Cellular and molecular insights into Hox protein action

René Rezsohazy1,*, Andrew J. Saurin2, Corinne Maurel-Zaffran2 and Yacine Graba2,*

ABSTRACT
Hox genes encode homeodomain transcription factors that control morphogenesis and have established functions in development and evolution. Hox proteins have remained enigmatic with regard to the molecular mechanisms that endow them with specific and diverse functions, and to the cellular functions that they control. Here, we review recent examples of Hox-controlled cellular functions that highlight their versatile and highly context-dependent activity. This provides the setting to discuss how Hox proteins control morphogenesis and organogenesis. We then summarise the molecular modalities underlying Hox protein function, in particular in light of current models of transcription factor function. Finally, we discuss how functional divergence between Hox proteins might be achieved to give rise to the many facets of their action.

KEY WORDS: Hox, Homeodomain, Morphogenesis

Introduction
Hox genes play fundamental roles in controlling the final morphology of bilaterian animals (Krumlauf, 1994; Pearson et al., 2005). Both loss and gain of Hox gene activity often result in homeotic transformations characterised by the appearance of an organ or structure that is normally formed elsewhere in the animal. Hox genes display several levels of evolutionary conservation (Fig. 1). First, at a structural level, they are often clustered in complexes, an organisation that is likely to reflect their phylogeny and regulatory constraints on their expression (Duboule, 2007; Mallo and Alonso, 2013). Second, at the molecular level, they all encode homeodomain transcription factors (Gehring et al., 1994). Third, in terms of function, they provide similar contributions in most animals and can even substitute the function of an orthologue in another species (McGinnis et al., 1990). Importantly, variation in Hox gene number, expression pattern and protein activity has played a major role in the evolution of the metazoan body plan (Gellon and Mallo, 2013). Hox genes have also been associated with a number of human diseases (Quinonez and Innis, 2014), highlighting the fact that Hox proteins are key factors in development, evolution and physiopathological processes.

Several decades of developmental studies have set the basis for a comprehensive, although not definitive, picture of the processes modulated by Hox genes. For example, initial studies recognised the patterning functions of Hox genes, showing that they provide axial positional information and contribute to defining cellular territories and establishing boundaries. More recent work showed that Hox genes also provide contributions to organogenesis per se (Hombria and Lovegrove, 2003) through the control of a variety of cellular functions including differentiation, proliferation, migration or death (Sanchez-Herrero, 2013). Although various molecular functions have been attributed to Hox proteins, including non-transcriptional functions such as replication and translation (Miotto and Graba, 2010; Rezsohazy, 2014) (Box 1), Hox proteins are best known as transcription factors. Consequently, it is thought that their functions rely on the selective activation (or repression) of downstream gene networks. In recent years, the search for Hox target genes has identified a wide range of downstream factors of diverse biological function (Choo and Russell, 2011; Graba et al., 1992; Hueber et al., 2007), including regulatory functions (e.g. signalling molecules or components of signalling pathways, transcription factors) and the so-called realisator genes that are more directly involved in morphogenetic processes.

The general path underlying Hox gene function – from the regulation of targets and gene regulatory networks to the control of patterning and cellular functions – has long been recognised (Fig. 2). However, there are several substantial gaps along this path. The molecular mechanisms underlying transcriptional regulation by Hox proteins are still poorly understood, and only a few Hox-dependent gene regulatory networks have been characterised. How these networks ultimately interface with the control of cellular functions also remains elusive. In addition, it is unclear how distinct or similar the paths for each Hox protein are. In this Review, we examine Hox transcription factor function at the cellular and molecular scale and discuss the extent to which this multiscale view allows a comprehensive understanding of how Hox proteins instruct differential morphogenesis.

An overview of the Hox gene family
Hox proteins have two distinctive and highly conserved features: the hexapeptide (HX) motif and the homeodomain (HD). The HX mediates contact with members of the PBC class of proteins, such as Extradenticle (Exd) in Drosophila (Mann and Chan, 1996). The HD is a widely used DNA-binding motif that is also found in non-Hox proteins. The HD folds into a triple helicoidal structure, with the HD N-terminal arm contacting the DNA in the minor groove, and helix 3, which is also termed the recognition helix, contacting the DNA in the major groove (Fig. 1A). Sequence conservation within Hox HDs is concentrated in helices 1 and 3, although some positions of the N-terminal arm and loops between the helices are also well conserved (Fig. 1A).

Hox genes are often clustered in the genome and are thought to originate from the last common ancestor of bilaterian animals, which is proposed to harbour a cluster of at least seven Hox genes (García-Fernández, 2005). During the Cambrian explosion, the number of Hox genes increased, notably through specific gene as well as genome duplications (Duboule, 2007). Vertebrates thus display thirteen parologue groups (PGs). Each PG comprises genes that occupy similar positions within genomic clusters (although with some exceptions, for example in Caenorhabditis elegans) and exhibit similar expression patterns and functions across species. Animals such as Drosophila melanogaster and C. elegans have a single Hox complement, whereas those that have undergone genome
duplications have several Hox gene clusters, such as four for mammals and seven in actinopterygians (Amores et al., 2004). Paralogous Hox proteins also display high levels of sequence conservation, mostly within the HD, with HDs from different animals but in the same PG displaying greater similarity than any other Hox HDs within the same animal. PGs are also clustered in an anterior, central and posterior classes, reflecting their domains of expression and action, but also correlating with the level of sequence conservation within the HD (Fig. 1B).

The cellular side of Hox activity

Recent reviews have summarised the many different cellular functions controlled by Hox proteins, including changes in cell shape and migration, proliferation or programmed cell death and differentiation (Cerdá-Esteban and Spagnoli, 2014; Philippidou and Dasen, 2013; Sanchez-Herrero, 2013; Shah and Sukumar, 2010; Taniguchi, 2014). Here, we review phenotypes associated with Hox gene dysfunction/manipulation that provide insights, at least at the phenomenological level, into how Hox genes achieve their functions. The examples chosen do not aim to provide an exhaustive overview but offer sufficient coverage to illustrate the diversity of cellular behaviours controlled by Hox proteins.

As early as during gastrulation in birds and mammals, Hox genes are expressed in an orderly temporal way in cells of the epiblast before their ingression into the primitive streak. It has been suggested that this expression controls the migratory properties of cells and the

Fig. 1. Hox genes and proteins. (A) The hexapeptide (HX) motif and the homeodomain (HD) are the most prominent sequence signatures of Hox proteins. The sequence of the HD is highly conserved, as illustrated here by a WebLogo (established using weblogo.berkeley.edu software) derived from a compilation of human/mouse and D. melanogaster sequences. The HD contains three alpha helices and contacts DNA in the major (via helix 3) and minor (via the N-terminal arm) grooves. The HX, a short motif located upstream of the HD, contacts protein partners of the PBC class, which are also HD-containing proteins, allowing the cooperative assembly of a Hox-PBC-DNA complex, here illustrated by the Ubx-Exd-DNA crystal structure (1B8I; Passner et al., 1999). (B) Representation of C. elegans, D. melanogaster, mouse and human Hox clusters. Paralogue groups (PGs) are classified as anterior, central and posterior classes. Hox genes are arranged in order along chromosomes according to their PG, with the exception of ceh-13 (PG1) and lin-39 (PG5) in C. elegans, the genomic locations of which are reversed (as indicated by dashed lines). WebLogos illustrating the HD conservation within each class highlights that the central class HDs display high conservation, whereas the anterior, and to a greater extent the posterior, class display less conservation. The subregions of the HD that are best conserved also vary within the different classes.
Hoxa5 in the mouse, which instructs epithelia from stromal cells in control of cell differentiation has also been well documented for expansion and differentiation (Chen and Capecchi, 1999). The controlling cell fate decisions along differentiation lineages. PG9 system (Castelli Gair Hombria et al., 2009).

