem to affect the conformation and the state of the chromatin in order to promote their cell-type specific transcriptional outputs.

3. Hox & Pioneer Function

Pioneer TFs, which are thought to regulate the first step of transcriptional initiation, have been described as major regulators of developmental processes, stem cell differentiation and hormonal regulation and their molecular mode of action has been extensively studied both *in vitro* and *in vivo* (reviewed in Zaret and Carroll, 2011). Consistent with the idea of a TF involved in the initial steps of gene expression, a pioneer TF has to bind to nucleosomal DNA and provide access for secondary factors to orchestrate their transcriptional program. They do so by acting as platform for TFs, histone modifiers and/or chromatin remodellers. The most prominent example of a pioneer TF is FoxA1, which is able to specifically promote or repress transcription

but binds nucleosomal DNA in a non-specific manner (Sekiya *et al.*, 2009; Sekiya and Zaret, 2007; Zaret and Carroll, 2011). In contrast to FoxA1, the *Drosophila* pioneer TF Zelda (Zld), a key TF for gene expression during early embryogenesis, recognizes and binds to a specific DNA-recognition motif (Foo *et al.*, 2014; Liang *et al.*, 2008; Sun *et al.*, 2015). Moreover, the well-described Hox-cofactors (Pbx/Meis) of the TALE family are also able to prime chromatin for the recruitment and subsequent transcriptional activity of secondary TFs (Berkes *et al.*, 2004 for mouse system, Maves *et al.*, 2007; Choe *et al.*, 2014 in zebrafish, Magnani *et al.*, 2011 for human system).

Hox TFs are candidates for pioneer function, as they are major regulators of development and potentially control DNA accessibility through chromatin regulation. Work by Beh *et al* (2016), described earlier, provides some of the evidences that support this assumption. It showed that AbdB interacted with closed chromatin *in vivo* on a specific set of target genes. Therefore, this data strongly suggests that AbdB may have pioneer functions and therefore, regulate chromatin accessibility. Similar genomic co-occurrence of closed chromatin regions and Hox binding events have been detected by comparing ChIP-seq experiments of Ubx and AbdA (in S2 *Drosophila* cells) to DNAse-seq data (Zouaz *et al.*, 2017). However, this observation is in contrast with the data of Beh *et al*, which showed that Ubx was able to access closed chromatin only in the presence of Hth (Exd/Hth). Thus, the potential pioneer role of Hox TFs still remains unclear.

In summary, although the function of Hox TFs as pioneer factors remains a matter of debate, it seems clear that Hox proteins play a role on the regulation of the chromatin environment, influencing histone mark deposition via recruitment of modifiers, nucleosome composition via the recruitment of histone variant, as well as on chromatin conformation such as nucleosome occupancy and chromatin looping between enhancers and promoters (figure 1).

B. Hox, from Promoter to mRNA

1. Hox & Polymerase Initiation Complex

Since *cis*-regulatory regions are determinants of the spatial and temporal regulation of gene expression, much emphasis has been placed to the study of CRMs directly regulated by Hox TFs. However, genome-wide analysis of the distribution of Hox binding sites led to the observation that despite being bound to CRMs, a considerable amount of binding sites is located at the proximal promoters of target genes (Zouaz *et al.*, 2017). Indeed, interactions between Hox TFs and components of the Polymerase Initiation Complex (PIC) have been known for some time (Johnson and Krasnow, 1990; Johnson and Krasnow, 1992.; Mortin *et al.*, 1992; Mortin and Lefevre, 1981). Johnson and Krasnow showed that *in vitro*, Ubx was able to activate transcription, most likely by stimulating the assembly of the PIC, either through the interaction with general transcription factors or with Pol II. This interaction with components of the PIC has also been reported for other Hox TFs. The *Drosophila* Hox proteins Antennapedia (Antp), Ubx and AbdB were shown to interact with the basal transcription factor TFIIEβ via their

DNA binding domain, the homeodomain, (for Antp and AbdB). This interaction, at least for AbdB, was required for transcriptional activity (Bondos *et al.*, 2006; Zhu and Kuziora, 1996). More recently, Prince *et al* performed yeast two hybrid analysis and reported an interaction between Antp and Bric à brac interacting protein 2 (BIP2). BIP2 is a TATA-binding protein associated factor (dTAFII3/II155), member of the TBP-associated TFIID complex of the PIC. This interaction was validated both *in vitro* and *in vivo* and additional experiments showed that BIP2 is a cofactor of Antp in promoting an eye-wing transformation (Prince *et al.*, 2008). Interestingly, this interaction depends on the well-conserved hexapeptide domain (HX, or YPWM) but is independent of the TALE TF Exd.

A major component of the PIC is the Mediator complex, a highly flexible multi-subunit complex important for establishing enhancer-promoter communication during gene expression. Boube *et al* identified a synergistic interaction between Sex comb reduced (Scr) and the *Drosophila* homologs of mediator-subunits Pap/dTRAP240 and dTRAP80 (Boube *et al*, 2000). Subsequently, they showed that the Mediator subunit Med19 genetically interacted with Ubx, Deformed (Dfd), AbdB and Antp (Boube *et al.*, 2014). Moreover, these interactions were shown to occur via the homeodomain. They observed that functional interaction between Ubx and Med19 is required for Ubx gene activation (but not repression) suggesting a target specific function of the Mediator-Hox complex.

2. Hox & Poised Promoter

Despite their association with several PIC components, the function of Hox TFs at the proximal promoters is still unclear. However, there is an increased number of studies linking Hox TFs with the release of Pol II from a poised/inactive to an active transcriptional state (Choe *et al.*, 2014; Chopra *et al.*, 2009; Zouaz *et al.*, 2017). Initiation of transcription is not an efficient process, as Pol II initiates transcription with the synthesis of several small RNA transcripts (less than 10 nt). Pol II clears the promoter only after synthesizing longer transcripts and subsequently transcribes the gene body. However, for a defined subset of inactive genes, Pol II arrests shortly after clearing the promoter. It is thought that this arrest not only marks these genes for activation but also promotes a rapid and synchronous transcriptional response upon activation (Boettiger and Levine, 2009). The promoter clearance of Pol II and the transition to transcription elongation are largely controlled through the phosphorylation of the Carboxy-Terminal domain (CTD) of its largest subunit. This domain contains tandem heptapeptide repeats (52 in mammals, 44 in *Drosophila* and 26 in yeast) with the consensus sequence YSPTSPS. The phosphorylation of serine residues 2 (Ser2) and 5 (Ser5) of these repeats by Cyclin dependent kinase 9 (Cdk9), a kinase subunit of P-TEFb, is a major determinant in promoter clearance and transition into elongation. Hence, Pol II is enriched in phosphorylation at Ser5 (Ser5P) when poised and Ser2 (Ser2P) during active transcriptional elongation. It has been shown that Cdk9 is necessary and sufficient to phosphorylate Ser2 and release Pol II from this paused, or so called poised state (reviewed in Price, 2008). Interestingly, the zebrafish

HoxB1b protein was shown to promote a phosphorylation switch of Pol II, from paused to the active elongating state by stimulating activity of P-TEFb (Choe *et al.*, 2014).

