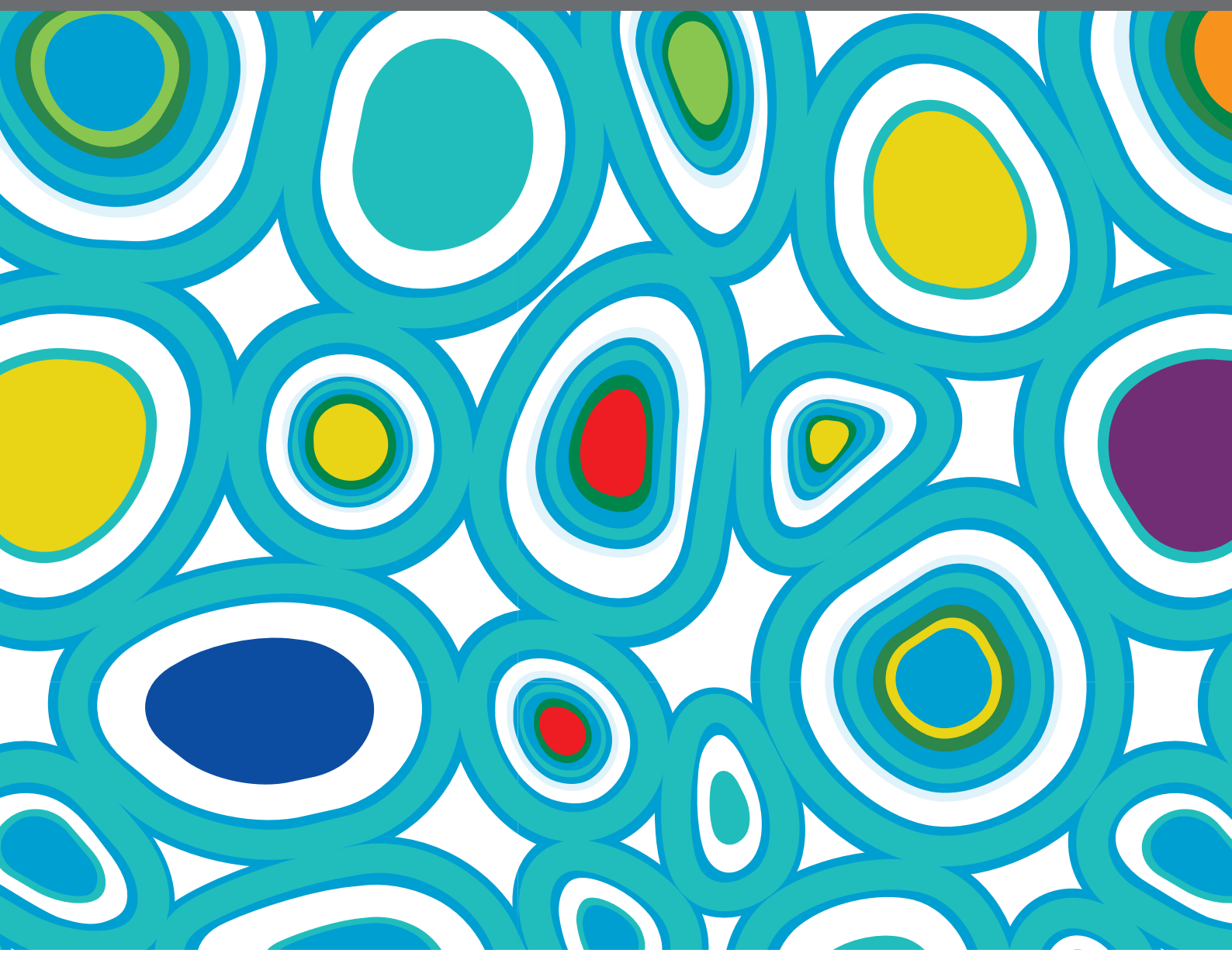
MECHANISMS OF HOX-DRIVEN PATTERNING AND MORPHOGENESIS

EDITED BY: Edwina McGlinn, Ernesto Sánchez-Herrero and Marie Kmita
PUBLISHED IN: Frontiers in Cell and Developmental Biology



frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-83250-215-0
DOI 10.3389/978-2-83250-215-0

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

MECHANISMS OF HOX-DRIVEN PATTERNING AND MORPHOGENESIS

Topic Editors:
Edwina McGlinn, Monash University, Australia
Ernesto Sánchez-Herrero, Centre for Molecular Biology Severo Ochoa, Spanish National Research Council (CSIC), Spain
Marie Kmita, Montreal Clinical Research Institute (IRCM), Canada

Citation: McGlinn, E., Sánchez-Herrero, E., Kmita, M., eds. (2022). Mechanisms of Hox-Driven Patterning and Morphogenesis. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-83250-215-0

Table of Contents

frontiers | Frontiers in Cell and Developmental Biology

TYPE Editorial
PUBLISHED 01 September 2022
DOI 10.3389/fcell.2022.992341

OPEN ACCESS

EDITED AND REVIEWED BY
Moises Mallo,
Gulbenkian Institute of Science (IGC), Portugal

*CORRESPONDENCE
Ernesto Sánchez-Herrero,
esherrero@cbm.csic.es

SPECIALTY SECTION
This article was submitted to Morphogenesis and Patterning, a section of the journal Frontiers in Cell and Developmental Biology

RECEIVED 12 July 2022
ACCEPTED 26 July 2022
PUBLISHED 01 September 2022

CITATION
Kmita M, McGlinn E and Sánchez-Herrero E (2022), Editorial: Mechanisms of hox-driven patterning and morphogenesis.
Front. Cell Dev. Biol. 10:992341.
doi: 10.3389/fcell.2022.992341

Editorial: Mechanisms of hox-driven patterning and morphogenesis

Marie Kmita[1,2,3], Edwina McGlinn[4] and Ernesto Sánchez-Herrero[5]*

[1]Département de Médecine, Faculté de Médecine, Université de Montréal, Montreal, QC, Canada, [2]Genetics and Development Research Unit, Institut de Recherches Cliniques de Montréal, Montreal, QC, Canada, [3]Department of Experimental Medicine, McGill University, Montreal, QC, Canada, [4]Australian Regenerative Medicine Institute, Monash University, Clayton, VI, Australia, [5]Centre for Molecular Biology Severo Ochoa, Spanish National Research Council (CSIC), Madrid, Spain

KEYWORDS

hox genes, drosophila, mouse, zebrafish, patterning, morphogenesis, evolution

Editorial on the Research Topic
Mechanisms of hox-driven patterning and morphogenesis

The Hox genes have attracted the interest of scientist for decades. Their organization in genetic clusters, their ordered chromosomal alignment and the correlation of this disposition with the evolutionary conserved gene expression along the anteroposterior axis have excited the curiosity of researchers and prompted countless investigations. Hox proteins have also been studied as examples of transcription factors that activate particular genes at certain positions. The present Research Topic Mechanisms of Hox-Driven Patterning and Morphogenesis gathers a series of reviews and original articles on Hox expression and function in different model organisms.

The seminal work of Ed Lewis characterized transformations in the *Drosophila* cuticle due to mutations in Bithorax Complex (BX-C) genes. Subsequent studies revealed that Hox genes are expressed and required also in muscles or nervous system. Two reviews focus on the role of Hox genes in these tissues. The manuscript by Poliacikova et al. reviews the role of Hox genes in the different steps of muscle specification in *Drosophila* and vertebrates. The authors describe Hox activity in somatic/skeletal, cardiac and visceral muscle development, explain the cooperation of Hox proteins with muscle-specific proteins and explain the role of these genes in neuromuscular circuits. These neuromuscular networks are detailed in the work of Joshi et al. They report muscle-motoneuron circuits in *Drosophila* and how these impact on development and behavior. The authors first summarize Hox role in specification of the central nervous system and then review recent advances in Hox determination of motoneuron morphology and physiology, focusing in two larval behaviors, locomotion and feeding, and two adult ones, egg-laying and self-righting.

Although it is well established that Hox genes are required during development, recent studies have uncovered a Hox function in adults, as described in the above review.

The original article by Li et al. demonstrates a role for Hox5 in adults. In a conditional triple mutant Hoxa5/Hoxb5/Hoxc5, in which Hox5 function is eliminated in the mesenchyme of adult mice, the investigators observed elimination of the elastin matrix, change in alveolar structures and an emphysema-like phenotype.

Two other research articles also unveil Hox5 function. In one of them (Mitchel et al.,) the activity of Hoxa5 in the development of the mouse sternum is analyzed. The authors first characterize in detail the origin and development of this structure. Then, they show Hoxa5 expression and requirement in the presternum, demonstrate the coordinated activity of other Hox genes in the formation of this structure and discuss the evolutionary implications of such function. In the second original article, by Howard et al., it is demonstrated that high levels of Hoxb5b in a restricted time window expand zebrafish neural crest cell number, increase the expression of the vagal neural crest cell markers *foxd3* and *phox2bb*, and extend the number of enteric neural progenitors; however, these progenitors do not expand later on to make enteric neurons along the gut. To explain the early expansion of neural crest cells but the later reduction of neurons, the authors argue that it can be due to the dynamic expression of Hox cofactors of the TALE family or to insufficiently timed signals needed for the continuous expansion and differentiation.

The most famous Hox mutation transforms halteres into wings in the fruitfly. Giraud et al. investigate the genetic basis of robustness and variation in haltere morphology, which is governed by the Hox gene *Ultrabithorax*. To this aim they carry out a genetic screen looking for haltere morphology changes in a wildtype or sensitized mutant background. Based on their results the authors postulate a self-sufficient mechanism whereby high levels of *Ultrabithorax* allow proper development of the haltere without the need for cofactors, thus ensuring its developmental stability.

The relationship between Hox gene function and evolution, originally proposed by Ed Lewis, is addressed in two reviews. In the first one, by Hombría et al., the authors describe the process of cephalization in vertebrates and arthropods and convincingly argue that the appearance of Hox cephalic genes predates the evolution of head structures in both groups of animals. As the formation of the cephalic region progressed during evolution, it incorporated trunk segments, and the Hox genes expressed in these segments acquired the control of specific genes to define cephalic structures. The second review (Krumlauf and Singh) deals with the role of Hox gene duplication and divergence in the development of morphology and in the emergence of new features in vertebrates. The authors discuss Hox gene duplication in evolution, showing examples of conservation and divergence of gene function, and explaining the role of the different domains of Hox proteins in the acquisition of specificity. They conclude that new specificity of Hox function can be achieved with just a few aminoacid changes in conserved regions and through interactions with proteins of the PBC class.

Cain and Gebelein discuss how the activity of different Hox proteins can determine the development of distinct structures by specific DNA binding and activation of particular genes. They summarize recent genomic and interactome data revealing how Hox proteins differ in the way they can bind closed chromatin, showing that some of them need PBC cofactors as pioneer factors. Then, they explain how the interaction of Hox proteins with cofactors and collaborators impinges on the way Hox proteins activate or repress gene expression, illustrating this differential Hox activity with the *Drosophila* Abdominal-A protein.

Finally, the comprehensive review by Hajirnis and Mishra discuss Hox organization and function. They first describe dispersion and clustering of Hox genes in different species, and then review the ordered disposition of cis-regulatory modules in the BX-C, and the opening of the BX-C chromatin, with emphasis on the organization of the *Abdominal-B* gene. They also review Hox function away from their traditional role of specifying particular structures and finally stress the importance of controlling Hox levels of expression by presenting examples of how increasing Hox protein levels can lead to cancer.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

frontiers in Cell and Developmental Biology

BRIEF RESEARCH REPORT
published: 23 July 2021
doi: 10.3389/fcell.2021.713282

Developmental Robustness: The Haltere Case in *Drosophila*

Guillaume Giraud[1†], Rachel Paul[1†], Marilyne Duffraisse[1], Soumen Khan[2], L. S. Shashidhara[2,3] and Samir Merabet[1]*

[1] IGFL, ENS Lyon, UMR 5242, Lyon, France, [2] Indian Institute of Science Education and Research (IISER), Pune, India, [3] Ashoka University, Sonipat, India

Developmental processes have to be robust but also flexible enough to respond to genetic and environmental variations. Different mechanisms have been described to explain the apparent antagonistic nature of developmental robustness and plasticity. Here, we present a "self-sufficient" molecular model to explain the development of a particular flight organ that is under the control of the Hox gene *Ultrabithorax* (*Ubx*) in the fruit fly *Drosophila melanogaster*. Our model is based on a candidate RNAi screen and additional genetic analyses that all converge to an autonomous and cofactor-independent mode of action for Ubx. We postulate that this self-sufficient molecular mechanism is possible due to an unusually high expression level of the Hox protein. We propose that high dosage could constitute a so far poorly investigated molecular strategy for allowing Hox proteins to both innovate and stabilize new forms during evolution.

Keywords: Hox, transcription, evolution, flight appendage, insects

OPEN ACCESS

Edited by:
Ernesto Sánchez-Herrero, Severo Ochoa Molecular Biology Center (CSIC-UAM), Spain

Reviewed by:
James C.-G. Hombría, Andalusian Center for Development Biology (CABD), Spain
David CRIBBS, Université Toulouse III Paul Sabatier, France

***Correspondence:**
Samir Merabet
samir.merabet@ens-lyon.fr

[†]These authors share first authorship

Specialty section:
This article was submitted to Morphogenesis and Patterning, a section of the journal Frontiers in Cell and Developmental Biology

Received: *22 May 2021*
Accepted: *30 June 2021*
Published: *23 July 2021*

Citation:
Giraud G, Paul R, Duffraisse M, Khan S, Shashidhara LS and Merabet S (2021) Developmental Robustness: The Haltere Case in Drosophila. Front. Cell Dev. Biol. 9:713282. doi: 10.3389/fcell.2021.713282

INTRODUCTION

Hox genes are well-known master developmental regulators that have extensively been exploited for diversifying animal body forms during evolution (Pearson et al., 2005; Pick and Heffer, 2012). Numerous cases of morphological diversification have been described as resulting from subtle modulations of the Hox gene expression profile in invertebrates (see for example Averof and Akam, 1995; Stern, 1998; Kittelmann et al., 2018) and vertebrates (see for example Gomez and Pourquié, 2009; Di-Poï et al., 2010; Mallo, 2018). Morphological innovations can also result from changes in the Hox protein sequence itself, as described for abdominal leg repression in arthropods (Galant and Carroll, 2002; Ronshaugen et al., 2002; Pearson et al., 2005; Saadaoui et al., 2011). Thus, despite a fundamental role during embryonic development, which involves a certain degree of stability for the underlying developmental programs, Hox genes remain tolerant for genetic variations and the evolution of phenotypic traits. Here we propose to directly tackle this apparent paradox by focusing on the flight appendage formation in insects in general, and in the fruit fly *Drosophila melanogaster* in particular.

Insects display an astonishing level of morphological diversity, as exemplified in their flight appendages, which differ from one order to the other. Ancestral insects had two pairs of wings on their second (T2, forewing) and third (T3, hindwing) thoracic segments (Carroll et al., 1995), most often of identical or highly similar morphology, as observed in damselflies and dragonflies (Odonata order). Forewings and hindwings can also be of different shape, size and/or color, as observed in the bees (Hymenoptera order) or butterflies (Lepidoptera order). In addition, wings

can also be strongly diverged into a new organ, as found in coleopterans, which have developed a protective envelope called elytron in place of the T2 wing, or in dipterans, which have developed a tiny dumbbell-shaped organ called haltere in place of the T3 wing.

What about the role of Hox genes in the morphological diversifications of flight appendages in insects? Most of our current understanding stands from studies in *Drosophila melanogaster* and the beetle *Tribolium castaneum*. Pioneer genetic work in *Drosophila* established the critical role of a single Hox gene, *Ultrabithorax* (*Ubx*), for the repression of anterior wing and the formation of posterior haltere on the third thoracic segment (Morata and Garcia-Bellido, 1976; Lewis, 1978; Bender et al., 1983; Carroll et al., 1995). A similar scenario was observed in *Tribolium*, where *Ubx* was shown to act by repressing the anterior elytron fate for ensuring posterior wing formation in the T3 segment (Tomoyasu et al., 2005).

Another striking feature relates to the relative morphological plasticity of halteres during dipteran evolution despite their critical role of flight. The role of halteres for flight behavior is well established (Dickinson, 1999; Hall et al., 2015): a fly without halteres cannot fly, and these balancing organs produce anti-phase beats and the inertial forces to stabilize the flight. Halteres display a certain level of morphological diversity amongst the different dipteran orders. Not only the shape but also the size (often but not systematically in correlation with the size of the adult insect) of the distal bulb part can fluctuate to some extent. In addition, some variation is observed in the region that connects the distal part of the haltere to the body, in particular in the arrangement of sensory elements (Agrawal et al., 2017). This level of morphological variation underlines that the Ubx-dependent haltere developmental program is not refringent to changes during insect evolution.

The ability of Ubx to specify different flight appendages during insect evolution while also ensuring a robust developmental program presents us with one of the most fascinating yet unsolved paradoxes. Our work aims at tackling this controversial issue, taking the opportunity of this special issue to present a brief research report that supports, together with a previous published work (Paul et al., 2021), a speculative model based on the Hox dosage.

RESULTS

From Ubx Expression Level to the Design of an RNAi Screen for Ubx Modulators in the Haltere Disc

Ubx is expressed in the entire haltere imaginal disc, but with distinct levels depending on the region: it is highly expressed in the so-called pouch region, which will give rise to the distal bulb called capitellum (**Figure 1A**), and less strongly expressed in the proximal regions that will give rise to the hinge (composed of the pedicellum and scabellum parts) and metanotum in the adult (**Figure 1A**; White and Wilcox, 1985; Delker et al., 2019). Early (Irvine et al., 1993), and recent (Delker et al., 2019) work showed that a negative autoregulatory loop contributes to the stabilization of distinct Ubx expression levels along the proximal-distal axis within the haltere imaginal disc of *Drosophila melanogaster*. Considering the robust regulation of Ubx levels within the haltere disc, we asked whether this expression profile was conserved in other *Drosophila* species. We observed the same proximal-distal bias in the third instar larval haltere imaginal discs of *Drosophila virilis* and *Drosophila simulans*, suggesting that the strong expression level of Ubx in the haltere pouch is not trivial (**Figure 1A**). Accordingly, changes in the Ubx expression level in the pouch have also been shown to increase (upon lower expression levels) or decrease (upon higher expression levels) the size of the capitellum (Crickmore et al., 2009). Interestingly, removing 40% of *Ubx* in a particular genetic background (a heterozygous context for the *abxbxpbx* mutation: **Figure 1B** (Casares et al., 1996; Paul et al., 2021) led to a significant increase of the size of the capitellum (**Figure 1B'**) and the apparition of a few wing-like bristles (**Figure 1B"**). Altogether these observations highlight that high Ubx levels in the pouch allow buffering against changes in the Ubx dose for ensuring a robust development of the haltere bulb. This feature was considered in our attempt to identify additional players that could participate in the Ubx-dependent haltere specification program in *Drosophila melanogaster*.

Our approach relied on a functional candidate RNAi screen that was performed in both the wild type (**Figure 1C**) and heterologous *abxbxpbx* (**Figure 1C′**) contexts. Each candidate RNAi was specifically expressed in the pouch with the *MS1096 driver*, with the rationale that they could affect upstream regulators of *Ubx* and/or Ubx cofactors (**Figures 1C,C′**). In any case, the *abxbxpbx*/+ background was considered as a sensitized context that could allow revealing phenotypes potentially buffered (and therefore not revealed) by the high level of Ubx in the wild type background.

A Candidate RNAi Screen Reveals Ubx-Autonomous Activity for Haltere Specification

In contrast to the wing, the haltere has never been the object of dedicated genetic screens and a large number of genes is more generally annotated for wing and not haltere development in *Drosophila*.[1] The master regulatory Ubx protein is known to specify the haltere in part by acting at several hierarchical levels to inhibit the wing developmental program (Akam, 1998; Mohit et al., 2003; Crickmore and Mann, 2006; de Navas et al., 2006a; Pallavi et al., 2006; Hersh et al., 2007; Makhijani et al., 2007; Pavlopoulos and Akam, 2011). Importantly, as mentioned above, ectopic expression of *Ubx* in the wing primordium is sufficient to transform the wing into a haltere, underlining that the wing-transformed tissue contains the set of molecular players that allow haltere development upon the Hox regulatory impulse. We thus decided to perform a candidate RNAi screen by targeting genes described to be expressed in the wing. More particularly, we focused on transcription factors encoding genes that have already

[1] https://flybase.org/

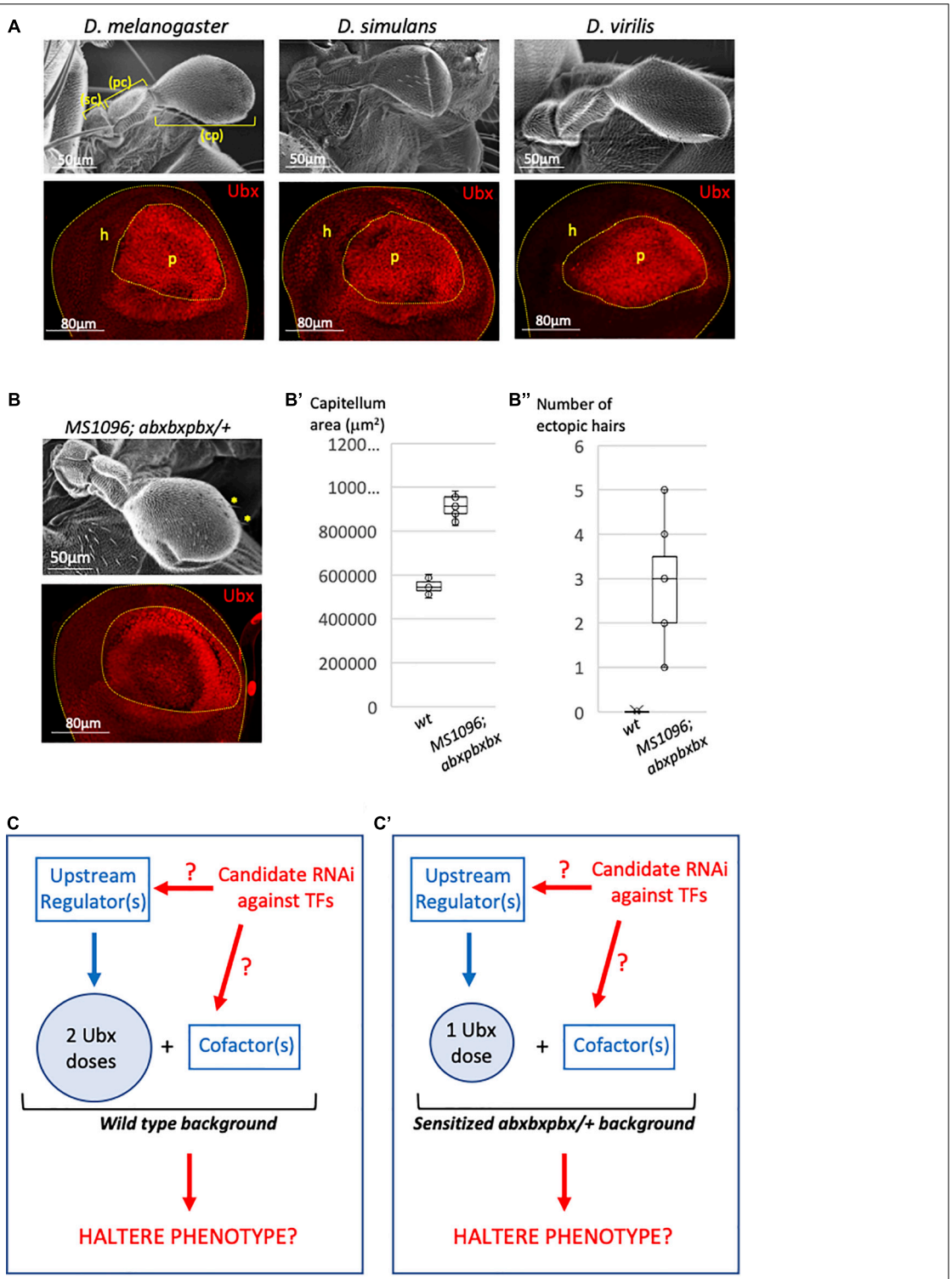


FIGURE 1 | Genetic model for the candidate Ubx regulators screening in the *Drosophila* haltere pouch. **(A)** Ubx (red) is expressed at high level in the pouch of the third instar larval haltere discs of *Drosophila melanogaster*, *Drosophila simulans,* and *Drosophila virilis*. The pouch (p) and hinge (h) regions will give rise to the capitellum (cp) and pedicellum (pc)+ scabellum (sc), respectively. Bottom images are illustrative confocal acquisition. Upper images are illustrative SEM acquisitions of adult halteres. **(B)** Haltere morphology (upper SEM picture) and Ubx expression profile (red, bottom confocal picture) in the haltere pouch in the *abxbxpbx* background context, as indicated. Yellow stars indicated ectopic wing-like bristles **(B')**. Boxplot representation of the quantification of the adult haltere size in the wild type (*wt*) and *abxbxpbx* backgrounds. **(B")** Boxplot representation of the quantification of the number of ectopic wing-like bristles in the wild type (*wt*) and *abxbxpbx* backgrounds. **(C–C')** Model for the candidate RNAi screen that could affect upstream *Ubx* regulators or Ubx cofactors in the *wt* **(C)** or sensitized **(C')** background. The RNAi screen is targeting transcription factors (TFs).

been tested in a RNAi screen for wing development (Schertel et al., 2015) and whose expression in the haltere imaginal disc was confirmed by RNA-seq (Khan et al., in preparation; see also section "Materials and Methods").

In total, we tested 117 genes in the wild type and *abxbxpbx*/+ context backgrounds (**Supplementary Table 1**). Two different phenotypes of the capitellum linked to a wing-like transformation, therefore to affected Ubx activity, were specifically analyzed: the size and the number of wing-like bristles. As expected, the control experiment with RNAi against *Ubx* led to a strong haltere-to-wing transformation phenotype while the haltere remained unchanged in the *MS1096*/+ background (**Figure 2A**). In contrast, RNAi against *Antennapedia* (*Antp*), which is expressed in a few cells of the hinge of the haltere disc (Paul et al., 2021), had no effects (**Supplementary Table 1**). Surprisingly, although the large majority of the 117 tested genes are known to play a role for wing development (Schertel et al., 2015), very few (7/117) affected the haltere capitellum in the wild type (*MS1096/ +*) background (**Figure 2B** and **Supplementary Table 2**). These genes correspond to different types of TF classes. Two genes, *engrailed* (*en*) and *cubitus interruptus* (*ci*) led to size increase and ectopic bristles (**Figure 2C** and **Supplementary Table 2**), highlighting a role for the Hedgehog (Hh) signaling pathway in the *Ubx* developmental program. Interestingly, the regulation of a direct Ubx target gene has been shown to rely on the integration of the Hox and Hh pathways in the haltere disc (Hersh and Carroll, 2005). Two other genes, *Polycomb* (*Pc*) and *Adh transcription factor 1* (*Adf-1*), led to size decrease with (*Pc*) or without (*Adf-1*) ectopic bristles (**Figure 2C** and **Supplementary Table 2**). Given the general role of Pc-G proteins as repressors of Hox gene expression (Kassis et al., 2017), we hypothesized that the effect observed with *Pc RNAi* could most likely result from increased *Ubx* expression in the haltere pouch. Finally, RNAi against *armadillo* (*arm*), *Distallless* (*Dll*), and *homothorax* (*hth*) led to ectopic bristles with no significant capitellum size defects (**Figure 2C** and **Supplementary Table 2**). Ectopic bristles were specifically observed in the pedicellum in the case of *hth* RNAi, which reflects the expression domain of *MS1096* outside the pouch in the imaginal disc (Paul et al., 2021). Previous work showed that Hth and Ubx are co-expressed in the hinge region and that Hth is downregulating *Ubx* to control its expression level outside the haltere pouch (Delker et al., 2019). The appearance of ectopic bristles on the pedicellum upon *hth* RNAi suggests that Hth could also act as a Ubx cofactor in this region. Accordingly, Ubx and Hth display a striking similarity in their genome-wide binding profiles in the haltere disc (Choo et al., 2011; Slattery et al., 2011) and the two proteins have been shown to interact *in vivo* (Bischof et al., 2018).

The other haltere phenotypes were observed in the *abxbxpbx*/+ background only. However, this sensitized background did not reveal a large number of positive genes (14/117 led to a phenotype: **Figure 2D** and **Supplementary Table 3**). Interestingly, 7 genes belong to the Trx group and one to Pc-G (**Supplementary Table 3**), highlighting that the *abxbxpbx* background is preferentially revealing upstream regulators of *Ubx*. The remaining 6 genes encode for different types of TFs (half of them coding for zinc-fingers containing TFs; **Supplementary Table 3**). None of these genes had a phenotype on both the size and bristles number upon RNAi, highlighting that their effects were moderate despite the sensitized genetic background.

In conclusion, the candidate RNAi screen revealed few potential cofactors of Ubx in the wild type (Arm, Ci, Dll, En, Hth) or sensitized (Apterous, Beadex, Deaf-1, Jumu, Mes-2, Zfh-2, Zf30c) background. This small number (12/117) was unexpected given the general tendency of Hox proteins to interact with many TFs in *Drosophila* (Baëza et al., 2015; Bischof et al., 2018).

A Minimal Form of Ubx Is Sufficient to Specify the Haltere Developmental Program in *Drosophila*

Results obtained from the candidate RNAi screen suggest that Ubx could trigger the haltere developmental program by interacting with an unexpectedly small number of transcriptional partners. In order to explore this molecular aspect further, we dissected the region(s) of the Ubx protein that could be necessary for its activity. The underlying hypothesis was to postulate that a large part of the protein sequence could potentially be dispensable because of a minimum number of interacting cofactors in the haltere disc. The role of the different regions in Ubx was assessed in the context of the genetic rescue of a mutant allelic combination where the haltere is transformed into a small wing in the adult fly (**Figure 3A**; Casares et al., 1996; Paul et al., 2021). These rescue assays are based on the allelic combination of *abxbxpbx* with an hypomorphic Gal4 insertion (allele Ubx^{LDN}; Casares et al., 1996; de Navas et al., 2006b) that allows to simultaneously express UAS constructs in this background (Brand and Perrimon, 1993). Here, rescue assays were performed with mutated and deleted forms that have previously been used to reveal an atypical nuclear export signal (NES) in Ubx (**Figure 3A**; Duffraisse et al., 2020).

As expected, the expression of wild type Ubx in the mutant allelic combination led to a complete rescue of the phenotype, with *de novo* formation of a normal haltere (**Figure 3B**). In contrast, expression of a HD-mutated form that cannot bind DNA (construct Ubx^{HD51}) led to a partial rescue, with the formation of a structure that displays both wing and haltere characteristics (a similar size to a small wing, with a mixture of wing-like and haltere-like hairs, together with the presence of a number of wing-like bristles on the margin: **Figure 3B**). This result underlines that the DNA-binding of Ubx is important for its correct activity in the haltere disc. In contrast, DNA-binding integrity was shown to be less critical for *Dll* repression in the epidermis (Sambrani et al., 2013). The first deleted form that we tested was truncated in the N- and C-terminal part (construct $Ubx^{dN130dC}$) and was not able to rescue the haltere-to-wing transformation phenotype (**Figure 3B**). This form has previously been shown to be constitutively exported, due to the absence of a NES inhibitory domain in the first 130 amino acids (Duffraisse et al., 2020). We confirmed that $Ubx^{dN130dC}$ was also constitutively exported in the haltere pouch, explaining that this deleted form was inactive in the rescue assay (**Figures 3C,C'**).

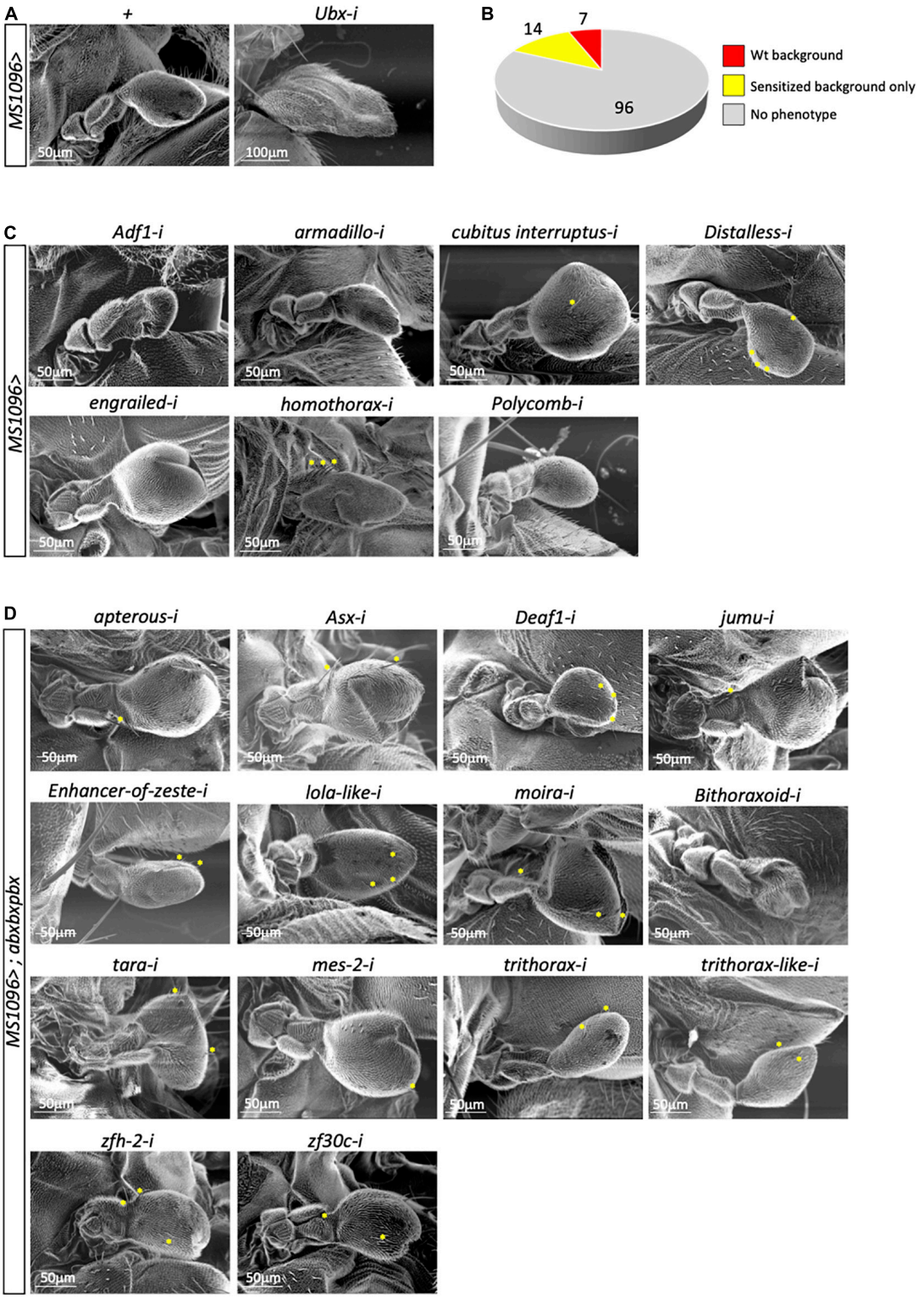


FIGURE 2 | A candidate RNAi screen for transcription factors revealed few potential Ubx cofactors in the haltere disc. **(A)** Control haltere phenotypes in the *MS1096/+* or *MS1096/+; UAS-UbxRNAi/+* backgrounds, as indicated. **(B)** Summary of the number of genes leading or not to a phenotype (size and/or ectopic wing-like bristles) in the wild type (*wt*, corresponding to MS1096/ +) or sensitized (*abxbxpbx/ +*) contexts. **(C)** Illustrative SEM pictures of phenotypes obtained with the different targeted genes in the wild type context, as indicated. **(D)** Illustrative SEM pictures of phenotypes obtained with the different targeted genes in the sensitized context, as indicated. Ectopic wing-like bristles are indicated when present (yellow stars). See also **Supplementary Tables 1**–**3**.