Shape changes and invagination during development of the posterior modifications in cell adhesion and cytoskeleton required for cell protein Abdominal B (AbdB) has been shown to coordinate the auditory circuits in the mouse (Di Bonito et al., 2013). In et al., 2013; Geisen et al., 2008), or in the development of central precerebellar system relaying inputs to the cerebellum (Di Meglio for example during the establishment of somatosensory topographic development in mice (Di-Po et al., 2007). During haematopoiesis, several Hox genes contribute to the control of cell stemness versus differentiation (Alharbi et al., 2013).

In this context, HOXB4 seems to act both in a cell-autonomous and non-cell-autonomous manner, possibly via effects on the establishment of the haematopoietic stem cell (HSC) niche (Jackson et al., 2012). Similarly, AbdB controls the stem cell niche in the Drosophila testis, notably by non-cell-autonomous regulation of centrosome orientation and the division rates of germline stem cells (Papagiannouli et al., 2014).

Hox proteins can also modulate cell proliferation and cell cycle progression. Among recently reported examples, Antennapedia (Antp) was shown to control cell cycle exit in Drosophila neuroblasts (Baumgardt et al., 2013). In zebrafish, Hoxb8a controls cell proliferation in the primordia that will contribute to the development of the lateral line (Breau et al., 2013). There are also several examples in which Hox proteins exhibit oncogenic activities, in particular in the context of leukaemia (Shah and Sukumar, 2010). Furthermore, in murine models of leukaemia, Hoxa9 has recently been shown to directly regulate cell cycle regulators together with CCAAT/enhancer-binding protein alpha (Cebpa) (Collins et al., 2014) and the methyltransferase G9a (Ehmt2) (Lehnertz et al., 2014).

Several instances suggest that Hox activity can impact on the decision to engage programmed cell death, for example during Hox-mediated regulation of the cell death gene reaper in Drosophila (Lohmann et al., 2002; Stöbe et al., 2009). The C. elegans LIN-39 Hox protein suppresses inappropriate apoptosis in male-specific serotonergic neuron precursors (Kalas et al., 2014), and MAB-5, another C. elegans Hox protein, directly controls apoptosis in pre-neural ventral blast cells (Liu et al., 2006). Hoxd gene deletions cause an up to 6-fold increase in apoptosis during kidney development in mice (Di-Poi et al., 2007). By contrast, Hoxa13 loss-of-function mouse mutants show defects in interdigital apoptosis (Post et al., 2000; Stadler et al., 2001), whereas HOXA5 stimulates apoptosis in human and mouse breast tissue by regulating p53 (TP53/TRP53), Twist and caspase 2 and 8 genes (Chen et al., 2004; Raman et al., 2000; Stasinopoulos et al., 2005).

Timing of inversion (limura and Pourquié, 2006). Later on, Hox genes assist cells in migration and, in the case of the central nervous system, in establishing networks and circuits (Gavalas et al., 1997), for example during the establishment of somatosensory topographic projections and wiring (Oury et al., 2006), in building up the precerebellar system relaying inputs to the cerebellum (Di Meglio et al., 2013; Geisen et al., 2008), or in the development of central auditory circuits in the mouse (Di Bonito et al., 2013). In C. elegans, Hox proteins also contribute to control neuroblast cell migration (Tamayo et al., 2013). Furthermore, in zebrafish, Hox genes have been highlighted as controlling collective cell migration in the development of the mechanoreceptive lateral line (Breau et al., 2013). In all these situations, Hox proteins control receptor/ligand expression leading to attractive/repulsive interactions between migrating cells and the environment that they travel through. Modulating cell shape is another way to control single or collective cell migration and tissue remodelling, and Hoxb1b has recently been shown to regulate microtubule dynamics during the process of neural tube formation in zebrafish (Zigman et al., 2014). Hox proteins also directly regulate cell shape, cell-to-cell communication and signalling to properly delimit functional and morphological borders between vertebrate hindbrain segments (Prin et al., 2014), and the Drosophila Hox protein Abdominal B (AbdB) has been shown to coordinate the modifications in cell adhesion and cytoskeleton required for cell shape changes and invagination during development of the posterior spiracle, which constitutes the posterior end of the larval respiratory system (Castelli Gair Hombria et al., 2009).

Hox proteins can also couple differentiation with morphogenesis, controlling cell fate decisions along differentiation lineages. PG9 genes, for example, are implicated in mammary gland maturation, expansion and differentiation (Chen and Capecechi, 1999). The control of cell differentiation has also been well documented for Hoxa5 in the mouse, which instructs epithelia from stromal cells in the lungs, gut and mammary gland (Aubin et al., 2002; Boucherat et al., 2012; Garin et al., 2006). In vertebrates, several Hox genes have been shown to either promote or inhibit vascular development or remodelling (Stoll and Kroll, 2012), supporting a general view that Hox proteins directly regulate effector genes involved in the control of cell-cell and cell-extracellular matrix interactions (Kachgal et al., 2012; Winnik et al., 2009). The skin is another organ in which several Hox proteins play roles in controlling differentiation, acting both as effector transcription factors, for example in the direct control of keratin genes in mammals, but also as ‘upper’ regulators, acting at a high level in the regulatory network hierarchy (Godwin and Capecechi, 1998; Johansson and Headon, 2014; La Celle and Polakowska, 2001; Potter et al., 2011; Rinn et al., 2008). Distinct Hox proteins are also expressed in the endometrium and appear functionally important for endometrial cell differentiation and receptivity (Lu et al., 2008; Xu et al., 2014). In Hoxd11-13 mutant mice, chondrocyte differentiation is arrested at an early stage, and shortly after birth these chondrocytes undergo rapid maturation and replacement by osteoblasts, thereby shortening bones (González-Martin et al., 2014). HOXA10 has been shown to sustain osteoblastogenesis (Gordon et al., 2011; Hassan et al., 2007).
The control of haematopoietic stemness provides another well-studied example of versatile and context-specific Hox protein function, and also highlights the existence of partnering with other transcription factors. Hoxb4 modulates HSC responsiveness to several extrinsic cues and displays a dose-dependent effect on HSC self-renewal, supporting normal differentiation versus perturbed differentiation (Klump et al., 2005; Schiedlmayer et al., 2007; Will et al., 2006). Context-specific HOXB4/Hoxb4 activity and promoter occupancy have been highlighted by genome-wide searches for targets, and these studies have revealed the dynamics of the Hoxb4 regulatory network in the embryonic stem to haematopoietic progenitor cell differentiation process (Fan et al., 2012; Oshima et al., 2011). Hoxb4 ChIP peaks correlate with known haematopoietic transcription factor binding sites or ChIP peaks, suggesting that Hoxb4 functions in a combinatorial fashion during haematopoiesis.

The functional versatility of Hox proteins is again exemplified during prostate cancer development. In prostate physiopathology, HOXB13 acts either as a tumour suppressor or as an oncogene, depending on the type of cancer cell. Some prostate cancers are known to grow in response to androgens, and HOXB13 is a bivalent regulator of the androgen receptor (AR) (Norris et al., 2009), positively or negatively regulating AR interactions with chromatin. This in turn modulates the cell response to androgens. In addition, in androgen-responsive prostate cancer cell models, HOXB13 activity results in a decrease in pRb (retinoblastoma 1) phosphorylation, which then leads to pRb-E2F complex stabilisation and growth inhibition (Hamid et al., 2014). By contrast, HOXB13 can be overexpressed in androgen-refractory prostate tumours, stimulating tumour progression rather than being an oncosuppressor (Kim et al., 2010). In these tumours, HOXB13 appears to inhibit p21 (CDKN1A), which then allows E2F activation and cell cycle progression.