A recent study suggested that Hox TFs might have a role in the release of paused Pol II (Zouaz *et al* 2017). Using ChIP-seq experiment in S2 *Drosophila* cells, Zouaz *et al* identified an enrichment of AbdA in proximal promoters, in the vicinity of the transcriptional start site (TSS). These regions were enriched in binding of the GAGA/Trl/GAF TF (see previous paragraph) and the transcription factor M1BP motif (Motif 1 binding protein). ChIP experiments highlighted that expression and binding of AbdA at M1BP bound promoters induced a decrease of poised Pol II, and an increase in the elongating form Ser2P-Pol II. Moreover, a correlation between these events and the decrease in PcG binding was observed , as seen by the decrease of binding of dRing as well as a small increase in H3K4me3.

A characteristic of poised genes is their bivalent state, identified by the presence of both active and repressive chromatin marks at promoters and/or enhancers (Gaertner *et al.*, 2012; Koenecke *et al.*, 2017). Thus, gene poising is an active process requiring both the regulation of Pol II and the chromatin landscape via histone marks and nucleosome remodelling (Boija *et al.*, 2017; Gaertner *et al.*, 2012; Levine, 2011). As previously stated, Antp interacts with Bip2, a component of the PIC. The observation that the human homologue of Bip2 interacts with a histone acetyltransferase complex raises the possibility that Hox proteins play a role in the regulation of chromatin marks deposited at poised promoters, a scenario further reinforced by the known interaction of Hox TFs and Nejire, the *Drosophila* homologue of the acetyltransferase CBP complex (Florence and McGinnis, 1998; Sorge *et al.*, 2012; Chariot *et al.*, 1999; Shen *et al.*, 2001).

In sum, poised Pol II is a fundamental checkpoint of transcription, during development. Throughout *Drosophila* embryogenesis, Pol II is poised and deposited *de novo* for the promotion of highly precise temporal regulation of gene expression (Gaertner *et al.*, 2012; Levine, 2011). In this context, Hox TFs appear as potential molecular regulators of poising via either regulation of the Pol II state or/and orchestrating the chromatin landscape (figure 1).

3. Hox & RNA Processing

mRNA processing occurs co-transcriptionally (Saldi *et al.*, 2016). In particular, transcription is concurrent with 5'capping, splicing and 3'end formation in order to generate mature mRNAs (Proudfoot *et al.*, 2002). Thus, along the gene body, elongation is not constant but subjected to variation of speed rates that favour exon skipping (positively or negatively) or alternative polyadenylation (Saldi *et al.*, 2016, Dujardin *et al*., 2014, Pinto *et al.*, 2011). It has been shown that mRNA processing depends on the recruitment of specific RNA-regulatory proteins and on the regulation of the chromatin landscape (Brown *et al.*, 2012). Indeed, exons and introns differ in nucleosome occupancy and in histone mark deposition (Saldi *et al.*, 2016). In addition, ATP-dependent chromatin remodelling complexes have been shown to regulate splicing through interaction with RNA-binding protein factors or nucleosome remodellers

(Waldholm *et al.*, 2011). ChIP-seq experiments performed on Hox TFs revealed that binding not only occurs on promoters and intergenic regions but also throughout the gene bodies (Beh *et al.*, 2016; Shlyueva *et al.*, 2016; Slattery *et al.*, 2011; Zouaz *et al.*, 2017). Thus, it is possible that Hox proteins affect mRNA processing, either via the regulation of the chromatin landscape or the recruitment of RNA-binding proteins. In fact, proteomic approaches have highlighted interactions between several RNA-binding proteins and Hox TFs, both in *Drosophila* and mammals (Bondos *et al.*, 2006; Shestakova *et al.*, 2017). For instance, it has been shown by yeast two-hybrid that Disconnected Interacting Protein 1 (DIP1) interacted with Ubx and that this interaction was dependent on the RNA-binding domain of DIP1 (Bondos *et al.*, 2004). DIP1 is known to synergize the Ubx mutant phenotype in flies and to impair its transcriptional activity *in vitro*.

Another example of the cross-talk between Hox and RNA processing is the RNA binding protein Split Ends (Spen). It modulates the function of the Hox TF Dfd (Gellon *et al.*, 1998) and cooperate with Antp to repress head-like structures in the thorax (Wiellette *et al.*, 1999), however, without evidence of direct interaction. Spen regulates RNA processing both at the level of alternative splicing (in human, Hiriart *et al.*, 2005) and polyadenylation as well as mRNA export (in plant, Hornyik *et al.*, 2010). In addition, Spen has been shown to have a homeotic function in *Drosophila*. Interestingly, it cooperates with Hox factors but does not affect the expression of the *Hox* genes themselves, suggesting instead a role of the Hox-Spen complex in mRNA processing.

In sum, these data suggest that Hox function is not restricted to the control of transcriptional initiation but that they may act on subsequent steps of transcription such as mRNA processing and mRNA export thereby providing new level of gene regulation by Hox factors (figure 1).

C. Hox & Enhancer Specificity

The specificity of Hox proteins is an issue that has long been studied. How the different Hox TFs recognize specific CRMs and activate gene expression in a spatial and temporal manner is still a poorly understood question. Considerable efforts have been put on the identification of cofactors such as TALE TFs, and their contribution in the specific activation of Hox targets (Ladam and Sagerström, 2014; Mann *et al.*, 2009; Merabet and Lohmann, 2015; Merabet and Mann, 2016). However, recent findings indicate that, in addition to Hox TFs and their cofactors, the nature of the Hox binding sites as well as the structure of the CRMs, play a major role in determining Hox specificity. In this section we will provide a general overview of some of the latest findings of enhancer regulation and its contribution to our understanding of the molecular mechanisms driving Hox specificity.