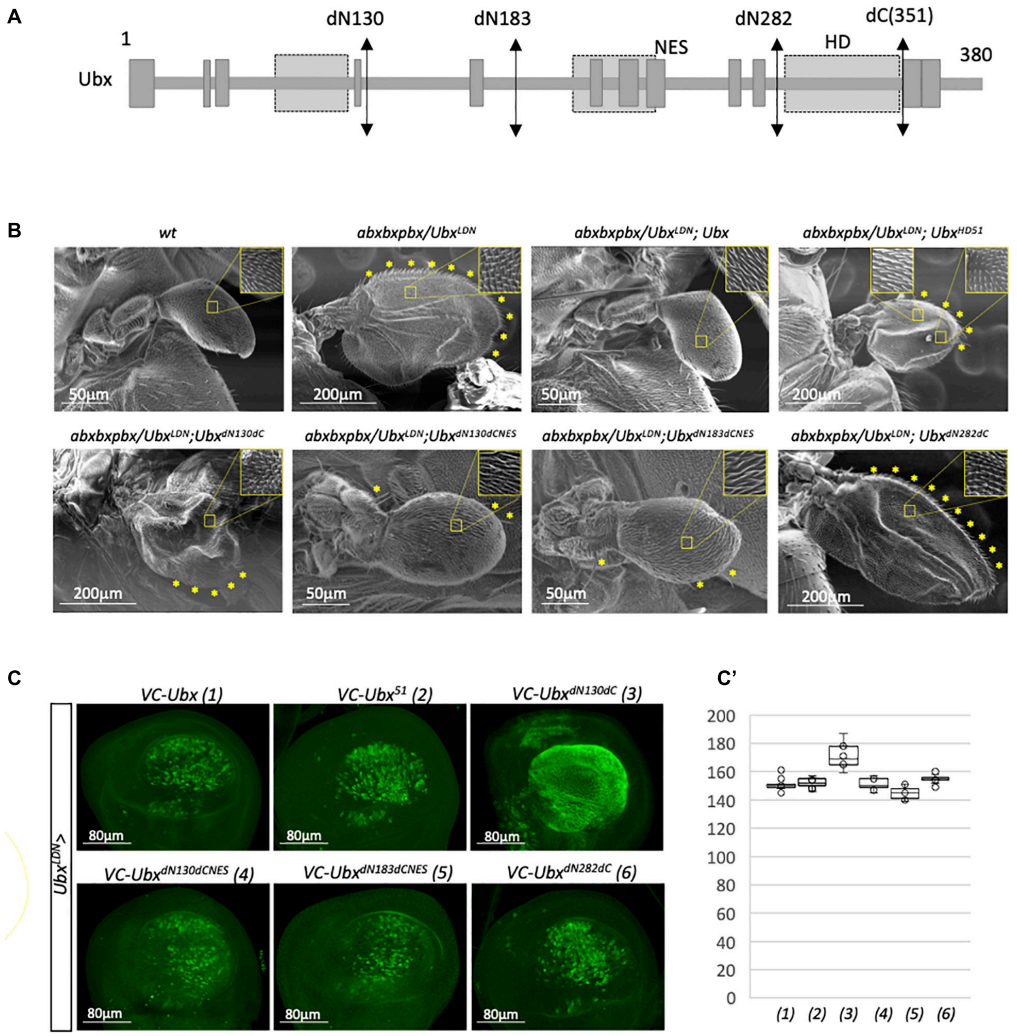


FIGURE 3 | A minimal form of Ubx can rescue the haltere-to-wing mutant phenotype. **(A)** Schematic representation of Ubx with the different sites of deletion. The emplacement of the homeodomain (HD), residue 51 of the HD and unconventional Nuclear Export Signal (NES) are indicated. The NES overlaps with the conserved hexapeptide (HX) motif of Ubx (Duffraisse et al., 2020). **(B)** Illustrative SEM pictures of the activity of the different deleted and mutated forms of Ubx in the rescue assay. The haltere-to-wing mutant context results from the combination of the *abxbxpbx* allele over the hypomorphic Ubx^{LDN} that corresponds to the insertion of a P*Gal4* in *Ubx* upstream regulatory sequences (Casares et al., 1996). The NES of Ubx has to be mutated for allowing the rescue with the deleted forms. Enlargements depict the haltere-like and/or wing-like hairs in the different genetic backgrounds. Wing-like bristles on the margin are also highlighted (yellow stars). Note that the Ubx^{HD51} construct leads to incomplete rescue with the formation of a structure that resembles to a wing in terms of size, hairs (with a mixture of wing-like and haltere-like hairs) and wing-like bristles, but with no obvious veins. In contrast, the $Ubx^{dN130dCNES}$ and $Ubx^{dN183dCNES}$ constructs lead to an almost complete rescue, although halteres are two times bigger than wild type halteres on average and contain few remaining wing-like bristles. All the phenotypes depicted in the pictures were robustly obtained from two independent experiments for each genetic background. **(C)** Expression of the various mutated and deleted forms of Ubx used in the rescue assay. Immunostaining was performed with an anti-GFP antibody that recognizes the C-terminal fragment of Cerulean (CC) fused to each construct (Duffraisse et al., 2020). The N-terminal deletions induced constitutive nuclear export except when the NES is mutated, as previously described (Duffraisse et al., 2020). **(C')** Boxplot representation of the quantification of the GFP immunostaining in the haltere pouch upon expression of the different Ubx constructs (1–6) with the *MS1096* driver.

We thus repeated the analysis with the additional mutation of the NES (construct $Ubx^{dN130dCNES}$; Duffraisse et al., 2020). In this context, the deleted form was able to rescue the mutant phenotype, confirming that the first 130 and last 29 residues of Ubx were not necessary for the haltere developmental program when the protein is properly addressed in the nucleus. An identical level of rescue was observed when deleting even more the N-terminal part in the context of the NES mutation (construct $Ubx^{dN183dCNES}$; **Figure 3B**). In contrast, using a minimal form of Ubx that corresponds to the HD only ($Ubx^{dN282dC}$) was not sufficient for rescuing the phenotype (**Figure 3B**), although it was correctly localized in the nucleus (**Figures 3C,C'**). Altogether, these results show that a large part of Ubx is dispensable for its DNA-binding dependent activity in the haltere disc and that the region included between the residues 183 and 351 is sufficient for ensuring the proper haltere developmental program.

DISCUSSION

The developmental program underlying haltere formation in *Drosophila* is highly robust, which is best exemplified by the fact that almost normal halteres are formed when 40% (*abxbxpbx*/+ background) or 50% (Ubx^1 null allele/+ background) of Ubx level is lost. The wing-like phenotypes (haltere size and ectopic wing-like bristles on the margin) obtained in these genetic backgrounds are weak, which is surprising given that no other Hox gene is expressed in the haltere pouch (there is therefore no redundancy, as it could be in other structures like the leg discs; Wirz et al., 1986; Paul et al., 2021). Strong haltere-to-wing transformation starts to be observed at 60% loss of Ubx (Paul et al., 2021), underlining that the high level of Ubx in the haltere pouch is a way to buffer against mutations that could affect *Ubx* expression. Interestingly, previous work with the Ubx^1 mutant allele also led to the conclusion that haltere development is under strong stabilizing selection (Gibson and Van Helden, 1997) and the notion of "potential variance" with a threshold-dependent response to Ubx haploinsufficency was proposed in this particular developmental context (Gibson et al., 1999). Thus, high Ubx level allows both a stable development of the haltere and subtle morphological changes upon variation. What are the molecular cues underlying Ubx transcriptional activity in this particular context?

Here, we propose that Ubx is working as a "self-sufficient" TF to regulate the set of its target genes and ensure the developmental robustness of the haltere capitellum.

First, a high dosage of Ubx could be used to recognize several monomeric binding sites on the different target enhancers (**Figure 4**). For example, Ubx has been shown to repress wing-promoting genes in the haltere disc through the recognition of several consensus monomeric binding sites in repressed target enhancers (Galant et al., 2002; Hersh and Carroll, 2005). Ubx has also been shown to recognize low-affinity binding sites, which are by now established as being critical for ensuring both the specificity and the robustness of Hox-controlled developmental enhancers (Crocker et al., 2015, 2016; Merabet and Mann, 2016). In this context, high doses of the Hox protein could be essential for efficient recognition and binding on these atypical sites *in vivo*. The multiplication of consensus and non-consensus monomeric binding sites could therefore, provide a certain degree of redundancy (**Figure 4**), making the expression of the enhancer very stable even when half of the monomeric binding sites are not occupied (due to mutations affecting the nucleotide sequence and/or Ubx levels).

Second, a high expression level could confer cofactor-independent activity to Ubx, explaining why our candidate RNAi screen was not successful in revealing phenotypes in the haltere capitellum, even in the sensitized *abxbxpbx* background. Some transcriptional partners might be involved in the case of the regulation of a few target genes and/or be required on a few binding sites in target enhancers, eventually leading to subtle phenotypes that were not captured in the screen. In any case, given that Ubx can perfectly bind to monomeric sites upon high level of expression, we hypothesize that such transcriptional partners could preferentially act as collaborators and modulate Ubx transcriptional activity rather than improving Ubx DNA-binding affinity and/or specificity. The presence of non-consensus binding sites could also be a way to increase monomeric DNA-binding specificity in the absence of cofactors, explaining that only Hox proteins with a similar HD to Ubx (Antennapedia, Antp and Abdominal-A, Abd-A) or strong monomeric DNA-binding activity (Abdominal-B, AbdB) could replace Ubx for haltere specification upon high Ubx-like expression level (Casares et al., 1996; Paul et al., 2021). This self-sufficient molecular model is reinforced by our observation that a minimal form of Ubx can perform the job of haltere specification. This model contrasts with other cofactors' and collaborator's based-models (Mann et al., 2009; Merabet and Mann, 2016; Sánchez-Higueras et al., 2019), illustrating the diversity of the molecular strategies that could be used by Hox proteins *in vivo*.

Third, a cofactor-independent mode of activity could allow stabilizing the Ubx developmental program against genetic variation. By definition, requiring fewer cofactors will diminish the number of mutations that could affect/modify the Hox function. Only mutations affecting Ubx levels at various extents could impact on haltere morphology (**Figure 4**). In this context, the comparison with the wing developmental program is interesting. Indeed, recent work showed that the Hox gene *Antp* is necessary for proper wing formation in *Drosophila* (Paul et al., 2021). However, in contrast to Ubx, Antp is expressed at low levels in specific regions of the pouch of the wing disc. We speculate that this low Hox dose background could serve as a genetic decanalization template for allowing more sensitive phenotypic variability of wings when compared to the haltere capitellum. At the molecular level, the activity of Antp could potentially be more dependent on the interaction with various transcriptional partners when compared to Ubx in the haltere disc. This cofactor-dependent mode of activity could make the wing developmental program more sensitive to genetic perturbations therefore, more plastic for phenotypic variation than the haltere capitellum (Parchem et al., 2007; Soto et al., 2008; Koshikawa, 2020). The same rational could potentially apply in the pedicellum of the haltere, which contains a lower expression level of Ubx when compared to the pouch, for example for varying the arrangement of sensory neurons (Agrawal et al., 2017). Interestingly, one of the rare strong candidate cofactor revealed in our screen (Hth) had a role in this particular region.

Whether a similar dose-dependent scenario could apply in other insect species remains to be investigated. Insect species with similar or dissimilar forewings and hindwings were shown to have similar or dissimilar Hox expression levels in their corresponding wing primordia (Paul et al., 2021). Hox level was also shown to be systematically higher in the hindwing primordia in insect species having different pairs of wings. Whether this differential expression profile is responsible for the phenotypic change is not known. Still, we speculate that a Hox dosage-based model (in contrast to a Hox-specific based model) could constitute a useful molecular mechanism for diversifying flight appendage morphologies in insects. This model is based on the finding that increasing the dose of Antp in the wing disc pouch was sufficient to transform the wing into a haltere

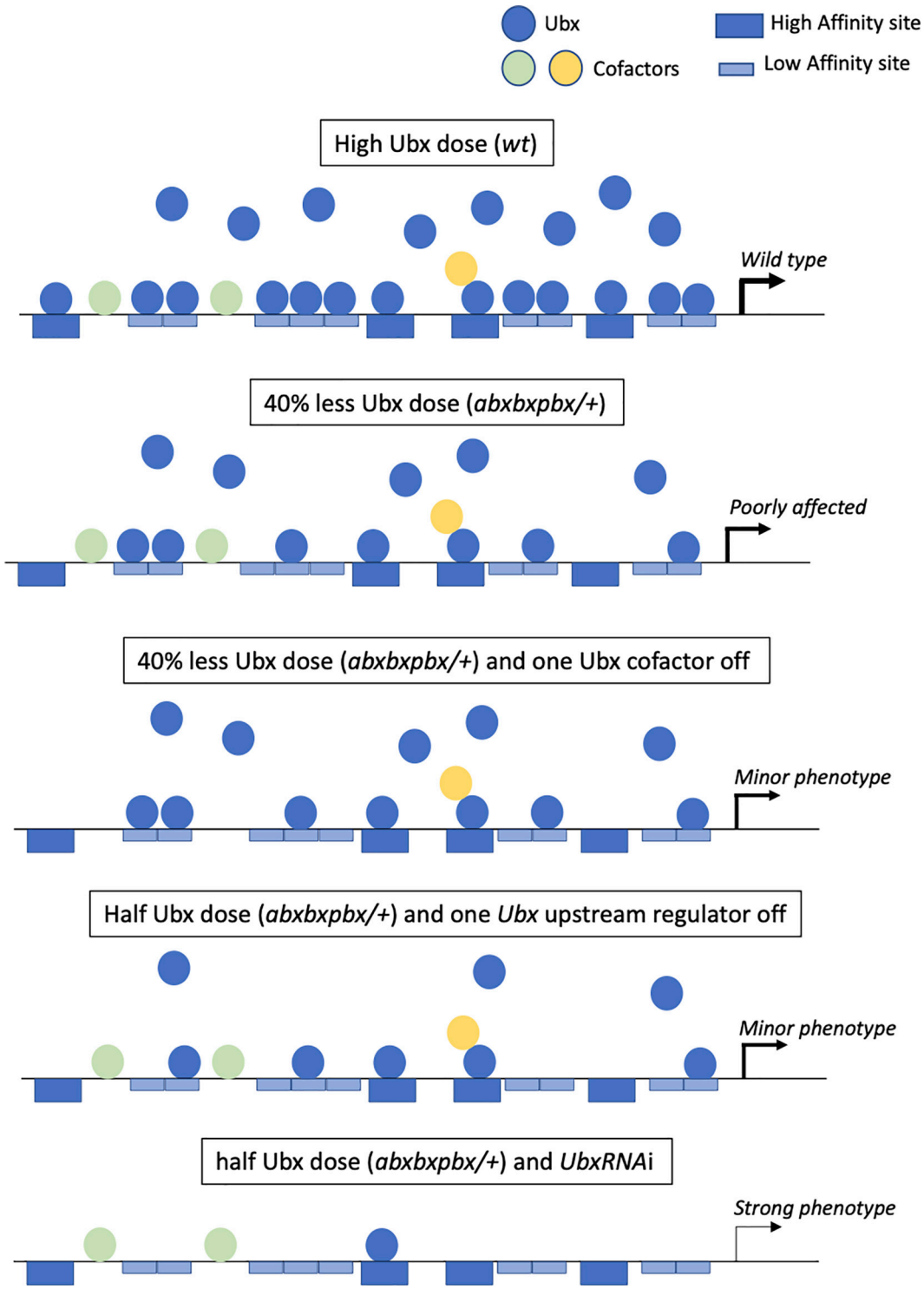


FIGURE 4 | The "self-sufficient" molecular model of Ubx in the haltere disc. Ubx (blue balls) recognizes several redundant high affinity (dark-blue large boxes) and low-affinity (light-blue large boxes) monomeric binding sites in target enhancers at normal (high) doses in the haltere disc. Few cofactors (colored balls) are acting as collaborators (without making dimeric DNA-bound Ubx/cofactor complexes) that modulate the transcriptional output of Ubx. In the context of the heterologous *abxbxpbx* mutant background, there is still enough Ubx molecules to bind on the majority of the redundant monomeric sites, allowing ensuring the haltere developmental program without major morphological variations. The loss of one cofactor in the wild type (not shown in the figure) or mutant context could potentially affect the regulation of some target genes, eventually leading to a subtle phenotype. The haltere morphology could also be moderately affected when targeting an upstream regulator of *Ubx*. Finally, only a strong decrease of Ubx expression level will affect the haltere developmental program, due to the inactivity of most monomeric binding sites in target enhancers.

(Paul et al., 2021). This phenotype recalls the controversy of flight appendages in Strepsiptera, which regroups endoparasitic insects with an inverted T2 haltere and T3 wing when compared to Diptera (Foottit et al., 2018). Thus, rather than resulting from *de novo* expression of Ubx in T2 as initially proposed (Whiting and Wheeler, 1994), the wing/haltere exchange could "simply" result from an inverted high and low expression level of Antp and Ubx, respectively.

MATERIALS AND METHODS

Drosophila Strains and Genetics Crosses

Drosophila strains were cultured following standard procedures at 25°C. Yellow white was used as a wild-type strain. All the RNAi TRiP lines were obtained from Bloomington (**Supplementary Table 1**). The following GAL4 were used: *MS1096-Gal4* (Bloomington, #8696) and *Ubx-GAL4*^{LDN}. UAS RNAi lines were crossed with *MS1096-Gal4* and *MS1096-Gal4; abxbxpbx* followed by incubation at 25°C and emerging flies were observed for the haltere phenotype.

RNA-Sequencing From Wing and Haltere Imaginal Discs

Wandering third instar larvae from *Drosophila melanogaster* (*CS* strain) were cut and inverted in PBS at 4°C. Wing and haltere imaginal discs were dissected and stored in Trizol separately. RNA extraction and sequencing were performed at Genotypic Technologies at Bangalore, India. Raw reads were filtered for adapter sequences and aligned to the dm6 genome using the HISAT2 software. The full sequencing data will be available in another work (Khan et al., in preparation).

Immunofluorescence Assay in *Drosophila* Imaginal Discs

Imaginal discs were fixed following dissection in 4% paraformaldehyde (methanol free) for 15 min. Washes were done with 1 × PBS 0.1%TritonX solution (PBTx). Samples were then blocked with 2% BSA solution for 2 h. Primary antibodies were incubated for ON at 4°C and then washed in PBTx and secondary antibodies incubated for 2 h at room temperature. Samples were then washed in PBTx and mounted in Vectashield (Vector laboratories) for confocal acquisition. Primary antibodies used were mouse anti-Ubx/ABD-A FP6.87 (1:20; DSHB) and rabbit anti-GFP PABG1 (1:500; Chromotek).

Imaging

The adult Drosophila appendage phenotype images were taken by Scanning electron microscope Hirox SH-3000. All the fluorescence microscopy images of haltere imaginal discs were captured using confocal Zeiss LSM 780. Images were captured at a 1,024 × 1,024 pixel resolution using 40x oil objective. The expression levels were quantified by measuring the intensity of GFP using the histogram function of the FIJI Software. The threshold was subjected to minute adjustment (using the « Image calculator» function) to create an image containing all positive nuclei (using the « Subtract » function) that were analyzed for fluorescence quantification (using the « analyze particles » function) and deduce the mean fluorescence intensity.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

GG: execution of the experiments, data analysis, and formal analysis. RP: conceptualization, execution of the experiments, data analysis, formal analysis, and writing. MD: execution of the experiments. SK and LS: conceptualization and writing. SM: conceptualization, data analysis, formal analysis, and writing. All authors contributed to the article and approved the submitted version.

FUNDING

Work in SM's laboratory was supported by the CNRS, ENS-Lyon, Fondation pour la Recherche Médicale (FRM 160896) and Centre Franco-Indien pour la Promotion de la Recherche Avancée (Cefipra N° 5503-P).

ACKNOWLEDGMENTS

We thank the Bloomington stock centre for providing the *Drosophila* fly lines, the Developmental Studies Hybridoma Bank for antibodies and the Arthrotools and PLATIM platforms of the UAR3444/US8 of Lyon for fly food and imaging facilities, respectively.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2021.713282/full#supplementary-material

Supplementary Table 1 | Full list of the candidate genes tested in the RNAi screen. The type of DNA-binding domain (DBD) and Bloomington RNAi line number are indicated. Boxes highlighted in yellow and orange denote genes that affect haltere formation in the wild type or sensitized background, respectively. Expression levels of each tested candidate gene analyzed from RNA-seq data are given in the standard RPKM (reads per kilobase per million) unit in the wing and haltere discs.

Supplementary Table 2 | List of genes leading to haltere phenotype upon RNAi in the wild type context. Haltere size and ectopic bristle quantifications are indicated. These phenotypes were reproducibly obtained from two independent experiments.

Supplementary Table 3 | List of genes leading to haltere phenotype upon RNAi in the sensitized context. Haltere size and ectopic bristle quantifications are indicated. These phenotypes were reproducibly obtained from two independent experiments.

REFERENCES

Agrawal, S., Grimaldi, D., and Fox, J. L. (2017). Haltere morphology and campaniform sensilla arrangement across Diptera. *Arthropod Struct. Dev.* 46, 215–229. doi: 10.1016/j.asd.2017.01.005

Akam, M. (1998). Hox genes: from master genes to micromanagers. *Curr. Biol.* 8, R676–R678.

Averof, M., and Akam, M. (1995). Hox genes and the diversification of insect and crustacean body plans. *Nature* 376, 420–423.

Baëza, M., Viala, S., Heim, M., Dard, A., Hudry, B., Duffraisse, M., et al. (2015). Inhibitory activities of short linear motifs underlie hox interactome specificity in vivo. *Elife* 4:e06034.

Bender, W., Akam, M., Karch, F., Beachy, P. A., Peifer, M., Spierer, P., et al. (1983). Molecular Genetics of the Bithorax Complex in *Drosophila melanogaster*. *Science* 221, 23–29. doi: 10.1126/science.221.4605.23

Bischof, J., Duffraisse, M., Furger, E., Ajuria, L., Giraud, G., Vanderperre, S., et al. (2018). Generation of a versatile BiFC ORFeome library for analyzing protein-protein interactions in live *Drosophila*. *Elife* 7:e38853.

Brand, A. H., and Perrimon, N. (1993). Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* 118, 401–415. doi: 10.1242/dev.118.2.401

Carroll, S. B., Weatherbee, S. D., and Langeland, J. A. (1995). Homeotic genes and the regulation and evolution of insect wing number. *Nature* 375, 58–61. doi: 10.1038/375058a0

Casares, F., Calleja, M., and Sánchez-Herrero, E. (1996). Functional similarity in appendage specification by the ultrabithorax and abdominal-a drosophila HOX genes. *EMBO J.* 15, 3934–3942. doi: 10.1002/j.1460-2075.1996.tb00767.x

Choo, S. W., White, R., and Russell, S. (2011). Genome-wide analysis of the binding of the hox protein ultrabithorax and the hox cofactor homothorax in *Drosophila*. *PLoS One* 6:e14778. doi: 10.1371/journal.pone.0014778

Crickmore, M. A., and Mann, R. S. (2006). Hox control of organ size by regulation of morphogen production and mobility. *Science* 313, 63–68. doi: 10.1126/science.1128650

Crickmore, M. A., Ranade, V., and Mann, R. S. (2009). Regulation of Ubx expression by epigenetic enhancer silencing in response to Ubx levels and genetic variation. *PLoS Genet.* 5:e1000633. doi: 10.1371/journal.pgen.1000633

Crocker, J., Abe, N., Rinaldi, L., McGregor, A. P., Frankel, N., Wang, S., et al. (2015). Low affinity binding site clusters confer hox specificity and regulatory robustness. *Cell* 160, 191–203. doi: 10.1016/j.cell.2014.11.041

Crocker, J., Preger-Ben Noon, E., and Stern, D. L. (2016). The soft touch: low-affinity transcription factor binding sites in development and evolution. *Curr. Top. Dev. Biol.* 117, 455–469.

de Navas, L., Foronda, D., Suzanne, M., and Sánchez-Herrero, E. (2006b). A simple and efficient method to identify replacements of P-lacZ by P-Gal4 lines allows obtaining Gal4 insertions in the bithorax complex of *Drosophila*. *Mech. Dev.* 123, 860–867. doi: 10.1016/j.mod.2006.07.010

de Navas, L. F., Garaulet, D. L., and Sánchez-Herrero, E. (2006a). The ultrabithorax hox gene of *Drosophila* controls haltere size by regulating the Dpp pathway. *Development* 133, 4495–4506. doi: 10.1242/dev.02609

Delker, R. K., Ranade, V., Loker, R., Voutev, R., and Mann, R. S. (2019). Low affinity binding sites in an activating CRM mediate negative autoregulation of the *Drosophila* hox gene ultrabithorax. *PLoS Genet.* 15:e1008444. doi: 10.1371/journal.pgen.1008444

Dickinson, M. H. (1999). Haltere-mediated equilibrium reflexes of the fruit fly, *Drosophila melanogaster*. *Philos. Trans. R. Soc. B Biol. Sci.* 354, 903–916. doi: 10.1098/rstb.1999.0442

Di-Poï, N., Montoya-Burgos, J. I., Miller, H., Pourquié, O., Milinkovitch, M. C., and Duboule, D. (2010). Changes in hox genes' structure and function during the evolution of the squamate body plan. *Nature* 464, 99–103. doi: 10.1038/nature08789

Duffraisse, M., Paul, R., Carnesecchi, J., Hudry, B., Banreti, A., Reboulet, J., et al. (2020). Role of a versatile peptide motif controlling hox nuclear export and autophagy in the *Drosophila* fat body. *J Cell Sci.* 133:jcs241943.

Foottit, R. G., Adler, P. H., and Kathirithamby, J. (2018). "Biodiversity of Strepsiptera," in *Insect Biodiversity*, eds R. G. Foottit and P. H. Adler (Hoboken, NJ: John Wiley & Sons Ltd), 673–704. doi: 10.1002/9781118945582.ch22

Galant, R., and Carroll, S. B. (2002). Evolution of a transcriptional repression domain in an insect hox protein. *Nature* 415, 910–913. doi: 10.1038/nature717

Galant, R., Walsh, C. M., and Carroll, S. B. (2002). Hox repression of a target gene: extradenticle-independent, additive action through multiple monomer binding sites. *Development* 129, 3115–3126. doi: 10.1242/dev.129.13.3115

Gibson, G., and Van Helden, S. (1997). Is function of the *Drosophila* homeotic gene ultrabithorax canalized? *Genetics* 147, 1155–1168. doi: 10.1093/genetics/147.3.1155

Gibson, G., Wemple, M., and van Helden, S. (1999). Potential variance affecting homeotic ultrabithorax and antennapedia phenotypes in *Drosophila melanogaster*. *Genetics* 151, 1081–1091. doi: 10.1093/genetics/151.3.1081

Gomez, C., and Pourquié, O. (2009). Developmental control of segment numbers in vertebrates. *J. Exp. Zool. B Mol. Dev. Evol.* 312, 533–544. doi: 10.1002/jez.b.21305

Hall, J. M., McLoughlin, D. P., Kathman, N. D., Yarger, A. M., Mureli, S., and Fox, J. L. (2015). Kinematic diversity suggests expanded roles for fly halteres. *Biol. Lett.* 11:20150845. doi: 10.1098/rsbl.2015.0845

Hersh, B. M., and Carroll, S. B. (2005). Direct regulation of knot gene expression by ultrabithorax and the evolution of cis-regulatory elements in *Drosophila*. *Development* 132, 1567–1577. doi: 10.1242/dev.01737

Hersh, B. M., Nelson, C. E., Stoll, S. J., Norton, J. E., Albert, T. J., and Carroll, S. B. (2007). The UBX-regulated network in the haltere imaginal disc of D. melanogaster. *Dev. Biol.* 302, 717–727. doi: 10.1016/j.ydbio.2006.11.011

Irvine, K. D., Botas, J., Jha, S., Mann, R. S., and Hogness, H. D. (1993). Negative autoregulation by ultrabithorax controls the level and pattern of its expression. *Development* 117, 397–399.

Kassis, J. A., Kennison, J. A., and Tamkun, J. W. (2017). Polycomb and trithorax group genes in *Drosophila*. *Genetics* 206, 1699–1725. doi: 10.1534/genetics.115.185116

Kittelmann, S., Buffry, A. D., Franke, F. A., Almudi, I., Yoth, M., Sabaris, G., et al. (2018). Gene regulatory network architecture in different developmental contexts influences the genetic basis of morphological evolution. *PLoS Genet.* 14:e1007375. doi: 10.1371/journal.pgen.1007375

Koshikawa, S. (2020). Evolution of wing pigmentation in *Drosophila*: diversity, physiological regulation, and cis-regulatory evolution. *Dev Growth Differ.* 62, 269–278. doi: 10.1111/dgd.12661

Lewis, E. B. (1978). A gene complex controlling segmentation in *Drosophila*. *Nature* 276, 565–570. doi: 10.1038/276565a0

Makhijani, K., Kalyani, C., Srividya, T., and Shashidhara, L. S. (2007). Modulation of decapentaplegic gradient during haltere specification in *Drosophila*. *Dev. Biol.* 302, 243–255. doi: 10.1016/j.ydbio.2006.09.029

Mallo, M. (2018). Reassessing the role of hox genes during vertebrate development and evolution. *Trends Genet.* 34, 209–217. doi: 10.1016/j.tig.2017.11.007

Mann, R. S., Lelli, K. M., and Joshi, R. (2009). Chapter 3 hox specificity:unique roles for cofactors and collaborators. *Curr. Top. Dev. Biol.* 88, 63–101. doi: 10.1016/s0070-2153(09)88003-4

Merabet, S., and Mann, R. S. (2016). To be specific or not: the critical relationship between hox and tale proteins. *Trends Genet.* 32, 334–347. doi: 10.1016/j.tig.2016.03.004

Mohit, P., Bajpai, R., and Shashidhara, L. S. (2003). Regulation of wingless and vestigial expression in wing and haltere discs of *Drosophila*. *Development* 130, 1537–1547. doi: 10.1242/dev.00393

Morata, G., and Garcia-Bellido, A. (1976). Developmental analysis of some mutants of the bithorax system of *Drosophila*. *Wilehm. Roux. Arch. Dev. Biol.* 179, 125–143. doi: 10.1007/bf00848298

Pallavi, S. K., Kannan, R., and Shashidhara, L. S. (2006). Negative regulation of Egfr/Ras pathway by ultrabithorax during haltere development in *Drosophila*. *Dev. Biol.* 296, 340–352. doi: 10.1016/j.ydbio.2006.05.035

Parchem, R. J., Perry, M. W., and Patel, N. H. (2007). Patterns on the insect wing. *Curr. Opin. Genet. Dev.* 17, 300–308.

Paul, R., Giraud, G., Domsch, K., Duffraisse, M., Marmigère, F., Khan, S., et al. (2021). Hox dosage contributes to flight appendage morphology in *Drosophila*. *Nat. Commun.* 12:2892.

Pavlopoulos, A., and Akam, M. (2011). Hox gene ultrabithorax regulates distinct sets of target genes at successive stages of *Drosophila* haltere morphogenesis. *Proc. Natl. Acad. Sci. U.S.A.* 108, 2855–2860. doi: 10.1073/pnas.1015077108

Pearson, J. C., Lemons, D., and McGinnis, W. (2005). Modulating hox gene functions during animal body patterning. *Nat. Rev. Genet.* 6, 893–904. doi: 10.1038/nrg1726

Pick, L., and Heffer, A. (2012). Hox gene evolution: multiple mechanisms contributing to evolutionary novelties. *Ann. N. Y. Acad. Sci.* 1256, 15–32. doi: 10.1111/j.1749-6632.2011.06385.x

Ronshaugen, M., McGinnis, N., and McGinnis, W. (2002). Hox protein mutation and macroevolution of the insect body plan. *Nature* 415, 914–917. doi: 10.1038/nature716

Saadaoui, M., Merabet, S., Litim-Mecheri, I., Arbeille, E., Sambrani, N., Damen, W., et al. (2011). Selection of distinct hox-extradenticle interaction modes fine-tunes hox protein activity. *Proc. Natl. Acad. Sci. U.S.A.* 108, 2276–2281. doi: 10.1073/pnas.1006964108

Sambrani, N., Hudry, B., Maurel-Zaffran, C., Zouaz, A., Mishra, R., Merabet, S., et al. (2013). Distinct molecular strategies for Hox-mediated limb suppression in *Drosophila*: from cooperativity to dispensability/antagonism in TALE partnership. *PLoS Genet.* 9:e1003307. doi: 10.1371/journal.pgen.1003307

Sánchez-Higueras, C., Rastogi, C., Voutev, R., Bussemaker, H. J., Mann, R. S., and Hombría, J. C. (2019). In vivo hox binding specificity revealed by systematic changes to a single cis regulatory module. *Nat. Commun.* 10:3597.

Schertel, C., Albarca, M., Rockel-Bauer, C., Kelley, N. W., Bischof, J., Hens, K., et al. (2015). A large-scale, in vivo transcription factor screen defines bivalent chromatin as a key property of regulatory factors mediating *Drosophila* wing development. *Genome Res.* 25, 514–523. doi: 10.1101/gr.181305.114

Slattery, M., Ma, L., Négre, N., White, K. P., and Mann, R. S. (2011). Genome-wide tissue-specific occupancy of the hox protein ultrabithorax and hox cofactor homothorax in *Drosophila*. *PLoS One* 6:e14686. doi: 10.1371/journal.pone.0014686

Soto, I. M., Carreira, V. P., Soto, E. M., and Hasson, E. (2008). Wing morphology and fluctuating asymmetry depend on the host plant in cactophilic *Drosophila*. *J. Evol. Biol.* 21, 598–609. doi: 10.1111/j.1420-9101.2007.01474.x

Stern, D. L. (1998). A role of ultrabithorax in morphological differences between *Drosophila* species. *Nature* 396, 463–466. doi: 10.1038/24863

Tomoyasu, Y., Wheeler, S. R., and Denell, R. E. (2005). Ultrabithorax is required for membranous wing identity in the beetle Tribolium castaneum. *Nature* 433, 643–647. doi: 10.1038/nature03272

White, R. A., and Wilcox, M. (1985). Distribution of ultrabithorax proteins in *Drosophila*. *EMBO J.* 4, 2035–2043. doi: 10.1002/j.1460-2075.1985.tb03889.x

Whiting, M. F., and Wheeler, W. C. (1994). Insect homeotic transformation. *Nature* 368:696. doi: 10.1038/368696a0

Wirz, J., Fessler, L. I., and Gehring, W. J. (1986). Localization of the antennapedia protein in *Drosophila* embryos and imaginal discs. *EMBO J.* 5, 3327–3334. doi: 10.1002/j.1460-2075.1986.tb04647.x

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Cell and Developmental Biology

REVIEW
published: 05 August 2021
doi: 10.3389/fcell.2021.718175

Anterior Hox Genes and the Process of Cephalization

James C.-G. Hombría, Mar García-Ferrés and Carlos Sánchez-Higueras*

Centro Andaluz de Biología del Desarrollo (Consejo Superior de Investigaciones Científicas/Junta de Andalucía/Universidad Pablo de Olavide), Seville, Spain

During evolution, bilateral animals have experienced a progressive process of cephalization with the anterior concentration of nervous tissue, sensory organs and the appearance of dedicated feeding structures surrounding the mouth. Cephalization has been achieved by the specialization of the unsegmented anterior end of the body (the acron) and the sequential recruitment to the head of adjacent anterior segments. Here we review the key developmental contribution of Hox1–5 genes to the formation of cephalic structures in vertebrates and arthropods and discuss how this evolved. The appearance of Hox cephalic genes preceded the evolution of a highly specialized head in both groups, indicating that Hox gene involvement in the control of cephalic structures was acquired independently during the evolution of vertebrates and invertebrates to regulate the genes required for head innovation.

Keywords: Hox genes, cephalogenesis, evolution, arthropods, vertebrates

OPEN ACCESS

Edited by:
Ernesto Sánchez-Herrero, Severo Ochoa Molecular Biology Center (CSIC-UAM), Spain

Reviewed by:
Anthony Percival-Smith, University of Western Ontario, Canada
Ramkumar Sambasivan, Indian Institute of Science Education and Research, Tirupati, India

****Correspondence:***
James C.-G. Hombría
jcashom@upo.es

Specialty section:
This article was submitted to Morphogenesis and Patterning, a section of the journal Frontiers in Cell and Developmental Biology

Received: *31 May 2021*
Accepted: *16 July 2021*
Published: *05 August 2021*

Citation:
Hombría JC-G, García-Ferrés M and Sánchez-Higueras C (2021) Anterior Hox Genes and the Process of Cephalization. Front. Cell Dev. Biol. 9:718175. doi: 10.3389/fcell.2021.718175

EVOLUTIONARY RELATIONSHIPS OF CEPHALIC HOX GENES

Hox genes are found in almost all animals, generally organized in large clusters of up to 15 genes. Their amino acid sequence has been used to classify Hox proteins in distinct homology groups and to infer how the cluster evolved (Zhang and Nei, 1996). Comparative analyses indicate that well developed clusters comprising at least seven Hox genes were already present more than 550 million years ago, suggesting their evolution was concurrent with the diversification of the main animal body plans that appeared during the Cambrian explosion (de Rosa et al., 1999).

The comparison of Hox complexes among living animal groups allows inferring the cluster's temporal evolution. No Hox genes have been found in simple animal forms like Sponges, Ctenophores, and Placozoa (Biscotti et al., 2014) but are present in Cnidarians (jellyfish and sea anemones), where they also control axial development (He et al., 2018). Cnidarian Hox genes are only related to the Anterior and the Posterior Hox groups, suggesting the Cnidarians diverged from other animals at an early stage in the complex's expansion (**Figure 1**). Slightly more diverse Hox complexes are present in the Acoels (extremely simple wormlike creatures), which probably constitute a sister group to all other bilaterians (Achatz et al., 2013). Acoels possess Hox proteins related to the Anterior, Central and Posterior paralogy groups. In contrast, all other animals studied to date have expanded clusters containing between 8 and 15 Hox genes that belong to seven defined groups (Biscotti et al., 2014). These include complex animals like the Chordates (which include the vertebrates), the Lophotrochozoa (which include the annelid worms and the molluscs); and the Ecdysozoa (which include the arthropods and the Onychophora).

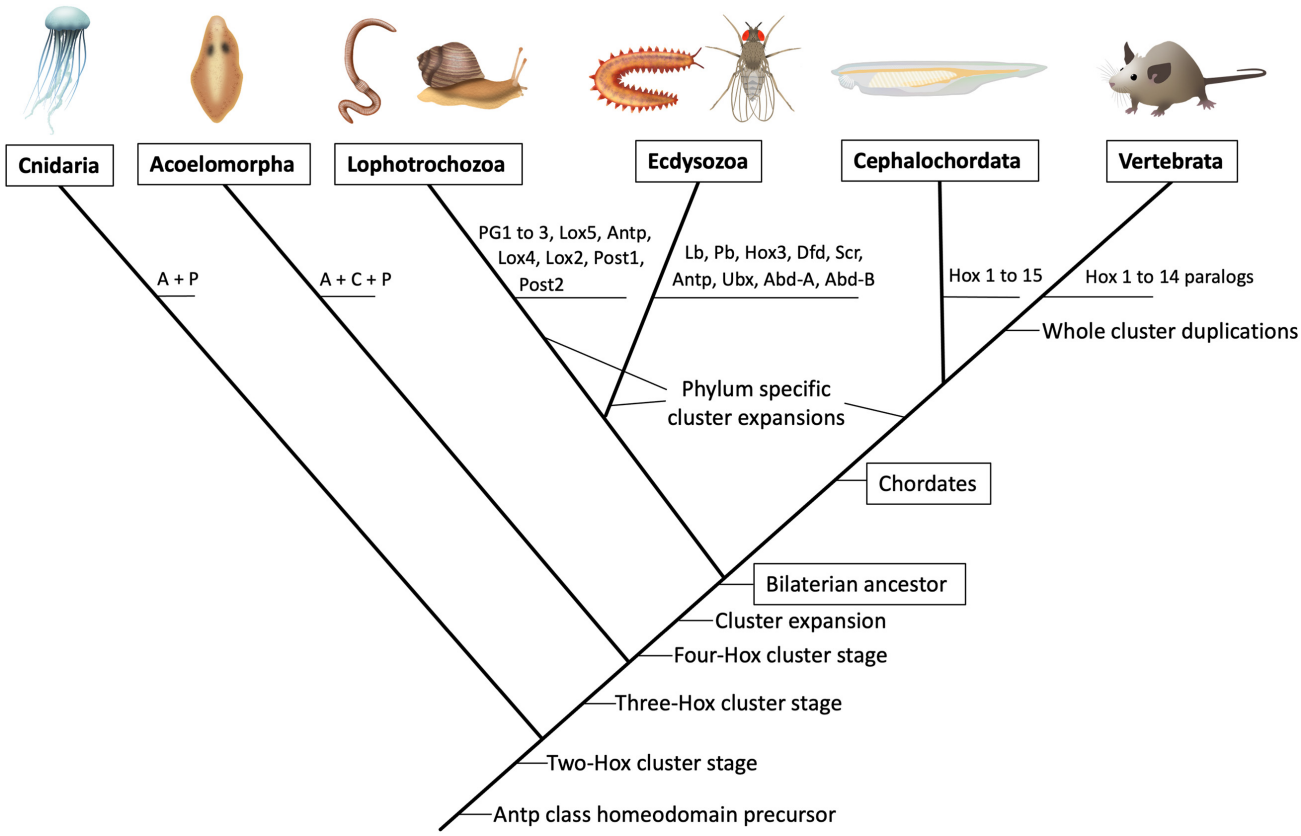


FIGURE 1 | Schematic phylogenetic tree showing the predicted evolution of the Hox paralogous groups. Illustrations of major extant animal groups are presented on top joined by their currently accepted branching points (not to temporal scale). The predicted appearance of particular Hox paralog gene precursors is depicted on the branching points. In the vertebrate lineage the whole cluster duplicated several times (not represented) giving rise to four paralogous Hox clusters named HoxA to HoxD in birds and mammals. In teleost fish a further duplication gave rise to eight Hox clusters. For a comprehensive analysis of Hox cluster evolution (see **Figure 2A**). A, anterior; C, central; P, posterior.