Finally, the versatility of Hox protein function may also result from post-translational modifications. Phospho isoforms of Drosophila Ultrabithorax (Ubx), for example, have long been recognised (Gavis and Hogness, 1991), and changes in Drosophila Antp and Sex combs reduced (Scr) protein activities upon phosphorylation have been demonstrated (Berry and Gehring, 2000; Jaffe et al., 1997). It was also suggested that phosphorylation has influenced the evolution of Ubx limb-repressing function in the arthropod lineage (Ronsenjaug et al., 2002). In vertebrates, HOXA10 expression accompanies myeloid differentiation and, in this context, HOXA10 acts both as an activator and a repressor at different stages of myelopoiesis, with its activity being regulated by tyrosine phosphorylation and dephosphorylation (Beil et al., 2007; Eklund et al., 2002). HOXB7 can be polyADP-ribosylated by poly(ADP-ribose) polymerase 1.

The control of haematopoietic stemness provides another well-studied example of versatile and context-specific Hox protein function, and also highlights the existence of partnering with other transcription factors. Hoxb4 modulates HSC responsiveness to several extrinsic cues and displays a dose-dependent effect on HSC self-renewal, supporting normal differentiation versus perturbed differentiation (Klump et al., 2005; Schiedlmayer et al., 2007; Will et al., 2006). Context-specific HOXB4/Hoxb4 activity and promoter occupancy have been highlighted by genome-wide searches for targets, and these studies have revealed the dynamics of the Hoxb4 regulatory network in the embryonic stem to haematopoietic progenitor cell differentiation process (Fan et al., 2012; Oshima et al., 2011). Hoxb4 ChIP peaks correlate with known haematopoietic transcription factor binding sites or ChIP peaks, suggesting that Hoxb4 functions in a combinatorial fashion during haematopoiesis.

The functional versatility of Hox proteins is again exemplified during prostate cancer development. In prostate physiopathology, HOXB13 acts either as a tumour suppressor or as an oncogene, depending on the type of cancer cell. Some prostate cancers are known to grow in response to androgens, and HOXB13 is a bivalent regulator of the androgen receptor (AR) (Norris et al., 2009), positively or negatively regulating AR interactions with chromatin. This in turn modulates the cell response to androgens. In addition, in androgen-responsive prostate cancer cell models, HOXB13 activity results in a decrease in pRb (retinoblastoma 1) phosphorylation, which then leads to pRb-E2F complex stabilisation and growth inhibition (Hamid et al., 2014). By contrast, HOXB13 can be overexpressed in androgen-refractory prostate tumours, stimulating tumour progression rather than being an oncosuppressor (Kim et al., 2010). In these tumours, HOXB13 appears to inhibit p21 (CDKN1A), which then allows E2F activation and cell cycle progression.

Finally, the versatility of Hox protein function may also result from post-translational modifications. Phospho isoforms of Drosophila Ultrabithorax (Ubx), for example, have long been recognised (Gavis and Hogness, 1991), and changes in Drosophila Antp and Sex combs reduced (Scr) protein activities upon phosphorylation have been demonstrated (Berry and Gehring, 2000; Jaffe et al., 1997). It was also suggested that phosphorylation has influenced the evolution of Ubx limb-repressing function in the arthropod lineage (Ronsenjaug et al., 2002). In vertebrates, HOXA10 expression accompanies myeloid differentiation and, in this context, HOXA10 acts both as an activator and a repressor at different stages of myelopoiesis, with its activity being regulated by tyrosine phosphorylation and dephosphorylation (Beil et al., 2007; Eklund et al., 2002). HOXB7 can be polyADP-ribosylated by poly(ADP-ribose) polymerase 1.
Hox proteins as transcription factors: the Hox-PBC partnership and modes of action

Understanding the diversity and specificity of Hox protein function as well as the versatile nature of Hox protein activity requires us to examine how Hox proteins act as transcription factors. Hox proteins share a highly conserved HD, raising the issue of how functional specificity is achieved (Hayashi and Scott, 1990). HDs are found in a large number of transcription factors, display a wide range of sequence variations, and in some instances are associated with other DNA-binding domains. A survey of the DNA binding properties of nearly all Drosophila and mouse HDs showed that divergent HDs recognise distinct DNA sequences but that, within a given group, high-affinity binding sites are shared by all members (Berger et al., 2008; Noyes et al., 2008). Furthermore, with the exception of the posterior Hox class (AbdB in Drosophila), all Hox proteins fall within a single HD class, the so-called Antennapedia (Antp) class, in which the HD is not associated with an additional DNA-binding domain, providing little if any apparent molecular ground for the distinct functions of individual Hox proteins.

Our current view of how Hox proteins act was profoundly influenced by the early identification of PBC class proteins, which function as general Hox co-factors (Fig. 3). Like Hox proteins, the PBC class comprises evolutionarily conserved HD-containing proteins. Initially identified as a leukemic fusion to E2, PBC proteins belong to the TALE class of HD proteins, which are characterised by a three amino acid loop extension (Bürghlin, 1998). It was shown that Exd, the single representative of the PBC class in flies, is required for proper Hox protein activity (Peifer and Wieschaus, 1990) and promotes Hox cooperative DNA binding (Chan et al., 1994; van Dijk and Murre, 1994). Similarly, vertebrate Pbx proteins have been shown to promote Hox cooperative DNA binding (Phelan et al., 1995), and functional studies in the field of cancer and developmental biology provided functionally relevant contexts for such associations (Moens and Selleri, 2006). Although providing a step forward in terms of understanding Hox protein specificity, the PBC partnership still raises the issue of how Hox proteins could adopt distinct functions by interacting with a member of the single Pbx class in vertebrates or a single protein (Exd) in Drosophila.

The consensus DNA sequence bound by Hox-PBC complexes is 5′-TGATNNATNN-3′; the first four nucleotides bind the PBC protein, while the remainder are used for Hox binding (Fig. 4). This sequence is significantly longer than that recognised by Hox monomers, improving de facto specificity, and in most instances is different from the monomeric binding site, indicating that the interaction with PBC proteins modifies Hox DNA binding. Extensive analysis of DNA binding sites recognised by all Drosophila Hox proteins in association with Exd using the Selex-Seq experimental platform (Slattery et al., 2011b) indicated that the presence of Exd uncovers Hox latent binding site preferences. The distinction between preferred DNA sequences mostly, but not uniquely, lies on the central NN nucleotides. Interestingly, these central nucleotides were previously shown to influence the binding site preference and functional specificity of the Drosophila Labial (Lab) and Deformed (Dfd) Hox proteins (Chan et al., 1997).

Hox-PBC interactions were shown to rely on a highly conserved interaction mode involving, on the Hox side, the HX motif that lies upstream of the HD, and, on the PBC side, the TALE peptide that lies between helices 1 and 2 of the HD (Johnson et al., 1995; Joshi et al., 2007; LaRonde-LeBlanc and Wolberger, 2003; Passner et al., 1999; Piper et al., 1999; Shen et al., 1996). Taken together with the recognised importance, mostly from chimeric protein analysis, of the N-terminal arm of the HD in Hox protein specificity, it was proposed that the HX, the linker region separating the HX from the HD, and the HD N-terminal arm constitute a specificity module (Joshi et al., 2010). This specificity module contacts the PBC partner through the HX motif, which in turn geometrically constrains the linker region to ultimately position the N-terminal arm of the HD within the DNA minor groove. This positioning relies on different, yet difficult to uncouple, mechanisms involving DNA minor groove shape recognition and base-specific recognition (Joshi et al., 2007).