Initial work from Richard Mann's laboratory has shown that the *Drosophila* TALE TFs Exd and Hth can form complexes with different Hox TFs (reviewed in Mann and Chan, 1996). More recently, they performed SELEX experiments, based on which they postulated binding

sites displaying variable affinities for different Hox/Exd/Hth complexes (Slattery, Riley, *et al.*, 2011). These differential affinities of Hox binding sites provide a simple explanation of how Hox targets are activated by specific Hox TFs along the anterior to posterior axis of the *Drosophila* embryo. These findings on binding site affinity by the Mann lab have been further supported by the work of Crocker *et al* (Crocker *et al.*, 2015). Using two Ubx regulated *shavenbaby* enhancers, E3N and 7H, they showed that these regulatory modules are controlled by a cluster of low-affinity Ubx/Exd/Hth binding sites that deviate from consensus sequence identified in the previous SELEX experiments (Crocker *et al.*, 2015). Optimizing the binding of Ubx/Exd/Hth by converting the sites into high-affinity sites led to a loss of specificity with ectopic activation of these enhancers by more anterior Hox TFs (Crocker *et al*., 2015). These studies showed that, in addition to *trans*-acting factors, the quality of the Hox binding sites is also a critical parameter in the specific activation of gene expression by Hox TFs. Moreover, it changes the initial view of the co-selective cooperative model that tried to address Hox specificity (reviewed in Biggin and Mcginnis, 1997). The identification of low-affinity binding sites shows that the Hox/Exd/Hth complexes seem to be able to recognize a higher diversity of binding sites and, consequently, a higher number than initially predicted. It may be that low-affinity binding sites are a consequence of enhancer evolution to accommodate equilibrium between robust activation of gene expression and cell-specificity. In this context, it seemed that a higher diversity of Hox binding sites, and the resulting sub-optimization with which Hox/Exd/Hth complexes bound to these sites, is required for cell-specific gene expression.

Supporting this view, recent data has shown that in addition to the nature of the Hox binding sites, the organization of these sites may play an important role in conferring specificity to Hox TFs (Farley *et al.*, 2015; Farley *et al.*, 2016). It is known for some time that the organization of TF binding sites is important for tissue specific expression (Swanson *et al.*, 2010). More recently, work from Mike Levine's laboratory showed that, in addition to spacing and orientation of TFs binding sites, their affinity is important to achieve not only the appropriate levels of gene expression but also tissue specificity (Farley *et al.*, 2015; Farley *et al.*, 2016). Through the study of synthetic enhancers, the authors highlighted the existence of a trade-off between the presence of optimal TFs binding sites and optimal distances between the sites. In enhancers containing optimal binding sites and spacing, higher levels of gene expression were observed. However, in these conditions, tissue specific expression was lost. Subsequently, specificity was recovered whenever the affinity of the binding sites or their spacing was suboptimal. Finally, this study also showed that the orientation of binding sites had an important effect on gene expression (Farley *et al.*, 2016). Taken together, these results suggest that the "sub-optimization" of the enhancer grammar can play an important role in Hox specificity.

Several studies suggest that the sub-optimization of *cis*-regulatory elements could play an important role in how TFs are organized in the nucleus and how this organization may provide a framework to modulate and regulate gene expression (Hnisz *et al.*, 2013; Izeddin *et al.*, 2014; Liu *et al.*, 2014). Work from Tsai *et al* revealed the existence of several nuclear

domains with high concentrations of Ubx proteins. These domains were shown to be transcriptionally active, as they co-localized with active Pol II and H3K4me3 chromatin marks and were excluded from repressive H3K27me3 marks (Tsai *et al.*, 2017). Furthermore, a correlation between low/high affinity binding sites and these Ubx enriched domains was detected in the study of E3N and E7 Ubx-regulated *shavenbaby* enhancers (Crocker *et al.*, 2015). Increasing the affinity of the low affinity binding sites resulted in reduced Ubx binding. Although the function of these Ubx enriched domains is still unclear, the authors suggested that high local concentrations of TFs could serve to overcome the weak and unstable interactions between TFs and these low-affinity sites. By increasing TF concentrations, TFs-DNA interactions occur at higher frequency and thus favour the establishment of cooperative interactions, resulting in higher transcriptional activation. The mechanism by which TF enriched- and transcriptionally active- domains are assembled is not clear. However, the observation that low-affinity enhancers are found within the same Ubx enriched domains when located in different chromosomal positions strongly supports the hypothesis that low-affinity TF binding sites orchestrate the concentration and stabilization of DNA-TF interactions (Izeddin *et al.*, 2014; Liu *et al.*, 2014; Tsai *et al.*, 2017). These observations are in line with the Phase-Separation Model proposed by Hnisz *et al* to explain some of the reported features of transcriptional control. According to this model, concentrating TFs into specific domains provides the condition for the establishment of cooperative interactions and the appropriate transcriptional complexes that promote gene transcription (Hnisz *et al.*, 2017).

In summary, the quality of the Hox binding sites has emerged in recent years as an important aspect of Hox specificity. In addition to the establishment of specific Hox protein complexes, the affinity of the Hox binding sites as well as how they are organized in the context of a CRM, seems to be crucial for Hox function and specificity.

Conclusion & Outlook

.

A point that will certainly be the focus of research is the role of the Hox TFs as regulators of the chromatin landscape. Although several observations suggest that the Hox proteins act as regulators of the chromatin landscape, there is still no clear evidence that Hox proteins may modulate the deposition of chromatin marks and the opening/closing of chromatin. Another important aspect is the nature of the Hox binding sites and its influence on the Hox specificity. By binding with different affinities to diverse binding sites, Hox complexes can activate specific targets in a highly context-dependent manner. However, some points still remain unclear regarding Hox specificity. Despite the differences in affinities, there are different Hox/Exd/Hth complexes that seem to recognize the same binding sites with similar affinity. Additionally, there are CRMs containing different binding sites that are regulated by the same Hox TF. It is possible that the organization of binding sites (Hox and cofactors) may provide the additional layer of specificity in these cases. Indeed, sub-optimization of the Hox binding sites

might have evolved to favour activation by the Hox TFs while preventing competition with other TFs (Noyes *et al.*, 2008). This might be especially true for the Homeodomain TFs family. This family comprises the second largest TFs family in *Drosophila* as well as in Humans (Tupler *et al*., 2001) and as expected, the members of this family were shown to recognize similar binding sites (Noyes *et al.*, 2008). It will be interesting to determine whether the existence of low-affinity binding sites in addition to provide a way to achieve a trade-off between the required transcriptional levels, tissue and/or cell specificity, might also provide a mechanism to prevent transcriptional interference by co-expressed Homeodomain TFs. This is even more relevant when taking into account the differential expression levels of the several Homeodomain TFs.