Several hypotheses based on sequence similarities have been proposed to explain how the Hox cluster expanded from a single original Antennapeadia class homeobox protein. This precursor protein duplicated to form an initial ProtoHox cluster composed of an Anterior and a Posterior gene (**Figure 2A**). After two consecutive duplications, or unequal crossovers, the Anterior Hox gave rise to new genes (Gehring et al., 2009). This resulted in a three gene-cluster composed of an Anterior, a Central and a Posterior Hox gene and then to a four gene-cluster encoding an Anterior, Group 3, Central and Posterior Hox. Sequence similarities with other non-Hox Antennapedia-class homeodomain proteins indicate that the ProtoHox gene cluster duplicated at some point during this initial expansion giving rise to the ParaHox cluster and the Hox cluster proper (Garcia-Fernandez, 2005a,b).

All known Hox genes can be related to one of these early four groups, suggesting that multiple tandem duplications expanded this primitive four Hox cluster prior to the divergence of the bilaterian animals. The duplication of the Anterior gene created the ancestors of Hox1/Labial/PG1 and the Hox2/Proboscipaedia/PG2 homologs in chordates, arthropods and molluscs, respectively (**Figure 2B**). Group three Hox proteins are represented today by a single Hox3/Hox3/PG3 homolog, although in insects this gene lost its typical Hox expression and got involved in extra-embryonic membrane specification and the establishment of the maternal antero-posterior axis (Zen, Zen2, and Bicoid) (Falciani et al., 1996; Hughes et al., 2004). A stepwise duplication of the Central gene gave rise to Hox4–8/Dfd, Scr, Antp, Ubx, abd-A/Antp, Lox5,4,2 (de Rosa et al., 1999; Li et al., 2020). Orthologs across species are difficult to assign for some of the Central group Hox genes. Central group sequence comparisons only cluster reliably Hox4 with Dfd and Hox5 with Scr, leaving uncertainty about when the Hox6–8/Antp, Ubx, abdA/Lox5, Antp, Lox4, Lox2 precursors formed. Thus, these must have arisen by independent duplications after the Chordate, Lophotrocozoa, and Ecdysozoa ancestors diverged or, if present before that time, their sequence has diverged so much that now it is impossible to confidently assign them to specific groups (Zhang and Nei, 1996; Hueber et al., 2010). Finally, duplication of the posterior gene created the ancestors of Hox9–14/Abd-B/Post 1–2.

In summary, the available data indicate that the Hox1–5 paralogy groups, that are expressed in cephalic regions, were already present in the bilaterian common ancestor of Chordates, Lophotrocozoa and Ecdysozoa, which may have resulted in

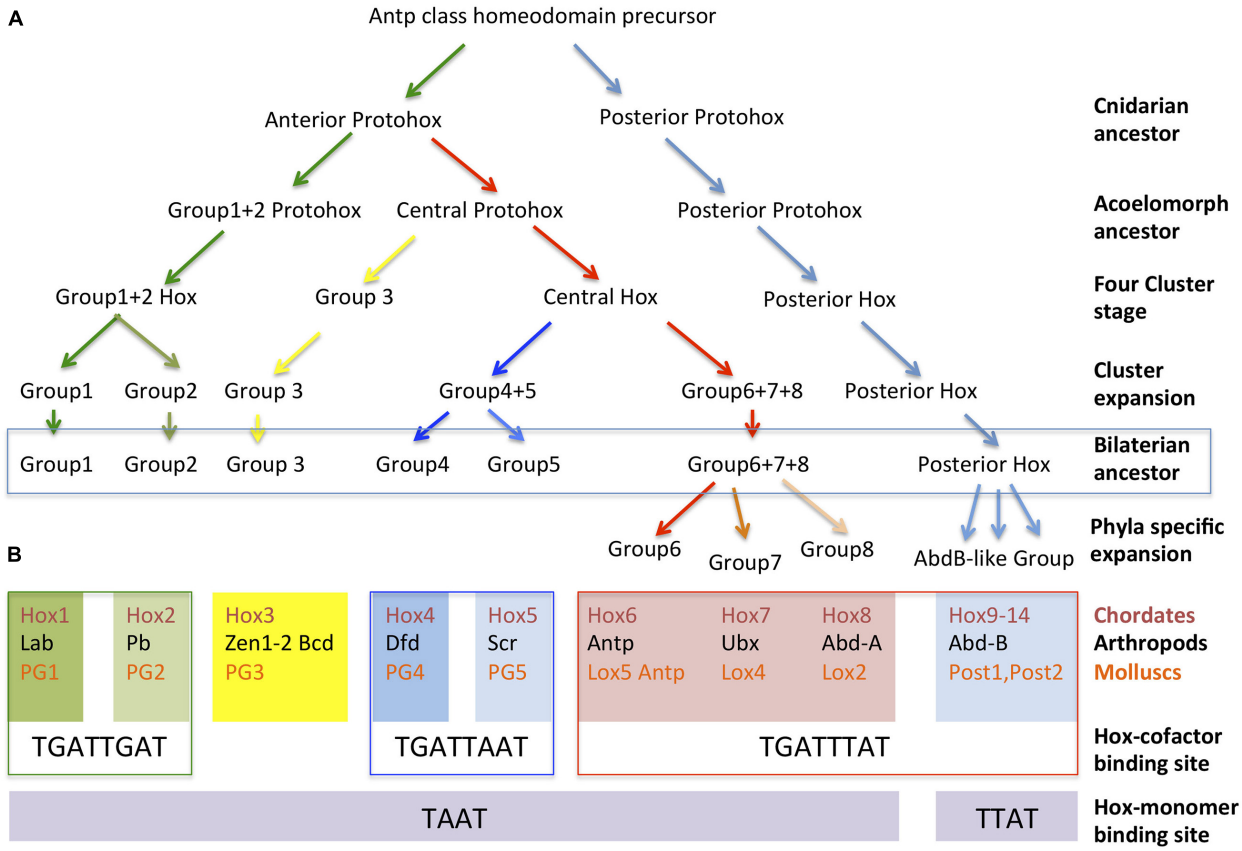


FIGURE 2 | Evolution of Hox protein clusters and correlation with their DNA binding preferences. **(A)** Inferred cluster evolution from sequential tandem duplications of an ancient Antp class homeobox gene. Three tandem duplications gave rise to a four-gene ProtoHox cluster. Duplication of the full ProtoHox cluster gave rise to the ParaHox cluster (not shown) and the Hox cluster proper. The duplication of specific genes in the four-gene Hox cluster resulted in its expansion in the common ancestor that gave rise to all bilateral animals. Phyla specific duplications resulted in the final expansion that resulted in the 10–15 Hox clusters found in modern species. **(B)** Classification of the Chordate, Arthropod, and Mollusc Hox proteins into paralogy groups according to their evolutionary relationships as inferred from their sequence similarities. Each of the seven groups is aligned over their preferred DNA binding sites either when in complex with TALE cofactors or as monomers. For simplicity, the Hox6-Hox8 genes have been organized in columns aligned with particular arthropod and mollusc equivalents, even though direct homology for genes in the column has not been demonstrated, see text.

these proteins sharing functional characteristics among these highly diverse phyla.

DNA BINDING PREFERENCES IN CEPHALIC HOX PROTEINS

Hox proteins can bind DNA as monomers or in complex with cofactor proteins. Two cofactor proteins are conserved and have been studied in vertebrates (Pbx and Meis) and in *Drosophila* (Exd and Hth). These cofactors contain an atypical DNA-binding homeodomain with a three aminio acid loop extension (TALE) that gives name to their class. Recent analyses have revealed that Hox proteins in complex with their cofactors show distinguishing conserved DNA binding specific preferences.

When binding to DNA as monomers, several studies have shown that Hox proteins recognize the same core sequence, TAAT, except the posterior AbdB-like Hox proteins which favor TTAT (Ekker et al., 1994; Noyes et al., 2008). Finding such a small DNA binding site and the fact that most Hox proteins bound the same sequence was at odds with their known *in vivo* regulatory specificity. Further analyses found that Hox proteins increase their target specificity using two alternative strategies: the use of clustered monomeric Hox sites, and the binding to DNA in complex with TALE cofactors. Hox-cofactor DNA binding results in the enlargement of the DNA recognition site and the increase in DNA binding affinity (van Dijk and Murre, 1994; Ryoo et al., 1999).

A thorough SELEX-seq analysis in *Drosophila melanogaster* of all possible Hox paralog complexes with Extradenticle (Exd) and Homothorax (Hth) has revealed that cofactor interaction uncovers a latent specificity present in the Hox protein that modulates its DNA binding preferences. Hox-cofactor DNA sequence preferences can be used to classify Hox proteins into three classes. Class 1 includes the Labial and Proboscipaedia proteins that present a higher affinity for TGATTGAT; Class 2 includes Dfd and Scr with preferential affinity for TGATTAAT; and Class 3 includes Antp, Ubx, Abd-A, and Abd-B whose

preferred binding site is TGATTTAT (Slattery et al., 2011; **Figure 2B**). Human HoxA1, HoxA5, and HoxA9 when in complex with the homologous vertebrate TALE cofactors have the same *in vitro* binding preferences as the *Drosophila* paralogs (Kribelbauer et al., 2017). In *Drosophila,* these *in vitro* differential binding preferences translate into a differential spatial specific downstream target activation, which can be also replicated by the *Amphioxus* Hox proteins (Sanchez-Higueras et al., 2019). Besides high affinity sites, cephalic Hox proteins may also activate their targets through low affinity binding sites and in some cases even share identical targets or functions with other Hox proteins (Hirth et al., 2001; Kribelbauer et al., 2019; Sanchez-Higueras et al., 2019). Despite this, not only the Hox protein sequence, but also the conservation of high affinity DNA binding preferences across distant species (**Figure 2B**) sets apart cephalic Hox proteins from other Hox proteins, underlying their ancient evolutionary relationship.

HOX PROTEIN CHARACTERISTICS INFLUENCING PARALOG DNA BINDING SPECIFICITY

Hox proteins present four regions with significant amino acid conservation: the hexapeptide, the linker region, the homeodomain and the C-terminal sequence adjacent to the homeodomain.

The *homeodomain* is the most conserved region. Although other transcription factors also possess homeodomains, four specific amino acids are common to all Hox proteins and probably confer the specific binding properties that distinguish this protein class from other homeodomain containing proteins. Besides the general Hox amino acids, a few amino acids are found to be present in specific paralogous groups providing them with specific characteristics (reviewed in Merabet et al., 2009). The so-called *hexapeptide* is a four to six amino acid sequence N-terminal to the homeodomain and separated from it by the linker region. The hexapeptide (HX) domain mediates the protein interaction with the TALE cofactors (Chang et al., 1995), although in some Hox proteins additional domains have also been found to establish physical Hox-Pbx interactions (Dard et al., 2018; Saurin et al., 2018). The *linker region* separates the HX from the homeodomain. It has a variable length that in many species has been shown to correlate with the paralogous group (In der Rieden et al., 2004; Merabet et al., 2009). The *C-terminal region* confers specific characteristics to paralogy group 1 Hox genes and can exert important regulatory functions (see below).

Although very few paralog specific characteristics have been studied in detail, some have been uncovered at the molecular level for the cephalic Hox1 and Hox5 proteins. The Labial/Hox1 paralog presents two particularities, the first one is that the hexapeptide, besides interacting with the Pbx/Exd co-factors, as in other Hox proteins, also has an inhibitory effect on the homeodomain preventing Lab/Hox1 binding to DNA. Inhibition is released when the Exd cofactor binds to the HX allowing the homeodomain to bind DNA (Chan et al., 1996). Whether Labial may still operate fully independent of Exd *in vivo* is still unclear. A second particularity has been found by cross-species functional analysis of Labial and its mouse orthologs HoxA1 and HoxB1. These studies uncovered how a six amino acid motif called CTM (C-terminal motif), located C-terminal to the homeodomain, modulates the Hox1-Pbx physical-interaction mediated by the HX. Although sequence conservation between HoxA1 and Lab is only about 30%, expression of the vertebrate protein can fully rescue fly *labial* mutant defects, indicating an ancestral function conserved by both paralogs. The conserved CTM motif present in both Labial and HoxA1 is required to retain an optimal physical interaction with Exd/Pbx1 as well as to perform the ancestral *in vivo* target regulation (Singh et al., 2020). In contrast, the HoxB1 paralog that has a divergent CTM sequence cannot rescue all *labial* mutant phenotypes in the fly. The divergent CTM motif reduces HoxB1′s interaction with Pbx1, preventing the formation of a ternary Hox-Pbx-Meis complex, which results in a different repertoire of genomic targets *in vivo.* This has been proposed to represent a case of Hox paralog neo-functionalization in brain and head tissues during development (Singh et al., 2020). *In vitro* SELEX-seq analysis of human HoxA1/Pbx shows similar DNA binding site preferences as its *Drosophila* counterpart (Kribelbauer et al., 2017). Additional structural and biochemical studies will help to decipher whether the CTM's modulatory effect on HoxB1-Pbx interactions either creates a new range of binding preferences, or forces Labial to operate preferentially as a monomer in a context-dependent manner.

Another case of paralog specificity has been described for the Scr/Hox5 paralog group, where it has been found that the interaction of Scr with their main DNA binding cofactor (Exd/Pbx) is not only necessary to recognize the DNA sequence, but also the specific DNA shape of its target DNA backbone, which has a strong impact on both DNA binding strength and specificity. A combination of structural studies, biochemistry and *in vivo* assays in *D. melanogaster* embryos showed that the specific binding of Scr-Exd heterodimers to a target site (AGATTAATCG) in the *fkh250* enhancer relies on the optimal interaction with the specific minor groove conformation. The width of the minor groove is determined by the DNA sequence, that in this case creates a narrowing of the groove in which two key Scr paralog-conserved basic amino acids (Arg3 and Arg5) from the homeodomain's N-terminal arm motif "RQR" and a Histydine from the adjacent linker region (His12) specifically interact (Joshi et al., 2007). These Hox-DNA interactions at the binding site's minor groove do not involve hydrogen bonds, unlike the direct homeodomain's third helix recognition of the bases in the major groove, and cannot take place without the Exd-YPWM interaction (Joshi et al., 2007; Slattery et al., 2011). Thus, it has been proposed that the binding preference of Dfd/Scr-Exd complexes for Class II core sites (TGATTAAT) exhibited in SELEX-seq experiments is strongly influenced by the conserved "RQR" motif and linker region (Joshi et al., 2007; Slattery et al., 2011). Moreover, high-throughput analyses including SELEX-seq as well as *in vivo* experiments in embryos showed that mutations of those Scr amino acids selecting DNA shape, bias the Scr-Exd binding preferences toward different core motifs including Class I (TGATTGAT) and Class III (TGATTTAT) (Abe et al., 2015). Interestingly, it was also shown that Scr-Exd DNA

binding specificity can be transferred to an Antp-Exd complex by mutating the residues from the N-terminal arm and linker to those involved in minor groove width recognition in Scr (Abe et al., 2015). Therefore, Hox-Exd complexes can discriminate differences in DNA minor groove shape through a reduced number of key side chain amino acids to establish a different set of functional binding specificities *in vivo*.

FUNCTION AND EXPRESSION OF VERTEBRATE CEPHALIC HOX PROTEINS

The functional analysis of vertebrate Hox genes is complicated by the existence of several paralogous Hox clusters due to successive duplications of the ancient Hox chordate cluster (Holland et al., 1994). In birds and mammals there are four Hox paralogous clusters named HoxA to HoxD containing paralogous genes that, in many cases, have redundant functions. Despite this, mutational analyses have demonstrated that vertebrate Hox proteins also control the morphological differentiation of repeated metameric structures along the antero-posterior axis. As in insects, spatial Hox expression in partially non-overlapping regions along the anterior-posterior body axis is key to confer each metamere with segment specific structures.

Experimental studies in mice and chick embryos, have shown Hox genes are required for the correct antero-posterior body axis segmental specification in the neural tissue, the branchial arch derivatives and the axial skeleton (Mallo et al., 2010; Philippidou and Dasen, 2013; Parker and Krumlauf, 2020). The requirement of Hox function is especially clear during embryonic development when studying the formation of the hindbrain rhombomeres, the neural crest cells and the somites, three structures with a transient segmental organization during early development.

Cephalic Hox Genes and Hindbrain Development

The hindbrain is the more posterior region of the vertebrate brain, giving rise to the pons, the medulla and the cerebellum. During early development the hindbrain becomes subdivided into eight segments known as rhombomeres (rh) that constitute lineage-restricted groups of cells that do not intermix (**Figure 3A**). Major nerves arise from different rhombomeres. Cephalic Hox genes are required for both hindbrain segmentation and for the specification of the motor nerves originating from the rhombomeres. In mice, abnormal rhombomeric segmentation is observed in mutations for Hox1 and Hox2 paralogy groups. Although single mutation of *Hoxa1* or *Hoxa2* genes already affect rhombomere segmentation, these defects increase in double mutants. Compound *Hoxa1/Hoxb1* mutants lack both rh4 and rh5 (Gavalas et al., 1998; Studer et al., 1998; Rossel and Capecchi, 1999). Compound *Hoxa2/Hoxb2* mutants lack boundaries between rh1 and rh4 (Davenne et al., 1999), and *Hoxa1/Hoxa2* double mutants completely lack rhombomere boundaries (Barrow et al., 2000).

The formation of specific hindbrain motoneurons also depends on cephalic Hox gene function. In mutants for *Hoxa2* the motoneurons of the trigeminal nerve, which forms in rh2–3, are disorganized and misrouted (Gavalas et al., 1997) while ectopic expression of *Hoxa2* in rh1 generates a trigeminal-like nerve (Jungbluth et al., 1999). In *Hoxb1* mutants, facial nerve motoneurons arising from rh4 acquire the characteristics of the trigeminal nerve leading to the loss of the facial nerve (Goddard et al., 1996; Studer et al., 1996; Gavalas et al., 2003). Moreover, ectopic expression of *Hoxb1* in rh1 can generate facial-like motoneurons, while its expression in rh2 transforms the trigeminal neurons into facial neurons (Bell et al., 1999). The abducens nerve originating in rh5 is absent in *Hoxa3/Hoxb3* double mutants and ectopic *Hoxa3* can induce its formation (Gaufo et al., 2003; Guidato et al., 2003). Although these results show the importance of Hox genes in the formation of particular nerve types, the cross-regulatory interactions among Hox gene expression and the requirement of more than one paralogous Hox group for nerve specification complicates the analysis.

Cephalic Hox Genes and Neural Crest Derivatives Development

The neural crest cells originate during development in different antero-posterior positions of the dorsal neural tube including the diencephalon, the mesencephalon, the rhombencephalon (the hindbrain) and the spinal cord. These cells lose their epithelial character and become migratory, giving rise to a variety of structures including cartilage, bones, pigment cells, peripheral neurons or glia depending on the segment where they are formed. The cranial neural crest cells (cNCC) originate from the anterior neural tube (**Figure 3A**, gray arrows), specified by Otx2 and the Hox1–4 genes (Minoux and Rijli, 2010). cNCC originating from the hindbrain migrate in separate streams to colonize the pharyngeal arches (PA). These cells participate in the formation of the ventral cranial bones and the nerve ganglia, and influence the migratory routes of the motoneurons growing from the rhombomeres. Cranial neural crest cells can give rise to cartilage while trunk neural crest cells do not. The cNCC originating from rh1 do not express any Hox gene while those from more posterior rhombomeres express Hox1–5 paralogs. *Hoxb1* (and *Hoxb2*) is expressed in rh4 and in the second pharyngeal arch (PA2). *Hoxa2* is active in rh3 and rh5 and in the neural crest cells colonizing PA2-A4. *Hoxa3* is expressed in rh5-rh6 (and *Hoxb3* is expressed in rh6-rh8) colonizing PA3 and PA4. *Hoxd4* is expressed in PA4 (Minoux and Rijli, 2010) (**Figure 3A**). *Hox5* is expressed in rh8 (Philippidou and Dasen, 2013) and at PA4 (Holland and Hogan, 1988; Kam and Lui, 2015). Mutation of the Hox1 paralog group results in the absence of all rh4 derived neural crest cells (Gavalas et al., 2001; McNulty et al., 2005). Mutation of the Hox2 paralogous genes result in the transformation of the PA2 derivatives into structures normally formed by PA1 in a typical homeotic transformation (Gendron-Maguire et al., 1993; Rijli et al., 1993; Hunter and Prince, 2002; Santagati et al., 2005). Ectopic Hox2 paralog gene expression in PA1 derivatives cause their transformation into structures normally formed by PA2 (Grammatopoulos et al., 2000; Pasqualetti et al., 2000;

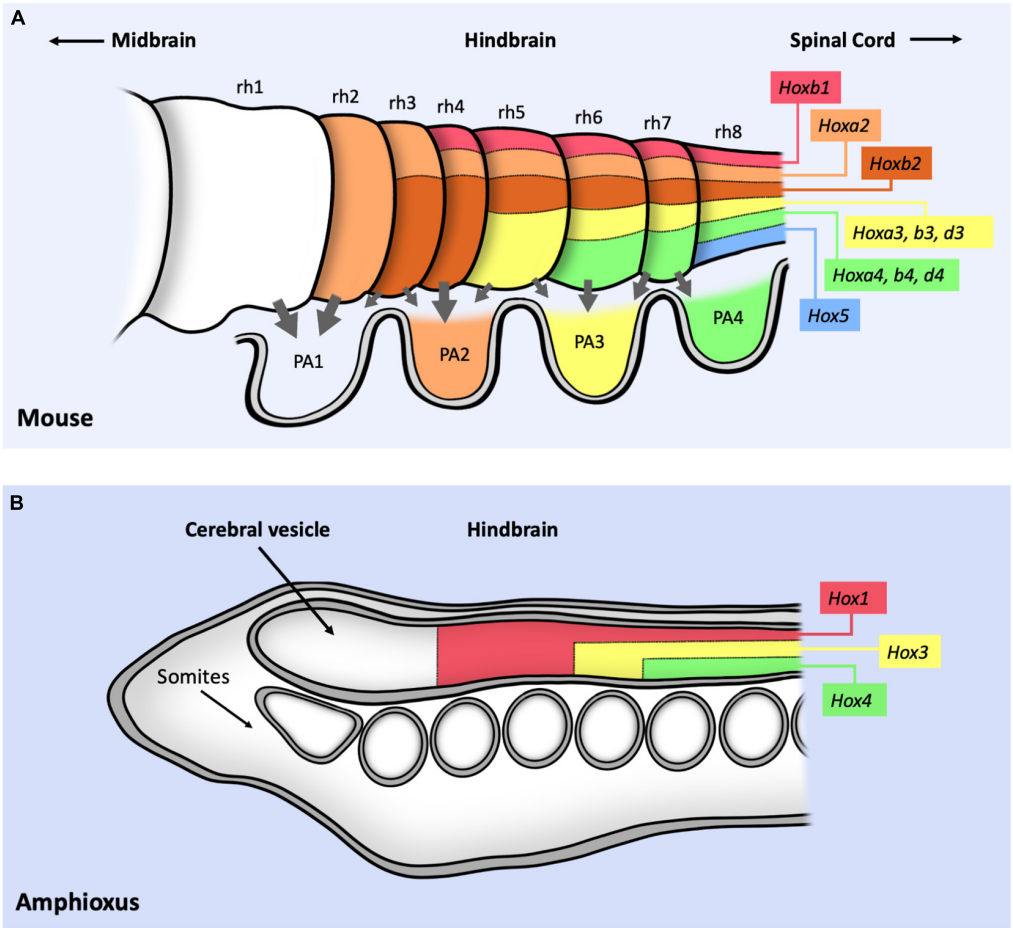


FIGURE 3 | Schematic representation of Hox expression in the cephalic mouse and Amphioxus central nervous systems. **(A)** Summary of Hox1–5 expression in the hindbrain, labeling the position of the eight different rhombomeres (rh1–8). Gray arrows represent the migratory movement of the cranial neural crest cells toward the pharyngeal arches (PA1–4). PA are colored following the pharyngeal Hox code: no Hox proteins are expressed in PA1; HoxA2 and B2 proteins are expressed from PA2 to PA4; HoxA3, B3 and D3 from PA3 to PA4; HoxA4, B4, D4 and Hox5 in PA4. **(B)** Expression of three Hox genes in the anterior part of the *Amphioxus* CNS. Note that the conserved relative position of Hox gene expression with respect to the mouse CNS suggest a positional correlation of both nervous systems despite the lack of rhombomeres or neural crest cells in *Amphioxus*. In both panels, information only concerns antero-posterior expression and not dorso-ventral expression. Figure based on data from references (Holland et al., 2008; Philippidou and Dasen, 2013; Parker et al., 2018).

Hunter and Prince, 2002; Kitazawa et al., 2015). Inactivation of Hox3 paralogs cause malformations of the PA3 and PA4 skeletal derivatives although these are not homeotic transformations (Condie and Capecchi, 1994; Manley and Capecchi, 1997). Interestingly, deletion of the HoxA cluster in cNCCs causes a partial homeotic transformation of PA3 and PA4 derivatives into PA1-like structures, indicating a requirement for both *Hoxa2* and *Hoxa3* (Minoux et al., 2009). Despite advances on our knowledge on Hox expression and function on the neural crest cells it is still unclear how the Hox genes integrate with the neural crest gene network (Parker and Krumlauf, 2020).

Cephalic Hox Genes and Axial Skeletal Development

Although not specifically cephalic structures, the cephalic Hox genes contribute to the specification of the cervical vertebrae during the development of the most anterior axial skeleton. The somites are transient embryonic structures produced by the segmental organization of the presomitic mesoderm that appear at both sides of the nerve cord prefiguring the axial skeleton. Besides the axial skeleton, somites give rise to several structures including dermis, tendons and muscles. The vertebrae originate from the ventral part of the somite, which experiences an Epithelial to Mesenchymal Transition (EMT) giving rise to the sclerotome.

Hoxa3/Hoxd3 are required for the development of the most anterior vertebrae: the atlas and the axis. Some Hox3 paralog mutants present homeotic transformations as well as defects that could be due to lack of vertebrae primordia cell proliferation. This defect is stronger in double *Hoxa3/Hoxd3* mutations, which present a complete deletion of the atlas (Condie and Capecchi, 1994). Loss of Hox4 paralogs cause transformations of cervical vertebrae (C) toward their anterior counterparts, the

atlas and the axis. In triple *Hoxa4/b4/d4* mutants, morphological characteristics of the atlas appear in C2–C5 vertebrae as well as some defects in C6 and C7 which resemble Hox5 mutant phenotypes (Horan et al., 1995).

Hox5 affects both the development of posterior cervical vertebrae and of anterior thoracic vertebrae and ribs (McIntyre et al., 2007). Hox5 mutants show transformations of the vertebrae toward the C2 (the axis). In this respect, Hox5 presents similar phenotypes to Hox6 mutants although Hox6 affect only from C7 to posterior segments. Ectopic Hox6 induces ectopic ribs in the cervical and lumbar regions.

FUNCTION AND EXPRESSION OF ARTHROPOD CEPHALIC HOX PROTEINS

Four Hox genes are expressed in the cephalic segments of *Drosophila*: *lab*, *pb*, *Dfd,* and *Scr* which are homologous to Hox1, Hox2, Hox4, and Hox5. The Hox3 gene homolog has lost its homeotic function in insects and acquired new functions in the specification of the extra-embryonic membranes and in anterior maternal specification. However, in other arthropods Hox3 expression fits with a homeotic activity (**Figure 4**).

The head of *Drosophila*, as that of other insects, is composed by a non-segmented region or acron followed by six segments, three pre-oral (labral, antennal, and intercalary) and three post-oral (mandibula, maxilla, and labium) (Juergens and Hartenstein, 1993). The acron, the labrum and antennal segments do not express any Hox gene although their development requires other homeobox containing genes like *otd*, *ems,* and *btd*, homologs of which are also used in vertebrates for the development of the most anterior head structures, suggesting a deep conservation of the anterior head's organization (Hirth et al., 2003).

labial (*lab*) is the Hox gene expressed most anteriorly, becoming activated in the intercalary segment, also known as tritocerebral segment (Hughes and Kaufman, 2002b). Embryonic *lab* loss of function in *Drosophila* results in larval phenotypes such as head involution defects and the absence of the H piece, but no homeotic transformations. Adult hypomorphs or mitotic clones show various head defects. In the anterior region of the head, there is a deletion of the vibrissae and the maxillary palps, and in the posterior region of the head, a transformation toward thoracic-like bristles and the appearance of thoracic spiracles, suggesting a transformation toward the mesothoracic segment (Diederich et al., 1989; Merrill et al., 1989). In pedipalp bearing arthropods, like spiders, *lab* disruption induces appendage loss from the tritocerebral segment (Pechmann et al., 2015). Several studies have implicated *lab* as an essential neuronal regulator. The tritocerebral neuromere, which corresponds to the most posterior part of the arthropod brain, is severely affected in *lab* mutants, showing loss of neuronal markers and axonal patterning defects (Hirth et al., 1998). Interestingly, this can be rescued by ectopic supply of any other Hox gene, except *Abd-B*, indicating that *cis*-regulatory elements confer the specificity of the interaction, rather than the Hox protein (Hirth et al., 2001). The role of *lab* in larval neuronal control is less explored nonetheless. Kuert et al. (2012) proposed that during the transition to the third larval instar (L3), *lab* induces apoptosis on two specific neuroblast lineages. By blocking apoptosis, they were able to rescue these two neuroblasts in L3, which are Lab positive.

proboscipedia (*pb*) embryonic expression in arthropods is highly variable, spanning from the pedipalp segment to the fourth leg segment in chelicerates to just half of the second antennal segment in crustaceans (Hughes and Kaufman, 2002b). Instead, among insects Pb expression seems to be conserved (Denell et al., 1996; Rogers and Kaufman, 1997). In *Drosophila* embryos, Pb expression in the gnathal segments is dependent on *Deformed* (*Dfd*) and *Sex combs reduced* (*Scr*) (Rusch and Kaufman, 2000), but *pb* null mutants show no apparent functional role during embryogenesis (Pultz et al., 1988). In contrast, adult *pb* null mutants show transformed labial palps into legs (Kaufman, 1978). Conversely, ectopic Pb expression in the leg primordia transforms legs into maxillary or labial palps (Aplin and Kaufman, 1997). In *Drosophila*, it has been proposed that Pb is a competence factor allowing Scr to switch from a T1 function into a proboscis function (Percival-Smith et al., 2013).

Deformed (*Dfd*) is expressed in the mandibular and maxillary segments of all arthropods (Hughes and Kaufman, 2002b). As shown by Lohmann et al. (2002), Dfd shapes the mandibular and maxillary boundary by controlling directly the proapoptotic gene *reaper*. In *Dfd* mutants, the border between the two gnathal segments is lost, and this can be rescued by restoring Reaper expression. In *Dfd* mutants, the mouth hooks and the sensory cirri do not develop and the maxillary sensory organ is disorganized. However, some of these defects could be caused indirectly by the defective head involution movements caused by these mutations. Other mouth parts are abnormal in *Dfd* mutants with a possible duplication of the cephalopharingeal plates (Regulski et al., 1987). The different developmental outcomes of Dfd activity in the mandibular and the maxillary segments have been attributed to the modulation of Dfd function in the mandible exerted by the Cap-n-collar (Cnc) protein. Isoform C of the Cnc basic leucine zipper protein in the mandible modulates the transcriptional regulation exerted by Dfd in that segment. In *cnc* null alleles mouth hooks and cirri, which are typical maxillary structures appear in the mandible with the disappearance of certain mandibular structures (McGinnis et al., 1998; Veraksa et al., 2000).

Dfd is also involved in neuronal specification, in a similar way to *lab*. Disruption of either one of them induces defects in axonal patterning, indicating that both of them play a role in the establishment of regional neuromere characteristics (Hirth et al., 1998). A specific pathway has been described, where *Dfd* controls autonomously the specification of maxillary neuroblasts by induction of the cell adhesion protein Amalgam. This pathway is redundant with a non-autonomous one controlled by *lab* and *Antp* (Becker et al., 2016).

Cephalic neuronal and endocrine specification can be translated into the control of specific behaviors, like feeding or molting. Friedrich et al. (2016) have shown that Dfd is expressed in the subesophageal ganglion, which innervates the muscles that control food intake. They observed that *Dfd* is required for the

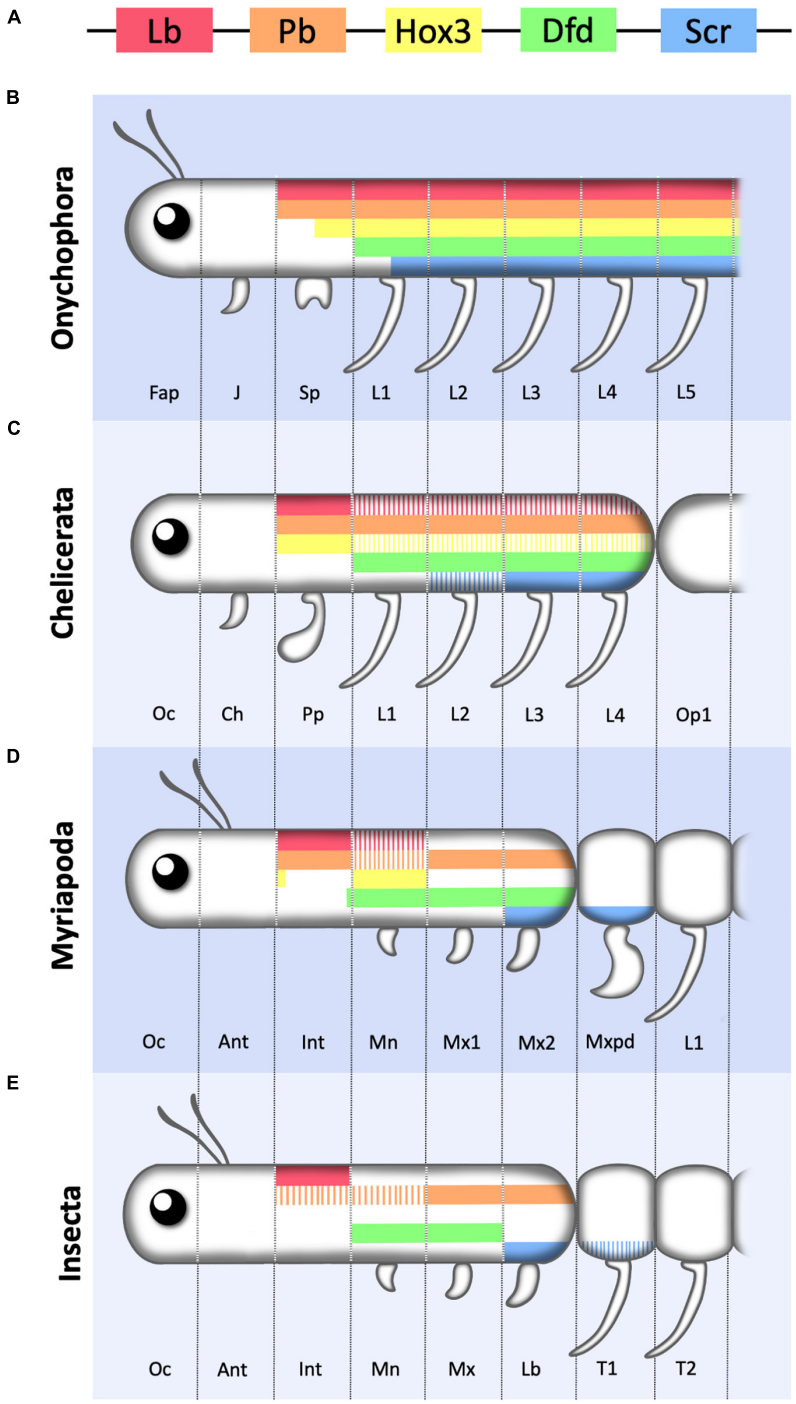


FIGURE 4 | Schematic representation of Hox1–5 expression in the anterior region of the Onychophora and various arthropods. **(A)** Color key of the five Hox genes considered in this review represented following the relative positions they occupy in the cluster. **(B)** Onychophora showing the three head segments (Fap, frontal appendage; J, jaw; Sp, slime papilla) and the first five leg (L) bearing segments. **(C)** Chelicerate showing the first three cephalic segments (Oc, ocular; Ch, chelicerae; Pp, pedipalp) and the four legged prosomal segments. **(D)** Myriapod showing the six cephalic segments (Oc, ocular; Ant, antennal; Int, intercalary; Mn, Mandibula; Mx1, first maxilla; Mx2, second maxilla) followed by the trunk leg (L) bearing segments of which the most anterior one forms a maxilliped (Mxpd). **(E)** Insect showing the six cephalic segments (Oc, ocular; Ant, antennal; Int, intercalary; Mn, Mandibula; Mx, maxilla; Lb, Labium) followed by two of the three thoracic leg (L) bearing segments. Hox3 has not been represented as in most insects it has lost its Hox function. Note that the most anterior cephalic segments lack Hox expression, and that anterior Hox genes required for the formation of cephalic structures in insects may be expressed in leg bearing segments in other phyla, indicating that there is no strict correlation between cephalization vs. trunk development and Hox1–5 expression. For simplicity Crustaceans are not included in the figure and coloring represents antero-posterior Hox expression only. Solid bars represent main or stronger expression compared to striped bars. Figure based on data from references (Hughes and Kaufman, 2002b; Janssen et al., 2014).

specification and maintenance of the feeding unit by regulating the synaptic stability protein Ankyrin2-XL.