This specificity module, which displays some paralogue specificity, and its role in DNA binding site selectivity provide important foundations for distinguishing the activity of Hox proteins belonging to different Pgs, and support the notion that sequence divergence of Hox PG proteins drives functional diversification. Yet it is insufficient to account for Hox protein specificity, as it has on its own insufficient predictive value (Ebner et al., 2005). For example, the Lab-Exd complex was shown to recognise Hox/PB composite sequences bearing different nucleotides at the junction between the Hox and Exd half-sites: GG, in an evolutionary conserved autoregulatory element of the Hoxb1 gene (Chan et al., 1997; Pöpperl et al., 1995); TA in a...
modified lab autoregulatory element (Grieder et al., 1997); and CA in a distinct Lab downstream target gene (Ebner et al., 2005). Conversely, Lab and Scr both recognise and act in vivo on sequences with identical nucleotides in the central position of the Hox/PBC consensus site. DNA sequence variations in the distal part of the Hox half-site, which are in a position to contact the Hox recognition helix, are also likely to play a role in defining Hox DNA binding specificity (Fig. 4). This highlights the existence of additional specificity mechanisms that remain to be elucidated, a prerequisite before we can reasonably predict the in vivo DNA binding site preferences of Hox paralogue proteins.

Additional protein partners, such as Homothorax (Hth) in *Drosophila* and the Meis class of proteins in vertebrates, are known to physically interact with, and influence the activity of, Hox-PBC-DNA regulatory complexes (Chan et al., 1997; Li-Kroeger et al., 2008; Mann and Affolter, 1998; Ryoo et al., 1999). Genome-wide analysis of the DNA binding profile of Hth in the haltere, a flight balancing appendage, and leg imaginal discs highlighted a significant, yet only partial, overlap with that of the Hox protein Ubx (Choo et al., 2011; Slattery et al., 2011a). Similarly, in mouse, the genomic distribution of Meis proteins overlaps with those of Hoxc9 and Hoxa2 (Penkov et al., 2013). In the case of branchial arches, it was recently proposed that Hox2 selectively enhances Meis binding to modify a Meis-dependent ground state into a second branchial arch identity, highlighting a role for a Hox protein in specifying the activity of another transcription factor (Amin et al., 2015). How Hth/Meis and Hox proteins reciprocally modify their DNA binding preferences and functional specificities remains unclear and awaits structural characterisation of the Hox-PBC-Meis-DNA complexes that have been biochemically and functionally characterised.

**Additional facets of Hox-PBC partnerships**

Although not as well documented, partnerships with PBC proteins serve functions other than conferring Hox DNA binding specificity. Illustrative of this is the role of Exd in buffering the monomeric binding potential of the Hox protein Lab. Based on a change in Lab proteolytic pattern upon Exd binding, together with increased Lab monomeric binding and in vivo hyperactivity following mutation of the HX motif, it was proposed that the binding of Exd to Lab overcomes an inhibitory role of the HX on Lab DNA binding (Chan et al., 1996). Together with the observation that mutation of the HX also increases the monomeric binding of other Hox proteins, such as *Drosophila* Antp and AbdB and mouse Hoxb8 (Hudry et al., 2012; Merabet et al., 2007), this suggests that Hox-PBC interactions might serve not only to promote binding selectivity, but also, at least for some Hox proteins, to concomitantly prevent the inappropriate binding of monomeric Hox proteins (Fig. 4).

The study of a Ubx-VP16 fusion protein that functionally mimics Antp led to the proposition that regulation of Hox activity contributes to the functional distinction of Hox paralogues (Li and McGinnis, 1999). Changes in Hox transcriptional activity were also proposed to underlie the evolution of morphological traits in arthropods (Galant and Carroll, 2002; Ronshaugen et al., 2002). How the regulation of activity relates to PBC partnership was addressed in the case of Dfd. Based on cell culture, in vitro biochemistry and functional studies in the *Drosophila* embryo, it was proposed that the HD inhibits Dfd transcriptional activation potential and that the interaction with Exd suppresses this inhibitory function (Li et al., 1999). In addition, the deletion of Dfd N-terminal domains in a Dfd-Scr chimera, in which the specificity module is that of Scr, resulted in a repression-to-activation switch in the regulation of the *forkhead* downstream target, further supporting a role for Exd in controlling the regulation of Hox activity (Joshi et al., 2010). How exactly the interaction with Exd impacts on Hox activity, and how general this function of Exd is, remain to be explored.

Although they generally provide a positive partnership with Hox proteins, Exd and Hth were recently found to antagonise the function of posterior Hox proteins. Using phenotypic and molecular analyses, it was shown that inappropriate maintenance of Exd/Hth expression in the posterior domain of the *Drosophila* embryo, where identity is specified by the posterior Hox gene *AbdB*, phenocopies AbdB loss of function (Rivas et al., 2013). Such an antagonistic relationship also applies to the repression of the limb-suppressing gene *Distalless* (*Dll*) by AbdB (Sambrani et al., 2013). This highlights a distinct Hox-PBC partnership, conveying antagonism instead of cooperation, and suggests that diversifying the nature of the PBC partnership might have played a role in distinguishing the function of posterior versus anterior/central Hox proteins. The basis for this antagonistic function remains enigmatic. It is interesting to note, however, that in the extensive Selex-Seq-based comparative characterisation of Hox monomer versus Hox-Exd binding properties, it was found that in the case of the anterior Lab and the central Ubx proteins, but not the posterior AbdB protein, Exd promotes the paralogue specialisation of DNA binding properties (Slattery et al., 2011b).
Finally, while the HX has long remained the only well-characterised PBC interaction mode, Hox-PBC interaction is likely to involve greater complexity, which in turn probably contributes to the diversity and specificity of Hox action. A survey of in vivo Hox-PBC interactions, covering most *Drosophila* and a few representatives of vertebrate Hox proteins, highlighted that alternatives to HX-mediated PBC interaction are common (Hudry et al., 2012), extending previous observations showing that Exd-dependent function can be achieved by HX-deficient Ubx and AbdA proteins (Galant et al., 2002; Merabet et al., 2003). In addition, although providing insights into how the N-terminal arm of the HD contacts the DNA minor groove (the proximal part of the Hox binding site), the HX interaction mode only provides limited insight into how association with Exd elicits changes occurring in the distal portion of the Hox binding sites (Fig. 4), which are likely to be mediated by the Hox recognition helix (Slattery et al., 2011b).

Studies on *Drosophila* Ubx and AbdA have raised interesting perspectives in this regard. They identified a paralogue-specific protein domain, UbdA, that is located immediately C-terminal to the HD and mediates Exd recruitment (Lelli et al., 2011; Merabet et al., 2007; Saadouli et al., 2011). This domain folds as a flexible extension of the HD recognition helix and defines Hox-PBC contacts that occur, when compared with those mediated by the HX motif, on the opposing side of the DNA double helix (Foos et al., 2015). This suggests that UbdA-mediated Exd contacts might directly fine-tune the positioning of the recognition helix within the DNA major groove and influence DNA contacts mediated by the distal part of the Hox binding site.

**Current models for transcription factor function: Hox proteins and beyond**

It was initially, and intuitively, assumed that all transcription factors act via a ‘surgical’ mode of action, binding to a limited number of genomic loci. However, current models of transcription factor function prompt us to reconsider this view. Work on transcription factors encompassing many classes of DNA-binding domains that pattern the early *Drosophila* embryo revealed that transcription factors with well-defined developmental functions show a surprisingly high number of genomic binding sites – up to 20,000 (Li et al., 2008; MacArthur et al., 2009). Even more surprising was the discovery that binding sites for these molecularly distinct transcription factors with diverse biological functions are highly overlapping (MacArthur et al., 2009). This highlights that a binding event is better seen as a probabilistic event, rather than as a qualitative on/off event. Such broad genomic binding is consistent with transcription factor nuclear abundance (Biggin, 2011), which has recently been re-evaluated (Li et al., 2014), and with thermodynamic models of genomic occupancy (Kaplan et al., 2011). Correlative analysis further supports the idea that genomic regions with high/frequent transcription factor occupancy support transcription (MacArthur et al., 2009), whereas regions of low/infrequent transcription factor occupancy have no transcriptional regulatory potential (Fisher et al., 2012). This led to a view of transcription factors acting within the context of a quantitative continua model, in which there are no clear limits between transcriptionally relevant and non-relevant binding events (Biggin, 2011; see below). The qualitative folding of transcription factor complexes might serve, beyond quantitative aspects, to functionally discriminate these binding events.