Hence, the link between the nature of the Hox binding sites, the syntax of the CRMs and the formation of different Hox complexes is crucial to understand the mode of action of Hox TFs. Likewise, the contribution of the CRMs to the nuclear organization of transcriptional complexes and how this impact on gene regulation are issues that need to be addressed in future studies. In this context, the effect of the Hox TFs on the PIC is still unclear. The observed interactions with basal transcription factors as well the stimulation of P-TEFb and subsequent activation of Pol II indicates a regulatory role of the Hox proteins in the activation of transcription. The presence of Hox TFs in paused/poised promoters further supports this idea. It will be interesting in future to determine if and how Hox proteins contribute to the release of paused Pol-II.

In sum, the diverse regulatory function of Hox TFs described in this review raises interesting challenges. Coordination of gene expression is undoubtedly a premising one. How do Hox proteins coordinate the different steps of gene expression to regulate their target genes? Does the involvement of Hox TFs at the several layers of gene expression constitute a general mode of action by which Hox proteins regulate their target genes? Or do these events take place in a gene specific manner? In other words, does the regulation of individual genes through the use of different regulatory mechanisms constitute an important aspect in Hox regulation? Moreover, does this provide the framework by which Hox proteins are able to establish and coordinate genetic networks?

Additionally, the well-known evolutionary conservation in term of protein domain and DNA-affinity of homeodomain containing TFs leads to several questions. For example, is the multi-step regulatory function of gene expression a general characteristic of homeodomain TFs? Is it evolutionary conserved in all bilaterian animals? It will be important to follow this line of research in depth to understand all the regulatory levels that allow Hox TF and other homeodomain-containing TFs to execute their diverse and highly context-dependent functions.

Acknowledgements

We would like to emphasize that as a result of the vast amount of published information as well as due to practical limitations (restriction on the number of citations) we tried to present the most recent and relevant data that would support our description of a multi-layer function of the Hox

TFs. Therefore, we would like to apologize to all those whose work, although also worthy of being cited, have, unfortunately, been left out.

References

AGELOPOULOS M, MCKAY DJ, MANN RS (2012). Developmental Regulation of Chromatin Conformation by Hox Proteins in Drosophila. *Cell Rep* 1: 350–359.

AGRAWAL P, HABIB F, YELAGANDULA R, SHASHIDHARA LS (2011). Genome-level identification of targets of Hox protein Ultrabithorax in Drosophila: novel mechanisms for target selection. *Sci Rep* 1:205.

BECCARI L, YAKUSHIJI-KAMINATSUI N, WOLTERING JM, NECSULEA A, LONFAT N, RODRÍGUEZ-CARBALLO E, MASCREZ B, YAMAMOTO S, KUROIWA A, DUBOULE D (2016). A role for HOX13 proteins in the regulatory switch between TADs at the *HoxD* locus. *Genes Dev*. 30: 1172-1186.

BEH CY, EL-SHARNOUBY S, CHATZIPLI A, RUSSELL S, CHOO SW, WHITE R (2016). Roles of cofactors and chromatin accessibility in Hox protein target specificity. *Epigenetics Chromatin* 9: 1.

BEI L, LU Y, BELLIS SL, ZHOU W, HORVATH E, EKLUND EA (2007). Identification of a HoxA10 Activation Domain Necessary for Transcription of the Gene Encoding β3 Integrin during Myeloid Differentiation. *J Biol Chem* 282: 16846–16859.

BENVENISTE D, SONNTAG H-J, SANGUINETTI G, SPROUL D (2014). Transcription factor binding predicts histone modifications in human cell lines. *Proc Natl Acad Sci* 111: 13367–13372.

BERKES CA, BERGSTROM DA, PENN BH, SEAVER KJ, KNOEPFLER PS, TAPSCOTT SJ (2004). Pbx Marks Genes for Activation by MyoD Indicating a Role for a Homeodomain Protein in Establishing Myogenic Potential. *Mol Cell* 14: 465–477.

BIGGIN MD, MCGINNIS W (1997). Regulation of segmentation and segmental identity by Drosophila homeoproteins: the role of DNA binding in functional activity and specificity. *Dev Camb Engl* 124: 4425–4433.

BOETTIGER AN, LEVINE M (2009). Synchronous and Stochastic Patterns of Gene Activation

in the Drosophila Embryo. *Science* 325: 471–473.

BOIJA A, MAHAT DB, ZARE A, HOLMQVIST P-H, PHILIP P, MEYERS DJ, COLE PA, LIS JT, STENBERG P, MANNERVIK M (2017). CBP Regulates Recruitment and Release of Promoter-Proximal RNA Polymerase II. *Mol Cell* 68: 491-503.e5.

BONDOS SE, CATANESE DJ, TAN X-X, BICKNELL A, LI L, MATTHEWS KS (2004). Hox Transcription Factor Ultrabithorax Ib Physically and Genetically Interacts with Disconnected Interacting Protein 1, a Double-stranded RNA-binding Protein. *J Biol Chem* 279: 26433–26444.

BONDOS SE, TAN X-X, MATTHEWS KS (2006). Physical and Genetic Interactions Link Hox Function with Diverse Transcription Factors and Cell Signaling Proteins. *Mol Cell Proteomics* 5: 824–834.

BOUBE M, FAUCHER C, JOULIA L, CRIBBS DL, BOURBON HM (2000). Drosophila homologs of transcriptional mediator complex subunits are required for adult cell and segment identity specification. *Genes Dev* 14: 2906–2917.

BOUBE M, HUDRY B, IMMARIGEON C, CARRIER Y, BERNAT-FABRE S, MERABET S, GRABA Y, BOURBON H-M, CRIBBS DL (2014). Drosophila melanogaster Hox Transcription Factors Access the RNA Polymerase II Machinery through Direct Homeodomain Binding to a Conserved Motif of Mediator Subunit Med19 Ed. GP Copenhaver. *PLoS Genet* 10: e1004303.

BROWN SJ, STOILOV P, XING Y (2012). Chromatin and epigenetic regulation of pre-mRNA processing. *Hum Mol Genet* 21: R90–R96.