Dfd is able to control endocrine primordia fate, by promoting the specification of the *Corpora Allata* (*CA*) in the maxilla (Sanchez-Higueras et al., 2014). The *CA* synthesizes juvenile hormone, which promotes the maintenance of the larval stage after molting (Hartenstein, 2006). Interestingly, in the larval termite soldier, *Dfd* seems to be able to respond to juvenile hormone levels to control mandible elongation by activating the *dachshund* transcription factor (Sugime et al., 2019). This indicates a reversed control mechanism where *Dfd* becomes downstream of the mechanisms it activated during embryogenesis.

Sex combs reduced (*Scr*) is expressed both in the last cephalic and in the first thoracic segments, a characteristic also observed for Hox5 in vertebrates. Embryonic Scr expression usually occupies the labial segment (the last cephalic segment) and the first thoracic segment (Hughes and Kaufman, 2002b). Similar to *Dfd*, *Scr* is implicated in endocrine organ formation during embryogenesis through the specification of the prothoracic gland (PG) primordia (Sanchez-Higueras et al., 2014). The PG synthesizes Ecdysone, which promotes the transition between larval stages or induces metamorphosis depending on the presence or absence of Juvenile Hormone (reviewed in Hartenstein, 2006). In *Drosophila*, Dfd and Scr are expressed transiently in the endocrine primordia where, together with STAT, mediate expression of the *snail* transcription factor to induce an EMT of these cells necessary for *CA* and PG development (Sanchez-Higueras et al., 2014). In the hemipteran insect bug *Oncopeltus fasciatus*, RNAi against *Scr* disrupts prothoracic gland fate, indicating that this gene network is conserved across insects (Hanna and Popadic, 2020). A recent study points out that Scr is activated in *Bombyx mori* larval prothoracic glands, where it negatively regulates the levels of Ecdysone to control the number of molts (Daimon et al., 2021). This suggests that the same circuit used to specify glands during embryogenesis is recruited at later stages to modulate endocrine gland function.

Besides the PG, the specification of the salivary glands in the labium also requires transient expression of *Scr*. Scr target genes in the salivary glands include the *fork head*, *sage* and *CrebA* transcription factors (Panzer et al., 1992; Andrew et al., 1994; Abrams and Andrew, 2005; Abrams et al., 2006). A similar control mechanism was proposed to mediate silk gland specification in *Bombyx mori* (Kokubo et al., 1997).

As mentioned before, Scr expression pattern is not restricted to the cephalic tagma, as it is also activated on the first thoracic (T1) segment. RNAi against *Scr* in *Oncopeltus fasciatus* is able to induce a small ectopic wing in T1, which is transcriptionally different from T2 wings. This data indicate that *Scr* could have an ancient function in the repression of wings (Medved et al., 2015).

HEAD EVOLUTION AND HOX GENES

As described above, Hox genes play an important role in the formation of the posterior cephalic segments in both vertebrates and arthropods. However, comparative analysis of cephalization among species that diverged early in evolution from vertebrates and arthropods, indicates that their ancestors already had a cluster containing the Hox1–5 genes before the formation of a complex head (Holland et al., 2008; Eriksson et al., 2010; Janssen et al., 2014). Due to the convergent recruitment of Hox proteins into cephalization after the split of both lineages, it is unlikely that they regulate similar downstream targets or gene networks in vertebrates and invertebrates. The fact that vertebrate Hox proteins can rescue the phenotypes caused by mutations in the homologous *Drosophila* Hox genes may be due to their capability of occupying similar binding sites present in the target genes of the other species, rather than the target genes themselves being the same.

Amphioxus, a chordate that separated from the lineage that gave rise to vertebrates about 500 million years ago is believed to have a similar body plan as their common ancestor (Garcia-Fernandez et al., 2009). *Amphioxus* has a simple nervous system composed of a cerebral vesicle followed by a nerve cord without any rhombomeric subdivisions (**Figure 3B**). Gene expression analysis shows that *Amphioxus* Hox genes are expressed in the anterior-posterior axis in a very similar relative position to that of their mouse homologs (**Figure 3A**). This spatial expression conservation allows comparing both CNSs despite their different morphologies, leading to the suggestion that the *Amphioxus* frontal cerebral vesicle corresponds to the vertebrate forebrain and that the hindbrain equivalent of *Amphioxus* has no rhombomeric subdivisions although it has a similar antero-posterior expression of Hox1–4 genes (Holland et al., 2008). Thus, the Hox1–4 spatial gene expression was set out in the primitive chordate nervous system before the evolution of rhombomeres or neural crest cells, implying that these *Hox* genes were recruited later during vertebrate evolution to regulate novel functions that increased cephalic complexity.

A very similar conclusion is reached when studying Hox expression and cephalization in evolutionary distant groups related to the arthropods. Onychophorans, the sister group of the Arthropods, are animals with a relatively simple head despite possessing a full set of cephalic Hox1–5 genes (Janssen et al., 2014; Mayer et al., 2015). The Onychophoran head consists of only three segments, contrasting with the six head segments present in insects, myriapods and crustaceans (Eriksson et al., 2010; **Figure 4**). In the Onychophorans, each cephalic segment has a modified limb structure which, from anterior to posterior give rise to an antenna, a jaw and a slime papilla. Posterior to the head, a trunk is formed by repeated metameres with a pair of legs of similar shape in each segment. The most anteriorly expressed Hox genes are *pb* and *lab* which are expressed from the slime papilla segment to the most posterior trunk segment; *Hox3* is expressed from the last few cells of the papilla segment posteriorly; and *Dfd* and *Scr* are expressed from the first leg segment backward (Janssen et al., 2014). All other Hox genes are expressed in progressively more posterior trunk segments. Therefore, compared to insects, in Onychophorans only Lab and Pb are expressed in segments with a distinct cephalic character, while all other Hox genes are expressed in segments with a similar trunk external shape. This suggests that *Hox* gene expression

in the anterior segments of the animal, especially *Dfd* and *Scr,* predates their involvement in cephalogenesis and, thus, they must have been recruited later in evolution to contribute to the formation of arthropod specific head structures. This prediction fits with the morphological diversity observed between spiders and other arthropods where *Hox* genes like *Dfd* and *Scr,* that contribute to head morphology in centipedes, crustaceans and insects, are expressed in leg bearing segments in the spiders (Damen et al., 1998; Hughes and Kaufman, 2002a; Schwager et al., 2007; Khadjeh et al., 2012).

An interesting observation has been made in centipedes, which lack a differentiated abdomen and have a large trunk made of externally similar segments. The head in these animals is composed of six segments (ocular, antennal, intercalary, mandibular, maxillary I and maxillary II) with the two most anterior ones lacking Hox expression. *labial* and *Proboscipedia* are expressed in the intercalary segment, *Hox3* is expressed mostly in the mandibular segment, *Dfd* is expressed from the mandible to the maxillary II and *Scr* in Maxillary II and the first trunk segment (Hughes and Kaufman, 2002a). Interestingly, the leg appendages in this first trunk segment have diverged in shape from those present in the rest of the trunk, having converted into poison containing fangs. This segment, called the maxilliped, due to its intermediate head and trunk morphology, expresses *Scr* and *Antp*. Although no functional data are available, this pattern of expression suggests that both genes control its novel morphology and thus, in centipedes, a novel cephalic segment may be in the process of being recruited to the head and with it *Antp*, a Hox gene that usually has trunk functions.

In summary, current studies suggest that an extended *Hox* gene cluster had already evolved in a primitive bilateral ancestor and that these genes were probably expressed differentially along the antero-posterior axis in a rather undifferentiated trunk. As the cephalic region evolved, it became more complex by sequentially adding adjacent trunk segments to the primitive head instead of duplicating existing cephalic segments. As a result, the *Hox* genes expressed in the recruited segments were adopted as the key transcriptional regulators modulating the expression of target genes that gave rise to the phylum specific cephalic structures. In both arthropods and vertebrates, anterior Hox genes were used to specify cephalic structures just because they were expressed differentially in the recruited segments adjacent to the primitive head. The fact that anterior Hox proteins had different DNA binding site preferences to those of posterior ones, probably facilitated the differential modulation of target genes in the head versus those activated in the trunk by more posteriorly expressed Hox proteins. Convergent cephalization occurring at different speeds in arthropods and vertebrates may have resulted in different numbers of anterior Hox genes been recruited to cephalic structures. Genes like Hox5 have a very marginal function in the mammalian hindbrain specification, while its Scr homolog is well established as a cephalic gene in insects, even though it is still involved in trunk development. This suggests that any other Hox gene could have been recruited to perform cephalic functions, a process that might still be occurring as is suggested by the recent evolution of a maxilliped segment with intermediate head-trunk morphology in the centipedes. Here an extra seventh cephalic segment could be evolving by recruiting Antp to the head, whilst in most animals paralog group 6 proteins are exclusively involved in trunk development.

AUTHOR CONTRIBUTIONS

JH researched, wrote, and organized the manuscript. MG-F and CS-H researched and wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by grant P20-00003 of the Junta de Andalucía, and grants BFU2016-76528-P and PID2019-104656GB-I00 of the Spanish MICINN/MINECO/FEDER to JH and by the MDM-2016-0687 María de Maeztu Research Excellence Award from the Agencia Estatal de Investigación to the Department of Gene Regulation and Morphogenesis at the CABD. MG-F was supported by FPI studentship BES-2017-081120 from MICINN.

REFERENCES

Abe, N., Dror, I., Yang, L., Slattery, M., Zhou, T., Bussemaker, H., et al. (2015). Deconvolving the recognition of DNA shape from sequence. *Cell* 161, 307–318. doi: 10.1016/j.cell.2015.02.008

Abrams, E. W., and Andrew, D. J. (2005). CrebA regulates secretory activity in the *Drosophila* salivary gland and epidermis. *Development* 132, 2743–2758. doi: 10.1242/dev.01863

Abrams, E. W., Mihoulides, W. K., and Andrew, D. J. (2006). Fork head and Sage maintain a uniform and patent salivary gland lumen through regulation of two downstream target genes, PH4alphaSG1 and PH4alphaSG2. *Development* 133, 3517–3527. doi: 10.1242/dev.02525

Achatz, J. G., Chiodin, M., Salvenmoser, W., Tyler, S., and Martinez, P. (2013). The Acoela: on their kind and kinships, especially with nemertodermatids and xenoturbellids (Bilateria incertae sedis). *Org. Divers. Evol.* 13, 267–286. doi: 10.1007/s13127-012-0112-4

Andrew, D. J., Horner, M. A., Petitt, M. G., Smolik, S. M., and Scott, M. P. (1994). Setting limits on homeotic gene function: restraint of Sex combs reduced activity by teashirt and other homeotic genes. *EMBO J.* 13, 1132–1144.

Aplin, A. C., and Kaufman, T. C. (1997). Homeotic transformation of legs to mouthparts by proboscipedia expression in *Drosophila* imaginal discs. *Mech. Dev.* 62, 51–60. doi: 10.1016/s0925-4773(96)00649-1

Barrow, J. R., Stadler, H. S., and Capecchi, M. R. (2000). Roles of Hoxa1 and Hoxa2 in patterning the early hindbrain of the mouse. *Development* 127, 933–944.

Becker, H., Renner, S., Technau, G. M., and Berger, C. (2016). Cell-Autonomous and non-cell-autonomous function of hox genes specify segmental neuroblast identity in the gnathal region of the Embryonic CNS in *Drosophila*. *PLoS Genet.* 12:e1005961. doi: 10.1371/journal.pgen.1005961

Bell, E., Wingate, R. J., and Lumsden, A. (1999). Homeotic transformation of rhombomere identity after localized Hoxb1 misexpression. *Science* 284, 2168–2171. doi: 10.1126/science.284.5423.2168

Biscotti, M. A., Canapa, A., Forconi, M., and Barucca, M. (2014). Hox and ParaHox genes: a review on molluscs. *Genesis* 52, 935–945. doi: 10.1002/dvg.22839

Chan, S. K., Popperl, H., Krumlauf, R., and Mann, R. S. (1996). An extradenticle-induced conformational change in a HOX protein overcomes an inhibitory function of the conserved hexapeptide motif. *EMBO J.* 15, 2476–2487.

Chang, C. P., Shen, W., Rozenfeld, S., Lawrence, H., Largman, C., and Cleary, M. (1995). Pbx proteins display hexapeptide-dependent cooperative DNA binding with a subset of Hox proteins. *Genes Dev.* 9, 663–674. doi: 10.1101/gad.9.6.663

Condie, B. G., and Capecchi, M. R. (1994). Mice with targeted disruptions in the paralogous genes hoxa-3 and hoxd-3 reveal synergistic interactions. *Nature* 370, 304–307. doi: 10.1038/370304a0

Daimon, T., Koyama, T., Yamamoto, G., Sezutsu, H., Mirth, C., and Shinoda, T. (2021). The number of larval molts is controlled by hox in caterpillars. *Curr. Biol. CB* 31, 884–891.e883. doi: 10.1016/j.cub.2020.11.017

Damen, W. G., Hausdorf, M., Seyfarth, E. A., and Tautz, D. (1998). A conserved mode of head segmentation in arthropods revealed by the expression pattern of Hox genes in a spider. *Proc. Natl. Acad. Sci. U.S.A.* 95, 10665–10670. doi: 10.1073/pnas.95.18.10665

Dard, A., Reboulet, J., Jia, Y., Bleicher, F., Duffraisse, M., Vanaker, J., et al. (2018). Human HOX proteins use diverse and context-dependent motifs to interact with TALE class cofactors. *Cell Rep.* 22, 3058–3071. doi: 10.1016/j.celrep.2018.02.070

Davenne, M., Maconochie, M., Neun, R., Pattyn, A., Chambon, P., Krumlauf, R., et al. (1999). Hoxa2 and Hoxb2 control dorsoventral patterns of neuronal development in the rostral hindbrain. *Neuron* 22, 677–691. doi: 10.1016/s0896-6273(00)80728-x

de Rosa, R., Grenier, J., Andreeva, T., Cook, C. E., Adoutte, A., Akam, M., et al. (1999). Hox genes in brachiopods and priapulids and protostome evolution. *Nature* 399, 772–776. doi: 10.1038/21631

Denell, R. E., Brown, S. J., and Beeman, R. W. (1996). Evolution of the organization and function of insect homeotic complexes. *Semin. Cell Dev. Biol.* 7, 527–538.

Diederich, R. J., Merrill, V. K., Pultz, M. A., and Kaufman, T. C. (1989). Isolation, structure, and expression of labial, a homeotic gene of the Antennapedia Complex involved in *Drosophila* head development. *Genes Dev.* 3, 399–414. doi: 10.1101/gad.3.3.399

Ekker, S. C., Jackson, D., Kessler, D. P., Sun, B., Young, K. E., and Beachy, P. (1994). The degree of variation in DNA sequence recognition among four *Drosophila* homeotic proteins. *EMBO J.* 13, 3551–3560.

Eriksson, B. J., Tait, N. N., Budd, G. E., Janssen, R., and Akam, M. (2010). Head patterning and Hox gene expression in an onychophoran and its implications for the arthropod head problem. *Dev. Genes Evol.* 220, 117–122. doi: 10.1007/s00427-010-0329-1

Falciani, F., Hausdorf, B., Schröder, R., Akam, M., Tautz, D., Denell, R., et al. (1996). Class 3 Hox genes in insects and the origin of zen. *Proc. Natl. Acad. Sci. U.S.A.* 93, 8479–8484. doi: 10.1073/pnas.93.16.8479

Friedrich, J., Sorge, S., Bujupi, F., Eichenlaub, M. P., Schulz, N. G., Wittbrodt, J., et al. (2016). Hox function is required for the development and maintenance of the *Drosophila* feeding motor unit. *Cell Rep.* 14, 850–860. doi: 10.1016/j.celrep.2015.12.077

Garcia-Fernandez, J. (2005a). Hox, ParaHox, ProtoHox: facts and guesses. *Heredity (Edinb)* 94, 145–152. doi: 10.1038/sj.hdy.6800621

Garcia-Fernandez, J. (2005b). The genesis and evolution of homeobox gene clusters. *Nat. Rev. Genet.* 6, 881–892. doi: 10.1038/nrg1723

Garcia-Fernandez, J., Jiménez-Delgado, S., Pascual-Anaya, J., Maeso, I., Irimia, M., Minguillón, C., et al. (2009). From the American to the European amphioxus: towards experimental Evo-Devo at the origin of chordates. *Int. J. Dev. Biol.* 53, 1359–1366. doi: 10.1387/ijdb.072436jg

Gaufo, G. O., Thomas, K. R., and Capecchi, M. R. (2003). Hox3 genes coordinate mechanisms of genetic suppression and activation in the generation of branchial and somatic motoneurons. *Development* 130, 5191–5201. doi: 10.1242/dev.00730

Gavalas, A., Davenne, M., Lumsden, A., Chambon, P., and Rijli, F. M. (1997). Role of Hoxa-2 in axon pathfinding and rostral hindbrain patterning. *Development* 124, 3693–3702.

Gavalas, A., Ruhrberg, C., Livet, J., Henderson, C. E., and Krumlauf, R. (2003). Neuronal defects in the hindbrain of Hoxa1, Hoxb1 and Hoxb2 mutants reflect regulatory interactions among these Hox genes. *Development* 130, 5663–5679. doi: 10.1242/dev.00802

Gavalas, A., Studer, M., Lumsden, A., Rijli, F., Krumlauf, R., and Chambon, P. (1998). Hoxa1 and Hoxb1 synergize in patterning the hindbrain, cranial nerves and second pharyngeal arch. *Development* 125, 1123–1136.

Gavalas, A., Trainor, P., Ariza-McNaughton, L., and Krumlauf, R. (2001). Synergy between Hoxa1 and Hoxb1: the relationship between arch patterning and the generation of cranial neural crest. *Development* 128, 3017–3027.

Gehring, W. J., Kloter, U., and Suga, H. (2009). Evolution of the Hox gene complex from an evolutionary ground state. *Current Top. Dev. Biol.* 88, 35–61. doi: 10.1016/S0070-2153(09)88002-2

Gendron-Maguire, M., Mallo, M., Zhang, M., and Gridley, T. (1993). Hoxa-2 mutant mice exhibit homeotic transformation of skeletal elements derived from cranial neural crest. *Cell* 75, 1317–1331. doi: 10.1016/0092-8674(93)90619-2

Goddard, J. M., Rossel, M., Manley, N. R., and Capecchi, M. R. (1996). Mice with targeted disruption of Hoxb-1 fail to form the motor nucleus of the VIIth nerve. *Development* 122, 3217–3228.

Grammatopoulos, G. A., Bell, E., Toole, L., Lumsden, A., and Tucker, A. S. (2000). Homeotic transformation of branchial arch identity after Hoxa2 overexpression. *Development* 127, 5355–5365.

Guidato, S., Prin, F., and Guthrie, S. (2003). Somatic motoneurone specification in the hindbrain: the influence of somite-derived signals, retinoic acid and Hoxa3. *Development* 130, 2981–2996. doi: 10.1242/dev.00496

Hanna, L., and Popadic, A. (2020). A hemipteran insect reveals new genetic mechanisms and evolutionary insights into tracheal system development. *Proc. Natl. Acad. Sci. U.S.A.* 117, 4252–4261. doi: 10.1073/pnas.1908975117

Hartenstein, V. (2006). The neuroendocrine system of invertebrates: a developmental and evolutionary perspective. *J. Endocrinol.* 190, 555–570. doi: 10.1677/joe.1.06964

He, S., Del Viso, F., Chen, C. Y., Ikmi, A., Kroesen, A. E., and Gibson, M. C. (2018). An axial Hox code controls tissue segmentation and body patterning in *Nematostella vectensis*. *Science* 361, 1377–1380. doi: 10.1126/science.aar8384

Hirth, F., Hartmann, B., and Reichert, H. (1998). Homeotic gene action in embryonic brain development of *Drosophila*. *Development* 125, 1579–1589.

Hirth, F., Kammermeier, L., Frei, E., Walldorf, U., Noll, M., and Reichert, H. (2003). An urbilaterian origin of the tripartite brain: developmental genetic insights from *Drosophila*. *Development* 130, 2365–2373. doi: 10.1242/dev.00438

Hirth, F., Loop, T., Egger, B., Miller, D. F., Kaufman, T. C., and Reichert, H. (2001). Functional equivalence of Hox gene products in the specification of the tritocerebrum during embryonic brain development of *Drosophila*. *Development* 128, 4781–4788.

Holland, L. Z., Holland, N. D., and Gilland, E. (2008). Amphioxus and the evolution of head segmentation. *Integr. Comp. Biol.* 48, 630–646. doi: 10.1093/icb/icn060

Holland, P. W. H., and Hogan, B. (1988). Spatially restricted patterns of expression of the homeobox-containing gene Hox 2.1. during mouse embryogenesis. *Development* 102, 159–174.

Holland, P. W., Garcia-Fernandez, J., Williams, N. A., and Sidow, A. (1994). Gene duplications and the origins of vertebrate development. *Dev. Suppl.* 125–133.

Horan, G. S., Ramírez-Solis, R., Featherstone, M., Wolgemuth, D., Bradley, A., and Behringer, R. (1995). Compound mutants for the paralogous hoxa-4, hoxb-4, and hoxd-4 genes show more complete homeotic transformations and a dose-dependent increase in the number of vertebrae transformed. *Genes Dev.* 9, 1667–1677. doi: 10.1101/gad.9.13.1667

Hueber, S. D., Weiller, G. F., Djordjevic, M. A., and Frickey, T. (2010). Improving Hox protein classification across the major model organisms. *PLoS One* 5:e10820. doi: 10.1371/journal.pone.0010820

Hughes, C. L., and Kaufman, T. C. (2002a). Exploring the myriapod body plan: expression patterns of the ten Hox genes in a centipede. *Development* 129, 1225–1238.

Hughes, C. L., and Kaufman, T. C. (2002b). Hox genes and the evolution of the arthropod body plan. *Evol. Dev.* 4, 459–499. doi: 10.1046/j.1525-142x.2002.02034.x

Hughes, C. L., Liu, P. Z., and Kaufman, T. C. (2004). Expression patterns of the rogue Hox genes Hox3/zen and fushi tarazu in the apterygote insect *Thermobia domestica*. *Evol. Dev.* 6, 393–401. doi: 10.1111/j.1525-142X.2004.04048.x

Hunter, M. P., and Prince, V. E. (2002). Zebrafish hox paralogue group 2 genes function redundantly as selector genes to pattern the second pharyngeal arch. *Dev. Biol.* 247, 367–389. doi: 10.1006/dbio.2002.0701

In der Rieden, P. M., Mainguy, G., Woltering, J. M., and Durston, A. J. (2004). Homeodomain to hexapeptide or PBC-interaction-domain distance: size apparently matters. *Trends Genet.* 20, 76–79. doi: 10.1016/j.tig.2003.12.001

Janssen, R., Eriksson, B. J., Tait, N. N., and Budd, G. E. (2014). Onychophoran Hox genes and the evolution of arthropod Hox gene expression. *Front. Zool.* 11:22. doi: 10.1186/1742-9994-11-22

Joshi, R., Passner, J., Rohs, R., Jain, R., Sosinsky, A., Crickmore, M. A., et al. (2007). Functional specificity of a Hox protein mediated by the recognition of minor groove structure. *Cell* 131, 530–543. doi: 10.1016/j.cell.2007.09.024

Juergens, G., and Hartenstein, V. (1993). "The terminal regions of the body pattern," in *The Development of Drosophila Melanogaster*, Vol. 1, eds M. Bate and A. Martinez-Arias (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press), 687–746.

Jungbluth, S., Bell, E., and Lumsden, A. (1999). Specification of distinct motor neuron identities by the singular activities of individual Hox genes. *Development* 126, 2751–2758.

Kam, M. K. M., and Lui, V. C. H. (2015). Roles of Hoxb5 in the development of vagal and trunk neural crest cells. *Dev. Groth Differ.* 57, 158–168.

Kaufman, T. C. (1978). Cytogenetic analysis of chromosome 3 in *DROSOPHILA* MELANOGASTER: isolation and characterization of four new alleles of the proboscipedia (pb) locus. *Genetics* 90, 579–596.

Khadjeh, S., Turetzek, N., Pechmann, M., Schwager, E. E., Wimmer, E., Damen, W., et al. (2012). Divergent role of the Hox gene Antennapedia in spiders is responsible for the convergent evolution of abdominal limb repression. *Proc. Natl. Acad. Sci. U.S.A.* 109, 4921–4926. doi: 10.1073/pnas.1116421109

Kitazawa, T., Fujisawa, K., Narboux-Nême, N., Arima, Y., Kawamura, Y., Inoue, T., et al. (2015). Distinct effects of Hoxa2 overexpression in cranial neural crest populations reveal that the mammalian hyomandibular-ceratohyal boundary maps within the styloid process. *Dev. Biol.* 402, 162–174. doi: 10.1016/j.ydbio.2015.04.007

Kokubo, H., Ueno, K., Amanai, K., and Suzuki, Y. (1997). Involvement of the Bombyx Scr gene in development of the embryonic silk gland. *Dev. Biol.* 186, 46–57. doi: 10.1006/dbio.1997.8578

Kribelbauer, J. F., Laptenko, O., Chen, S., Martini, G. D., Freed-Pastor, W. A., Prives, C., et al. (2017). Quantitative analysis of the DNA methylation sensitivity of transcription factor complexes. *Cell Rep.* 19, 2383–2395. doi: 10.1016/j.celrep.2017.05.069

Kribelbauer, J. F., Rastogi, C., Bussemaker, H. J., and Mann, R. S. (2019). Low-Affinity binding sites and the transcription factor specificity paradox in eukaryotes. *Annu. Rev. Cell Dev. Biol.* 35, 357–379. doi: 10.1146/annurev-cellbio-100617-062719

Kuert, P. A., Bello, B. C., and Reichert, H. (2012). The labial gene is required to terminate proliferation of identified neuroblasts in postembryonic development of the *Drosophila* brain. *Biol. Open* 1, 1006–1015. doi: 10.1242/bio.20121966

Li, Y., Nong, W., Baril, T., Yip, H. Y., Swale, T., Hayward, A., et al. (2020). Reconstruction of ancient homeobox gene linkages inferred from a new high-quality assembly of the Hong Kong oyster (*Magallana hongkongensis*) genome. *BMC Genomics* 21:713. doi: 10.1186/s12864-020-07027-6

Lohmann, I., McGinnis, N., Bodmer, M., and McGinnis, W. (2002). The *Drosophila* Hox gene deformed sculpts head morphology via direct regulation of the apoptosis activator reaper. *Cell* 110, 457–466. doi: 10.1016/s0092-8674(02)00871-1

Mallo, M., Wellik, D. M., and Deschamps, J. (2010). Hox genes and regional patterning of the vertebrate body plan. *Dev. Biol.* 344, 7–15. doi: 10.1016/j.ydbio.2010.04.024

Manley, N. R., and Capecchi, M. R. (1997). Hox group 3 paralogous genes act synergistically in the formation of somitic and neural crest-derived structures. *Dev. Biol.* 192, 274–288. doi: 10.1006/dbio.1997.8765

Mayer, G., Franke, F. A., Treffkorn, S., Gross, V., and de Sena Oliveira, I. (2015). "Onychophora," in *Evolutionary Developmental Biology of Invertebrates 3: Ecdysozoa I: Non-Tetraconata*, Vol. 3, ed. A. Wanninger (Vienna: Springer-Verlag), 53–98.

McGinnis, N., Ragnhildstveit, E., Veraksa, A., and McGinnis, W. (1998). A cap 'n' collar protein isoform contains a selective Hox repressor function. *Development* 125, 4553–4564.

McIntyre, D. C., Rakshit, S., Yallowitz, A., Loken, L., Jeannotte, L., Capecchi, M., et al. (2007). Hox patterning of the vertebrate rib cage. *Development* 134, 2981–2989. doi: 10.1242/dev.007567

McNulty, C. L., Peres, J. N., Bardine, N., van den Akker, W. M., and Durston, A. J. (2005). Knockdown of the complete Hox paralogous group 1 leads to dramatic hindbrain and neural crest defects. *Development* 132, 2861–2871. doi: 10.1242/dev.01872

Medved, V., Marden, J., Fescemyer, H., Der, J., Liu, J., Mahfooz, N. S., et al. (2015). Origin and diversification of wings: insights from a neopteran insect. *Proc. Natl. Acad. Sci. U.S.A.* 112, 15946–15951. doi: 10.1073/pnas.1509517112

Merabet, S., Hudry, B., Saadaoui, M., and Graba, Y. (2009). Classification of sequence signatures: a guide to Hox protein function. *Bioessays* 31, 500–511. doi: 10.1002/bies.200800229

Merrill, V. K., Diederich, R. J., Turner, F. R., and Kaufman, T. C. (1989). A genetic and developmental analysis of mutations in labial, a gene necessary for proper head formation in *Drosophila* melanogaster. *Dev. Biol.* 135, 376–391. doi: 10.1016/0012-1606(89)90187-5

Minoux, M., and Rijli, F. M. (2010). Molecular mechanisms of cranial neural crest cell migration and patterning in craniofacial development. *Development* 137, 2605–2621. doi: 10.1242/dev.040048

Minoux, M., Antonarakis, G. S., Kmita, M., Duboule, D., and Rijli, F. M. (2009). Rostral and caudal pharyngeal arches share a common neural crest ground pattern. *Development* 136, 637–645. doi: 10.1242/dev.028621

Noyes, M. B., Christensen, R., Wakabayashi, A., Stormo, G., Brodsky, M., and Wolfe, S. (2008). Analysis of homeodomain specificities allows the family-wide prediction of preferred recognition sites. *Cell* 133, 1277–1289. doi: 10.1016/j.cell.2008.05.023

Panzer, S., Weigel, D., and Beckendorf, S. K. (1992). Organogenesis in *Drosophila* melanogaster: embryonic salivary gland determination is controlled by homeotic and dorsoventral patterning genes. *Development* 114, 49–57.

Parker, H. J., and Krumlauf, R. (2020). A Hox gene regulatory network for hindbrain segmentation. *Curr. Top. Dev. Biol.* 139, 169–203. doi: 10.1016/bs.ctdb.2020.03.001

Parker, H. J., Pushel, I., and Krumlauf, R. (2018). Coupling the roles of Hox genes to regulatory networks patterning cranial neural crest. *Dev. Biol.* 444, S67–S78. doi: 10.1016/j.ydbio.2018.03.016

Pasqualetti, M., Ori, M., Nardi, I., and Rijli, F. M. (2000). Ectopic Hoxa2 induction after neural crest migration results in homeosis of jaw elements in *Xenopus*. *Development* 127, 5367–5378.

Pechmann, M., Schwager, E. E., Turetzek, N., and Prpic, N. M. (2015). Regressive evolution of the arthropod tritocerebral segment linked to functional divergence of the Hox gene labial. *Proc. Biol. Sci.* 282:20151162. doi: 10.1098/rspb.2015.1162

Percival-Smith, A., Sivanantharajah, L., Pelling, J. J., and Teft, W. A. (2013). Developmental competence and the induction of ectopic proboscises in *Drosophila* melanogaster. *Dev. Genes Evol.* 223, 375–387. doi: 10.1007/s00427-013-0454-8

Philippidou, P., and Dasen, J. S. (2013). Hox genes: choreographers in neural development, architects of circuit organization. *Neuron* 80, 12–34. doi: 10.1016/j.neuron.2013.09.020

Pultz, M. A., Diederich, R. J., Cribbs, D. L., and Kaufman, T. C. (1988). The proboscipedia locus of the Antennapedia complex: a molecular and genetic analysis. *Genes Dev.* 2, 901–920. doi: 10.1101/gad.2.7.901

Regulski, M., McGinnis, N., Chadwick, R., and McGinnis, W. (1987). Developmental and molecular analysis of Deformed; a homeotic gene controlling *Drosophila* head development. *EMBO J.* 6, 767–777.

Rijli, F. M., Mark, M., Lakkaraju, S., Dierich, A., Dollé, P., and Chambon, P. (1993). A homeotic transformation is generated in the rostral branchial region of the head by disruption of Hoxa-2, which acts as a selector gene. *Cell* 75, 1333–1349. doi: 10.1016/0092-8674(93)90620-6

Rogers, B. T., and Kaufman, T. C. (1997). Structure of the insect head in ontogeny and phylogeny: a view from *Drosophila*. *Int. Rev. Cytol.* 174, 1–84. doi: 10.1016/s0074-7696(08)62115-4

Rossel, M., and Capecchi, M. R. (1999). Mice mutant for both Hoxa1 and Hoxb1 show extensive remodeling of the hindbrain and defects in craniofacial development. *Development* 126, 5027–5040.

Rusch, D. B., and Kaufman, T. C. (2000). Regulation of proboscipedia in *Drosophila* by homeotic selector genes. *Genetics* 156, 183–194.

Ryoo, H. D., Marty, T., Casares, F., Affolter, M., and Mann, R. S. (1999). Regulation of Hox target genes by a DNA bound Homothorax/Hox/Extradenticle complex. *Development* 126, 5137–5148.

Sanchez-Higueras, C., Rastogi, C., Voutev, R., Bussemaker, H. J., Mann, R. S., and Hombría, J. C. (2019). In vivo Hox binding specificity revealed by systematic

changes to a single cis regulatory module. *Nat. Commun.* 10:3597. doi: 10.1038/s41467-019-11416-1

Sanchez-Higueras, C., Sotillos, S., and Castelli-Gair Hombria, J. (2014). Common origin of insect trachea and endocrine organs from a segmentally repeated precursor. *Curr. Biol. CB* 24, 76–81. doi: 10.1016/j.cub.2013.11.010

Santagati, F., Minoux, M., Ren, S. Y., and Rijli, F. M. (2005). Temporal requirement of Hoxa2 in cranial neural crest skeletal morphogenesis. *Development* 132, 4927–4936. doi: 10.1242/dev.02078

Saurin, A. J., Delfini, M. C., Maurel-Zaffran, C., and Graba, Y. (2018). The generic facet of hox protein function. *Trends Genet.* 34, 941–953. doi: 10.1016/j.tig.2018.08.006

Schwager, E. E., Schoppmeier, M., Pechmann, M., and Damen, W. G. (2007). Duplicated Hox genes in the spider *Cupiennius salei*. *Front. Zool.* 4:10. doi: 10.1186/1742-9994-4-10

Singh, N. P., Kumar, B. D., Paulson, A., Parrish, M., Zhang, Y., Florens, L., et al. (2020). A six-amino-acid motif is a major determinant in functional evolution of HOX1 proteins. *Genes Dev.* 34, 1680–1696. doi: 10.1101/gad.342329.120

Slattery, M., Riley, T., Liu, P., Abe, N., Gomez-Alcala, P., Dror, I., et al. (2011). Cofactor binding evokes latent differences in DNA binding specificity between Hox proteins. *Cell* 147, 1270–1282. doi: 10.1016/j.cell.2011.10.053

Studer, M., Gavalas, A., Marshall, H., Ariza-McNaughton, L., Rijli, F., Chambon, P., et al. (1998). Genetic interactions between Hoxa1 and Hoxb1 reveal new roles in regulation of early hindbrain patterning. *Development* 125, 1025–1036.

Studer, M., Lumsden, A., Ariza-McNaughton, L., Bradley, A., and Krumlauf, R. (1996). Altered segmental identity and abnormal migration of motor neurons in mice lacking Hoxb-1. *Nature* 384, 630–634. doi: 10.1038/384630a0

Sugime, Y., Oguchi, K., Gotoh, H., Hayashi, Y., Matsunami, M., Shigenobu, S., et al. (2019). Termite soldier mandibles are elongated by dachshund under hormonal and Hox gene controls. *Development* 146:dev171942. doi: 10.1242/dev.171942

van Dijk, M. A., and Murre, C. (1994). extradenticle raises the DNA binding specificity of homeotic selector gene products. *Cell* 78, 617–624. doi: 10.1016/0092-8674(94)90526-6

Veraksa, A., McGinnis, N., Li, X., Mohler, J., and McGinnis, W. (2000). Cap 'n' collar B cooperates with a small Maf subunit to specify pharyngeal development and suppress deformed homeotic function in the *Drosophila* head. *Development* 127, 4023–4037.

Zhang, J., and Nei, M. (1996). Evolution of Antennapedia-class homeobox genes. *Genetics* 142, 295–303.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Cell and Developmental Biology

REVIEW
published: 11 August 2021
doi: 10.3389/fcell.2021.718308

Homeotic Genes: Clustering, Modularity, and Diversity

Nikhil Hajirnis[1] and Rakesh K. Mishra[1,2,3]**

[1] CSIR – Centre for Cellular and Molecular Biology (CCMB), Hyderabad, India, [2] AcSIR – Academy of Scientific and Innovative Research, Ghaziabad, India, [3] Tata Institute for Genetics and Society (TIGS), Bangalore, India

Hox genes code for transcription factors and are evolutionarily conserved. They regulate a plethora of downstream targets to define the anterior-posterior (AP) body axis of a developing bilaterian embryo. Early work suggested a possible role of clustering and ordering of *Hox* to regulate their expression in a spatially restricted manner along the AP axis. However, the recent availability of many genome assemblies for different organisms uncovered several examples that defy this constraint. With recent advancements in genomics, the current review discusses the arrangement of *Hox* in various organisms. Further, we revisit their discovery and regulation in *Drosophila melanogaster*. We also review their regulation in different arthropods and vertebrates, with a significant focus on *Hox* expression in the crustacean *Parahyale hawaiensis*. It is noteworthy that subtle changes in the levels of Hox gene expression can contribute to the development of novel features in an organism. We, therefore, delve into the distinct regulation of these genes during primary axis formation, segment identity, and extra-embryonic roles such as in the formation of hair follicles or misregulation leading to cancer. Toward the end of each section, we emphasize the possibilities of several experiments involving various organisms, owing to the advancements in the field of genomics and CRISPR-based genome engineering. Overall, we present a holistic view of the functioning of *Hox* in the animal world.