To ascertain whether the quantitative continua model also applies to Hox proteins, it is necessary to appraise their genomic distribution. However, data are available in only a few instances: for two *Drosophila* proteins (Dfd and Ubx), three worm proteins (MAB-5, EGL-5 and LIN-39) and for a few vertebrate proteins (Hoxa2, Hoxa9/A9, HOXA13, Hoxb4/B4, Hoxc6/C6, Hoxc9/C9 and HOXD13; Table 1). The number of binding sites for each of these proteins in their respective genomes ranges from 500 to nearly 30,000. Whether these differences relate to intrinsic differences in DNA binding behaviour is difficult to assess, as the data were generated using different methodologies (ChIP-chip, ChIP-Seq) and were analysed using distinct bioinformatic approaches and statistical treatments. In particular, it is difficult to determine the appropriate stringency of analysis to use, especially when considering the recent finding that low-affinity binding sites mediate Ubx function in patterning the larval cuticle (Crocker et al., 2015).

Consideration of the characteristics of the starting biological material is also important when appreciating the significance of the binding events detected. In this context, two features – cellular heterogeneity and the developmental time window covered – are particularly important, and it should be noted that the binding events detected represent a sum of events occurring in different cell types and/or at different time points. Although generally applicable, these limitations hold especially true for Hox proteins, which are known to pattern fields of cells that encompass diverse cell types in a very dynamic manner by regulating distinct sets of target genes at different time points (Fan et al., 2012; Pavlopoulos and Akam, 2011; Sorge et al., 2012), possibly associated with dynamic genome binding (Fan et al., 2012; Niu et al., 2011). With these limitations in mind, and taking into account the often low false discovery rate (FDR) applied, the numbers (see Table 1) seem comparable to those found for other transcription factors, suggesting that Hox DNA binding might also be widespread.

Evaluating the genomic binding ability of Hox proteins also allows comparisons to be made, thereby enabling us to question how distinct the binding properties of different Hox paralogue proteins are and, in the longer term, to assess how evolutionarily conserved these features are. In the case of *Drosophila* Ubx and Dfd (Sorge et al., 2012) it was found that, despite sharing identical DNA binding sequences, the proteins bind to non-overlapping genomic regions, indicating that *in vivo* binding is regulated beyond the intrinsic DNA binding potential of Hox proteins. Thus, at least in this single example, it seems that, although widespread, Hox proteins might not share the overlapping genomic binding property that other transcription factors exhibit (MacArthur et al., 2009).

Another important issue when considering the quantitative continua model is to characterise the relationship between binding events and the transcription of nearby genes. Although transcriptomic data are available for a large number of Hox proteins (as recently summarised by Sanchez-Herrero, 2013), in only a few instances have experiments been conducted on identical or similar biological material allowing genomic and transcriptomic data to be linked (Table 1 and Fig. 5, solid red and green boxes). In all cases, it was found that only a small fraction of binding events are associated with Hox-mediated transcriptional control (Table 1), supporting the quantitative continua model. This conclusion should be taken with caution, however, as dynamic DNA binding might equally account for these observations.

**Extending the Hox framework: additional transcription factors, chromatin regulators and links to the transcription machinery**

Genetic studies in *Drosophila* have identified Exd-independent Hox functions, for example during heart patterning and specification of
the haltere (Galant et al., 2002; Perrin et al., 2004). Consistent with this, genomic data highlight poor overlap between Ubx and Exd binding events (Negre et al., 2011), as well as between mouse Hoxa2/Hoxc9 and Pbx1 (Penkov et al., 2013), collectively supporting the idea that Hox proteins can also act without the help of PBC proteins. In line with this, sequence analyses of Hox-precipitated genomic DNA have identified DNA binding motifs for known transcription factors that thus emerge as potential additional Hox protein partners (Table 1). This provides impetus to extend our view of the architecture of the Hox response element (Mann et al., 2009; Pearson et al., 2005). However, these findings must be correlated with functional studies or proteome-wide interactomic

### Table 1. Evaluating the genomic binding and gene regulatory effects of Hox proteins

<table>
<thead>
<tr>
<th>Genomic binding</th>
<th>Biological context</th>
<th>Method</th>
<th># peaks</th>
<th>Genes regulated</th>
<th>Biological context</th>
<th># genes</th>
<th>Partners predicted from enriched DNA binding motifs*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hoxa2</td>
<td>Mouse II BA (E11.5) [1]</td>
<td>ChiP-Seq</td>
<td>8245</td>
<td>Mouse II BA (E11.5) [1]</td>
<td></td>
<td>489</td>
<td>CREB, MYB, CAUDAL, ETS, MYC, STAT</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Rhombomeres (E9.5)</td>
<td></td>
<td>220</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Human leukemoid cell lines [3]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Human cord blood cells [4]</td>
<td></td>
<td>&gt;1981†</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Mouse HPCa [5]</td>
<td></td>
<td>&gt;76‡</td>
<td>C/EBPα</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Human primary fibroblasts [9]</td>
<td></td>
<td>176</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HOXA13</td>
<td>Mouse embryonic fibroblasts [7]</td>
<td>ChiP library</td>
<td>NA</td>
<td>Mouse embryonic fibroblasts [8]</td>
<td></td>
<td>68</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Three stages of mouse ESC to HSC differentiation [12]</td>
<td>ChiP-chip</td>
<td>3632, 7232, 29,313</td>
<td>Four stages of mouse ESC to HSC differentiation [12]</td>
<td></td>
<td>5780</td>
<td>FLI1, RUNX1, SCL</td>
</tr>
<tr>
<td>Hoxc9/HOXC9</td>
<td>Mouse thoracic spinal cord (E12.5) [14]</td>
<td>ChiP-chip</td>
<td>NA</td>
<td>–</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HOXD13</td>
<td>Human chordroblast cell line [16]</td>
<td>ChiP-chip</td>
<td>248</td>
<td>–</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dfd</td>
<td>Drosophila embryo (5-9.5 h) [17]</td>
<td>ChiP-chip</td>
<td>1079</td>
<td>Drosophila embryo (5-9.5 h) [17]</td>
<td></td>
<td>244</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Drosophila embryo (0-12 h and 8-12 h) [19]</td>
<td>ChiP-chip</td>
<td>889</td>
<td>Drosophila embryo (5-9.5 h) [18]</td>
<td></td>
<td>213</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Drosophila haltere imaginal disc (L3) [20-22]</td>
<td>ChiP-chip</td>
<td>519, 4590, 1147</td>
<td>Genetically modified haltere and wing (Vg, Cbx) [23]</td>
<td></td>
<td>202-724</td>
<td>GAF, MAD, PC, TRX, AEF1, TOP2 [20]; LEF1, STAT1, E2F2, MEF2, IRX2, GRAYHEAD, GATA, SOX7, TBP [24]</td>
</tr>
<tr>
<td></td>
<td>Expression in wing and genetically modified haltere [26]</td>
<td>ChiP-chip</td>
<td>94405</td>
<td>–</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAB-5</td>
<td>C. elegans [27]</td>
<td>ChiP-chip</td>
<td>4493</td>
<td>–</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EGL-5</td>
<td>C. elegans [27]</td>
<td>ChiP-Seq</td>
<td>4688</td>
<td>–</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LIN-39</td>
<td>C. elegans [27]</td>
<td>ChiP-Seq</td>
<td>7493</td>
<td>–</td>
<td>–</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

II BA, second branchial arches; ESC, embryonic stem cell; HSC, haematopoietic stem cell; HPC, haematopoietic progenitor cell.