CAO Q, WANG X, ZHAO M, YANG R, MALIK R, QIAO Y, POLIAKOV A, YOCUM AK, LI Y, CHEN W, CAO X, JIANG X, DAHIYA A, HARRIS C, FENG FY, KALANTRY S, QIN ZS, DHANASEKARAN SM, CHINNAIYAN AM (2014). The central role of EED in the orchestration of polycomb group complexes. *Nat Commun* 5.

CARNESECCHI J, FORCET C, ZHANG L, TRIBOLLET V, BARENTON B, BOUDRA R, CERUTTI C, BILLAS IML, SÉRANDOUR AA, CARROLL JS, BEAUDOIN C, VANACKER J-M (2017). ERRα induces H3K9 demethylation by LSD1 to promote cell invasion. *Proc Natl Acad Sci U S A* 114: 3909–3914.

CASTELLI-GAIR HOMBRÍA J., SÁNCHEZ-HIGUERAS C, SÁNCHEZ-HERRERO E (2016). Control of Organogenesis by Hox Genes. In *Organogenetic Gene Networks* (Eds James

Castelli-Gair Hombría and P Bovolenta). Springer International Publishing, Cham, pp. 319–373.

CHARIOT A, VAN LINT C, CHAPELIER M, GIELEN J, MERVILLE M-P, BOURS V (1999). CBP and histone deacetylase inhibition enhance the transactivation potential of the HOXB7 homeodomain-containing protein. *Oncogene* 18: 4007–4014.

CHOE S-K, LADAM F, SAGERSTRÖM CG (2014). TALE Factors Poise Promoters for Activation by Hox Proteins. *Dev Cell* 28: 203–211.

CHOE S-K, LU P, NAKAMURA M, LEE J, SAGERSTRÖM CG (2009). Meis Cofactors Control HDAC and CBP Accessibility at Hox-Regulated Promoters during Zebrafish Embryogenesis. *Dev Cell* 17: 561–567.

CHOO SW, WHITE R, RUSSELL S (2011). Genome-Wide Analysis of the Binding of the Hox Protein Ultrabithorax and the Hox Cofactor Homothorax in Drosophila Ed. G Gibson. *PLoS ONE* 6: e14778.

CHOPRA VS, HONG J-W, LEVINE M (2009). Regulation of Hox Gene Activity by Transcriptional Elongation in Drosophila. *Curr Biol* 19: 713.

CROCKER J, ABE N, RINALDI L, MCGREGOR AP, FRANKEL N, WANG S, ALSAWADI A, VALENTI P, PLAZA S, PAYRE F, MANN RS, STERN DL (2015). Low Affinity Binding Site Clusters Confer Hox Specificity and Regulatory Robustness. *Cell* 160: 191–203.

DOMSCH K, PAPAGIANNOULI F, LOHMANN I (2015). The HOX-Apoptosis Regulatory Interplay in Development and Disease. *Curr Top Dev Biol* 114: 121–158.

DUBOULE D (1992). The vertebrate limb: a model system to study the Hox/HOM gene network during development and evolution. *BioEssays News Rev Mol Cell Dev Biol* 14: 375–384.

DUJARDIN G, LAFAILLE C, DE LA MATA M, MARASCO LE, MUÑOZ MJ, LE JOSSIC-CORCOS C, CORCOS L, KORNBLIHTT AR (2014). How Slow RNA Polymerase II Elongation Favors Alternative Exon Skipping. *Mol Cell* 54: 683–690.

FARLEY EK, OLSON KM, ZHANG W, BRANDT AJ, ROKHSAR DS, LEVINE MS (2015). Suboptimization of developmental enhancers. *Science* 350: 325–328.

FARLEY EK, OLSON KM, ZHANG W, ROKHSAR DS, LEVINE MS (2016). Syntax compensates for poor binding sites to encode tissue specificity of developmental enhancers. *Proc Natl Acad Sci* 113: 6508–6513.

FLORENCE B, MCGINNIS W (1998). A Genetic Screen of the Drosophila X Chromosome for Mutations That Modify Deformed Function. *Genetics* 150: 1497-1511.

FOO SM, SUN Y, LIM B, ZIUKAITE R, O'BRIEN K, NIEN C-Y, KIROV N, SHVARTSMAN SY, RUSHLOW CA (2014). Zelda Potentiates Morphogen Activity by Increasing Chromatin Accessibility. *Curr Biol* 24: 1341–1346.

GAERTNER B, JOHNSTON J, CHEN K, WALLASCHEK N, PAULSON A, GARRUSS AS, GAUDENZ K, DE KUMAR B, KRUMLAUF R, ZEITLINGER J (2012). Poised RNA Polymerase II Changes over Developmental Time and Prepares Genes for Future Expression. *Cell Rep* 2: 1670–1683.

GARAULET DL, FORONDA D, CALLEJA M, SANCHEZ-HERRERO E (2008). Polycomb-dependent Ultrabithorax Hox gene silencing induced by high Ultrabithorax levels in Drosophila. *Development* 135: 3219–3228.

GELLON G, HARDING KW, MCGINNIS N, MARTIN MM, MCGINNIS W (1997). A genetic screen for modifiers of Deformed homeotic function identifies novel genes required for head development. *Development* 124: 3321-3331.

GORDON JAR, HASSAN MQ, SAINI S, MONTECINO M, VAN WIJNEN AJ, STEIN GS, STEIN JL, LIAN JB (2010). Pbx1 Represses Osteoblastogenesis by Blocking Hoxa10-Mediated Recruitment of Chromatin Remodeling Factors. *Mol Cell Biol* 30: 3531–3541.

DE GRAAF CA, VAN STEENSEL B (2013). Chromatin organization: form to function. *Curr Opin Genet Dev* 23: 185–190.

HASSAN MQ, TARE R, LEE SH, MANDEVILLE M, WEINER B, MONTECINO M, VAN WIJNEN AJ, STEIN JL, STEIN GS, LIAN JB (2007). HOXA10 Controls Osteoblastogenesis by Directly Activating Bone Regulatory and Phenotypic Genes. *Mol Cell Biol* 27: 3337–3352.

HIRIART E, GRUFFAT H, BUISSON M, MIKAELIAN I, KEPPLER S, MERESSE P, MERCHER T, BERNARD OA, SERGEANT A, MANET E (2005). Interaction of the Epstein-Barr Virus mRNA Export Factor EB2 with Human Spen Proteins SHARP, OTT1, and a Novel Member of the Family, OTT3, Links Spen Proteins with Splicing Regulation and mRNA Export. *J Biol Chem* 280: 36935–36945.