Keywords: hox, evolution, patterning, gene regulation, bithorax complex, vertebrate hox, modularity and adaptability, homeotic transformation

OPEN ACCESS

Edited by:
Ernesto Sánchez-Herrero, Severo Ochoa Molecular Biology Center (CSIC-UAM), Spain

Reviewed by:
Samir Merabet, Centre National de la Recherche Scientifique (CNRS), France
Leslie Pick, University of Maryland, College Park, United States

***Correspondence:**
Nikhil Hajirnis
nikhil.hajirnis@ccmb.res.in
Rakesh K. Mishra
mishra@ccmb.res.in

Specialty section:
This article was submitted to Morphogenesis and Patterning, a section of the journal Frontiers in Cell and Developmental Biology

Received: *31 May 2021*
Accepted: *22 July 2021*
Published: *11 August 2021*

Citation:
Hajirnis N and Mishra RK (2021) Homeotic Genes: Clustering, Modularity, and Diversity. Front. Cell Dev. Biol. 9:718308. doi: 10.3389/fcell.2021.718308

INTRODUCTION

The development of an animal progresses three-dimensionally across anterior-posterior (AP), dorsal-ventral (DV), and left-right (LR) axes. A combination of various transcription factors, epigenetic regulators, cell receptors, and signaling molecules are involved in the overall development of an organism (François et al., 1994; Beddington and Robertson, 1999; Levin, 2005; Peel et al., 2005; Dequéant and Pourquié, 2008; Basson, 2012; Perrimon et al., 2012; Coutelis et al., 2013; Berenguer et al., 2020). Homeotic genes or *Hox* are one of the significant contributors to bilaterian development and are evolutionarily conserved. They are often present in clusters and code for transcription factors (HOX) that act on the downstream genes to provide identity to developing segments along the AP axis of a bilaterian embryo (Akam et al., 1988; Akam, 1998; Lewis, 1998).

A series of genetic crossings and recombinations lead to the discovery of *Hox* in *Drosophila melanogaster*. Interestingly, all the genes were mapped to the right arm of the third chromosome in the fruit fly. Further, the genes were clustered together in two complexes of ∼300 Kb

each – Antennapedia complex (ANT-C) and bithorax complex (BX-C), named after the initial phenotypes obtained for different alleles in both complexes. Strikingly, the arrangement of these genes on the chromosome followed an order that was collinear to the segments affected in mutants from anterior to posterior direction. Thus, in the fly, the concept of spatial collinearity was established for *Hox* expression (Lewis, 1978, 1998; Kaufman et al., 1990). This expression is under the control of several *cis*-regulatory elements (CREs) that cluster together to form *cis*-regulatory modules (CRMs) (Peifer and Bender, 1986; Peifer et al., 1987; Celniker et al., 1989; Martin et al., 1995; Maeda, 2009; Chopra, 2011; Bekiaris et al., 2018). Series of discoveries toward the turn of the 20th century showed the presence of *Hox* in all bilaterians and even in cnidarians (Ferrier et al., 2000; Kourakis and Martindale, 2001; Ferrier and Minguillón, 2003; Ikuta et al., 2004; Duboule, 2007; Mooi and David, 2008; Mallo et al., 2010; Ikuta, 2011; Janssen et al., 2014; Fritsch et al., 2015; Schiemann et al., 2017; Wanninger and Wollesen, 2019; Nong et al., 2020). The transcription factors coded by these genes have a conserved helix-turn-helix motif-containing DNA binding domain. The domain binds to DNA in a sequence-specific manner and is called the homeodomain due to its discovery in the factors coded by *Hox*. Many transcription factors in addition to HOX across animals, plants, and fungi have the homeodomain (McGinnis et al., 1984b,a; Scott and Weiner, 1984; Suzuki and Yagi, 1994; Williams, 1998; Holland, 2001; Holland et al., 2007; Son et al., 2020). Therefore, all Hox genes are homeobox genes, but all homeobox genes are not Hox genes. The complex interplay of HOX proteins with other players in the system contributes toward diversity in the animal kingdom (Akam et al., 1988; Akam, 1998; Lewis, 2007; Holland, 2015; Rogers, 2020).

In the current review, we discuss the arrangement and copies of Hox genes in different organisms. We then revisit their discovery and regulation in *D. melanogaster,* subsequently commenting upon their *cis*-regulation in vertebrates. Further, the review highlights the presence of these genes in other arthropods and their expansion in vertebrates, with a significant focus on *Hox* expression in *Parhyale hawaiensis*. The crustacean is an emerging model organism with established gene-editing techniques such as CRISPR-Cas9 to decipher the role of *Hox*, adding them to the league of other classical models, including fruit fly, zebrafish, or mouse (Martin et al., 2016; Sun and Patel, 2019). Notably, the function of these genes is not limited to segment identity and homeotic transformations (Castelli-Gair Hombría and Lovegrove, 2003). We also underline many upcoming reports that describe their role in tissue homeostasis, embryonic cell fate determination, organogenesis including abdominal epithelium in flies or hair follicles in mammals, maintenance of stem cells niche, and misregulation leading to cancer (Lewis, 2000; Awgulewitsch, 2003; Shah and Sukumar, 2010; Estacio-Gómez and Díaz-Benjumea, 2014; Singh and Mishra, 2014; Domsch et al., 2019). Toward the end of each section, we emphasize the possibilities of novel experiments to understand the regulation and functioning of Hox genes in different organisms. This largely owes to the recent advances in genomics and genome editing technologies, including CRISPR-Cas9. We thus present a bird' eye view of the Hox field and prospective investigations required to understand their role in various organisms.

HOX CLUSTERS: ARRANGEMENT, POSITIONING, AND DUPLICATIONS

The animal kingdom has diverse body forms, symmetries, and developmental axes. *Hox* are one of the key contributors to this diversity as they provide identity to different segments during embryonic development, are involved in tissue homeostasis and organ positioning, and help in maintaining cellular identities post-embryonic development (Lewis, 2000; Castelli-Gair Hombría and Lovegrove, 2003; Lovegrove et al., 2006; Mallo et al., 2010; Sánchez-Herrero, 2013; Papagiannouli and Lohmann, 2015; Hrycaj and Wellik, 2016; Domsch et al., 2020). They are present in cnidarians with ancestral elements of the anterior and posterior Hox genes (Chourrout et al., 2006; Ikuta, 2011; Gaunt, 2018; Rentzsch and Holstein, 2018; Nong et al., 2020). During evolution, bilaterians acquired another set of central Hox genes and formed a complete set of genes responsible for the animal development across the anterior-posterior body axis (Chourrout et al., 2006; Hrycaj and Wellik, 2016). Usually, Hox genes are present in a cluster and exhibit spatial collinearity; the genes present in one end of the cluster are expressed in the anterior-most regions (or segments) of the developing embryo. At the same time, the genes present on the opposite end are responsible for posterior development (Gaunt, 2015). However, this is not universally true.

The Hox genes of California two-eyed octopus, *Octopus bimaculoides*, are completely dispersed across the genome (Albertin et al., 2015). Other than the octopus, most other bilaterians show clustering of at least two Hox genes in *cis*-. For example, in *Parhyale hawaiensis*, a crustacean, some of the *Hox* are arranged in clusters of two and four genes. However, the detailed arrangement of all Hox genes in *Parhyale* remains elusive due to the absence of long contigs (Serano et al., 2016). Even in a marine chordate, *Ciona intestinalis*, *Hox* appear to be present in an exceptionally dispersed cluster, or they could even be disseminated across the genome (Spagnuolo et al., 2003; Ikuta et al., 2004). *D. melanogaster* has a partially contiguous arrangement of *Hox*. As mentioned earlier, the Hox genes cluster of *Drosophila* is split into two complexes with 5 and 3 *Hox* in them. Both complexes are around 300 Kb in length and are separated by a distance of ∼9 Mb (Dessain and McGinnis, 1993; Rogers, 2020). Other than *Drosophila*, the red fluor beetle, *Tribolium castaneum*, has been a subject of extensive studies for patterning and evolution in insects. Both the insects have similar expressions of *Hox* orthologs in anterior-posterior segments. However, their arrangement is quite different in the genome. *T. castaneum Hox* are organized in a single tight cluster as opposed to the split found in *Drosophila* (Beeman, 1987; Telford, 2000; Brown et al., 2002; Shippy et al., 2008). In other organisms such as the starfish, *Acanthaster planci*, and sea urchin *Strongylocentrotus purpuratus*, *Hox* are present in a cluster. Still, either their orientation is altered, or they have re-ordered arrangement when compared to the majority

of bilaterians that follow collinearity. In *S. purpuratus*, posterior *Hox* (*Hox11/13*) have relocated to positions analogous to central *Hox* and vice-versa (Howard-Ashby et al., 2006; Baughman et al., 2014). In contrast, the Hox genes of cephalochordate amphioxus, *Branchiostoma floridae*, are present as a single intact cluster in the order of their evolutionary homologs, along the anterior-posterior body axis. It is the most cohesive cluster of *Hox* discovered in the animal kingdom from *Hox1* to *Hox14*. Later, analysis of the region between *Hox14* and *EvxA – EvxB* led to the finding of another paralogous posterior Hox gene called *Hox15*. Thus, the cluster of Hox genes in amphioxus is by far the largest intact cluster in terms of the number of Hox genes and spans 470 Kb (Holland et al., 2008). The above examples suggest that animals have varied arrangements of Hox genes as they underwent multiple combinations of convergent and divergent evolutionary processes throughout the tree of life (**Figure 1A**).

In several organisms, Hox genes are present in multiple copies of paralogous genes. For instance, the cluster of Hox genes in the annelid *Helobdella robusta* appear fragmented along with varying copies of different homeotic genes. This is especially true for anterior and central *Hox* orthologs such as *Deformed* (*Dfd*) and *Sex combs reduced* (*Scr*) present in two and five copies, respectively. The *Leech homeobox gene* (*Lox4*) is also present in two copies, whereas orthologs like *Proboscipedia* (*Pb*), *Ultrabithorax* (*Ubx*), and *abdominal-A* (*abd-A*) appear completely absent. The posterior Hox gene, *Post2*, is also present in three copies (Kourakis and Martindale, 2001).

Similarly, the Chinese scorpion, *Mesobuthus martensii,* has two sets of Hox genes, with one of the sets being more clustered than the other. Interestingly, the duplication of Hox genes in scorpions is associated with variation in and extension of the posterior-most segments of the animal, including telson (Di et al., 2015). Vertebrates have at least four different paralogous complexes of Hox genes clusters. Each complex has a different number of Hox homologs arranged in a tight cluster of ∼100 Kb. The clustering also follows spatial collinearity like its invertebrate counterparts. In addition to that, vertebrate Hox genes are also expressed in a temporally collinear manner. The genes present in one end of the complex are expressed earlier during embryonic development and vice-versa. The different complexes work independently, as well as in concert, to fine-tune the growth of a developing embryo (**Figure 1B**; Burke et al., 1995; Medina-Martínez et al., 2000; Suemori and Noguchi, 2000; Spitz et al., 2001; Kmita et al., 2005; Tschopp et al., 2009; Yamada et al., 2021).

In conclusion, the Hox genes are present in different positions and numbers across the genome, from an atomized and dispersed manner in octopus to cleanly clustered complexes in vertebrates (**Figures 1A,B**). It, therefore, becomes important to understand the significance of clustering and ordering of Hox genes in different organisms. Many of the available genome sequences still lack chromosome level assemblies. With the advancement of long-read nanopore sequencing and the use of proximity ligation assays like Hi-C, it is possible to achieve chromosome level assemblies (Wang et al., 2014; Kadota et al., 2020). The ongoing earth biogenome project shall benefit from these techniques, and analysis of Hox genes arrangement in different animals will help us better understand their clustering and ordering throughout the tree of life (Lewin et al., 2018). An in-depth overview of known Hox genes clusters and their arrangement across different organisms is nicely covered in a review by Stephen Gaunt (2018).

DROSOPHILA HOX COMPLEX: A SPLIT THAT UNIFIED THE FIELD

Homeotic genes were discovered by Ed Lewis in *D. melanogaster* in the latter half of the 20th century (Lewis, 1978). There are two clusters of these genes in the fruit fly, the Antennapedia complex (ANT-C) and the bithorax complex (BX-C). The ANT-C is responsible for the identity of anterior segments of the fly from the head through thoracic segment 2 (T2) and has five Hox genes. In the proximo-distal arrangement concerning centromere-telomere, these genes are ordered as *labial* (*lab*), *Proboscipedia* (*Pb*), *Deformed* (*Dfd*), *Sex comb reduced* (*Scr*), and *Antennapedia* (*Antp*). The BX-C has three genes in the centromeric proximo-distal order of *Ultrabithorax* (*Ubx*), *abdominal-A* (*abd-A*), and *Abdominal-B* (*Abd-B*). These genes provide identity to the posterior two-thirds of the fly's body axis from T3 to abdominal segment 8/9 (A8/9), which is the terminal segment in the fly (**Figure 2A**; Lewis, 1978, 1998; Kaufman et al., 1990; Dessain and McGinnis, 1993). It is noteworthy that there are various CRMs for each Hox gene in the fly. These CRMs consist of numerous regulatory elements, including enhancers, initiators, insulators or boundary elements (BE), Polycomb/Trithorax response elements (P/TRE), and promoter tethering sequences (PTS), that together orchestrate the segment-specific expression of these genes (**Figures 2B,D**; Celniker et al., 1989, 1990; Simon et al., 1990; Sánchez-Herrero, 1991; Castelli-gair et al., 1992; Muller and Bienz, 1992; Mishra and Karch, 1999; Bender and Hudson, 2000; Calhoun and Levine, 2003; Lin et al., 2003; Mihaly et al., 2006; Iampietro et al., 2010; Li et al., 2015).

There are nine CRMs of the three BX-C genes in the order *anterobithorax/bithorax* (*abx/bx*) and *bithoraxoid/postbithorax* (*bxd/pbx*) for *Ubx*, *infra-abdominal2* (*iab2*), *iab3* and *iab4* for *abd-A*, and *iab5*, *iab6*, *iab7,* and *iab8/9* for *Abd-B*. Each of the CRM drives segment-specific expression of the associated gene in embryonic parasegment 5 (PS5) through PS14, corresponding to segments T3 through A8/9 in the adult fly. Deletions of CRMs cause loss of function (LoF) phenotypes for the associated Hox genes and lead to anteriorization of respective segments. For example, deletion of *iab5* causes homeotic transformation of A5 to A4. The mutant has two copies of A4 after A3 that follow the normal occurrence of A6, A7, and genitalia (A8/9) (**Figure 2E**; Martinez-Arias and Lawrence, 1985; Peifer and Bender, 1986; Turner and Kaufman, 1987; Galloni et al., 1993; Casares and Sanchez-Herrero, 1995; Hendrickson and Sakonju, 1995; Martin et al., 1995; Bender and Hudson, 2000; Bae et al., 2002; Estrada et al., 2002; Deutsch, 2004; Mihaly et al., 2006; Starr et al., 2011). Further, chromatin domain boundaries separate the CRMs of the BX-C. These include Front-ultraabdominal (Fub) that separates *bxd/pbx* from *iab2*, Mis-cadastral pigmentation (MCP) separating *iab4* and *iab5*, Frontabdominal6 (Fab6) between *iab5* and *iab6*, (Fab7) demarcating the domains of *iab6* and *iab7*, followed by (Fab8), which is present between *iab7* and *iab8/9*

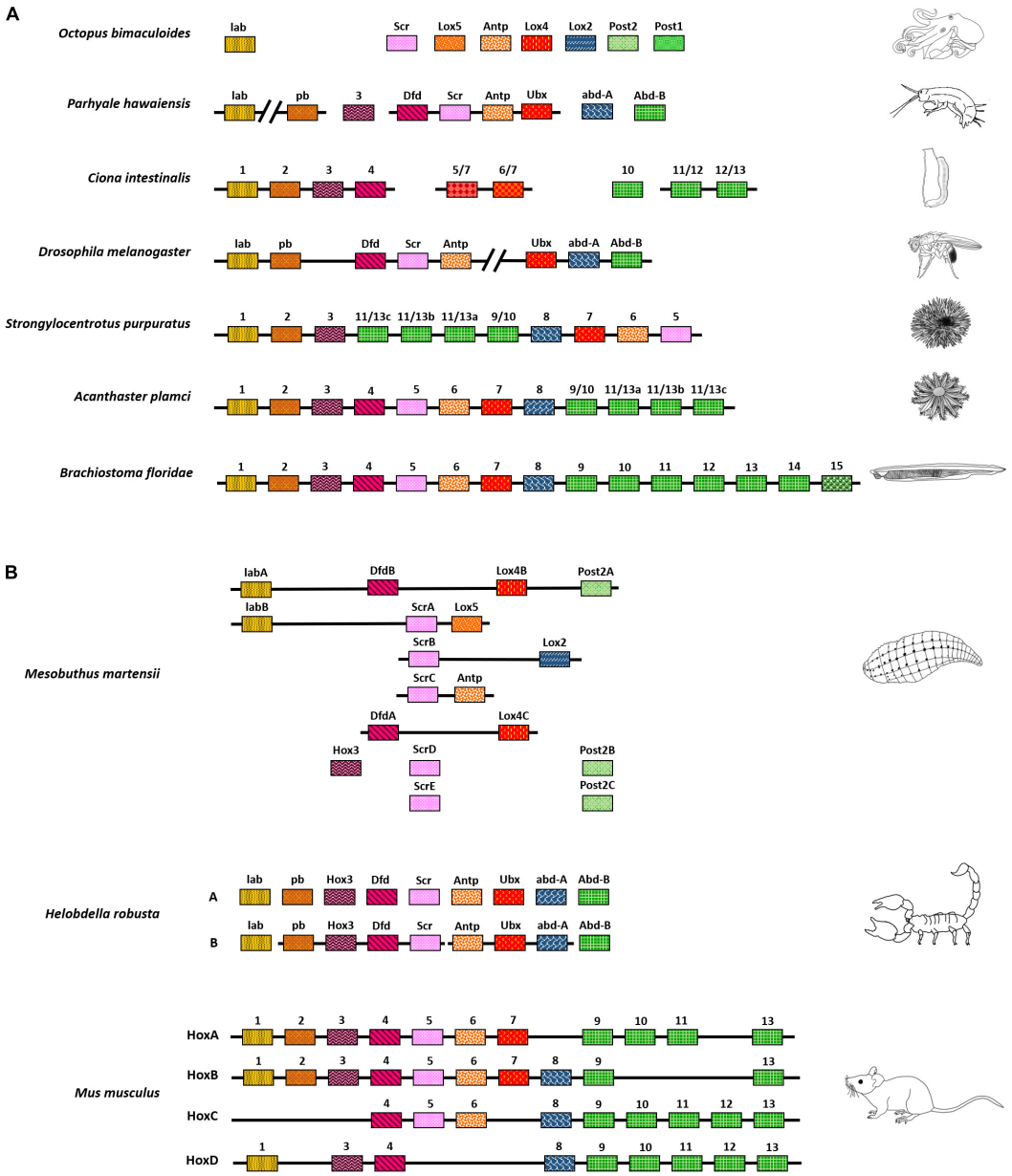


FIGURE 1 | Chromosomal arrangement of Hox genes. **(A)** Variety of *Hox* arrangement observed in different animals, from a completely scattered arrangement in octopus to a fully intact *Hox* cluster in Amphioxus. **(B)** *Hox* have also undergone duplication in several organisms. They have fragmented arrangements in animals such as the freshwater leech or are present in an intact cluster, as seen in vertebrates. The depiction here is independent of phylogeny and represents the order of clustering. Also, the organisms shown here do not represent their respective phyla.

(**Figure 2B**). These BEs maintain the autonomous domains of functioning for different CRMs and genes. In contrast to the LoF phenotypes of CRM deletions, the deletions of BEs cause gain of function phenotypes for the associated Hox genes leading to posteriorization of the related segments. This phenotype is due to the ectopic activation of posterior CRM and its prevalence over the anterior one. For instance, deletion of the chromatin domain boundary, Fab7 leads to the homeotic transformation of A6 to A7 as depicted in **Figure 2E** (Simon et al., 1990; Karch et al., 1994; Hagstrom et al., 1996; Zhou et al., 1996; Mihaly et al., 1998; Mishra and Karch, 1999; Muller et al., 1999; Barges et al., 2000; Schweinsberg et al., 2004; Bender and Lucas, 2013; Postika et al., 2018, 2021). Furthermore, multiple P/TREs adjacent to the BEs and inside CRMs maintain the repressed or activated states of associated CRMs. A combination of boundaries and PREs maintain the distinct autonomy of CRMs wherein the PREs are known to function via DNA kissing (Simon et al., 1993; Chan et al., 1994; Mishra et al., 2001; Vazquez et al., 2006; Lanzuolo et al., 2007; Bantignies et al., 2011; Négre et al., 2011; Singh and Mishra, 2015). The CRMs of BX-C are also present in a spatially

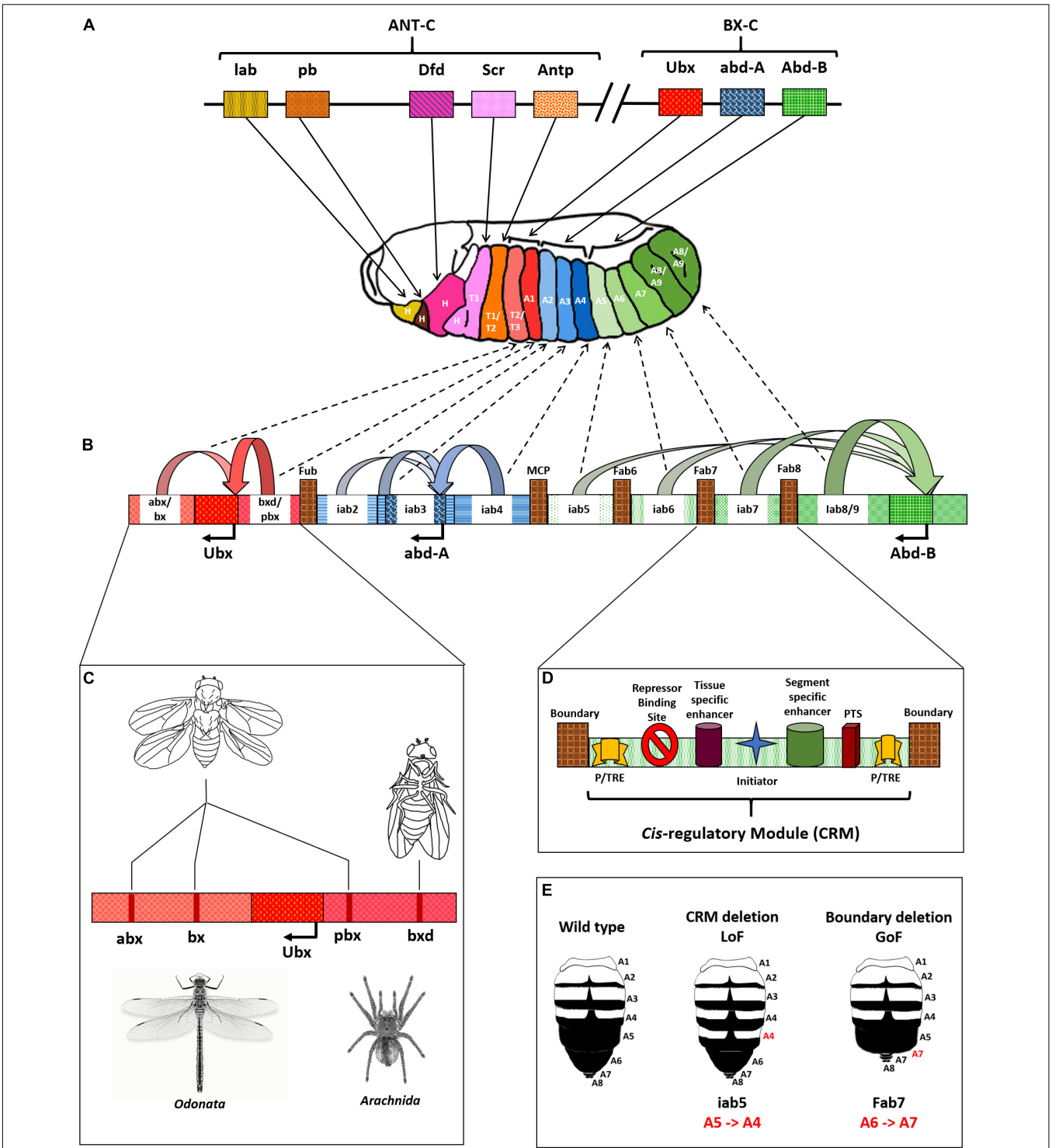


FIGURE 2 | *Drosophila* Hox complex and regulation of BX-C. **(A)** *Drosophila* Hox genes are split into two complexes, Antennapedia complex (ANT-C), and bithorax complex (BX-C), as shown. Each gene is responsible for providing identity to a specific segment, as indicated by bold arrows. **(B)** *Cis*-regulatory modules (CRMs) of the BX-C cause differential expression of Hox genes in a segment-specific manner. The genes influenced by their CRMs are shown as curved arrows with respective colors, and the dotted arrows indicate the segments they influence. **(C)** Representation of *cis*-regulatory module with different elements including boundary/insulator, Polycomb/Trithorax Response Elements (P/TRE), and promoter tethering sequences (PTS). **(D)** Deletion mutations in CRMs of *Ubx* leading to phenotypes that look similar to an odonate, like dragonfly or an arachnid, like spider. **(E)** Loss and gain of function mutations in the abdominal region of *Drosophila* due to deletions of CRMs or BEs.

collinear manner in tune with their associated genes (Lewis, 1978, 1998; Maeda, 2006). **Figure 2** summarizes the arrangement of *D. melanogaster Hox* and the elements of the bithorax complex. Notably, the significance of positioning of CRMs in a collinear manner is still elusive. A significant merit could be the sequential regulation of the Hox genes by upstream regulators as proposed in the open for business model of BX-C regulation (see next section) (Maeda and Karch, 2007, 2015; Kyrchanova et al., 2015).

Since Hox genes provide identities to a developing segment, altering the levels of these genes can tip the scale in favor of distinct traits gained or lost during evolution, albeit they are not the sole drivers of the process (Ho et al., 2009; Holland, 2015). For example, mutations in the CRMs of the *D.mel Ubx* gene manifest fascinating phenotypes. A triple deletion mutant for *abx*, *bx,* and *pbx* causes homeotic transformation of T3 into a copy of T2. The T3 of flies possesses a pair of rudimentary wings called halteres that help maintain balance during flight (Lewis, 1978; Dickinson et al., 1999; Yarger and Fox, 2016). In the triple mutant, the halteres get completely transformed into wings, and the posterior thorax attains the morphology of the anterior one resulting in a fly with four wings, compared to a pair of wings and halteres in normal conditions (**Figure 2C**). Since the CRMs maintain required levels of *Ubx* in T3, their absence leads to a lack of expression of the gene. This loss of function causes T3 to anteriorize into a copy of T2 (Little et al., 1990; Martínez-Laborda et al., 1996). It was a remarkable achievement for two reasons – (1) All three mutations were within a span of 100 kb of each other and were therefore extremely difficult to obtain in *cis*-through the genetic crossing. (2) The fruit fly, a dipteran, looks strikingly similar to an odonate like dragonfly or damselfly with four distinct wings (Lewis, 1978). A combination of three intergenic mutations led to the development of body morphology that diverged almost 500 million years before the arrival of dipterans (**Figure 2C**; Misof et al., 2014). Similarly, flies hemizygous for *bxd* have a partial transformation of A1 into a copy of T3, resulting in a fly with four pairs of legs instead of three. This feature is similar to an arachnid that includes spiders, scorpions, and ticks (**Figure 2C**; Shultz, 1989).

The presence of intact CRMs juxtaposed with genes would ensure that they provide coordinated expression during embryonic development. This is evident from the case of *Drosophila buzzati*, where the gene *labial* (*lab*), an anterior gene, is relocated to a position nearer to *abd-A* and *Abd-B*, the genes that define the posterior development of the fly. Nevertheless, the expression pattern for all Hox genes remains similar to *D. melanogaster*. The rearrangement of the *lab* locus was attributed to the presence of two transposable elements, ISBu2 and ISBu3, that stabilized over generations. These transposons together flank the gene *lab* and its associated non-coding elements. So, the overall arrangement of transposons, associated non-coding elements, and absence of any other coding gene indicate the functional intactness of the locus (Negre et al., 2003). The *D. buzzati lab*, hence, still expresses in the anterior part of the body, unlike its neighbors *abd-A* and *Abd-B* (**Figure 3A**).

The process of *cis*-regulation can be effectively carried out even in the presence of a non-related DNA element in between. For example, few of the *cis*-regulators of the *Dmel Scr* gene are present after a non-homeotic gene, *ftz* (Gindhart et al., 1995; Calhoun and Levine, 2003). The *ftz* gene is, however, flanked by two strong boundary elements SF1 and SF2, that presumably loop out the gene and its associated regulators, thus, facilitating proper interaction of *Scr* enhancer, T1, with the *Scr* gene (**Figure 3B**; Nègre et al., 2010; Li et al., 2015). This process is similar to insulator bypass events, observed in BX-C in the presence of boundary elements like MCP, Fab-7, or Fab-8 (Sipos et al., 1998; Mishra and Karch, 1999; Kyrchanova et al., 2011, 2015, 2019).

Though the genes and CRMs can together relocate to various positions across the genome or be reorganized by chromatin domain boundaries, an arbitrary split in the middle of CRMs is deleterious. This is apparent from the famous *Antennapedia* mutant, $Antp^{73b}$. The *Dmel Antp* gene has two promoters, P1 and P2. A breakpoint of 45 Kb upstream of P2 separates it from P1 and results in an inversion that repositioned P2 around 160 Kb away from its original locus. The inversion also leads to repositioning a non-specific promoter of an uncharacterized gene, *responsible for dominant* phenotype (*rfd*), in the *Antp* locus. This promoter (P^{rfd}) causes ectopic expression of *Antp,* leading to a gain of function phenotype, characterized by the homeotic transformation of antennae and arista in the fly into a pair of legs (**Figure 3C**). Embryos homozygous for $Antp^{73b}$ die early during development. These findings support the theory that ectopic promoters can drive the expression of nearby genes in a non-specific manner (Laughon et al., 1986; Schneuwly et al., 1987). Along with gaining insights into the regulation of BX-C, *Scr* locus, and *Antp* associated dominant phenotype, the understanding of the Hox complex in *Drosophila* was pivotal for dissecting the embryonic development of an organism and also led to a better understanding of crucial facets of gene regulation (**Figure 3**). A plethora of subsequent studies in the following decades after the discovery of *Hox* revealed their existence in all bilaterians as well as cnidarians (Burke et al., 1995; Brooke et al., 1998; Peterson et al., 2000; Ferrier and Holland, 2001; Samadi and Steiner, 2010; Ikuta, 2011; Gaunt, 2018). Deciphering the functioning of the *Drosophila* Hox genes complex, in particular, the BX-C, led to a better understanding of embryonic development, molecular biology, patterning, and evolution. Welcome Bender rightly proposed that the regulation of BX-C should enter textbooks at par with *lac* operon, phage transcription, and yeast mating-type (Bender, 2020).

CLUSTERING, *CIS*-REGULATION, AND REMOTE CONTROLS OF *HOX* EXPRESSION

Segment-specific activation and expression of Hox genes are important for segment identity. Transcription factor coding genes including Gap, Pair-rule, and segment polarity genes act upstream of Hox genes and regulate their expression via associated CRMs in insects (Capdevila and Garcia-bellido, 1981; Reinitz and Levine, 1990; Kornberg and Tabata, 1993; Casares and Sanchez-Herrero, 1995; Drewell et al., 2014). As mentioned earlier, there are nine CRMs in the BX-C that direct expression

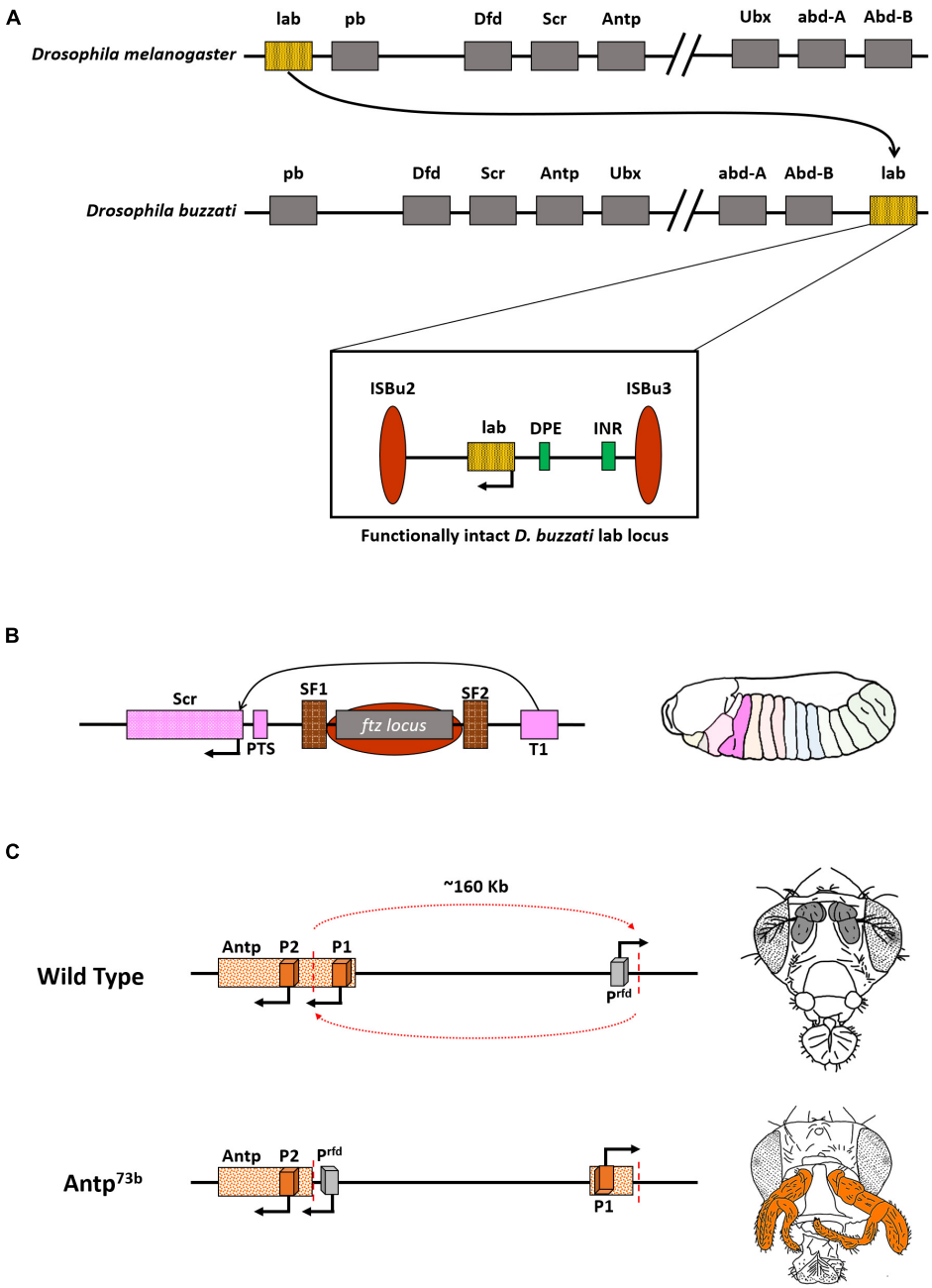


FIGURE 3 | Functional CRMs can accommodate changes in gene order. **(A)** Representation of Hox in *D. melanogaster* and *D. buzzati*. Note that the gene *labial* (*lab*) is present next to *Abdominal-B* (*Abd-B*), and the split is between *Ubx* and *abd-A,* wherein *Ubx* is now a part of ANT-C. **Inset:** Functional intactness of *lab* locus despite rearrangement owing to two transposons, ISBu2/3, that flank the regulatory cassette of the gene. **(B)** Illustration depicting regulation of Scr gene bypassing ftz locus. The ftz gene and regulators are flanked by boundary elements SF1 and SF2. T1 is the enhancer for Scr. **Right:** Domain expression of Scr in a normal condition. **(C)** Inversion involving *Antp* (*Antp*^{73b}) spanning ∼160 Kb juxtaposes the gene to an ectopic promoter that drives its expression in antennae and arista of the flies (colored bold gray in wild type fly). This leads to the homeotic transformation of antennae to legs (highlighted in orange).

levels and patterns of *Ubx*, *abd-A,* and *Abd-B* in a segment-specific manner.