*TALE co-factors excluded.

†This number corresponds to genes similarly upregulated or downregulated by both HOXA9 and HOXA10.

‡This number corresponds to genes co-regulated by Hoxa9 and C/EBPα.

screensings that identify Hox-interacting transcription factors (Bondos et al., 2006; Lambert et al., 2012).

Understanding the function of Hox transcription factors also requires insight into whether they modulate and respond to various features of chromatin. Within the context of the quantitative continua model, the finding that overlapping genomic binding correlates with chromatin accessibility (Li et al., 2011) suggests that such accessibility might explain the genomic distribution of most transcription factors. In addition, predictions of transcription factor genomic binding sites by computational modelling (based on transcription factor nuclear concentration and binding preferences) fit better with actual in vivo genomic distributions when chromatin accessibility is taken into account (Kaplan et al., 2011). This view also accommodates the recent discovery of highly occupied target (HOT) regions (Kvon et al., 2012). How the pattern of chromatin accessibility is defined remains to be understood. In one case, that of the regulation of Dll by the abdominal Hox proteins Ubx and AbdA, it was shown that Hox proteins modulate chromatin conformation; relaxation of a compacted configuration was found in cells devoid of Ubx and AbdA, which are likely to act by controlling the association of GAGA-associated factor (GAF; also known as Trl) and histone H2A\(\text{V}\)-containing nucleosomes (Agelopoulos et al., 2012). This suggests a possible role for Hox proteins in defining the chromatin features crucial for gene regulation.

The connection of Hox proteins with chromatin regulators is also supported by a number of other findings. For example, binding motifs for GAF, Trithorax (Trx) and Polycomb group (PcG) proteins are overrepresented in Ubx-precipitated genomic fragments (Agrawal et al., 2011), and that of the CBP/p300 (Crebbp/Ep300) histone acetyltransferase (HAT) is found in Hoxa9 ChipPed fragments (Huang et al., 2012). In addition, a Hoxb1-Pbx-Meis complex plays a role in controlling hoxb1a transcription in zebrafish, which was shown to rely on defining the level of histone H4 acetylation by controlling the recruitment of histone deacetylase (HDAC)/CBP at the hoxb1a promoter (Choe et al., 2009). Finally, the repressive action that Pbx1 exerts on the osteoblastogenic action of Hoxa10 was shown to involve Runx2 and CBP/p300 recruitment and a decrease in histone H3K9 methylation (Gordon et al., 2011).

Together, the available data indicate that the interplay between Hox proteins and chromatin regulators and features is key to Hox-mediated control of gene regulation. Understanding how diverse this interplay is, how specific it may be for each Hox protein, and how it contributes to generate specificity in Hox transcriptional responses remains a major challenge for the future.

It is also essential to uncover how Hox proteins interface with the general transcription machinery. Early genetic data identifying mutations in RNA polymerase (Pol) II subunits that phenocopy Ubx haploinsufficiency (Mortin et al., 1992), together with the in vitro reconstruction of Ubx transcriptional potential based on S2 or Kc nuclear extracts (Johnson and Krasnow, 1990, 1992), suggested that Ubx might directly interact with general components of the transcription machinery, possibly by regulating assembly of the pre-initiation complex. More recently, additional links have emerged. The HX motif of Antp was shown to establish functional contacts with Bip2 (TAF3), a TATA-binding protein-associated factor (Prince et al., 2008). The Mediator complex, which is a major component of the RNA Pol II machinery, has also been linked to Hox proteins, with Med13 and Med19 subunits being required for proper Hox gene function (Boube et al., 2000, 2014). Direct physical interaction was further identified for Med19, which appears dedicated to gene activation (Boube et al., 2014). The precise contribution of these interactions to transcriptional control by Hox proteins remains to be elucidated.

Mechanistic insight into how Hox proteins control transcription was recently gained by further exploring hoxb1a regulation in the zebrafish embryo (Choe et al., 2014). Studying the dynamics of Hoxb1, Pbx and Meis proteins at the hoxb1a promoter, it was found that Pbx/Meis association precedes that of Hoxb1 but is insufficient for transcriptional activation, even though RNA Pol II is already recruited. The subsequent recruitment of Hoxb1 releases poised RNA Pol II, allowing efficient transcription. This role in controlling RNA Pol II pausing contrasts with the recognised role for Hox proteins in pre-initiation complex assembly, suggesting that Hox proteins may control several aspects of transcription.

Thus, although we are still in the early stages, functional data allow for a larger framing of Hox protein function. A survey of Hox interactomic data has identified a large number of often as yet unexplored sequence-specific transcription factors, chromatin regulators and components of the general transcription machinery (Fig. 6 and Table 2). This extends the perspectives provided by functional studies and provides ground for more broadly exploring the interplay between Hox proteins and additional transcription factors, as well as their interface with chromatin regulators and the general transcription machinery.
In some instances, Hox paralogues display mixed behaviour. This is evident in the shared and specific functions of Hox PG9 paralogues during spinal motoneuron specification (Jung et al., 2014). Spinal cord motoneurones are organised into columns and pools, and their specification depends on the activity of Hox genes belonging to PGs 6, 8 and 9, which show mutually repressive activities (Datsen & Jessell, 2009). Hoxc9, in particular, is crucial for specifying motoneurones targeting the thoracic region, and it displays a global repressive activity toward anterior Hox genes (Jung et al., 2010). Each mouse Hox PG9 protein appears to be able to repress Hox genes specifying brachial, upper limb-innervating motoneurones. However, only Hoxa9 and Hoxc9 are able to repress this brachial motoneuron fate properly and promote motoneuron columns involved in thoracic innervation. This means that Hoxa9 and Hoxc9 are functionally equivalent in promoting thoracic motoneuron specification, and also that the shared ability of PG9 Hox proteins to repress ‘brachial’ Hox genes is not sufficient in this respect, such that Hoxb9/d9 are not functionally interchangeable with Hoxa9/c9 in governing motoneuron specification (Jung et al., 2014).

Hoxa2 and Hoxb2 provide an additional example of mixed behaviour. They share the ability to bind and regulate certain target genes and enhancers in the hindbrain (Alexander et al., 2009; Lampe et al., 2008; Matis et al., 2007), but they display opposite activities in regulating oligodendrocyte precursor cells in this territory. Indeed, Hoxa2 has been shown to inhibit whereas Hoxb2 promotes oligodendrogenesis at its early steps. The lack of either Hoxa2 or Hoxb2 therefore results in opposite phenotypes in terms of oligodendrocyte patterning (Miguez et al., 2012). However, to what extent these opposing activities rely on intrinsic and specific properties of Hoxa2 and Hoxb2, or whether it is the rhombomere-specific cell context combined with Hoxa2 versus Hoxb2 expression levels that actually contributes to the differential effect exerted by these proteins on oligodendrogenesis, needs to be investigated further.

The concomitant neo- or subfunctionalisation and functional equivalence shared by PG2 or PG9 genes suggests that the variation in sequence within a PG might be at the origin of protein neo- and subfunctionalisation and that, depending on the context, the same protein may exhibit functional differences. This contextual activity is reminiscent of the latent specificity model, which was proposed to account for the binding selectivity elicited through interaction with Exd in an in vitro Selex-Seq setting (Slattery et al., 2011b). This model proposes that, whereas Hox proteins acting alone have little specificity, interaction with Exd unveils latent intrinsic Hox specificity. In the case of paralogous Hox protein functions, it then suggests that divergence alone between paralogous Hox proteins is insufficient to make them functionally different, but that association with other proteins might unveil latent functional differences that then result in neo- or subfunctionalisation.