HNISZ D, ABRAHAM BJ, LEE TI, LAU A, SAINT-ANDRÉ V, SIGOVA AA, HOKE HA, YOUNG

RA (2013). Super-Enhancers in the Control of Cell Identity and Disease. *Cell* 155: 934–947.

HNISZ D, SHRINIVAS K, YOUNG RA, CHAKRABORTY AK, SHARP PA (2017). A Phase Separation Model for Transcriptional Control. *Cell* 169: 13–23.

HOMBRÍA JC-G, LOVEGROVE B (2003). Beyond homeosis—HOX function in morphogenesis and organogenesis. *Differentiation* 71: 461–476.

HORNYIK C, TERZI LC, SIMPSON GG (2010). The Spen Family Protein FPA Controls Alternative Cleavage and Polyadenylation of RNA. *Dev Cell* 18: 203–213.

HUANG Y, SITWALA K, BRONSTEIN J, SANDERS D, DANDEKAR M, COLLINS C, ROBERTSON G, MACDONALD J, CEZARD T, BILENKY M, THIESSEN N, ZHAO Y, ZENG T, HIRST M, HERO A, JONES S, HESS JL (2012). Identification and characterization of Hoxa9 binding sites in hematopoietic cells. *Blood* 119: 388–398.

HUISINGA KL, BROWER-TOLAND B, ELGIN SCR (2006). The contradictory definitions of heterochromatin: transcription and silencing. *Chromosoma* 115: 110–122.

IZEDDIN I, RÉCAMIER V, BOSANAC L, CISSÉ II, BOUDARENE L, DUGAST-DARZACQ C, PROUX F, BÉNICHOU O, VOITURIEZ R, BENSAUDE O, DAHAN M, DARZACQ X (2014). Single-molecule tracking in live cells reveals distinct target-search strategies of transcription factors in the nucleus. *eLife* 3.

JERKOVIĆ I, IBRAHIM DM, ANDREY G, HAAS S, HANSEN P, JANETZKI C, GONZÁLEZ NAVARRETE I, ROBINSON PN, HECHT J, MUNDLOS S (2017). Genome-Wide Binding of Posterior HOXA/D Transcription Factors Reveals Subgrouping and Association with CTCF Ed. EM Mendenhall. *PLOS Genet* 13: e1006567.

JOHNSON FB, KRASNOW MA (1992). Differential regulation of transcription preinitiation complex assembly by activator and repressor homeo domain proteins. *Genes Dev* 6: 2177-2189.

JOHNSON FB, KRASNOW MA (1990). Stimulation of transcription by an Ultrabithorax protein in vitro. *Genes Dev* 4: 1044–1052.

KINGSTON RE, TAMKUN JW (2014). Transcriptional Regulation by Trithorax-Group Proteins. *Cold Spring Harb Perspect Biol* 6: a019349–a019349.

KOENECKE N, JOHNSTON J, HE Q, MEIER S, ZEITLINGER J (2017). Drosophila poised

enhancers are generated during tissue patterning with the help of repression. *Genome Res* 27: 64-74.

KOUZARIDES T (2007). Chromatin Modifications and Their Function. *Cell* 128: 693–705.

KWONG C, ADRYAN B, BELL I, MEADOWS L, RUSSELL S, MANAK JR, WHITE R (2008). Stability and Dynamics of Polycomb Target Sites in Drosophila Development Ed. B van Steensel. *PLoS Genet* 4: e1000178.

LADAM F, SAGERSTRÖM CG (2014). Hox regulation of transcription: More complex(es): Hox Transcription Complexes. *Dev Dyn* 243: 4–15.

LEVINE M (2011). Paused RNA Polymerase II as a Developmental Checkpoint. *Cell* 145: 502–511.

LIANG H-L, NIEN C-Y, LIU H-Y, METZSTEIN MM, KIROV N, RUSHLOW C (2008). The zinc-finger protein Zelda is a key activator of the early zygotic genome in Drosophila. *Nature* 456: 400–403.

LIU Z, LEGANT WR, CHEN B-C, LI L, GRIMM JB, LAVIS LD, BETZIG E, TJIAN R (2014). 3D imaging of Sox2 enhancer clusters in embryonic stem cells. *eLife* 3.

LORCH Y, MAIER-DAVIS B, KORNBERG RD (2014). Role of DNA sequence in chromatin remodeling and the formation of nucleosome-free regions. *Genes Dev* 28: 2492–2497.

LU Y, GOLDENBERG I, BEI L, ANDREJIC J, EKLUND EA (2003). HoxA10 Represses Gene Transcription in Undifferentiated Myeloid Cells by Interaction with Histone Deacetylase 2. *J Biol Chem* 278: 47792–47802.

MAGNANI L, BALLANTYNE EB, ZHANG X, LUPIEN M (2011). PBX1 Genomic Pioneer Function Drives ERα Signaling Underlying Progression in Breast Cancer Ed. BE Clurman. *PLoS Genet* 7: e1002368.

MANN RS, CHAN SK (1996). Extra specificity from extradenticle: the partnership between HOX and PBX/EXD homeodomain proteins. *Trends Genet TIG* 12: 258–262.

MANN RS, LELLI KM, JOSHI R (2009). Chapter 3 Hox Specificity. In *Current Topics in Developmental Biology* Elsevier, pp. 63–101.

MAVES L, WASKIEWICZ AJ, PAUL B, CAO Y, TYLER A, MOENS CB, TAPSCOTT SJ (2007). Pbx homeodomain proteins direct Myod activity to promote fast-muscle differentiation.

Development 134: 3371–3382.

MEIER K, BREHM A (2014). Chromatin regulation: How complex does it get? *Epigenetics* 9: 1485–1495.

MERABET S, HUDRY B (2011). On the border of the homeotic function: Re-evaluating the controversial role of cofactor-recruiting motifs: The role of cofactor-recruiting motifs in conferring Hox evolutionary flexibility may critically depend on the protein environment. *BioEssays* 33: 499–507.

MERABET S, LOHMANN I (2015). Toward a New Twist in Hox and TALE DNA-Binding Specificity. *Dev Cell* 32: 259–261.

MERABET S, MANN RS (2016). To Be Specific or Not: The Critical Relationship Between Hox And TALE Proteins. *Trends Genet* 32: 334–347.