These regions are tightly regulated. Probing the chromatin landscapes of Hox locus has shed some light on their mode of regulation. Segment-specific ChIP-seq for H3K27me3 repressive marks on *Drosophila* BX-C has pinpointed regions that were sequentially de-methylated from anterior to posterior segment in the fly embryo. For instance, in the head, the BX-C is marked with H3K27me3, coinciding with the absence of expression of all the three genes in the complex. While in A1, the *Ubx* domain lacked H3K27me3 marks corroborating with the expression status of *Ubx* in the segment. However, the other two genes of

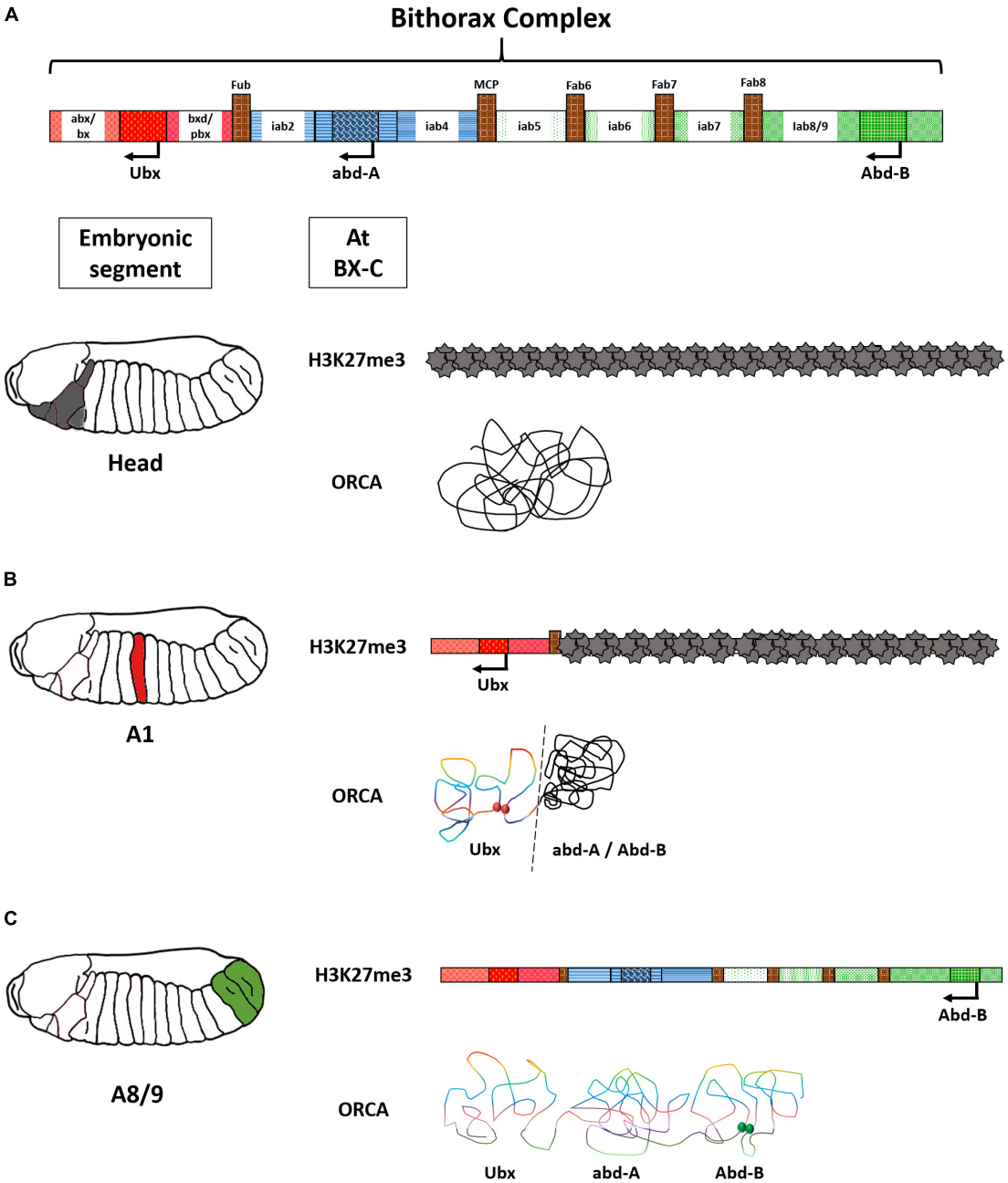


FIGURE 4 | Open for business model of the bithorax complex. **(A)** H3K27me3 marks on the bithorax complex in the head reveal complete repression of the locus. The same is corroborated by a representation of the ORCA image that shows clustering of the entire BX-C in one domain. **(B)** In A1, only the *Ubx* domain is derepressed and forms a distinct loop. Red spheres indicate contact points of Ubx enhancers with its promoter. **(C)** In A8/9, the entire BX-C is de-repressed and forms multiple loop domains. Green spheres indicate contact points of Abd-B enhancers with its regulators. The above image is a conceptual representation of data published by Bowman et al. (2014) and Mateo et al. (2019).

the same complex, *abd-A*, and *Abd-B*, were still carrying the repressive marks (**Figures 4A–C**). This indicated a segment-specific "opening" of BX-C CRMs as one would move from the anterior to the posterior regions in the fly axis and was aptly called the open for business model of the bithorax complex (Bowman et al., 2014; Maeda and Karch, 2015). This model was later reinforced by visualization of chromatin landscape of the BX-C using the optical reconstruction of chromatin architecture (ORCA) technique. It deploys sequential probing of the region of interest on a chromosome, which in this case was ∼320 Kb of the BX-C, by fluorescent probes. The probes are hybridized and washed in a series. They are then coupled with continuous imaging using two customized microscope platforms optimized for HiLo illumination (Mateo et al., 2019). ORCA

is advantageous over conventional confocal microscopy due to the single-molecule resolution possible using the said platform. Like conventional imaging, samples are uniformly illuminated, but a high pass filter rejects the illuminated regions that are outside focus. The extracted data is fused with low-frequency in-focus illumination to render a spatially resolved, sharp image (Ford et al., 2012). ORCA of BX-C revealed interactions of segment-specific enhancers with the associated promoters in an *in vivo* context. Regions devoid of repressive marks were forming a distinct loop, while the ones that remained repressed were forming another closed loop domain (**Figures 4A–D**; Mateo et al., 2019). The clustering of the CRMs and genes in a relatively short region can be an efficient way to moderate Hox levels. In the unusual cases of Hox arrangement like octopus, sea star, or sea urchin, deciphering 3D genome architecture would provide crucial insights into their functioning.

The range of available model organisms limits our current knowledge. Nevertheless, *in silico* and synthetic biology approaches can help in designing experiments of physiological and evolutionary relevance. Crocker et al. (2017) modeled functional enhancers based on the binding sites of various transcription factors across different species of *Drosophila*. They produced several synthetic enhancers which could be validated *in vivo* in a developing fly embryo. However, only a limited number of the predicted enhancers could emulate the expression ability of the native ones (Crocker et al., 2017). This could be because of additional factors like insulators and Polycomb/Trithorax response elements (P/TREs) that contribute to regulatory aspects of the genome. Toward this, Srinivasan et al. developed *in silico* tools to predict chromatin domain boundaries and P/TREs in *Drosophila* and other insects (Srinivasan and Mishra, 2012, 2020). With the ever-expanding availability of genome sequences, such tools can be extended to model regulation of genes, including *Hox,* in a diverse set of organisms (Lewin et al., 2018).

Notably, many of the regulatory elements of the genome, like enhancers and insulators, are known to interact with regions that are several Mbs apart (Long et al., 2016). Despite that, the clustering of CRMs and Hox genes in complex organisms suggests a very strong functional consequence. It is speculated that the order of genes within the Hox complex is important for proper body axis development. However, it may be the order of CRMs that might be equally important.

An intriguing region to understand the significance of relative positioning of CRMs can be the *Abd-B* locus in the BX-C. Each of the *iab*s (CRM) in the region is demarcated by chromatin domain boundaries (BEs) (**Figures 2B,D,E**). For example, *iab5* specifies PS10 (A5) identity and is followed by a BE, Fab6. The BE separates *iab5* from the next CRM *iab6*, which specifies PS11 (A6) of the fly embryo, thus ensuring autonomous domains of the two CRMs (Galloni et al., 1993; Lewis, 2007; Bender et al., 2011; Postika et al., 2021). Together, the four *iab*s of the *Abd-B* locus provide identities to four abdominal segments in the fly from A5 to A8 (terminalia). Hence, the number of these CRMs and their relative positioning in the genome is collinear with the segment they provide identity (Maeda, 2006; Lewis, 2007; Kyrchanova et al., 2015; Maeda and Karch, 2015). Furthermore, many of the BEs are known to function in an orientation-dependent manner. However, most of these studies are done in a transgenic context or a narrow region within the BX-C (Galloni et al., 1993; Martin et al., 1995; Bender and Hudson, 2000; Kyrchanova et al., 2016, 2019). In principle, the iterative arrangement of CRMs and BEs in the *Abd-B* locus is a compelling case to decipher their role in complex systems. An interesting experiment would be to generate targeted inversions and duplications of CRMs in the BX-C and examine the resulting novel phenotypes. The re-arrangements should be developed in a manner that does not affect binding sites for transcription factors, repressors, or chromatin remodelers obtained from existing ChIP data in the modENCODE consortium (Celniker et al., 2009; Nègre et al., 2010; Négre et al., 2011). Moreover, these re-arrangements should not fuse the domains of two nearby genes or known regulators, as indicated in **Figure 5A**. One could harness the potential of Cre-LoxP or FLP-FRT systems to bring about these changes. The recombinase recognition sequences can be knocked in at specific sites using CRISPR/Cas9 (Li et al., 2020). For instance, a reorganized locus with MCP followed by *iab7*, *iab6,* and *iab5* will offer a new playground for *cis-*/*trans-* factors to regulate *Abd-B*. The rearrangement would render *iab7* flanked by MCP and Fab7 in opposite directions, whereas Fab6 and Fab8 boundaries will flank *iab5*. Although the relative positioning of *iab6* would remain the same, but, according to the open for business model of BX-C regulation, either *iab5* and *iab8* will become accessible to *Abd-B* promoter, or *iab7* will be accessible irrespective of re-ordering in PS11 (prospective A6). Such an experiment can unfold the aspects of directionality, ordering, and relative positioning of CRMs within the particular Hox gene locus. Similarly, generating duplications of CRMs like *iab5* and *iab6* will provide a better understanding of the significance of the number of CRMs required to specify a segment (**Figure 5A**). Inversions in several *cis*-regulators in vertebrates have revealed the significance of positioning distal enhancers concerning Hox genes (Kmita et al., 2000; Zakany et al., 2004). Site-specific rearrangements and deletions of vertebrate *cis*-regulators revealed modularity associated with their arrangements and caused changes in the topologically associated domains (TADs) in which they reside. This leads to the ectopic expression of Hox genes in non-specific regions of the limb, thereby suggesting a significant role of the positions of CREs (Fabre et al., 2017). Since BX-C has a spatially collinear arrangement of the CRMs with a clear understanding of their components, the re-engineered locus will provide a deeper understanding of the evolution of CRM positioning and copy number variations (CNVs).

Overall, the situation is perhaps a bit more complex in vertebrates. They have a minimum of 4 *Hox* complexes distributed across different chromosomes. Each complex has its own set of regulators. Their embryonic expression follows spatio-temporal collinearity. This means that the genes present toward one end of the cluster are expressed earlier in the anterior regions. The genes present toward the other end of the cluster are expressed later in time in the posterior regions. So, there is an added temporal aspect of regulation in addition to the pre-existing spatial one. Furthermore, the clustering of *Hox* is more pronounced in vertebrates, with no non-homeotic genes present

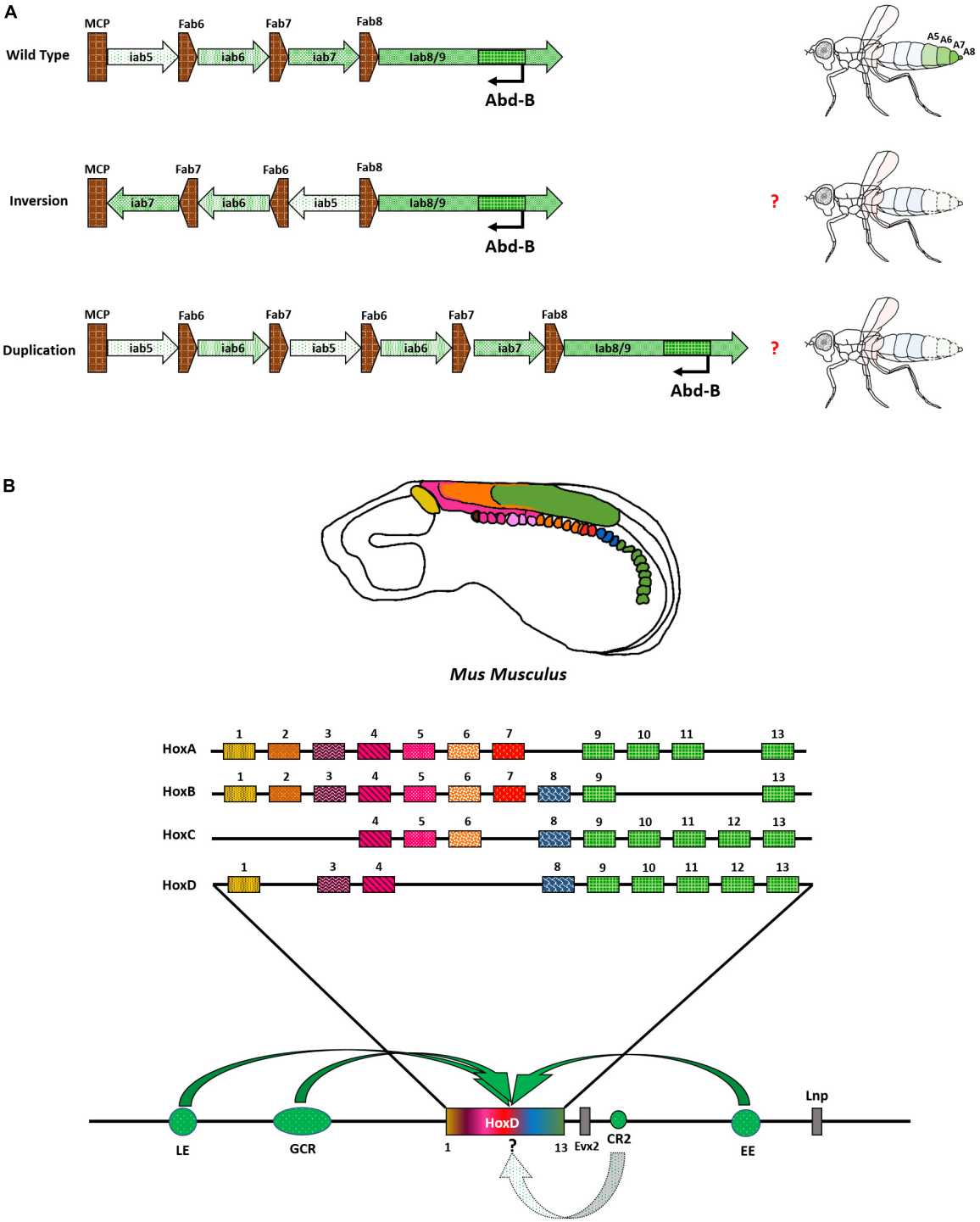


FIGURE 5 | Re-arrangement of *Abd-B* locus and vertebrate Hox complex with CRMs. **(A)** Representation of proposed experiment to re-arrange the CRMs of Abd-B locus in the BX-C. **(B)** Hox genes are expressed as indicated in *Mus musculus.* HoxD locus is shown as a representation of regulatory modules associated with Hox complexes. Bold, curved arrows indicate their approximate presence and interaction with Hox complex (not to scale and point precisely on a particular Hox). The role of CR2 in modulating Hox is still unknown and is represented by a dotted arrow.

in the complex. The intergenic distance is also drastically reduced and the entire *Hox* complex resides within a span of ∼100 Kb. In contrast, both *Hox* complexes in *Drosophila* are larger than 300 Kb. The reduction in the size of the complex can be attributed to the positioning of CRMs of vertebrate *Hox*, outside the cluster on either end, several Kbs away. These regions constitute

the global control regions (GCRs), early enhancers (EE), late enhancers (LE), and many other uncharacterized regulatory elements (Soshnikova, 2014). The tight clustering of *Hox* in vertebrates might also help in robust regulation during secondary axis formation in the limbs when the collinearity is replayed (Soshnikova and Duboule, 2009; Mallo et al., 2010; Mallo and Alonso, 2013; Soshnikova, 2014). Some studies have shown several regions that are ultra-conserved near the *HoxD/Evx* locus of vertebrates. One of these regions, called conserved region 2 (CR2), was shown to have an early enhancer but late repressor activity in a transgenic context in zebrafish, *Danio rerio* (Sabarinadh et al., 2004; Kushawah and Mishra, 2017). The exact mechanism and mode of function of these elements are still unknown. It is also not known whether these regions have an impact on Hox genes. Deletions of these regions in several combinations can help us dissect their significance (**Figure 5B**).

The spatio-temporal regulation of Hox genes in vertebrates has some fascinating offshoots. Marsupials like Tammar wallaby, *Macropus eugenii*, have delayed expression of posterior Hox genes, *HoxA13* and *HoxD13*, attributed to weaker hind limbs in newborn animals. The forelimbs are relatively stronger, which helps them to climb the brood pouch of their parent. The delayed expression of the posterior *Hox* is yet another example of modularity and differential expression, possibly due to differences in clustering and accessibility of CRMs which can be accessed via the genome sequence available for marsupials (Chew et al., 2012; Deakin, 2012).

Similar variations of spatio-temporal regulation can be observed in simpler chordates like amphioxus. Despite being in a tight cluster, the spatially collinear expression of Hox genes is perturbed in *Branchiostoma floridae*. *Hox6*, a central Hox gene, expresses almost ubiquitously across the neural tube, posterior to the cerebral vesicle. While *Hox14*, a posterior Hox gene, is expressed in the most anterior cerebral vesicle. Furthermore, *Hox14* mRNA is also detected in the pharyngeal endoderm. Interestingly, levels of *Hox6* vary greatly in closely related species. Unlike *B. floridae Hox6,* which shows a uniform expression throughout the neural tube, the *B. lanceolatum* homolog expresses in a spatially restricted manner (Pascual-Anaya et al., 2012). This indicates subtle modulations of HOX levels in closely related species and is similar to changes observed in invertebrates. Deep sequencing of flanking regions of Hox loci in multiple organisms along with a Bag-of-Motif analysis to understand protein-DNA interactions can shed light on putative regulatory mechanisms associated with the clustering of CRMs.

In simpler organisms like annelids or mollusks, the arrangement of *Hox* thus seems to be dispensable, but with the evolution of complexity, clustering becomes a necessity for co-regulation.

MODULATING HOX IN ARTHROPODS

The property of a system to separate and re-integrate its components to form a viable system is called modularity. Subtle changes in Hox expression can quickly orchestrate the evolutionary modularity. The studies are not limited to fruit flies. In an amphipod crustacean, *Parhyale hawaiensis*, the interplay between various Hox genes and their ability to act independently was comprehended by a series of sophisticated experiments involving manipulation of Hox levels (Liubicich et al., 2009; Martin et al., 2016; Sun and Patel, 2019).

The amphipod is bilaterally symmetrical and has multiple segments with specialized appendages. A group of metameric segments evolved to perform a common function is called tagma, and the associated evolutionary process is called tagmatization (Abzhanov and Kaufman, 2000). The arrangement of appendages in the order of their occurrence from anterior to posterior segments in *Parhyale* is as follows – feeding appendages (mandible, Mn and maxillipeds Mx, or, gnathopods), claws (T2–T3), forward (T4–T5), and reverse (T6–T8) walking legs (pereopods), swimming appendages (pleopods or swimmerets) in the segments A1 to A3, and appendages for holding substrates (uropods) formed in A4–A6. A simple representation of *P. hawaiensis* tagmatization is depicted in **Figure 6A**.

Recent advances in CRISPR-based gene editing have allowed researchers to perform knock-out experiments in *P. hawaiensis*. Martin et al. (2016) knocked out *Ph Ubx* by CRISPR/Cas9 and observe that the locomotor appendages acquire the identity of feeding appendages (T2–T5 → Mn/Mxp). Further, knocking out a posterior Hox gene *Ph abd-A*, which would otherwise be responsible for forming reverse walking legs in the crustacean, now has them transformed into a copy of forward walking legs (T6–T8 → T4); **Figure 6B**. This was as expected from previous studies in *Drosophila* that indicate the anteriorization of segments in the absence of posterior *Hox*; a property called the posterior prevalence of Hox genes. However, another class of legs called pleopods or swimming appendages was transformed into a copy of posterior appendages, uropods (A1–A3 → A4), in the Ph abd-A knock-out animals. This was an apparent effect of the additive function of *Ph abd-A* to regulate segment identity in either direction along the AP axis. In the anterior regions, *Ph abd-A* works with *Ph Ubx* to develop segments with reverse walking legs (T6–T8). In the posterior regions, *Ph abd-A* functions with *Ph Abd-B* to develop swimming appendages (A1–A3) as depicted in **Figure 6C**. Strikingly, *knock-out* of an even more posterior gene *Ph Abd-B* displayed a spectacular non-linear transformation of uropods into copies of forward walking legs but not swimmerets (A4–A6 → T5; not A3); **Figure 6D**. This suggested that the ABD-B represses *Ph Ubx* in the posterior segments of the *Parhyale* from A1 to A6, whereas *Ph abd-A* expression is independent of the ABD-B levels. *Ph Abd-B* knock-out animals had derepression of *Ph Ubx* in the posterior segments, leading to extreme transformation into forward walking legs. The segment with overlapping domains of *Ph Ubx* and *Ph abd-A* developed reverse walking legs in *Ph Abd-B* knock-outs while swimmerets were altogether absent from the organism (A1–A3 → T8); **Figures 6A,D**. The studies from the crustacean suggest that alongside collinear expression of *Hox*, the co-regulation, inter-regulation, and cross-talk between different HOX cause varying phenotypes. The interplay between these genes brings about diversity in the animal kingdom (Martin et al., 2016).

In addition to the AP axis, a handful of studies also show the role of Hox genes in LR and DV axis determination

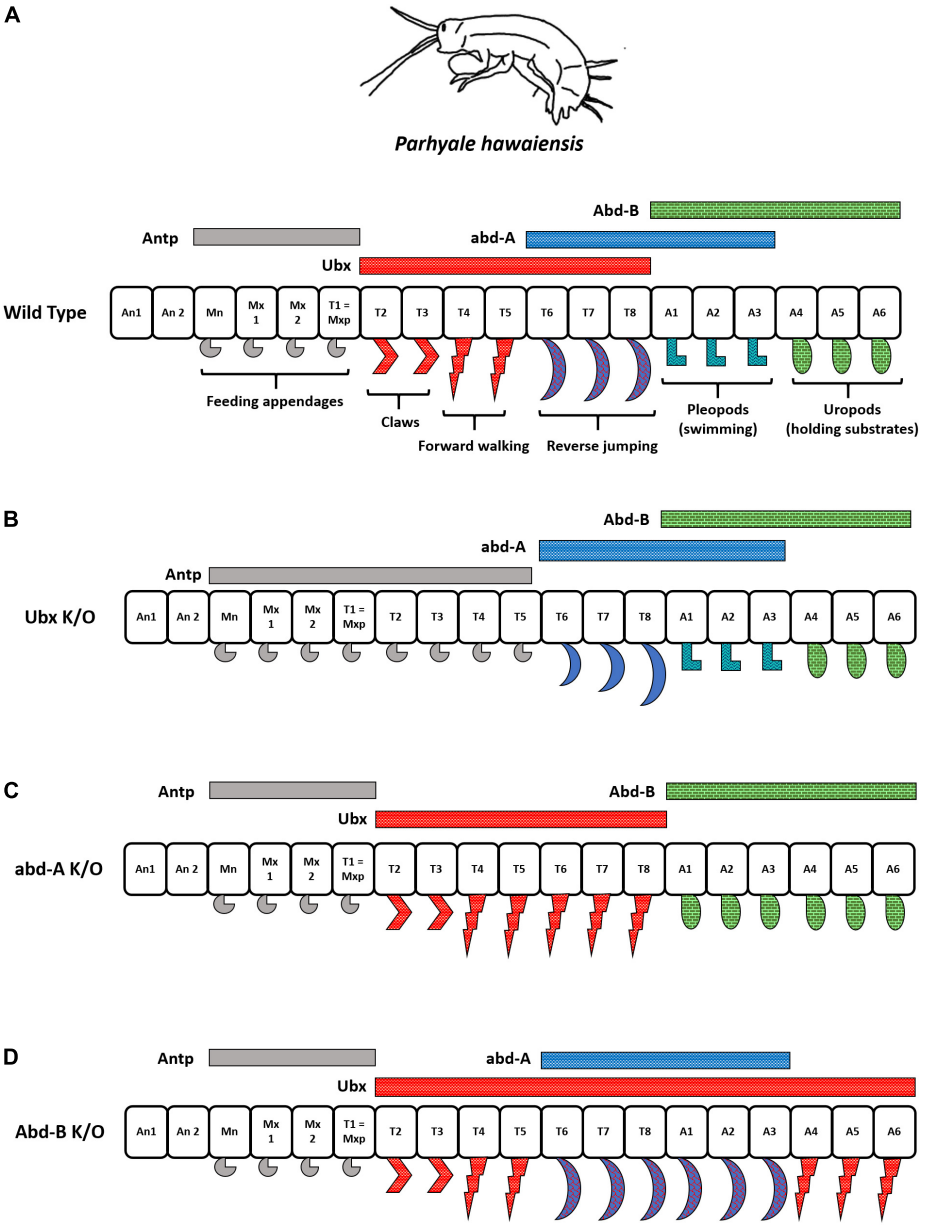


FIGURE 6 | Interplay of different Hox genes in *Parahyale hawaiensis*. **(A)** Hox expression in *Parhyale hawaiensis*. Note that *abd-A* has overlapping regions of functioning with *Ubx* as well as *Abd-B*. **(B)** *Ubx* knock-out animals show anteriorization of forward walking legs. **(C)** abd-A knock-out animals show anteriorization of reverse walking legs and posteriorization of swimming appendages. **(D)** Abd-B knock-out animals show anteriorization of swimming appendages into reverse walking legs and uropods in forward walking legs.

(Thickett and Morgan, 2002; Mohit et al., 2006; Coutelis et al., 2013). An exemplar is the *Dmel Abd-B* regulating *MyosinID* (*MyoID*), a protein responsible for complete dextral (clockwise 360°) rotation of spermiduct around hindgut during metamorphosis. *Abd-B* knocked down flies show partial sinistral (anti-clockwise) rotation to varying degrees that causes male sterility due to rotation of external genitalia (Spéder et al., 2006; Coutelis et al., 2013). Crustaceans such as *P. hawaiensis* develop symmetrically along the LR axis, and early knock-down of *Ph Ubx* in one of the sides causes asymmetrical homeotic transformation of segments, including appendage formation. This was done by injecting morpholinos for *Ph Ubx* knockdown in one of two-celled stage embryo cells. Each cell follows its fate separately across the LR axis of development. Although the system was utilized to compare wild type versus knockdown phenotypes in the same organism (Browne et al., 2005; Liubicich et al., 2009; Pavlopoulos et al., 2009), the study also implies asymmetrical differentiation of body segments upon differential expression of Hox genes. In tune with this, in the *Xenopus* embryo, *HoxC8* expresses asymmetrically along the left-right

axis of development in the lateral plate mesoderm (Thickett and Morgan, 2002). One interesting organism worth probing for Hox genes regulation and determination of the LR axis is the fiddler crab. It is a natural example of left-right asymmetry in appendage formation. The female fiddler crabs have similar-sized left and right feeding appendages. In comparison, males have one of their feeding appendages extraordinarily enlarged. They use this appendage to fight competitors during mating and display handedness (Pardo et al., 2020). A detailed understanding of Hox expression in these organisms can shed light upon the formation of segments in AP and LR axes of development.

Another example of a modified and rather intriguing appendage is the scorpion's tail, including the terminal telson. Scorpions have undergone duplications of Hox genes, which are correlated to the heteronomy of the posterior segments (Sharma et al., 2014). Arizona bark scorpion, *Centruroides sculpturatus*, has 19 Hox genes instead of 10 in its sister groups. The dual copies are expressed in varying degrees from antero-central to telson. These include *Antp*, *Ubx*, *abd-A,* and *Abd-B*. In *C. sculpturatus*, extended-expression of the two copies of *Cs Antp* and *Cs Ubx* is corroborated with enlarged telson in a distinct shape for an appendage. Notedly, the telson is formed posterior to terminalia (anus). It would be interesting to delete one or multiple copies of each of these Hox genes and observe the changes in body patterning. The tagmatization could be affected to the extent that the body form might become less elongated, as is the case with Opiliones, harvestmen, or instigated to form a telson-less scorpion (Sharma et al., 2012). The opposite spectrum of body formation is seen in Tardigrades, in which deletion of several Hox genes correlates with their compact body plan with simpler, repetitive, and less (four) number of segments (Smith et al., 2016).

Other than the levels of HOX, structural modifications in the transcription factors can help in diverse functions. Recent experiments with flies provided evidence of functional conservation of mouse Hox genes. Singh et al. (2020) replaced *labial*, the anterior-most gene in *Drosophila* Hox complex, with *Hox1* from *Mus musculus*. Interestingly, out of the three copies of *Hox1* in the form of *HoxA1*, *HoxB1*, and *HoxD1*, only *HoxA1* could rescue the *labial* knock-out phenotype completely. They also developed animals with chimeric HOX proteins and discovered a six-amino acid C-terminal motif in *HoxA1* essential for its interaction with PBX. The ortholog-specific interaction leads to differential occupancy of *HoxA1* across the genome. This study strongly supports the notion of evolutionary modularity in Hox complex by causing structural changes in HOX that lead to similar yet functionally divergent protein products (Singh et al., 2020).

An ordered arrangement of *Hox* could have played an important role in their sequential co-regulation along the AP axis, as indicated by our understanding of BX-C regulation. One can consider Hox genes as switches to control different electrical equipment at home. They can be present anywhere across the house and can still function, as is the case of an octopus. But clustering on a switchboard gives quick, precise, and perhaps, robust control over the spatio-temporal regulation of Hox genes. This modularity could have been one reason for arthropods to surpass mollusks as the richest bio-diverse species on our planet (Benton, 2010). Many genes are co-regulated in different organisms (Snel et al., 2004). Overall, clustering is more abundant in vertebrates than invertebrates (Elizondo et al., 2009; Ferrier, 2016). Nevertheless, in addition to clustering, the ordering is an important property of Hox complexes that need to be pondered upon. The past decade has witnessed rapid advancements in our understanding of epigenetic factors, inter-genic regulators, and chromatin organization (Narlikar and Ovcharenko, 2009; Hübner et al., 2013; Allis and Jenuwein, 2016; Hug and Vaquerizas, 2018). Understanding them in the context of gene clusters, including Hox complexes, will be riveting. The Hox genes have a tremendous potential to modulate diversity by teaming up with multiple partners and setting a stage for downstream players in various axes. Different combinations of *cis*- and *trans*- regulators together bring about manifold changes that can drive evolution.

HOX GENES: MASTER REGULATORS BEYOND EMBRYOGENESIS AND HOMEOSIS

Classically, mutations in Hox genes are associated with the homeotic transformation of one body segment into another, a process called homeosis (Lewis, 1994). These mutations transformed embryonic segments, and therefore the Hox genes were established as the regulators during embryonic development (Pradel and White, 1998). However, even during embryonic development, Hox genes can still play a non-homeotic role by specifically affecting tissue homeostasis and organogenesis (Castelli-Gair Hombría and Lovegrove, 2003).

Recent studies opened new horizons to understand the role of Hox genes in an organism. A rising number of articles suggest their role beyond homeotic functions and embryonic development (Wang et al., 2009; Estacio-Gómez and Díaz-Benjumea, 2014; Gummalla et al., 2014; Rux and Wellik, 2017). In *D. melanogaster*, prolonged expression of Hox genes beyond embryogenesis is observed in developing larva and pupa (Wang et al., 2009). The three genes of the bithorax complex, *Ubx*, *abd-A*, and *Abd-B,* have defined anterior limits of expression in *Drosophila* larvae. The larva undergoes metamorphosis during pupal stages of development, ultimately eclosing as adults. One key event during this process is autophagy of most of the larval tissues, including the fat body, salivary glands, and trachea. This is further coupled with the differentiation of adult tissues that goes on till eclosion. Interestingly, all the three genes of BX-C, *Ubx*, *abd-A,* and *Abd-B* are expressed in the larval fat body (Marchetti et al., 2003). Down regulation of *Ubx* is accompanied by developmental and starvation-induced autophagy, whereas sustained expression of the Hox gene inhibits autophagy and delays metamorphosis (Banreti et al., 2014).

Like the larval fat body, larval epithelial cells (LECs) also undergo apoptosis during metamorphosis. Further, another group of cells called histoblast nest cells (HNCs) differentiates to form adult abdominal epithelial cells during pupation. Posterior BX-C genes *abd-A* and *Abd-B* have overlapping expressions in the

LECs. Loss of *abd-A* impairs the apoptotic pathway in LECs and cannot be rescued by *Abd-B* alone. Moreover, HNC proliferation is hindered by *abd-A* down regulations, and the cells fail to form a complete epithelium in *abd-A* knocked down pupae. Thus ABD-A is required for both, apoptosis of LECs as well as the proliferation of HNCs to form mature abdominal epithelium in adults. The study showed that ABD-A was present in the LECs and contributed toward development together with the posterior Hox gene product ABD-B, therefore defying the property of posterior prevalence (Singh and Mishra, 2014). The study also contributed to our understanding of Hox genes' modular capacity in an extra-homeotic and extra-embryonic manner.

Similar reports for *Abd-B* were observed in testis development, where it remains active in pre-meiotic spermatocytes. Tissue-specific knockdown of *Abd-B* in adult testes leads to a loss of maintenance of the stem cell niche required to produce normal sperms. This is because ABD-B has direct binding sites on *src42A* and *sec63*, members of *Boss* signaling involved in testes formation and sperm differentiation. *Abd-B* also has an extended effect on the orientation of centrosomes and the division rates of germline stem cells (Papagiannouli and Lohmann, 2015).

Obtaining tissue-specific cells for further studies of Hox was a Herculean task a couple of years back, as one had to do neck-breaking dissections to get ounces of desirable material. Although now, endogenous tagging of Hox genes has solved a lot of such problems. Cell sorting of fluorescently labeled HOX expressing tissues followed by multi-*omics* experiments can help us understand the genome-wide effects of HOX in adult tissues. Domsch et al. (2019) reported an endogenously tagged line for *Ubx* with GFP at the N-terminal. They utilized this resource to establish the role of *Ubx* as a major repressor of factors involved in alternate fate development in mesodermal cells. Sorting GFP expressing cells followed by ChIP and Co-IP experiments helped in a deeper understanding of modalities of *Ubx* functioning. This revealed UBX's ability to cause repression by constantly interacting with a member of Polycomb Repressive Complex protein Pleiohomeotic (PHO) (Domsch et al., 2019). In their recent work, Paul et al. (2021) showed that not only the presence of HOX but also their dosage determines the formation of appendages – in their case, wing appendages.

The extraembryonic roles of *Hox* are more distinct in vertebrates. As early as 2003, it was evident that Hox genes play a role in non-homeotic fashion owing to the near-complete loss of hair formation in mice deficient for *HoxC13*. Although the mouse also had patterning defects, hair growth was uniformly reduced across the body (Awgulewitsch, 2003).

Recent reports showed several *HoxC* genes in the dermal papilla and associated it with regional follicle variation. In a mutant mouse line called *Koala* mutant, a 1 Mb inversion encompassed disintegration of *HoxC4* to *HoxC13* from the main complex leading to their misexpression. CTCF ChiP-seq revealed changes in levels of CTCF binding within the *HoxC* complex and perturbation of topologically associated domains (TADs) (Millar, 2018). Similar deletion studies have identified the role of HoxA genes in mammary gland formation during specific transition developmental periods (Lewis, 2000).

Owing to their multifaceted roles during and after development, levels of Hox proteins need to be tightly regulated. Misexpression of these genes has been observed in various cancers like breast cancer, melanoma, bone cancer, blood cancer, and colorectal cancer (Shah and Sukumar, 2010). Central and posterior Hox genes, *HoxA5* and *HoxD9*, have been implicated in esophageal squamous cell carcinoma. Strikingly, they were found to localize more in the cytoplasm of the mucosa cells in esophageal cancer than in the nucleus in normal cellular conditions (Takahashi et al., 2007). Similarly, ectopic expression of *HoxC6*, *HoxC11*, *HoxD1,* and *HoxD8* are observed in different cases of neuroblastoma (Manohar et al., 1996; Zhang et al., 2007). Overexpression of posterior Hox genes, particularly HoxA9-11, HoxB13, and HoxC10, is linked to the onset and tumor progression of ovarian, cervical, and prostate cancers (Jung et al., 2004; Cheng et al., 2005; Miao et al., 2007; Zhai et al., 2007). Misexpression of *HoxA9*, *HoxA10,* and *HoxC6* was also reported in cases of Leukemia caused by translocations of *mixed-lineage leukemia Mll gene*. MLL is the vertebrate homolog of *Drosophila* Trithorax (TRX) protein and helps maintain an active state of *Hox* expression in required domains (Armstrong et al., 2002; Ono et al., 2005). Hox-associated cancer is not limited to genetic mutations. Rauch et al. (2007) showed increased methylation of *HoxA7* and *HoxA9* associated CpG islands. The study highlighted epigenetic misregulation as a putative cause for Hox-related lung tumors. Likewise, promoter methylation of *HoxA5* and downregulation of *HoxA10* are associated with progressive breast carcinoma. The disease can also be caused by overexpression of *HoxB7* and *HoxB13* in these tissues (Raman et al., 2000; Chu et al., 2004; Chen et al., 2008; Jerevall et al., 2008). Misexpression studies in *Drosophila* confirmed the causal effect and established flies as a model to study Hox-associated oncogenesis. The outcome of the study was the ability of *Dfd*, *Ubx*, and *abd-A* genes to be leukemogenic when overexpressed in fat body and lamellocytes (Ponrathnam et al., 2021).

Detailed understanding of Hox genes expression and interaction during embryogenesis, tissue formation, organogenesis, and cellular homeostasis is required to delineate their functional modalities. Due to their overarching involvement in multiple processes of body formation, patterning, and evolution, Hox genes occupy a prime position in our quest toward understanding these processes in depth.

CONCLUDING REMARKS

A long-debated topic in the field of Hox genes was their presence in the form of clusters and the property of spatio-temporal collinearity. Some recent developments also demonstrated the functioning of Hox independent of clustering. However, coordinated functioning is better when they are clustered together, as implied by the open for business model of the bithorax complex. Alterations of CRMs throughout the Hox led to a myriad of homeotic transformations. Similar genomic alterations across evolution might have experimented with Hox

modules and their expression to bring about the enormous diversity we see today. Individual notes are pleasant to hear, but it's the symphony that conveys the melody. *Hox* come together to set up the primary and secondary axes and provide constant inputs in different tissues, therefore orchestrating the developmental design sublimely. *In vivo* experiments with the latest genome editing tools and a better understanding of non-coding DNA become important for comprehending the conductors of this symphony.

AUTHOR CONTRIBUTIONS

RKM and NH conceived the design of the article and edited the manuscript. NH wrote the manuscript with inputs from RKM, and conceptualized and drew illustrations upon discussion with RKM. Both authors contributed to the article and approved the submitted version.

FUNDING

NH is a fellow of the Department of Biotechnology, Govt. of India, and thanks the agency for providing timely financial support. RKM is recipient of JC Bose National fellowship, India. The authors would also like to thank CSIR, India, JC Bose National Fellowship, India and Tata Institute for Genetics and Society, India for financial support at various levels.

ACKNOWLEDGMENTS

The authors acknowledge the members of the RM group, especially, Shubhanshu Pandey, Ashish Bihani, M. S. Soujanya, Sonu Yadav, and Lorraine Teron for carefully reviewing the manuscript. The authors would also like to extend gratitude toward Surbhi Lambhate for continuous discussions and suggestions while preparing the manuscript.

REFERENCES

Abzhanov, A., and Kaufman, T. C. (2000). Crustacean (malacostracan) Hox genes and the evolution of the arthropod trunk. Development 127, 2239–2249. doi: 10.1242/dev.127.11.2239

Akam, M. (1998). Hox genes, homeosis and the evolution of segment identity: No need for hopeless monsters. *Int. J. Dev. Biol.* 42, 445–451. doi: 10.1387/ijdb.9654030

Akam, M., Dawson, I., and Tear, G. (1988). Homeotic genes and the control of segment diversity. *Development* 133, 123–133.