Conclusions

Examining Hox protein function at both a cellular and molecular scale allows us to draw interesting correlations between the molecular modalities of Hox proteins and their functional properties during development. This certainly suggests that we are on the correct track to decipher principles and mechanisms beyond the specific and diverse functions that Hox proteins perform. It is also clear that large-scale genomic, transcriptomic and proteomic/interactomic approaches can provide further insights into Hox protein functions and modes of action. However, our current genomic and transcriptomic view is restricted to just a few Hox proteins, and the available data can only rarely be reliably integrated.
<table>
<thead>
<tr>
<th>Transcription factors</th>
<th>Chromatin regulators</th>
<th>General transcription factors</th>
<th>Other interactors</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HOX1</strong> FHL5, GRN, HMG1, HOXA1, HOXB1, HOXC9, HOXD3, HSFY1, IKG2, LDB1, LFXN, MDF1, MEIS1, NABP2L2, NR3C1, NAR4A, OGT, PAX6, PAX9, PBX1, PITX2, PKNOX1, PKNOX2, PLSCR, PLSCR2, RBC1, TCF3, TRIP6, ZBTB16, ZBTB32, ZNF390</td>
<td>KDM1A, KDM5, PRMT6, SUV39H1, CREBBP, EP300</td>
<td>ABHD17A, ADAMTSL4, ADCK4, AGPAT1, BLCL2, C10orf94, CCDC33, CERS2, CHIC2, COX1, CRACR2A, DKK1, DUSP22, EFEMP2, FAM154A, G9P, GPRIN2, HSDB37, KRT81, KRTAP26-1, KRTAP3-2, KRTAP3-3, KRTAP4-12, KRTAP5-9, LGAL13, LIM5, LN2X, MGAT5B, PCSK5, PDCD6IP, PDLIM7, PFKM, PIK3R1, PRNP, PRRC2B, RAB33A, RBPM5, RGS17, RGS20, SERPIN3, SIRT1, SMOC1, SPRY1, SPRY2, TRAF1, TRAF2, TRAPP3A, TRIM23</td>
<td></td>
</tr>
<tr>
<td><strong>HOX2</strong> CHD4, LFI205B, MEIS1, RBM39, TLE6</td>
<td>BRCA1, CREBBP, EP300</td>
<td>CIITA</td>
<td>ASB12, CERS2, ERCC3, MND1, PSMA3, PSMB2, RAI14, RCHY1, SOD1, YBX3, ZDHHC16, ZDHHC6</td>
</tr>
<tr>
<td><strong>HOX3</strong> ALX4, E2F4, HMG1, HOXA1, ME2O2, SOX10, TLE6</td>
<td>CREBBP, EP300</td>
<td>CERS2, CSDE1, DDB1, POT1, PWP1, SEC23B, TERF1, TIN2F, ZDHHC17, YWHAE</td>
<td></td>
</tr>
<tr>
<td><strong>HOX4</strong> ALX4, ETV4, HOXB13, MEIS1, NF2, PBX1, PBX3, POU2F1, SAL4, SNA1, SOX10, SOX8, ZBTB32</td>
<td>CREBBP, EP300, PRMT5, WHSC1L1, ZMYND11</td>
<td>ASB48, CNOT7, CUL4A, DDB1, ELMD2, HIPK1, HNRNPAB, HTT, LYN, PRKD2, RBX1, RNF32, RNPS1, RXR, TIGD4, XRC5, XRC6, XRC6, ZSCAN18</td>
<td></td>
</tr>
<tr>
<td><strong>HOX5</strong> CDX4, DlT13, FOXA2, FOXO1, MEIS1, MYC, PBX1, PBX3, SMAD1, SOX2, TWIST1, ZNF408</td>
<td>APP, CREBBP, EP300</td>
<td>CSN2K2B, KRT15, SAT1, UBC</td>
<td></td>
</tr>
<tr>
<td><strong>HOX6</strong> ALX4, HMG1, HOXB7, NRF1, PBX1, PBX3, PKNOX2, PLIN1C, POU2F1, TOF2, TSC22D3</td>
<td>APP, CREBBP, EP300, PAR1P, GMNN</td>
<td>CSN2K2A1, IRAK3, PEX5, PRKDC, UBC</td>
<td></td>
</tr>
<tr>
<td><strong>HOX7</strong> HOXB6, JUNB, MEIS1, NFKBIA, NPM1, PBX1, PKNOX</td>
<td>APP, CREBBP, EP300, PAR1P, GMNN, KMT2A, RCC1</td>
<td>CSN1K2A1, IRAK3, PEX5, PRKDC, UBC</td>
<td></td>
</tr>
<tr>
<td><strong>HOX8</strong> DLX2, DLX5, EYA3, HMG1, HOMEZ, IRF7, IRF9, JUN, MEIS1, MSX1, NR3C1, PBX1, PBX2, PBX3, POU2F1, SMAD1, SMAD4, SMAD6, STAT5B, TSC22D3, ZFP30</td>
<td>APP, CREBBP, EP300, PAR1P, GMNN</td>
<td>CSN2K2A1, IRAK3, PEX5, PRKDC, UBC</td>
<td></td>
</tr>
<tr>
<td><strong>HOX9</strong> BARX1, ETV5, HMG1B, HOPX, HOXC13, HOXD9, IFO4, IRF9, JUN, MDF1, MEIS1, MEIS2, MYBBP1A, NFKBIA, NKP2-1, NMI, PAX6, PBX1, PBX2, PBX3, PHTF1, PKNOX1, PLIN1C, POU2F1, SMAD2, TBX2, TBX5, SPZ1, TAL1, TRIP6, ZFP329, ZFP364, ZNF408</td>
<td>APP, CREBBP, EP300, PAR1P, GMNN</td>
<td>CSN2K2A1, IRAK3, PEX5, PRKDC, UBC</td>
<td></td>
</tr>
<tr>
<td><strong>HOX10</strong> ALX4, EMX1, FOXO1, HMG1, HOXA10, MEIS1, MYC, PBX1, PBX2, PBX3, SPI1, ZFP292</td>
<td>APP, CREBBP, EP300, PAR1P, GMNN, HDAC12, ING4, KMT2A, PGC2F2, PHF17, PRMT5, SUZ12</td>
<td>CSN2K2A1, IRAK3, PEX5, PRKDC, UBC</td>
<td></td>
</tr>
<tr>
<td><strong>HOX11</strong> ALX4, EMX1, FOXO1, HMG1, HOXA10, MEIS1, MYC, PBX1, PBX2, PBX3, SPI1, ZFP292</td>
<td>APP, CREBBP, EP300, PAR1P, GMNN, HDAC12, ING4, KMT2A, PGC2F2, PHF17, PRMT5, SUZ12</td>
<td>CSN2K2A1, IRAK3, PEX5, PRKDC, UBC</td>
<td></td>
</tr>
<tr>
<td><strong>HOX12</strong> ALX4, JUN, MAF, MAFB, MAFG, MAFG, MAFK, MEIS1, MEIS2, MEIS2, MEIS3, NR2E3, OTX2, POU2F1, RHOXF2, SMAD1, SMAD2, SMAD5, SOX5, ZNF390</td>
<td>APP, CREBBP, EP300, PAR1P, GMNN, HDAC12, ING4, KMT2A, PGC2F2, PHF17, PRMT5, SUZ12</td>
<td>CSN2K2A1, IRAK3, PEX5, PRKDC, UBC</td>
<td></td>
</tr>
<tr>
<td><strong>HOX13</strong> ALX4, DLX1, DLX5, ELF1, ELK1, HAND2, HOXC9, HOXD4, IRF4, MEIS1, MEIS2, MEIS3, NR2E3, OTX2, POU2F1, RHOXF2, SMAD1, SMAD2, SMAD5, SOX5, ZNF390</td>
<td>APP, CREBBP, EP300, PAR1P, GMNN, HDAC12, ING4, KMT2A, PGC2F2, PHF17, PRMT5, SUZ12</td>
<td>CSN2K2A1, IRAK3, PEX5, PRKDC, UBC</td>
<td></td>
</tr>
</tbody>
</table>

Continued
into a biologically relevant picture, as data most often concern different cellular material and developmental stages. Expanding the realm of large-scale genomic and transcriptomic data is thus essential, and will be of maximum benefit if generated in a manner that will optimise cross-comparisons.