MORTIN MA, LEFEVRE G (1981). An RNA polymerase II mutation in Drosophila melanogaster that mimics ultrabithorax. *Chromosoma* 82: 237–247.

MORTIN MA, ZUERNER R, BERGER S, HAMILTON BJ (1992). Mutations in the second-largest subunit of Drosophila RNA polymerase II interact with Ubx. *Genetics* 131: 895–903.

NOYES MB, CHRISTENSEN RG, WAKABAYASHI A, STORMO GD, BRODSKY MH, WOLFE SA (2008). Analysis of Homeodomain Specificities Allows the Family-wide Prediction of Preferred Recognition Sites. *Cell* 133: 1277–1289.

PEARSON JC, LEMONS D, MCGINNIS W (2005). Modulating Hox gene functions during animal body patterning. *Nat Rev Genet* 6: 893–904.

PETRUK S (2001). Trithorax and dCBP Acting in a Complex to Maintain Expression of a Homeotic Gene. *Science* 294: 1331–1334.

PINTO PAB, HENRIQUES T, FREITAS MO, MARTINS T, DOMINGUES RG, WYRZYKOWSKA PS, COELHO PA, CARMO AM, SUNKEL CE, PROUDFOOT NJ, MOREIRA A (2011). RNA polymerase II kinetics in *polo* polyadenylation signal selection: RNA polymerase II kinetics and alternative polyadenylation. *EMBO J* 30: 2431–2444.

PRICE DH (2008). Poised Polymerases: On Your Mark…Get Set…Go! *Mol Cell* 30: 7–10.

PRINCE F, KATSUYAMA T, OSHIMA Y, PLAZA S, RESENDEZ-PEREZ D, BERRY M, KURATA S, GEHRING WJ (2008). The YPWM motif links Antennapedia to the basal

transcriptional machinery. *Development* 135: 1669–1679.

PROUDFOOT NJ, FURGER A, DYE MJ (2002). Integrating mRNA Processing with Transcription. *Cell* 108: 501–512.

REZSOHAZY R, SAURIN AJ, MAUREL-ZAFFRAN C, GRABA Y (2015). Cellular and molecular insights into Hox protein action. *Development* 142: 1212–1227.

ROUMENGOUS S, ROUSSET R, NOSELLI S (2017). Polycomb and Hox Genes Control JNK-Induced Remodeling of the Segment Boundary during Drosophila Morphogenesis. *Cell Rep* 19: 60–71.

ROY SS, MUKHERJEE AK, CHOWDHURY S (2018). Insights about genome function from spatial organization of the genome. *Hum Genomics* 12: 8.

SALDI T, CORTAZAR MA, SHERIDAN RM, BENTLEY DL (2016). Coupling of RNA Polymerase II Transcription Elongation with Pre-mRNA Splicing. *J Mol Biol* 428: 2623–2635.

SALEH M, RAMBALDI I, YANG XJ, FEATHERSTONE MS (2000). Cell signaling switches HOX-PBX complexes from repressors to activators of transcription mediated by histone deacetylases and histone acetyltransferases. *Mol Cell Biol* 20: 8623–8633.

SÁNCHEZ-HERRERO E (2013). Hox Targets and Cellular Functions. *Scientifica* 2013: 1–26.

SCHUETTENGRUBER B, BOURBON H-M, DI CROCE L, CAVALLI G (2017). Genome Regulation by Polycomb and Trithorax: 70 Years and Counting. *Cell* 171: 34–57.

SEKIYA T, MUTHURAJAN UM, LUGER K, TULIN AV, ZARET KS (2009). Nucleosome-binding affinity as a primary determinant of the nuclear mobility of the pioneer transcription factor FoxA. *Genes Dev* 23: 804–809.

SEKIYA T, ZARET KS (2007). Repression by Groucho/TLE/Grg proteins: genomic site recruitment generates compacted chromatin in vitro and impairs activator binding in vivo. *Mol Cell* 28: 291–303.

SHEN W -f., KRISHNAN K, LAWRENCE HJ, LARGMAN C (2001). The HOX Homeodomain Proteins Block CBP Histone Acetyltransferase Activity. *Mol Cell Biol* 21: 7509–7522.

SHESTAKOVA EA, BOUTIN M, BOURASSA S, BONNEIL E, BIJL JJ (2017). Identification of proteins associated with transcription factors HOXA9 and E2A-PBX1 by tandem affinity purification. *Mol Biol* 51: 432–444.

SHLYUEVA D, MEIRELES-FILHO ACA, PAGANI M, STARK A (2016). Genome-Wide Ultrabithorax Binding Analysis Reveals Highly Targeted Genomic Loci at Developmental Regulators and a Potential Connection to Polycomb-Mediated Regulation Ed. M Mallo. *PLOS ONE* 11: e0161997.

SLATTERY M, MA L, NÉGRE N, WHITE KP, MANN RS (2011). Genome-Wide Tissue-Specific Occupancy of the Hox Protein Ultrabithorax and Hox Cofactor Homothorax in Drosophila Ed. G Gibson. *PLoS ONE* 6: e14686.

SLATTERY M, RILEY T, LIU P, ABE N, GOMEZ-ALCALA P, DROR I, ZHOU T, ROHS R, HONIG B, BUSSEMAKER HJ, MANN RS (2011). Cofactor Binding Evokes Latent Differences in DNA Binding Specificity between Hox Proteins. *Cell* 147: 1270–1282.

SORGE S, HA N, POLYCHRONIDOU M, FRIEDRICH J, BEZDAN D, KASPAR P, SCHAEFER MH, OSSOWSKI S, HENZ SR, MUNDORF J, RÄTZER J, PAPAGIANNOULI F, LOHMANN I (2012). The *cis* -regulatory code of Hox function in *Drosophila*: The *cis* -regulatory code of Hox function in *Drosophila*. *EMBO J* 31: 3323–3333.

SUN Y, NIEN C-Y, CHEN K, LIU H-Y, JOHNSTON J, ZEITLINGER J, RUSHLOW C (2015). Zelda overcomes the high intrinsic nucleosome barrier at enhancers during *Drosophila* zygotic genome activation. *Genome Res* 25: 1703–1714.

SWANSON CI, EVANS NC, BAROLO S (2010). Structural Rules and Complex Regulatory Circuitry Constrain Expression of a Notch- and EGFR-Regulated Eye Enhancer. *Dev Cell* 18: 359–370.