Albertin, C. B., Simakov, O., Mitros, T., Wang, Z. Y., Pungor, J. R., Edsinger-Gonzales, E., et al. (2015). The octopus genome and the evolution of cephalopod neural and morphological novelties. *Nature* 524, 220–224. doi: 10.1038/nature14668

Allis, C. D., and Jenuwein, T. (2016). The molecular hallmarks of epigenetic control. *Nat. Rev. Genet.* 17, 487–500. doi: 10.1038/nrg.2016.59

Armstrong, S. A., Staunton, J. E., Silverman, L. B., Pieters, R., Den Boer, M. L., Minden, M. D., et al. (2002). MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia. *Nat. Genet.* 30, 41–47. doi: 10.1038/ng765

Awgulewitsch, A. (2003). Hox in hair growth and development. *Naturwissenschaften* 90, 193–211. doi: 10.1007/s00114-003-0417-4

Bae, E., Calhoun, V. C., Levine, M., Lewis, E. B., and Drewell, R. A. (2002). Characterization of the intergenic RNA profile at abdominal-A and abdominal-B in the Drosophila bithorax complex. *Proc. Natl. Acad. Sci. U. S. A.* 99, 16847–16852. doi: 10.1073/pnas.222671299

Banreti, A., Hudry, B., Sass, M., Saurin, A. J., and Graba, Y. (2014). Hox Proteins Mediate Developmental and Environmental Control of Autophagy. *Dev. Cell* 28, 56–69. doi: 10.1016/j.devcel.2013.11.024

Bantignies, F., Roure, V., Comet, I., Leblanc, B., Schuettengruber, B., Bonnet, J., et al. (2011). Polycomb-dependent regulatory contacts between distant hox loci in drosophila. *Cell* 144, 214–226. doi: 10.1016/j.cell.2010.12.026

Barges, S., Mihaly, J., Galloni, M., Hagstrom, K., Müller, M., Shanower, G., et al. (2000). The Fab-8 boundary defines the distal limit of the bithorax complex iab-7 domain and insulates iab-7 from initiation elements and a PRE in the adjacent iab-8 domain. *Development* 127, 779–790.

Basson, M. A. (2012). Signaling in Cell Differentiation and Morphogenesis. *Cold Spring Harb. Perspect. Biol.* 4, 1–21.

Baughman, K. W., Mcdougall, C., Cummins, S. F., Hall, M., Degnan, B. M., Satoh, N., et al. (2014). Genomic organization of Hox and ParaHox clusters in the echinoderm, Acanthaster planci. *Genesis* 52, 952–958. doi: 10.1002/dvg.22840

Beddington, R. S. P., and Robertson, E. J. (1999). Axis development and early asymmetry in mammals. *Cell* 96, 195–209. doi: 10.1016/S0092-8674(00)80560-7

Bekiaris, P. S., Tekath, T., Staiger, D., and Danisman, S. (2018). Computational exploration of cis -regulatory modules in rhythmic expression data using the " Exploration of Distinctive CREs and CRMs " (EDCC) and " CRM Network Generator " (CNG) programs. *PLoS One* 13:e0190421. doi: 10.1371/journal.pone.0190421

Beeman, R. W. (1987). A homoeotic gene cluster in the red flour beetle. *Nature* 327, 3–5.

Bender, W., Akam, M., Karch, F., Beachy, P. A., Peifer, M., Spierer, P., et al. (2011). Molecular Genetics of the Bithorax Complex in Drosophila melanogaster. *Sciecne* 221, 23–29.

Bender, W., and Hudson, A. (2000). P element homing to the Drosophila bithorax complex. *Development* 127, 3981–3992.

Bender, W., and Lucas, M. (2013). The Border Between the Ultrabithorax and abdominal-A Regulatory Domains. *Genetics* 193, 1135–1147. doi: 10.1534/genetics.112.146340

Bender, W. W. (2020). Molecular Lessons from the Drosophila Bithorax Complex. *Genetics* 216, 613–617.

Benton, M. J. (2010). The origins of modern biodiversity on land. *Philos. Trans. R. Soc. B Biol. Sci.* 365, 3667–3679. doi: 10.1098/rstb.2010.0269

Berenguer, M., Meyer, K. F., Yin, J., and Duester, G. (2020). Discovery of genes required for body axis and limb formation by global identification of retinoic acid-regulated epigenetic marks. *PLoS Biol.* 18:e3000719. doi: 10.1371/journal.pbio.3000719

Bowman, S. K., Deaton, A. M., Domingues, H., Wang, P. I., Sadreyev, R. I., Kingston, R. E., et al. (2014). H3K27 modifications define segmental regulatory domains in the Drosophila bithorax complex. *Elife* 3, 1–13. doi: 10.7554/eLife.02833

Brooke, N. M., Garcia-Fernàndez, J., and Holland, P. W. H. (1998). The ParaHox gene cluster is an evolutionary sister of the Hox gene cluster. *Nature* 392, 920–922. doi: 10.1038/31933

Brown, S. J., Fellers, J. P., Shippy, T. D., Richardson, E. A., Maxwell, M., Stuart, J. J., et al. (2002). Sequence of the Tribolium castaneum homeotic complex: the region corresponding to the Drosophila melanogaster antennapedia complex. *Genetics* 1074, 1067–1074.

Browne, W. E., Price, A. L., Gerberding, M., and Patel, N. H. (2005). Stages of embryonic development in the amphipod crustacean, Parhyale hawaiensis. *Genesis* 42, 124–149. doi: 10.1002/gene.20145

Burke, A. C., Nelson, C. E., Morgan, B. A., and Tabin, C. (1995). Hox genes and the evolution of vertebrate axial morphology. *Development* 121, 333–346. doi: 10.1242/dev.121.2.333

Calhoun, V. C., and Levine, M. (2003). Long-range enhancer-promoter interactions in the Scr-Antp interval of the Drosophila Antennapedia complex. *Proc. Natl. Acad. Sci. U. S. A.* 100, 9878–9883. doi: 10.1073/pnas.1233791100

Capdevila, M., and Garcia-bellido, A. (1981). Genes Involved in the Activation of the Bithorax Complex in Drosophila. *Wilehm Roux Arch. Dev. Biol.* 190, 339–350.

Casares, F., and Sanchez-Herrero, E. (1995). Regulation of the infraabdominal regions of the bithorax complex of Drosophila by gap genes. *Development* 121, 1855–1866. doi: 10.1242/dev.121.6.1855

Castelli-gair, J. E., Capdevila, M., Micol, J., and Garcia-bellido, A. (1992). Positive and negative cis-regulatory elements in the bithoraxoid region of the Drosophila Ultrabithorax gene. *Mol. Genet. Genomics* 234, 177–184.

Castelli-Gair Hombría, J., and Lovegrove, B. (2003). Beyond homeosis - HOX function in morphogenesis and organogenesis. *Differentiation* 71, 461–476. doi: 10.1046/j.1432-0436.2003.7108004.x

Celniker, S. E., Dillon, L. A. L., Gerstein, M. B., Gunsalus, K. C., Henikoff, S., Karpen, G. H., et al. (2009). Unlocking the secrets of the genome. *Nature* 459, 927–930. doi: 10.1038/459927a

Celniker, S. E., Keelan, D. J., and Lewis, E. B. (1989). The molecular genetics of the bithorax complex of Drosophila: characterization of the products of the Abdominal-B domain. *Genes Dev.* 3, 1424–1436. doi: 10.1101/gad.3.9.1424

Celniker, S. E., Sharma, S., Keelan, D. J., and Lewis, E. B. (1990). The molecular genetics of the bithorax complex of Drosophila: cis-regulation in the Abdominal-B domain. *EMBO J.* 9, 4277–4286. doi: 10.1002/j.1460-2075.1990.tb07876.x

Chan, C. S., Rastelli, L., and Pirrotia, V. (1994). A Polycomb response element in the Ubx gene that determines an epigenetically inherited state of repression. *EMBO J.* 13, 2553–2564. doi: 10.1002/j.1460-2075.1994.tb06545.x

Chen, H., Ji, S. L., Liang, X., Zhang, H., Zhu, T., Zhang, Z., et al. (2008). Hoxb7 inhibits transgenic HER-2/neu-induced mouse mammary tumor onset but promotes progression and lung metastasis. *Cancer Res.* 68, 3637–3644. doi: 10.1158/0008-5472.CAN-07-2926

Cheng, W., Liu, J., Yoshida, H., Rosen, D., and Naora, H. (2005). Lineage infidelity of epithelial ovarian cancers is controlled by HOX genes that specify regional identity in the reproductive tract. *Nat. Med.* 11, 531–537. doi: 10.1038/nm1230

Chew, K. Y., Yu, H., Pask, A. J., Shaw, G., and Renfree, M. B. (2012). HOXA13 and HOXD13 expression during development of the syndactylous digits in the marsupial *Macropus eugenii*. *BMC Dev. Biol.* 12:2. doi: 10.1186/1471-213X-12-2

Chopra, V. S. (2011). Chromosomal organization at the level of gene complexes. *Cell. Mol. Life Sci.* 68, 977–990. doi: 10.1007/s00018-010-0585-2

Chourrout, D., Delsuc, F., Chourrout, P., Edvardsen, R. B., Rentzsch, F., Renfer, E., et al. (2006). Minimal ProtoHox cluster inferred from bilaterian and cnidarian Hox complements. *Nature* 442, 684–687. doi: 10.1038/nature04863

Chu, M. C., Selam, F. B., and Taylor, H. S. (2004). HOXA10 regulates p53 expression and matrigel invasion in human breast cancer cells. *Cancer Biol. Ther.* 3, 568–572. doi: 10.4161/cbt.3.6.848

Coutelis, J. B., Géminard, C., Spéder, P., Suzanne, M., Petzoldt, A. G., and Noselli, S. (2013). Drosophila Left/Right Asymmetry Establishment Is Controlled by the Hox Gene Abdominal-B. *Dev. Cell* 24, 89–97. doi: 10.1016/j.devcel.2012.11.013

Crocker, J., Tsai, A., and Stern, D. L. (2017). A Fully Synthetic Transcriptional Platform for a Multicellular Eukaryote. *Cell Rep.* 18, 287–296. doi: 10.1016/j.celrep.2016.12.025

Deakin, J. E. (2012). Marsupial genome sequences: providing insight into evolution and disease. *Scientifica* 2012:543176. doi: 10.6064/2012/543176

Dequéant, M. L., and Pourquié, O. (2008). Segmental patterning of the vertebrate embryonic axis. *Nat. Rev. Genet.* 9, 370–382. doi: 10.1038/nrg2320

Dessain, S., and McGinnis, W. (1993). Drosophila Homeobox Genes. *Adv. Dev. Biochem.* 2, 1–55. doi: 10.1016/s1064-2722(08)60035-3

Deutsch, J. S. (2004). Segments and parasegments in arthropods: a functional perspective. *BioEssays* 26, 1117–1125. doi: 10.1002/bies.20111

Di, Z., Yu, Y., Wu, Y., Hao, P., He, Y., Zhao, H., et al. (2015). Genome-wide analysis of homeobox genes from Mesobuthus martensii reveals Hox gene duplication in scorpions. *Insect Biochem. Mol. Biol.* 61, 25–33. doi: 10.1016/j.ibmb.2015.04.002

Dickinson, M. H., Trans, P., Lond, R. S., and Dickinson, M. H. (1999). Haltere - mediated equilibrium reflexes of the fruit fly, Drosophila melanogaster. *Philos. Trans. R Soc. Lond. B Biol. Sci.* 354, 903–916. doi: 10.1098/rstb.1999.0442

Domsch, K., Carnesecchi, J., Disela, V., Friedrich, J., Trost, N., Ermakova, O., et al. (2019). The Hox Transcription Factor Ubx stabilizes Lineage Commitment by Suppressing Cellular Plasticity. *Elife* 8:e42675.

Domsch, K., Schröder, J., Janeschik, M., Schaub, C., and Lohmann, I. (2020). The Hox transcription factor Ubx ensures somatic myogenesis by suppressing the mesodermal master regulator Twist. *BioRxiv* 1–29. doi: 10.1101/2020.02.24.963231

Drewell, R. A., Nevarez, M. J., Kurata, J. S., Winkler, L. N., Li, L., and Dresch, J. M. (2014). Deciphering the combinatorial architecture of a Drosophila homeotic gene enhancer. *Mech. Dev.* 131, 68–77. doi: 10.1016/j.mod.2013.10.002

Duboule, D. (2007). The rise and fall of Hox gene clusters. *Development* 134, 2549–2560. doi: 10.1242/dev.001065

Elizondo, L., Jafar-Nejad, P., Clewing, J., and Boerkoel, C. (2009). Gene Clusters, Molecular Evolution and Disease: a Speculation. *Curr. Genomics* 10, 64–75. doi: 10.2174/138920209787581271

Estacio-Gómez, A., and Díaz-Benjumea, F. J. (2014). Roles of Hox genes in the patterning of the central nervous system of Drosophila. *Fly* 8, 26–32. doi: 10.4161/fly.27424

Estrada, B., Casares, F., Busturia, A., and Sánchez-herrero, E. (2002). Genetic and molecular characterization of a novel iab-8 regulatory domain in the Abdominal-B gene of Drosophila melanogaster. *Development* 5204, 5195–5204.

Fabre, P. J., Leleu, M., Mormann, B. H., Lopez-Delisle, L., Noordermeer, D., Beccari, L., et al. (2017). Large scale genomic reorganization of topological domains at the HoxD locus. *Genome Biol.* 18, 1–15. doi: 10.1186/s13059-017-1278-z

Ferrier, D. E. K. (2016). Evolution of homeobox gene clusters in animals: the Giga-cluster and Primary vs. secondary clustering. *Front. Ecol. Evol.* 4:36. doi: 10.3389/fevo.2016.00036

Ferrier, D. E. K., and Holland, P. W. H. (2001). Ancient Origin of the Hox Gene Cluster. *Nat. Rev. Genet.* 2, 33–8.

Ferrier, D. E. K., and Minguillón, C. (2003). Evolution of the Hox/ParaHox gene clusters. *Int. J. Dev. Biol.* 47, 605–611.

Ferrier, D. E. K., Minguillón, C., Holland, P. W. H., and Garcia-Fernàndez, J. (2000). The amphioxus Hox cluster: deuterostome posterior flexibility and Hox14. *Evol. Dev.* 2, 284–293. doi: 10.1046/j.1525-142X.2000.00070.x

Ford, T. N., Lim, D., and Mertz, J. (2012). Fast optically sectioned fluorescence HiLo endomicroscopy. *J. Biomed. Opt.* 17:021105. doi: 10.1117/1.jbo.17.2.021105

François, V., Solloway, M., O'Neill, J. W., Emery, J., and Bier, E. (1994). Dorsal-ventral patterning of the Drosophila embryo depends on a putative negative growth factor encoded by the short gastrulation gene. *Genes Dev.* 8, 2602–2616. doi: 10.1101/gad.8.21.2602

Fritsch, M., Wollesen, T., De Oliveira, A. L., and Wanninger, A. (2015). Unexpected co-linearity of Hox gene expression in an aculiferan mollusk. *BMC Evol. Biol.* 15:151. doi: 10.1186/s12862-015-0414-1

Galloni, M., Gyurkovics, H., Schedl, P., and Karch, F. (1993). The bluetail transposon: evidence for independent cis-regulatory domains and domain boundaries in the bithorax complex. *EMBO J.* 12, 1087–1097. doi: 10.1002/j.1460-2075.1993.tb05750.x

Gaunt, S. J. (2015). The significance of Hox gene collinearity. *Int. J. Dev. Biol.* 59, 159–170. doi: 10.1387/ijdb.150223sg

Gaunt, S. J. (2018). Hox cluster genes and collinearities throughout the tree of animal life. *Int. J. Dev. Biol.* 62, 673–683. doi: 10.1387/ijdb.180162sg

Gindhart, J. G., King, A. N., and Kaufman, T. C. (1995). Characterization of the cis-Regulatory Region of the Drosophila Homeotic Gene Sex combs reduced. *Genetics* 139, 781–795.

Gummalla, M., Galetti, S., Maeda, R. K., and Karch, F. (2014). Hox gene regulation in the central nervous system of Drosophila. *Front. Cell. Neurosci.* 8:96. doi: 10.3389/fncel.2014.00096

Hagstrom, K., Muller, M., and Schedl, P. (1996). Fab-7 functions as a chromatin domain boundary to ensure proper segment specification by the Drosophila bithorax complex. *Genes Dev.* 10, 3202–3215. doi: 10.1101/gad.10.24.3202

Hendrickson, J. E., and Sakonju, S. (1995). Cis and trans interactions between the iab regulatory regions and abdominal-A and abdominal-B in Drosophila melanogaster. *Genetics* 139, 835–848.

Ho, M. C. W., Johnsen, H., Goetz, S. E., Schiller, B. J., Bae, E., Tran, D. A., et al. (2009). Functional evolution of cis-regulatory modules at a homeotic gene in Drosophila. *PLoS Genet.* 5:e1000709. doi: 10.1371/journal.pgen.1000709

Holland, L. Z., Albalat, R., Azumi, K., Benito-Gutiérrez, È, Blow, M. J., Bronner-Fraser, M., et al. (2008). The amphioxus genome illuminates vertebrate origins and cephalochordate biology. *Genome Res.* 18, 1100–1111. doi: 10.1101/gr.073676.107

Holland, P. W. H. (2001). Beyond the Hox: how widespread is homeobox gene clustering?. *J. Anat.* 199, 13–23. doi: 10.1017/S0021878201008275

Holland, P. W. H. (2015). Did homeobox gene duplications contribute to the Cambrian explosion?. *Zool. Lett.* 1:1. doi: 10.1186/s40851-014-0004-x

Holland, P. W. H., Booth, H. A. F., and Bruford, E. A. (2007). Classification and nomenclature of all human homeobox genes. *BMC Biol.* 5:47. doi: 10.1186/1741-7007-5-47

Howard-Ashby, M., Materna, S. C., Brown, C. T., Chen, L., Cameron, R. A., and Davidson, E. H. (2006). Identification and characterization of homeobox transcription factor genes in Strongylocentrotus purpuratus, and their expression in embryonic development. *Dev. Biol.* 300, 74–89. doi: 10.1016/j.ydbio.2006.08.039

Hrycaj, S. M., and Wellik, D. M. (2016). Hox genes and evolution. *F1000Res.* 5:859. doi: 10.12688/f1000research.7663.1

Hübner, M. R., Eckersley-Maslin, M. A., and Spector, D. L. (2013). Chromatin organization and transcriptional regulation. *Curr. Opin. Genet. Dev.* 23, 89–95. doi: 10.1016/j.gde.2012.11.006

Hug, C. B., and Vaquerizas, J. M. (2018). The Birth of the 3D Genome during Early Embryonic Development. *Trends Genet.* 34, 903–914. doi: 10.1016/j.tig.2018.09.002

Iampietro, C., Gummalla, M., Mutero, A., and Maeda, R. K. (2010). Initiator Elements Function to Determine the Activity State of BX-C Enhancers. *PLoS Genet.* 6:e1001260. doi: 10.1371/journal.pgen.1001260

Ikuta, T. (2011). Evolution of invertebrate deuterostomes and Hox/ParaHox genes. *Genom. Proteom. Bioinf.* 9, 77–96. doi: 10.1016/S1672-0229(11)60011-9

Ikuta, T., Yoshida, N., Satoh, N., and Saiga, H. (2004). Ciona intestinalis Hox gene cluster: its dispersed structure and residual colinear expression in development. *Proc. Natl. Acad. Sci. U. S. A.* 101, 15118–15123.

Janssen, R., Eriksson, B., Tait, N. N., and Budd, G. E. (2014). Onychophoran Hox genes and the evolution of arthropod Hox gene expression. *Front. Zool.* 11:22. doi: 10.1186/1742-9994-11-22

Jerevall, P. L., Brommesson, S., Strand, C., Gruvberger-Saal, S., Malmström, P., Nordenskjöld, B., et al. (2008). Exploring the two-gene ratio in breast cancer-independent roles for HOXB13 and IL17BR in prediction of clinical outcome. *Breast Cancer Res. Treat.* 107, 225–234. doi: 10.1007/s10549-007-9541-8

Jung, C., Kim, R. S., Zhang, H. J., Lee, S. J., and Jeng, M. H. (2004). HOXB13 induces growth suppression of prostate cancer cells as a repressor of hormone-activated androgen receptor signaling. *Cancer Res.* 64, 9185–9192. doi: 10.1158/0008-5472.CAN-04-1330

Kadota, M., Nishimura, O., Miura, H., Tanaka, K., Hiratani, I., and Kuraku, S. (2020). Multifaceted Hi-C benchmarking: what makes a difference in chromosome-scale genome scaffolding?. *Gigascience* 9:giz158. doi: 10.1093/gigascience/giz158

Karch, F., Galloni, M., Sipos, L., Gausz, J., Gyurkovics, H., and Schedl, P. (1994). Mcp and Fab-7: molecular analysis of putative boundaries of cis-regulatory domains in the bithorax complex of Drosophila melanogaster. *Nucleic Acids Res.* 22, 3138–3146. doi: 10.1093/nar/22.15.3138

Kaufman, T. C., Seeger, M. A., and Olsen, G. (1990). Molecular and Genetic Organization of the Antennapedia Gene Complex of Drosophila melanogaster. *Adv. Genet.* 27, 309–362.

Kmita, M., Kondo, T., and Duboule, D. (2000). Targeted inversion of a polar silencer within the HoxD complex re-allocates domains of enhancer sharing. *Nat. Genet.* 26, 451–454. doi: 10.1038/82593

Kmita, M., Tarchini, B., Zàkàny, J., Logan, M., Tabin, C. J., and Duboule, D. (2005). Early developmental arrest of mammalian limbs lacking HoxA/HoxD gene function. *Nature* 435, 1113–1116. doi: 10.1038/nature03648

Kornberg, T. B., and Tabata, T. (1993). Segmentation of the Drosophila embryo. *Curr. Opin. Genet. Dev.* 3, 585–593. doi: 10.1016/0959-437X(93)90094-6

Kourakis, M. J., and Martindale, M. Q. (2001). Hox gene duplication and deployment in the annelid leech Helobdella. *Evol. Dev.* 3, 145–153. doi: 10.1046/j.1525-142X.2001.003003145.x

Kushawah, G., and Mishra, R. K. (2017). Ultraconserved sequences associated with HoxD cluster have strong repression activity. *Genome Biol. Evol.* 9, 2134–2139. doi: 10.1093/gbe/evx148

Kyrchanova, O., Ivlieva, T., Toshchakov, S., Parshikov, A., Maksimenko, O., and Georgiev, P. (2011). Selective interactions of boundaries with upstream region of Abd-B promoter in Drosophila bithorax complex and role of dCTCF in this process. *Nucleic Acids Res.* 39, 3042–3052. doi: 10.1093/nar/gkq1248

Kyrchanova, O., Mogila, V., Wolle, D., and Deshpande, G. (2016). Functional Dissection of the Blocking and Bypass Activities of the Fab-8 Boundary in the Drosophila Bithorax Complex. *PLoS Genet.* 12:e1006188. doi: 10.1371/journal.pgen.1006188

Kyrchanova, O., Mogila, V., Wolle, D., Paolo, J., White, R., Georgiev, P., et al. (2015). The boundary paradox in the Bithorax complex. *Mech. Dev.* 138, 122–132. doi: 10.1016/j.mod.2015.07.002

Kyrchanova, O., Sabirov, M., Mogila, V., Kurbidaeva, A., Postika, N., Maksimenko, O., et al. (2019). Complete reconstitution of bypass and blocking functions in a minimal artificial Fab-7 insulator from Drosophila bithorax complex. *Proc. Natl. Acad. Sci. U. S. A.* 116, 13462–13467. doi: 10.1073/pnas.1907190116

Lanzuolo, C., Roure, V., Dekker, J., Bantignies, F., and Orlando, V. (2007). Polycomb response elements mediate the formation of chromosome higher-order structures in the bithorax complex. *Nat. Cell Biol.* 9, 1167–1174. doi: 10.1038/ncb1637

Laughon, A., Boulet, A. M., Bermingham, J. R., Laymon, R. A., and Scott, M. P. (1986). Structure of transcripts from the homeotic Antennapedia gene of Drosophila melanogaster: two promoters control the major protein-coding region. *Mol. Cell. Biol.* 6, 4676–4689. doi: 10.1128/mcb.6.12.4676

Levin, M. (2005). Left-right asymmetry in embryonic development: a comprehensive review. *Mech. Dev.* 122, 3–25. doi: 10.1016/j.mod.2004.08.006

Lewin, H. A., Robinson, G. E., Kress, W. J., Baker, W. J., Coddington, J., Crandall, K. A., et al. (2018). Earth BioGenome Project: sequencing life for the future of life. *Proc. Natl. Acad. Sci. U. S. A.* 115, 4325–4333. doi: 10.1073/pnas.1720115115

Lewis, E. B. (1978). A gene complex controlling segmentation in Drosophila. *Nature* 276, 565–570.

Lewis, E. B. (1994). Homeosis: the first 100 years. *Trends Genet.* 10, 341–343.

Lewis, E. B. (1998). The bithorax complex: the first fifty years. *Int. J. Biomed. Sci.* 42, 403–415. doi: 10.1007/978-1-4020-6345-9_32

Lewis, E. B. (2007). "A gene complex controlling segmentation in Drosophila," in *Genes, Development, and Cancer*, ed. H. D. Lipshitz (Dordrecht: Springer), 229–242. doi: 10.1007/978-1-4020-6345-9_10

Lewis, M. T. (2000). Homeobox genes in mammary gland development and neoplasia. *Breast Cancer Res.* 2, 158–169. doi: 10.1186/bcr49

Li, J., Li, Y., Pawlik, K. M., Napierala, J. S., and Napierala, M. (2020). A CRISPR-Cas9, Crelox, and Flp-FRT Cascade Strategy for the Precise and Efficient Integration of Exogenous DNA into Cellular Genomes. *CRISPR J.* 3, 470–486. doi: 10.1089/crispr.2020.0042

Li, M., Ma, Z., Liu, J. K., Roy, S., Patel, S. K., Lane, D. C., et al. (2015). An Organizational Hub of Developmentally Regulated Chromatin Loops in the Drosophila Antennapedia Complex. *Mol. Cell. Biol.* 35, 4018–4029. doi: 10.1128/MCB.00663-15

Lin, Q., Wu, D., and Zhou, J. (2003). The promoter targeting sequence facilitates and restricts a distant enhancer to a single promoter in the Drosophila embryo. *Development* 130, 519–526. doi: 10.1242/dev.00227

Little, J. W., Byrd, C. A., and Brower, D. L. (1990). Effect of abx, bx and pbx mutations on expression of homeotic genes in Drosophila larvae. *Genetics* 124, 899–908.

Liubicich, D. M., Serano, J. M., Pavlopoulos, A., Kontarakis, Z., Protas, M. E., Kwan, E., et al. (2009). Knockdown of Parhyale Ultrabithorax recapitulates evolutionary changes in crustacean appendage morphology. *Proc. Natl. Acad. Sci. U. S. A.* 106, 13892–13896. doi: 10.1073/pnas.0903105106

Long, H. K., Prescott, S. L., and Wysocka, J. (2016). Ever-Changing Landscapes: transcriptional Enhancers in Development and Evolution. *Cell* 167, 1170–1187. doi: 10.1016/j.cell.2016.09.018

Lovegrove, B., Simões, S., Rivas, M. L., Sotillos, S., Johnson, K., Knust, E., et al. (2006). Coordinated Control of Cell Adhesion, Polarity, and Cytoskeleton Underlies Hox-Induced Organogenesis in Drosophila. *Curr. Biol.* 16, 2206–2216. doi: 10.1016/j.cub.2006.09.029

Maeda, R. K. (2006). The ABC of the BX-C: the bithorax complex explained. *Development* 133, 1413–1422. doi: 10.1242/dev.02323

Maeda, R. K. (2009). The Bithorax Complex of Drosophila: an Exceptional Hox Cluster. *Curr. Top. Dev. Biol.* 88, 1–33. doi: 10.1016/S0070-2153(09)88001-0

Maeda, R. K., and Karch, F. (2007). Making connections: boundaries and insulators in Drosophila. *Curr. Opin. Genet. Dev.* 17, 394–399. doi: 10.1016/j.gde.2007.08.002

Maeda, R. K., and Karch, F. (2015). The open for business model of the bithorax complex in Drosophila. *Chromosoma* 124, 293–307. doi: 10.1007/s00412-015-0522-0

Mallo, M., and Alonso, C. R. (2013). The regulation of Hox gene expression during animal development. *Development* 140, 3951–3963. doi: 10.1242/dev.068346

Mallo, M., Wellik, D. M., and Deschamps, J. (2010). Hox genes and regional patterning of the vertebrate body plan. *Dev. Biol.* 344, 7–15. doi: 10.1016/j.ydbio.2010.04.024

Manohar, C. F., Salwen, H. R., Furtado, M. R., and Cohn, S. L. (1996). Up-Regulation of HOXC6, HOXD1 and HOXD8 Homeobox Gene Expression in Human Neuroblastoma Cells following Chemical Induction of Differentiation. *Tumor Biol.* 17, 34–47.

Marchetti, M., Fanti, L., Berloco, M., and Pimpinelli, S. (2003). Differential expression of the Drosophila BX-C in polytene chromosomes in cells of larval fat bodies: a cytological approach to identifying *in vivo* targets of the homeotic Ubx, Abd-A and Abd-B proteins. *Development* 130, 3683–3689. doi: 10.1242/dev.00587

Martin, A., Serano, J. M., Jarvis, E., Bruce, H. S., Wang, J., Ray, S., et al. (2016). CRISPR/Cas9 Mutagenesis Reveals Versatile Roles of Hox Genes in Crustacean Limb Specification and Evolution. *Curr. Biol.* 26, 14–26. doi: 10.1016/j.cub.2015.11.021

Martin, C. H., Mayeda, C. A., Davis, C. A., Ericsson, C. L., Knafels, J. D., Mathog, D. R., et al. (1995). Complete sequence of the bithorax complex of Drosophila. *Proc. Natl. Acad. Sci. U. S. A.* 92, 8398–8402. doi: 10.1073/pnas.92.18.8398

Martinez-Arias, A., and Lawrence, P. A. (1985). Parasegments and compartments in the Drosophila embryo. *Nature* 313, 639–642. doi: 10.1038/313639a0

Martínez-Laborda, A., Serrano-Cartagena, J., and Micol, J. L. (1996). A genetic analysis of bx bxd cis double mutants in the Drosophila Ultrabithorax gene. *Mol. Gen. Genet.* 250, 540–546. doi: 10.1007/s004380050107

Mateo, L. J., Murphy, S. E., Hafner, A., Cinquini, I. S., Walker, C. A., and Boettiger, A. N. (2019). Visualizing DNA folding and RNA in embryos at single-cell resolution. *Nature* 568, 49–54. doi: 10.1038/s41586-019-1035-4

McGinnis, W., Garber, R. L., Wirz, J., Kuroiwa, A., and Gehring, W. J. (1984a). A homologous protein-coding sequence in drosophila homeotic genes and its conservation in other metazoans. *Cell* 37, 403–408. doi: 10.1016/0092-8674(84)90370-2

McGinnis, W., Levine, M. S., Hafen, E., Kuroiwa, A., and Gehring, W. J. (1984b). A conserved DNA sequence in homoeotic genes of the Drosophila Antennapedia and bithorax complexes. *Nature* 308, 428–433. doi: 10.1038/308428a0

Medina-Martínez, O., Bradley, A., and Ramírez-Solis, R. (2000). A large targeted deletion of Hoxb1-Hoxb9 produces a series of single- segment anterior homeotic transformations. *Dev. Biol.* 222, 71–83. doi: 10.1006/dbio.2000.9683

Miao, J., Wang, Z., Provencher, H., Muir, B., Dahiya, S., Carney, E., et al. (2007). HOXB13 promotes ovarian cancer progression. *Proc. Natl. Acad. Sci. U. S. A.* 104, 17093–17098. doi: 10.1073/pnas.0707938104

Mihaly, J., Barges, S., Sipos, L., Maeda, R., Cléard, F., Hogga, I., et al. (2006). Dissecting the regulatory landscape of the Abd-B gene of the bithorax complex. *Development* 133, 2983–2993. doi: 10.1242/dev.02451

Mihaly, J., Hogga, I., Barges, S., Galloni, M., Mishra, R. K., Hagstrom, K., et al. (1998). Chromatin domain boundaries in the Bithorax complex. *Cell. Mol. Life Sci.* 54, 60–70. doi: 10.1007/s000180050125

Millar, S. E. (2018). Hox in the Niche Controls Hairy-geneity. *Cell Stem Cell* 23, 457–458. doi: 10.1016/j.stem.2018.09.012

Mishra, R. K., and Karch, F. (1999). Boundaries that demarcate structural and functional domains of chromatin. *J. Biosci.* 24, 377–399. doi: 10.1007/bf02941252

Mishra, R. K., Mihaly, J., Barges, S., Spierer, A., Karch, F., Hagstrom, K., et al. (2001). The iab-7 polycomb response element maps to a nucleosome-free region of chromatin and requires both GAGA and pleiohomeotic for silencing activity. *Mol. Cell. Biol.* 21, 1311–1318. doi: 10.1128/MCB.21.4.1311-1318.2001

Misof, B., Jun Tong, K., Duchêne, S., Ho, S. Y. W., and Lo, N. (2014). Phylogenomics resolves the timing and pattern of insect evolution. *Science* 346:763. doi: 10.1126/science.aaa5460

Mohit, P., Makhijani, K., Madhavi, M. B., Bharathi, V., Lal, A., Sirdesai, G., et al. (2006). Modulation of AP and DV signaling pathways by the homeotic gene Ultrabithorax during haltere development in Drosophila. *Dev. Biol.* 291, 356–367. doi: 10.1016/j.ydbio.2005.12.022

Mooi, R., and David, B. (2008). Radial Symmetry, the Anterior / Posterior Axis, and Echinoderm Hox Genes. *Annu. Rev. Ecol. Evol. Syst.* 39, 43–62. doi: 10.1146/annurev.ecolsys.39.110707.173521

Muller, J., and Bienz, M. (1992). Sharp anterior boundary of homeotic gene expression conferred by the fushi tarazu protein. *EMBO J.* 11, 3653–3661. doi: 10.1002/j.1460-2075.1992.tb05450.x

Muller, M., Hagstrom, K., Gyurkovics, H., Pirrotta, V., and Schedl, P. (1999). The Mcp Element From the Drosophila melanogaster Bithorax Complex Mediates Long-Distance Regulatory Interactions. *Genetics* 153, 1333–1356.

Narlikar, L., and Ovcharenko, I. (2009). Identifying regulatory elements in eukaryotic genomes. *Brief. Funct. Genomics Proteomics* 8, 215–230. doi: 10.1093/bfgp/elp014

Negre, B., Ranz, J. M., Casals, F., Cáceres, M., and Ruiz, A. (2003). A New Split of the Hox Gene Complex in Drosophila: relocation and Evolution of the Gene labial. *Mol. Biol. Evol.* 20, 2042–2054. doi: 10.1093/molbev/msg238

Négre, N., Brown, C. D., Ma, L., Bristow, C. A., Miller, S. W., Wagner, U., et al. (2011). A cis-regulatory map of the Drosophila genome. *Nature* 471, 527–531. doi: 10.1038/nature09990

Nègre, N., Brown, C. D., Shah, P. K., Kheradpour, P., Morrison, C. A., Henikoff, J. G., et al. (2010). A comprehensive map of insulator elements for the Drosophila genome. *PLoS Genet.* 6:e1000814. doi: 10.1371/journal.pgen.1000814

Nong, W., Cao, J., Li, Y., Qu, Z., Sun, J., Swale, T., et al. (2020). Jellyfish genomes reveal distinct homeobox gene clusters and conservation of small RNA processing. *Nat. Commun.* 11:3051. doi: 10.1038/s41467-020-16801-9

Ono, R., Nosaka, T., and Hayashi, Y. (2005). Roles of a trithorax group gene. MLL, in hematopoiesis. *Int. J. Hematol.* 81, 288–293. doi: 10.1532/IJH97.04196

Papagiannouli, F., and Lohmann, I. (2015). Stage-specific control of stem cell niche architecture in the Drosophila testis by the posterior Hox gene Abd-B. *Comput. Struct. Biotechnol. J.* 13, 122–130. doi: 10.1016/j.csbj.2015.01.001

Pardo, J. C. F., Stefanelli-Silva, G., Christy, J. H., and Costa, T. M. (2020). Fiddler crabs and their above-ground sedimentary structures: a review. *J. Ethol.* 38, 137–154. doi: 10.1007/s10164-020-00647-1

Pascual-Anaya, J., Adachi, N., Álvarez, S., Kuratani, S., D'Aniello, S., and Garcia-Fernàndez, J. (2012). Broken colinearity of the amphioxus Hox cluster. *Evodevo* 3:28. doi: 10.1186/2041-9139-3-28

Paul, R., Giraud, G., Domsch, K., Duffraisse, M., Marmigère, F., Khan, S., et al. (2021). Hox dosage contributes to flight appendage morphology in Drosophila. *Nat. Commun.* 12:2892. doi: 10.1038/s41467-021-23293-8

Pavlopoulos, A., Kontarakis, Z., Liubicich, D. M., Serano, J. M., Akam, M., Patel, N. H., et al. (2009). Probing the evolution of appendage specialization by Hox gene misexpression in an emerging model crustacean. *Proc. Natl. Acad. Sci. U. S. A.* 106, 13897–13902. doi: 10.1073/pnas.0902804106

Peel, A. D., Chipman, A. D., and Akam, M. (2005). Arthropod segmentation: beyond the Drosophila paradigm. *Nat. Rev. Genet.* 6, 905–916. doi: 10.1038/nrg1724

Peifer, M., and Bender, W. (1986). The anterobithorax and bithorax mutations of the bithorax complex. *EMBO J.* 5, 2293–2303.

Peifer, M., Karch, F. F., and Bender, W. (1987). The bithorax complex: control of segment identity. *Genes Dev.* 1, 891–898.