Besides specific functions, evidence is now also emerging that proteins from distinct PGs can perform similar or even identical functions. This is illustrated in *Drosophila* by the inhibition of autophagy that is driven similarly by most *Drosophila* Hox proteins (Banreti et al., 2014) and by the repression of the limb-promoting gene *Dll* by the Ubx, AbdA and posterior AbdB central proteins (Sambrani et al., 2013). Similar situations exist in vertebrates too: the functional equivalence of Hox5-8, which converge in specifying brachial motoneurons in the mouse spinal cord (Lacombe et al., 2013); the promotion of haematopoietic stem and progenitor cell expansion and inhibition of differentiation by Hoxb3, Hoxb4/B4 and Hoxb6 (reviewed by Alharbi et al., 2013); and the control of embryo implantation in mammals by HOXA9-11 and HOXD10 (Lu et al., 2008; Xu et al., 2014). It is tempting to speculate that such generic Hox functions might result from non-discriminative modes of action, consistent with widespread Hox genomic DNA binding, possibly without or with limited help from assisting protein partners. These generic functions might alternatively be seen as the result of functional convergence arising from distinct molecular modalities. Although less intuitive, this is supported by cooperative versus antagonistic partnership with Exd for Ubx- and AbdB-mediated repression of *Dll* (Sambrani et al., 2013). Investigating more broadly and understanding the molecular modalities of such generic Hox functions is an important objective and should uncover novel aspects of Hox protein modes of action.

There is also much to be learned from considering Hox protein function within a larger framework, including a wider partnering potential with other sequence-specific transcription factors, the general transcription machinery and chromatin regulators. It will also be crucial to investigate possible links between Hox proteins and the nuclear architecture, which is known to impact transcriptional regulation (Schneider and Grosschedl, 2007). Proteomic/interactomic data are available for a set of Hox proteins, but candidate Hox partnerships can rarely be traced back to specific cellular and developmental contexts. The full potential of these data will only be grasped once the functional relevance of protein-protein physical associations have been established. This will also clarify how such partnerships contribute to functional versatility, and how versatility in Hox protein function relates to cellular contexts. Integrating such large-scale data sets will be essential for determining whether the prevailing concepts and mechanisms are sufficient to explain the many facets of Hox protein function.

**Acknowledgements**

We apologise to the authors of primary research work that could not be cited owing to space limitation.

---

**Table 2. Continued**

<table>
<thead>
<tr>
<th>Transcription factors</th>
<th>Chromatin regulators</th>
<th>General transcription factors</th>
<th>Other interactors</th>
</tr>
</thead>
<tbody>
<tr>
<td>DFD</td>
<td>ASE, BCD</td>
<td></td>
<td>ALD, CDC2C, CDK4, CDK5, CG13996, CG15140, CG1812, CG3748, CG4078, CG8128, CG9927, CKS85A, CUL-3, CYCG, KP78B, MAD2, ORB2, ORC4, PGAP5, PNG, POLO, RECQ4, RUX, TRIP1</td>
</tr>
<tr>
<td>ANTP</td>
<td>AL, ALY, APT, ARM, CYCK, H, NOC, SMOX, ZF30C</td>
<td>DMI-2, DIP1</td>
<td>14-3-3-α, BRD, CBP80, CG11164, CG13474, CG6455, CYCK, DSH, EF1γ, EF2, FZO, GLYP, HSC70-4, MRPL44, MS(3)76CC, NMO, NRT, OTU, P120CTN, PK17E, RAD23, REF1, RPL22, RPN6, RPS13, TERM, TRN, YURI, ZN72D</td>
</tr>
<tr>
<td>UBX</td>
<td>AL, ALY, APT, ARM, CYCK, H, NOC, SMOX, ZF30C</td>
<td>DMI-2, DIP1</td>
<td>CYCG, CD5, CG13728, CG1486, CSN6, EF2B, GINT3, GKT, PK17E, PXB, RPS12, DHS-14, F53F1.4, K10D6.4, O44719_CAEEL, Q21423_CAEEL, Q22262_CAEEL, Q9XM6_CAEEL, RAD26</td>
</tr>
<tr>
<td>ABDA</td>
<td>CYCG</td>
<td></td>
<td>C34B2.4, DSH-2, FGSG4U7_CAEEL, O01527_CAEEL, O44949_CAEEL, YQJ3_CAEEL</td>
</tr>
<tr>
<td>ABDB</td>
<td>CYCG</td>
<td></td>
<td>CELE_F47G4.4, CCCH-2, GCN-2, GIPC-1, GST-7, LGC-37, PEPT-3, PIK-1, PSR-1, Q965P0_CAEEL, RPS-10, VAV-1</td>
</tr>
<tr>
<td>LIN-39</td>
<td>PAX-3, CEH-27, CEH-51, DMD-8, EFL2, FKH-6, HLH-26, KLF-2, MED-2, NHR-67, POP-1, SPE-44, TAB-1, TBX-9, ZIP-10, ZLF-16, ZLF-2</td>
<td>TFIIEB</td>
<td>CELE_F47G4.4, CCCH-2, GCN-2, GIPC-1, GST-7, LGC-37, PEPT-3, PIK-1, PSR-1, Q965P0_CAEEL, RPS-10, VAV-1</td>
</tr>
<tr>
<td>MAB-5</td>
<td>CEH-23, ELF-1, EYA-1, ZLF-16</td>
<td>TFIIEB</td>
<td>CELE_F47G4.4, CCCH-2, GCN-2, GIPC-1, GST-7, LGC-37, PEPT-3, PIK-1, PSR-1, Q965P0_CAEEL, RPS-10, VAV-1</td>
</tr>
<tr>
<td>EML-5</td>
<td>EML-5, POP-1</td>
<td></td>
<td>F57G12.2, ACT-5, ICD-1, RPL-17, RPL-18</td>
</tr>
<tr>
<td>PHP-3</td>
<td>CEH-51, CRT-1, EFL-2, FKH-2, FKH-6, KLF-2, LIN-48, MDT-17, NHR-47, SPE-44, UNC-86</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NOB-1</td>
<td>ASH-1, BCD</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Data on Hox protein-protein interactions with transcription factors, general transcription factors, chromatin regulators and other interactors were collected from protein interaction databases as described by Rezsohazy (2014).
Competing interests
The authors declare no competing or financial interests.

Funding
Work in the Y.G. laboratory is supported by the Centre National de la Recherche Scientifique (CNRS), Aix Marseille Université (AMU), Agence Nationale de la Recherche (ANR), Centre Franco-Indien pour la Promotion de la Recherche Avancée (CEFIPRA), ARC and the A*MIDEX project [ANR-11-IDEX-0001-02]. Work in the R.R. laboratory is supported by Fonds National de la Recherche Scientifique (FRNS, FRIA grants), an Action de Recherche Concertée [ARC grant 12/17-041], Fonds Spéciaux de Recherche (FSR, Université Catholique de Louvain) and the Belgian Foundation Against Cancer.

References