TIE F, BANERJEE R, FU C, STRATTON CA, FANG M, HARTE PJ (2016). Polycomb inhibits histone acetylation by CBP by binding directly to its catalytic domain. *Proc Natl Acad Sci* 113: E744–E753.

TIE F, BANERJEE R, SAIAKHOVA AR, HOWARD B, MONTEITH KE, SCACHERI PC, COSGROVE MS, HARTE PJ (2014). Trithorax monomethylates histone H3K4 and interacts directly with CBP to promote H3K27 acetylation and antagonize Polycomb silencing. *Development* 141: 1129–1139.

TIE F, BANERJEE R, STRATTON CA, PRASAD-SINHA J, STEPANIK V, ZLOBIN A, DIAZ MO, SCACHERI PC, HARTE PJ (2009). CBP-mediated acetylation of histone H3 lysine 27

antagonizes Drosophila Polycomb silencing. *Development* 136: 3131–3141.

TUPLER R, PERINI G, GREEN MR (2001). Expressing the human genome. *Nature* 409: 832–833.

TSAI A, MUTHUSAMY AK, ALVES MR, LAVIS LD, SINGER RH, STERN DL, CROCKER J (2017). Nuclear microenvironments modulate transcription from low-affinity enhancers. *eLife* 6.

WALDHOLM J, WANG Z, BRODIN D, TYAGI A, YU S, THEOPOLD U, FARRANTS AK, VISA N (2011). SWI/SNF regulates the alternative processing of a specific subset of pre-mRNAs in Drosophila melanogaster. *BMC Mol Biol* 12: 46.

WIELLETTE EL, HARDING KW, MACE KA, RONSHAUGEN MR, WANG FY, MCGINNIS W (1999). spen encodes an RNP motif protein that interacts with Hox pathways to repress the development of head-like sclerites in the Drosophila trunk. *Development* 126: 5373-5385.

WORKMAN JL (2006). Nucleosome displacement in transcription. *Genes Dev* 20: 2009–2017.

ZARET KS, CARROLL JS (2011). Pioneer transcription factors: establishing competence for gene expression. *Genes Dev* 25: 2227–2241.

ZHU A, KUZIORA MA (1996). Homeodomain interaction with the beta subunit of the general transcription factor TFIIE. *J Biol Chem* 271: 20993–20996.

ZOUAZ A, AURADKAR A, DELFINI MC, MACCHI M, BARTHEZ M, ELA AKOA S, BASTIANELLI L, XIE G, DENG W, LEVINE SS, GRABA Y, SAURIN AJ (2017). The Hox proteins Ubx and AbdA collaborate with the transcription pausing factor M1BP to regulate gene transcription. *EMBO J* 36: 2887–2906.

Figure Legends

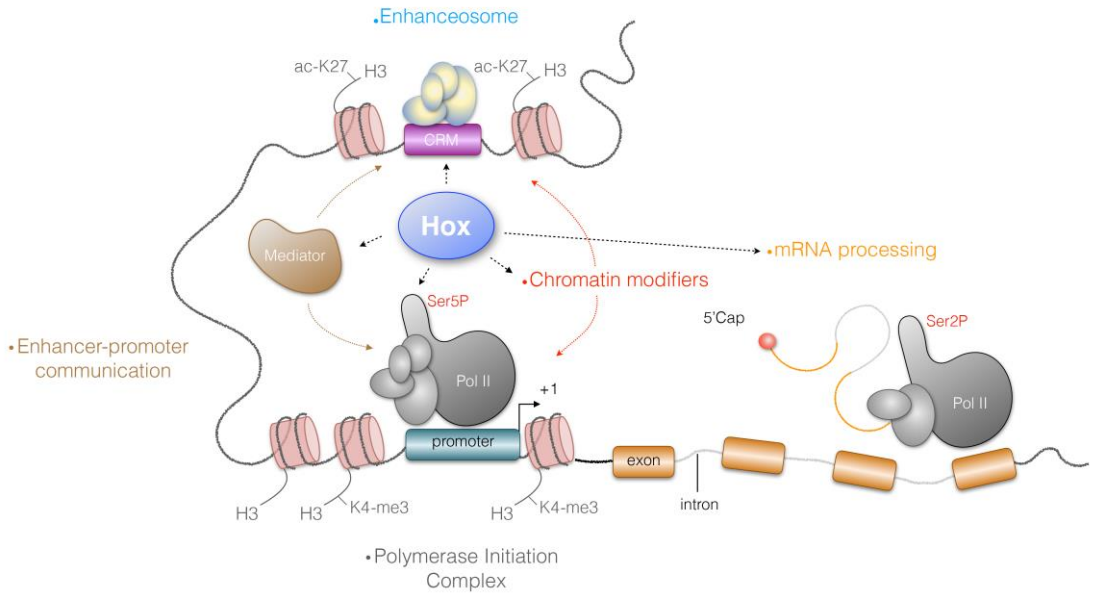

Figure 1: Hox TFs are multi-step players of gene expression. Representative picture of the potential multi-level regulatory function by Hox factors in the chromatin landscape (red), at enhanceosome (blue), enhancer-promoter communication (brown), polymerase initiation complex (grey) and mRNA processing levels (orange). Specifically, Hox TFs act: i), on the chromatin state by the regulation of histone acetylation (H3K27ac) and methylation (H3K4me3, H3K27me3) via interplay between TrxG, CBP, and PcG, HDAC ii), on nucleosome occupancy via SWI/SNF nucleosome remodelling complex, iii), on the chromatin conformation via the regulation of enhancer-promoter loop (CTCF, histone variant, Mediator complex), iv), on Polymerase Initiation Complex (PIC) and the switch from paused Ser5P-Pol II to active Ser2P-Pol II, v), on mRNA processing and vi), at the enhanceosome level characterised by defined enhancer syntax. They do so by the establishment of multi-regulatory complexes that will orchestrate their function *in vivo* in order to deploy their specific and diverse morphogenesis functions. H3: histone 3, ac-K27: H3K27ac, K4-me3: H3K4me3, H2av: histone 2A variant, CRM: cis-regulatory module, SWI/SNF: SWItch/Sucrose Non-Fermentable, 5'cap: 5' RNA-capping, pol II: RNA Polymerase II, Ser2P: serine 2 phosphorylated of Pol II, Ser5P: serine 5 phosphorylated of Pol II**,** +1: first nucleosome after the transcriptional start site (TSS).