Perrimon, N., Pitsouli, C., and Shilo, B. Z. (2012). Signaling mechanisms controlling cell fate and embryonic patterning. *Cold Spring Harb. Perspect. Biol.* 4:a005975. doi: 10.1101/cshperspect.a005975

Peterson, K. J., Cameron, R. A., and Davidson, E. H. (2000). Bilaterian Origins: significance of New Experimental Observations. *Dev. Biol.* 17, 1–17. doi: 10.1006/dbio.1999.9475

Ponrathnam, T., Saini, R., Banu, S., and Mishra, R. K. (2021). Drosophila Hox genes induce melanized pseudo-tumors when misexpressed in hemocytes. *Sci. Rep.* 11:1838. doi: 10.1038/s41598-021-81472-5

Postika, N., Metzler, M., Affolter, M., Müller, M., Schedl, P., Georgiev, P., et al. (2018). Boundaries mediate long-distance interactions between enhancers and

promoters in the Drosophila Bithorax complex. *PLoS Genet.* 14:e1007702. doi: 10.1371/journal.pgen.1007702

Postika, N., Schedl, P., Georgiev, P., and Kyrchanova, O. (2021). Mapping of functional elements of the Fab-6 boundary involved in the regulation of the Abd-B hox gene in Drosophila melanogaster. *Sci. Rep.* 11:4156. doi: 10.1038/s41598-021-83734-8

Pradel, J., and White, R. A. H. (1998). From selectors to realizators. *Int. J. Dev. Biol.* 421, 417–421.

Raman, V., Martenser, S. A., Reisman, D., Evron, E., Odenwald, W. F., Jaffee, E., et al. (2000). Compromised HOXA5 function can limit p53 expression in human breast tumours. *Nature* 405, 974–978. doi: 10.1038/35016125

Rauch, T., Wang, Z., Zhang, X., Zhong, X., Wu, X., Lau, S. K., et al. (2007). Homeobox gene methylation in lung cancer studied by genome-wide analysis with a microarray-based methylated CpG island recovery assay. *Proc. Natl. Acad. Sci. U. S. A.* 104, 5527–5532. doi: 10.1073/pnas.0701059104

Reinitz, J., and Levine, M. (1990). Control of the initiation of homeotic gene expression by the gap genes giant and tailless in Drosophila. *Dev. Biol.* 140, 57–72. doi: 10.1016/0012-1606(90)90053-L

Rentzsch, F., and Holstein, T. W. (2018). Making head or tail of cnidarian hox gene function. *Nat. Commun.* 9, 9–11. doi: 10.1038/s41467-018-04585-y

Rogers, B. T. (2020). *Drosophila Embryo: Homeotic Genes in Specification of the Anterior–Posterior Axis*. Hoboken: Wiley Online Library.

Rux, D. R., and Wellik, D. M. (2017). Hox genes in the adult skeleton: novel functions beyond embryonic development. *Dev. Dyn.* 246, 310–317. doi: 10.1002/dvdy.24482

Sabarinadh, C., Subramanian, S., Tripathi, A., and Mishra, R. K. (2004). Extreme conservation of noncoding DNA near HoxD complex of vertebrates. *BMC Genomics* 5:75. doi: 10.1186/1471-2164-5-75

Samadi, L., and Steiner, G. (2010). Expression of Hox genes during the larval development of the snail, Gibbula varia (L.)-further evidence of non-colinearity in molluscs. *Dev. Genes Evol.* 220, 161–172. doi: 10.1007/s00427-010-0338-0

Sánchez-Herrero, E. (1991). Control of the expression of the bithorax complex genes abdominal-A and abdominal-B by cis-regulatory regions in Drosophila embryos. *Development* 111, 437–449.

Sánchez-Herrero, E. (2013). Hox Targets and Cellular Functions. *Scientifica* 2013:738257. doi: 10.1155/2013/738257

Schiemann, S. M., Martín-durán, J. M., Aina, B., Vellutini, B. C., and Passamaneck, Y. J. (2017). Clustered brachiopod Hox genes are not expressed collinearly and are associated with lophotrochozoan novelties. *Proc. Natl. Acad. Sci. U. S. A.* 7, 1913–1922. doi: 10.1073/pnas.1614501114

Schneuwly, S., Kuroiwa, A., and Gehring, W. J. (1987). Molecular analysis of the dominant homeotic Antennapedia phenotype. *EMBO J.* 6, 201–206. doi: 10.1002/j.1460-2075.1987.tb04739.x

Schweinsberg, S., Hagstrom, K., Gohl, D., Schedl, P., Kumar, R. P., Mishra, R., et al. (2004). The enhancer-blocking activity of the Fab-7 boundary from the Drosophila bithorax complex requires GAGA-factor-binding sites. *Genetics* 168, 1371–1384. doi: 10.1534/genetics.104.029561

Scott, M. P., and Weiner, A. J. (1984). Structural relationships among genes that control development: sequence homology between the antennapedia, ultrabithorax, and fushi tarazu loci of Drosophila. *Proc. Natl. Acad. Sci. U. S. A.* 81, 4115–4119. doi: 10.1073/pnas.81.13.4115

Serano, J. M., Martin, A., Liubicich, D. M., Jarvis, E., Bruce, H. S., La, K., et al. (2016). Comprehensive analysis of Hox gene expression in the amphipod crustacean Parhyale hawaiensis. *Dev. Biol.* 409, 297–309. doi: 10.1016/j.ydbio.2015.10.029

Shah, N., and Sukumar, S. (2010). The Hox genes and their roles in oncogenesis. *Nat. Rev. Cancer* 10, 361–371. doi: 10.1038/nrc2826

Sharma, P. P., Schwager, E. E., Extavour, C. G., and Giribet, G. (2012). Hox gene expression in the harvestman Phalangium opilio reveals divergent patterning of the chelicerate opisthosoma. *Evol. Dev.* 14, 450–463. doi: 10.1111/j.1525-142X.2012.00565.x

Sharma, P. P., Schwager, E. E., Extavour, C. G., and Wheeler, W. C. (2014). Hox gene duplications correlate with posterior heteronomy in scorpions. *Proc. Biol. Sci.* 281:20140661. doi: 10.1098/rspb.2014.0661

Shippy, T. D., Ronshaugen, M., Cande, J., He, J., Beeman, R. W., Levine, M., et al. (2008). Analysis of the Tribolium homeotic complex: insights into mechanisms constraining insect Hox clusters. *Dev. Genes Evol.* 218, 127–139. doi: 10.1007/s00427-008-0213-4

Shultz, J. W. (1989). Morphology of locomotor appendages in Arachnida: evolutionary trends and phylogenetic implications. *Zool. J. Linn. Soc.* 97, 1–55. doi: 10.1111/j.1096-3642.1989.tb00552.x

Simon, J., Chiang, A., Bender, W., Shimell, M. J., and Connor, M. O. (1993). Elements of the Drosophila Bithorax Complex That Mediate Repression by Polycomb Group Products. *Dev. Biol.* 158, 131–144. doi: 10.1006/dbio.1993.1174

Simon, J., Peifer, M., Bender, W., and O'Connor, M. (1990). Regulatory elements of the bithorax complex that control expression along the anterior-posterior axis. *EMBO J.* 9, 3945–3956. doi: 10.1002/j.1460-2075.1990.tb07615.x

Singh, N. P., de Kumar, B., Paulson, A., Parrish, M. E., Zhang, Y., Florens, L., et al. (2020). A six-amino-acid motif is a major determinant in functional evolution of HOX1 proteins. *Genes Dev.* 34, 1680–1696. doi: 10.1101/gad.342329.120

Singh, N. P., and Mishra, R. K. (2014). Role of abd-A and Abd-B in Development of Abdominal Epithelia Breaks Posterior Prevalence Rule. *PLoS Genet.* 10:e1004717. doi: 10.1371/journal.pgen.1004717

Singh, N. P., and Mishra, R. K. (2015). Specific combinations of boundary element and Polycomb response element are required for the regulation of the Hox genes in Drosophila melanogaster. *Mech. Dev.* 138, 141–150. doi: 10.1016/j.mod.2015.07.016

Sipos, L., Mihály, J., Karch, F., Schedl, P., Gausz, J., and Gyurkovics, H. (1998). Transvection in the Drosophila Abd-B domain: extensive upstream sequences are involved in anchoring distant cis-regulatory regions to the promoter. *Genetics* 149, 1031–1050.

Smith, F. W., Boothby, T. C., Giovannini, I., Rebecchi, L., Jockusch, E. L., and Goldstein, B. (2016). The Compact Body Plan of Tardigrades Evolved by the Loss of a Large Body Region. *Curr. Biol.* 26, 224–229. doi: 10.1016/j.cub.2015.11.059

Snel, B., van Noort, V., and Huynen, M. A. (2004). Gene co-regulation is highly conserved in the evolution of eukaryotes and prokaryotes. *Nucleic Acids Res.* 32, 4725–4731. doi: 10.1093/nar/gkh815

Son, S. H., Son, Y. E., Cho, H. J., Chen, W., Lee, M. K., Kim, L. H., et al. (2020). Homeobox proteins are essential for fungal differentiation and secondary metabolism in Aspergillus nidulans. *Sci. Rep.* 10:6094. doi: 10.1038/s41598-020-63300-4

Soshnikova, N. (2014). Hox genes regulation in vertebrates. *Dev. Dyn.* 243, 49–58. doi: 10.1002/dvdy.24014

Soshnikova, N., and Duboule, D. (2009). Epigenetic temporal control of mouse hox genes *in vivo*. *Science* 324, 1321–1323. doi: 10.1126/science.1171468

Spagnuolo, A., Ristoratore, F., Di Gregorio, A., Aniello, F., Branno, M., and Di Lauro, R. (2003). Unusual number and genomic organization of Hox genes in the tunicate Ciona intestinalis. *Gene* 309, 71–79. doi: 10.1016/S0378-1119(03)00488-8

Spéder, P., Ádám, G., and Noselli, S. (2006). Type ID unconventional myosin controls left-right asymmetry in Drosophila. *Nature* 440, 803–807. doi: 10.1038/nature04623

Spitz, F., Gonzalez, F., Peichel, C., Vogt, T. F., Duboule, D., and Zákány, J. (2001). Large scale transgenic and cluster deletion analysis of the HoxD complex separate an ancestral regulatory module from evolutionary innovations. *Genes Dev.* 15, 2209–2214. doi: 10.1101/gad.205701

Srinivasan, A., and Mishra, R. K. (2012). Chromatin domain boundary element search tool for Drosophila. *Nucleic Acids Res.* 40, 4385–4395. doi: 10.1093/nar/gks045

Srinivasan, A., and Mishra, R. K. (2020). Genomic organization of Polycomb Response Elements and its functional implication in Drosophila and other insects. *J. Biosci.* 45:12. doi: 10.1007/s12038-019-9975-2

Starr, M. O., Ho, M. C. W., Gunther, E. J. M., Tu, Y. K., Shur, A. S., Goetz, S. E., et al. (2011). Molecular dissection of cis-regulatory modules at the Drosophila bithorax complex reveals critical transcription factor signature motifs. *Dev. Biol.* 359, 176–189. doi: 10.1016/j.ydbio.2011.07.028

Suemori, H., and Noguchi, S. (2000). Hox C cluster genes are dispensable for overall body plan of mouse embryonic development. *Dev. Biol.* 220, 333–342. doi: 10.1006/dbio.2000.9651

Sun, D. A., and Patel, N. H. (2019). The amphipod crustacean Parhyale hawaiensis: an emerging comparative model of arthropod development, evolution, and regeneration. *Wiley Interdiscip. Rev. Dev. Biol.* 8:e355. doi: 10.1002/wdev.355

Suzuki, M., and Yagi, N. (1994). DNA recognition code of transcription factors in the helix-turn-helix, probe helix, hormone receptor, and zinc finger families. *Proc. Natl. Acad. Sci. U. S. A.* 91, 12357–12361. doi: 10.1073/pnas.91.26.12357

Takahashi, O., Hamada, J., Abe, M., Hata, S., Asano, T., Takahashi, Y., et al. (2007). Dysregulated expression of HOX and ParaHOX genes in human esophageal squamous cell carcinoma. *Oncol. Rep.* 17, 753–760.

Telford, M. J. (2000). Evidence for the derivation of the Drosophila fushi tarazu gene from a Hox gene orthologous to lophotrochozoan Lox5. *Curr. Biol.* 10, 349–352. doi: 10.1016/S0960-9822(00)00387-0

Thickett, C., and Morgan, R. (2002). Hoxc-8 expression shows left-right asymmetry in the posterior lateral plate mesoderm. *Gene Expr. Patterns* 2, 5–6. doi: 10.1016/S0925-4773(02)00353-2

Tschopp, P., Tarchini, B., Spitz, F., Zakany, J., and Duboule, D. (2009). Uncoupling time and space in the collinear regulation of Hox genes. *PLoS Genet.* 5:e1000398. doi: 10.1371/journal.pgen.1000398

Turner, F. R., and Kaufman, T. C. (1987). A Developmental Genetic Analysis of the Gene Regulator of postbithorax in Drosophila melanogaster. *Dev. Biol.* 432, 418–432.

Vazquez, J., Muller, M., Pirrotta, V., and Sedat, J. W. (2006). The Mcp Element Mediates Stable Long-Range Chromosome–Chromosome Interactions in Drosophila. *Mol. Biol. Cell* 17, 2158–2165. doi: 10.1091/mbc.E06

Wang, K. C., Helms, J. A., and Chang, H. Y. (2009). Regeneration, repair and remembering identity: the three Rs of Hox gene expression. *Trends Cell Biol.* 19, 268–275. doi: 10.1016/j.tcb.2009.03.007

Wang, Y., Yang, Q., and Wang, Z. (2014). The evolution of nanopore sequencing. *Front. Genet.* 5:449. doi: 10.3389/fgene.2014.00449

Wanninger, A., and Wollesen, T. (2019). The evolution of molluscs. *Biol. Rev.* 94, 102–115. doi: 10.1111/brv.12439

Williams, R. W. (1998). Plant homeobox genes: many functions stem from a common motif. *BioEssays* 20, 280–282. doi: 10.1002/(SICI)1521-1878(199804)20:4<280::AID-BIES2<3.0.CO;2-U

Yamada, K., Maeno, A., Araki, S., Kikuchi, M., Suzuki, M., Ishizaka, M., et al. (2021). An atlas of seven zebrafish hox cluster mutants provides insights into sub/neofunctionalization of vertebrate hox clusters. *Development* 148:dev198325. doi: 10.1242/DEV.198325

Yarger, A. M., and Fox, J. L. (2016). Dipteran Halteres: perspectives on Function and Integration for a Unique Sensory Organ. *Integr. Comp. Biol.* 56, 865–876. doi: 10.1093/icb/icw086

Zakany, J., Kmita, M., and Duboule, D. (2004). A Dual Role for Hox Genes in Limb Anterior-Posterior. *Science* 51, 1669–1672.

Zhai, Y., Kuick, R., Nan, B., Ota, I., Weiss, S. J., Trimble, C. L., et al. (2007). Gene expression analysis of preinvasive and invasive cervical squamous cell carcinomas identifies HOXC10 as a key mediator of invasion. *Cancer Res.* 67, 10163–10172. doi: 10.1158/0008-5472.CAN-07-2056

Zhang, X., Hamada, J. I., Nishimoto, A., Takahashi, Y., Murai, T., Tada, M., et al. (2007). HOXC6 and HOXC11 increase transcription of S100β gene in BrdU-induced *in vitro* differentiation of GOTO neuroblastoma cells into Schwannian cells. *J. Cell. Mol. Med.* 11, 299–306. doi: 10.1111/j.1582-4934.2007.00020.x

Zhou, J., Barolo, S., Szymanski, P., and Levine, M. (1996). The Fab-7 element of the bithorax complex attenuates enhancer-promoter interactions in the Drosophila embryo. *Genes Dev.* 10, 3195–3201. doi: 10.1101/gad.10.24.3195

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Cell and Developmental Biology

REVIEW
published: 18 October 2021
doi: 10.3389/fcell.2021.731996

Hox Proteins in the Regulation of Muscle Development

*Gabriela Poliacikova, Corinne Maurel-Zaffran, Yacine Graba and Andrew J. Saurin**

Aix-Marseille University, CNRS, IBDM, UMR 7288, Marseille, France

Hox genes encode evolutionary conserved transcription factors that specify the anterior–posterior axis in all bilaterians. Being well known for their role in patterning ectoderm-derivatives, such as CNS and spinal cord, Hox protein function is also crucial in mesodermal patterning. While well described in the case of the vertebrate skeleton, much less is known about Hox functions in the development of different muscle types. In contrast to vertebrates however, studies in the fruit fly, *Drosophila melanogaster*, have provided precious insights into the requirement of Hox at multiple stages of the myogenic process. Here, we provide a comprehensive overview of Hox protein function in *Drosophila* and vertebrate muscle development, with a focus on the molecular mechanisms underlying target gene regulation in this process. Emphasizing a tight ectoderm/mesoderm cross talk for proper locomotion, we discuss shared principles between CNS and muscle lineage specification and the emerging role of Hox in neuromuscular circuit establishment.

Keywords: Hox, muscle, patterning, mesoderm, neuromuscular

OPEN ACCESS

Edited by:
Edwina McGlinn,
Monash University, Australia

Reviewed by:
Carlos Estella,
Centro de Biología Molecular Severo Ochoa (CSIC-UAM), Spain
Alain Vincent,
FR3743 Centre de Biologie Intégrative (CBI), France

***Correspondence:**
Andrew J. Saurin
andrew.saurin@univ-amu.fr

Specialty section:
This article was submitted to Morphogenesis and Patterning, a section of the journal Frontiers in Cell and Developmental Biology

Received: *28 June 2021*
Accepted: *28 September 2021*
Published: *18 October 2021*

Citation:
Poliacikova G, Maurel-Zaffran C, Graba Y and Saurin AJ (2021) Hox Proteins in the Regulation of Muscle Development. Front. Cell Dev. Biol. 9:731996. doi: 10.3389/fcell.2021.731996

INTRODUCTION

Homeotic or *Hox* genes are highly conserved homeodomain transcription factors that play a fundamental role in bilaterian animal body patterning (Pearson et al., 2005). Several characteristics are at the core of Hox gene function. First, Hox genes are differentially expressed along the anterior–posterior axis of the embryo allocating distinct morphological identities to each body part. Second, manipulation of their expression often results in spectacular homeotic transformations, where the morphology of one given body part is transformed into that of another. Third, Hox genes are grouped in clusters: one split cluster in *Drosophila*, the *Antennapedia* complex, ANT-C, harboring the Hox genes *lab*, *pb*, *Dfd*, *Scr*, *Antp* and the *Bithorax* complex, BX-C, harboring the Hox genes *Ubx*, *abd-A*, *Abd-B*; and at least four clusters in vertebrates (*HoxA-HoxD*) with each cluster harboring 1–13 Hox genes. Genomic clustering of Hox genes in such a way is essential for their correct spatio-temporal expression, which is controlled by important regulatory elements located within and around these clusters. Fourth, they all bind to a very similar set of "AT"-rich DNA-binding sites, achieving functional specificity by cooperating with transcriptional cofactors, the best characterized being PBC proteins (Extradenticle/Exd in *Drosophila*) and MEIS proteins (Homothorax/Hth in *Drosophila*), also encoding for homeodomain transcription factors. However, a large number of Hox-PBC/Meis independent functions have been reported and reciprocally, PBC/Meis proteins can function without binding to Hox (Mallo and Alonso, 2013; Rezsohazy et al., 2015).

Besides their canonical role in providing spatial coordinates that pattern the developing embryo, Hox genes play "non-homeotic" roles where they are involved in the regulation of virtually all basic cellular processes including cell death, cell proliferation, migration and differentiation, as well as in the development of whole structures/organs (Hombría and Lovegrove, 2003; Sánchez-Herrero, 2013). Yet relatively little is known into how much Hox proteins play a role in specification and development of vertebrate muscles, undoubtedly in part due to the large muscle diversification that exists in vertebrates, the large number of Hox genes present (more than 30), and a significant functional redundancy between paralogous group members. In *Drosophila melanogaster* however, there is now abundant evidence for Hox involvement in the patterning of several mesodermal derivatives, including somatic, cardiac and visceral muscles. It is now widely accepted that *Drosophila* myogenesis recapitulates, even though at different scales, all major muscle developmental events that are also at work in vertebrates, such as progenitor specification, myoblast migration and fusion, muscle attachment to tendons cells and assembly of contractile sarcomeres. Furthermore, a part of gene regulatory networks crucial for correct myogenesis such as *twist*, *mef2*, *lbx/ladybird* and the fusion machinery are well conserved (Schnorrer and Dickson, 2004; Taylor, 2006). Here, we focus specifically on Hox function in muscle precursor specification and in patterning of mesodermal derivatives, highlighting Hox target genes and their regulatory mechanisms, when available. We separate the topic in three sections, somatic/skeletal, cardiac and visceral muscles and in each section, we review available data from first *Drosophila* and then from vertebrates. In the somatic/skeletal section we further discuss an emerging role of Hox in the establishment of proper muscle-motoneuron connections. We distinguish Hox specific and non-specific, so-called generic functions, the first referring to functions performed by a single Hox that cannot be assumed by any other and the latter to functions that can be performed by several Hox genes (reviewed in Saurin et al., 2018).

SOMATIC/SKELETAL MUSCULATURE

Basics on Somatic/Skeletal Muscle Development

The somatic muscle development in *Drosophila* as well as the functional conservation with vertebrate muscles have been extensively reviewed (Taylor, 2006; Dobi et al., 2015; Schulman et al., 2015; Gunage et al., 2017; Poovathumkadavil and Jagla, 2020). Briefly, the *Drosophila* life cycle comprises two mobile stages, the larval stage where the crawling movements enable larval feeding and the adult stage where flies can fly, jump and walk. Distinct sets of muscles, produced by two rounds of myogenesis are used during each stage, with larval muscles being produced during embryogenesis and the adult muscles during metamorphosis. Interestingly, both groups develop from mesoderm-derived somatic muscle progenitors marked by high *twist* expression, even though at different developmental timepoints (Baylies and Bate, 1996). Embryonic muscle progenitors are singled out from so-called equivalence groups or promuscular clusters in each hemisegment by lateral inhibition via Notch signaling at stages 11–12 (Carmena et al., 2002). The remaining cells of the cluster become fusion-competent myoblasts, providing mass to the growing muscle by fusing with it. After specification, muscle progenitors undergo a symmetrical division producing two muscle founder cells (FCs), or an asymmetrical division producing either one FC and one adult muscle progenitor (AMP) or one FC and one pericardial progenitor (Ruiz Gómez and Bate, 1997; Carmena et al., 1998). FCs seed individual embryonic muscles and express a unique combination of identity transcription factors (iTFs), controlling all muscle characteristics, including muscle size, position, innervation and attachment. During embryonic stages 12–15 fusion competent myoblasts fuse with FCs assuring muscle growth. Until larval hatching, muscles are properly oriented, attached to tendons and innervated.

On the other hand, AMPs/myoblasts that retain high *twist* expression do not differentiate directly after their specification but are set aside in a quiescent state and associated with imaginal disks and nerves. During larval stages, myoblasts proliferate extensively until the beginning of metamorphosis where they migrate, fuse and differentiate at an appropriate body position to constitute the adult body musculature. For the development of the most prominent adult muscles, the indirect flight muscles (IFMs) composed of dorsal longitudinal muscles (DLM) and dorsal ventral muscles (DVM), two different strategies are employed: The DLMs that span the thorax antero-posteriorly develop by the fusion of myoblasts with three larval scaffolds that escape the histolysis process during first hours of pupal development; the DVMs, spanning the thorax dorso-ventrally, develop by *de novo* myoblast fusion without any scaffold, like in the case of the larval musculature (Fernandes et al., 1991). The larval and adult myogenic process are thus intimately linked since muscle progenitors giving rise to both larval and adult muscles are specified during embryogenesis (even though they do not differentiate at the same time), the fusion process seems to involve the same molecular players and certain larval muscles are reused for the development of a group of adult muscles, like the DLM flight muscles (**Figure 1A**).

In vertebrates, like in *Drosophila*, muscles develop from mesoderm-derived muscle progenitors (reviewed in Endo, 2015; Musumeci et al., 2015; Chal and Pourquié, 2017). Briefly, skeletal muscles develop from the paraxial mesoderm, bilaterally flanking the neural tube. The paraxial mesoderm in the trunk region transiently and progressively subdivides into somites, which are themselves compartmentalized into the dorsal epithelial dermomyotome (giving rise to muscles, dermis, and brown fat) and ventral mesenchymal sclerotome (giving rise to the axial skeleton, cartilage, and tendons). A layer of muscle precursors/myocytes called myotome forms beneath the dermomyotome at E8 in the mouse and E2.5 in the chick. Muscle precursors after their differentiation give rise to the body wall muscles and some undergo long-range migration toward future limb buds where they proliferate and differentiate into limb muscles. In the cranial region, paraxial mesoderm forms the cranial mesoderm that will give rise to the muscles,

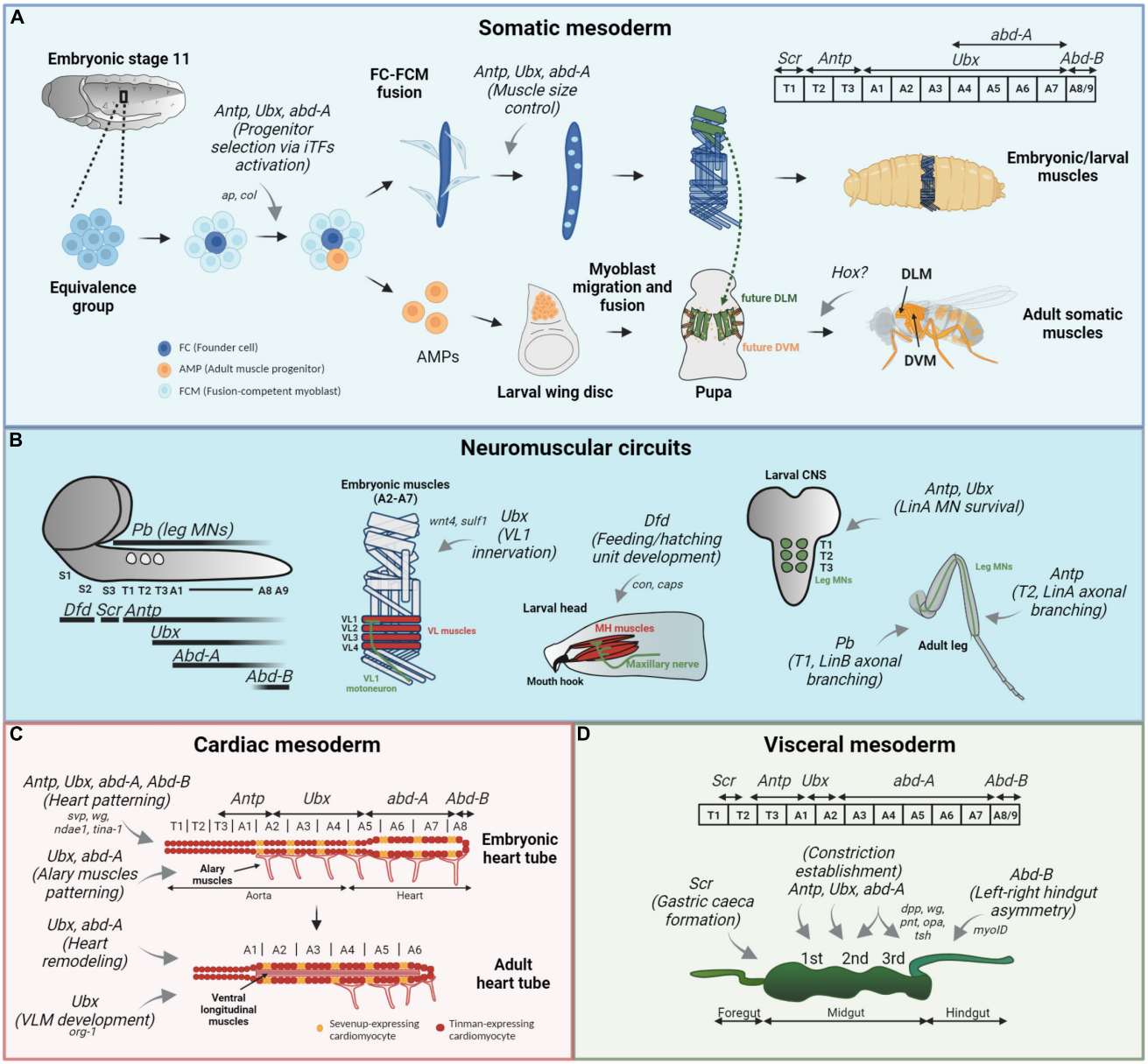


FIGURE 1 | Hox functions in *Drosophila* muscle patterning. Hox functions in the somatic **(A)** mesodermal patterning, neuromuscular circuit establishment **(B)** and cardiac **(C)** and visceral **(D)** mesodermal patterning are depicted by gray arrows, and their target genes are specified, if available. Their expression patterns are showed for each mesodermal derivative. **(A)** In the somatic muscle development, Antp, Ubx, and AbdA regulate progenitor selection from different equivalence groups (by binding to a muscle specific *cis*-regulatory region of *ap* and *col*, in LT and DA3 muscles, respectively) and later control muscle size by specifying the number of FCM (light blue cells) that will fuse with the FC (dark blue cell). The role for Hox in adult musculature, derived from AMPs (orange cells) is not clearly established. **(B)** Hox expression pattern in the larval CNS is depicted on the left, with Pb expression in leg MNs (circles). In this context, Ubx establishes proper connection between embryonic VL1 muscle and VL1-MN, by regulating *wnt4* and *sulf1* expression in the VL2 muscle. Dfd is responsible for the feeding and hatching motor unit development and maintenance in the larva, a process involving target genes *con* and *caps*. Antp and Ubx play a role in LinA lineage MN survival during larval development, and Antp further regulates correct morphology of this neurons in the adult T2 leg. Finally, Pb controls branching of some LinB MNs in T1 adult leg. **(C)** The heart tube is composed of *Svp*-expressing (orange cells) and *Tin*-expressing (red cells) cardioblasts and surrounding pericardial cells (not represented). Antp, Ubx, and AbdA are all involved in the patterning of embryonic heart tube, by controlling the expression of *svp* in the segment-specific manner. Furthermore, AbdA with Svp regulate *wg* expression in cardioblasts expressing Svp. AbdB on the other hand is a suppressor of cardiogenesis. Ubx and AbdA pattern alary muscles present in the abdominal segments. For proper heart remodeling, Ubx repression and AbdA function modulation are required. Finally, Ubx is involved in adult ventral longitudinal muscle (VLM) development. **(D)** In the embryonic visceral mesoderm, Antp establishes the first, Ubx and AbdA the second and AbdA the third midgut constriction. A complex genetic network involving *dpp*, *wg*, *pnt*, *opa*, and *tsh* is involved in the middle constriction development. Scr is involved in the formation of gastric caeca in the anterior midgut (caeca are not represented) and AbdB is responsible for the correct left-right embryonic hindgut asymmetric morphology via *myoID* regulation.

connective tissue and skeleton of the face and skull. The myogenic process can be subdivided into primary/embryonic myogenesis (E10.5–E12.5 in mice and E3–E7 in the chick) and secondary/fetal myogenesis (E14.5–E17.5 in mice and E8+ in the chick). In vertebrates, individual muscles are not seeded by a single founder myoblast like in *Drosophila*, but by scaffolds of primary fibers composed of several fused muscle precursors. Subsequently, additional fibers are added along them to assure the muscle growth.

Hox Control of *Drosophila* Embryonic/Larval Somatic Musculature

In the embryonic somatic mesoderm, *Hox* gene expression has a spatially restricted pattern, where *Sex combs reduced* (*Scr*) is expressed in the first thoracic segment, T1, *Antennapedia* (*Antp*) in T2 and T3, *Ultrabithorax* (*Ubx*) in abdominal segments A1–A7, *abdominal*-A (*abd-A*) in A4–A8 and *Abdominal-B* (*AbdB*) in A8–A9 (Michelson, 1994; Enriquez et al., 2010; **Figure 1A**). Their implication in the subdivision of somatic mesoderm began with the description of mutant alleles of *AbdB* (Lawrence and Johnston, 1984) and *Ubx* (Hooper, 1986) shown to transform larval muscles into more anterior ones. Conversely, the overexpression of Ubx or AbdA in the embryonic mesoderm led to the abdominal transformation of thoracic muscle precursors (Greig and Akam, 1993; Michelson, 1994). Collectively, these studies demonstrated a clear role for Hox genes in muscle patterning and argued toward a cell-autonomous role of Hox in the somatic mesoderm, meaning that Hox genes would confer a specific identity to the muscle precursors and account for segment-specific differences in the muscle pattern. Further evidence for cell-autonomous Hox function came from a pioneering study showing that Antp directly regulates the expression of the *apterous* muscle identity gene in the somatic mesoderm, specifying a subset of lateral transverse muscles (LT1–4) (**Figure 1A**). Importantly, it was shown that this regulation occurs via an *apterous* muscle-specific enhancer, providing a molecular mechanism by which a Hox protein can directly regulate its muscle target gene expression and contribute to the establishment of segment-specific muscle patterns via progenitor selection (Capovilla et al., 2001). Whether Ubx/AbdA specify abdominal LT1–4 muscles and/or regulate *apterous* expression was not addressed.

Building on the same principle, it was found that Antp, Ubx, and AbdA regulate progenitor selection via distinct *cis*-regulatory modules in a segment-specific manner (**Figure 1A**). Focusing on the dorsal acute DA3 muscle lineage, specified by the combinatorial activity of *nautilus* (*nau*, *Drosophila* ortholog of mammalian *MyoD*) and *collier* (*col*, ortholog of mammalian early B-cell factors, EBFs) it was shown that the progenitor specification via Hox-mediated regulation of muscle iTFs is superimposed on the A-P and D-V positional information. Furthermore, the precise timing of Hox activity was traced to the progenitor stage, where first, Hox activity controls the number of progenitors that express *col* and *nau* in a segment-specific manner and then Hox proteins interplay with iTFs to allocate a correct number of nuclei assigned to the DA3 muscle (Enriquez and Vincent, 2010; Enriquez et al., 2010; **Figure 1A**). Interestingly, this study highlighted a Hox generic function, where Antp, Ubx, and AbdA can all provide DA3 progenitor selection (reviewed in Saurin et al., 2018). This generic function is then followed by a Hox specific function, where each Hox determines the final size of the muscle in a segment-specific manner. These studies provided an important insight into the function of Hox genes in myogenesis, showing that Hox inputs were crucial in the process of progenitor selection, which gives rise to "founder" cells, seeding the formation of syncytial muscle fibers.

It was recently shown that Ubx is involved in the muscle differentiation process by directly repressing the master mesodermal regulator *twist* (*twi*) (Domsch et al., 2021). Here, a number of interesting points can be highlighted: (1) Ubx mesodermal downregulation interferes with abdominal embryonic muscle development because of failure to repress *twi*; (2) *twi* repression is mediated by direct Ubx binding to the *twi* promoter region, competing with binding of the transcription factors Tinman (Tin) and Muscle enhancer factor 2 (Mef2), providing another promoter-based molecular evidence about the Hox muscle patterning mechanism; (3) even though Ubx binding displaces Tin on the *twi* promoter, Tin must be bound to the promoter for Ubx to be recruited, once again demonstrating a tight link between Hox and tissue-specific transcription factors for the correct patterning establishment.

Collectively, these studies provide precious insight into the way muscle progenitors are patterned. Hox proteins activate and collaborate with some of the muscle iTFs in a segment-specific fashion and by doing so they are responsible for the specification of different muscle types.

Is Patterning of the T2 Mesoderm Hox Independent?

As the understanding of muscle development evolved, an important role in muscle patterning was attributed to the ectoderm. A large number of studies argued toward a Hox non-autonomous role in myogenesis, giving importance to Hox-controlled signals coming from the overlying epidermis and nervous systems (Roy and Raghavan, 1997; Roy et al., 1997; Rivlin et al., 2001; Dutta et al., 2010). For a long time, the patterning of mesoderm in T2 was considered as being non-autonomous, because of the lack of any Hox gene expression in this segment. In complement to this observation, it was shown that overexpression of Antp (highly expressed in T2 and T3 epidermis and nervous system) in the ectoderm of *Antp* mutant embryos completely rescued the disorganized T2 muscle pattern provoked by Antp loss-of-function (Roy et al., 1997). This was taken as evidence for a non-autonomous role of Antp in T2 muscle pattern establishment. However, Antp was later described as being expressed in the T2 mesoderm, although at a much weaker level than in the T3 mesoderm and to autonomously specify a subset of muscles in the T2 (Capovilla et al., 2001). Thus, the default in T2 muscle patterning in *Antp* mutant embryos is either pointing toward an autonomous role of Antp in the mesoderm or toward a combination of autonomous activity and inductive cues coming from the T2 ectoderm. It was also shown that Antp mesodermal overexpression did not modify T3 muscle

pattern, but was sufficient to transform the T2 into a T3 muscle pattern, which was interpreted as another evidence for the lack of Antp T2 expression (Roy et al., 1997). Knowing that Antp expression in the T2 mesoderm is weaker compared to that of T3, higher levels of Antp in T2 may convert its muscle pattern into a more posterior one, without excluding the presence of Antp in T2 mesoderm.

Hox Control of *Drosophila* Adult Musculature

As mentioned earlier, *Drosophila* adult musculature develops from pools of myoblasts associated with imaginal tissues and nerves. The most prominent adult muscles, thoracic flight muscles arise from myoblasts associated with the wing imaginal disks. Knowing that T2 mesoderm was mistakenly considered as a Hox-free region because of lack of *Antp* expression, it is possible that T2 wing disk-associated myoblasts also express Antp. In support of this, it was reported that Antp is transcribed from two distinct promoters, termed P1 and P2, with transcripts from P1 being localized to the anterior part of the wing disk epithelium and P2 transcripts in the presumptive notum region containing myoblasts (Jorgensen and Garber, 1987). Recently, antibody stainings have confirmed the presence of Antp protein in the myoblast-containing region of the wing disk (Paul et al., 2021). If indeed the expression of Antp is confirmed by appropriate markers in adult muscle precursors [such as Mef2, Holes in muscles (Him), or Twi], then it is highly possible that Antp is directly involved in T2 adult muscle development, including muscles used for flight and leg muscles.

Early myoblast transplantation experiments showed that T2 myoblasts, associated with the wing imaginal disks, can contribute to a vast population of adult muscles (Lawrence and Brower, 1982). Since T2 wing-associated myoblasts were considered as a Hox-free region (see above), it was proposed that they do not require any positional Hox input for their migration and fusion. Supporting this view, mesodermal overexpression of Ubx does not perturb thoracic myoblast migration pattern, demonstrating that the myoblast migration process is likely independent of Hox positional input. These myoblasts however fail to give rise to a proper adult flight muscle, because of their inability to activate Act88F, a fibrillar muscle differentiation marker, that is repressed due to Ubx overexpression (Roy and Raghavan, 1997).

Concerning myoblast fusion, the importance of Hox cues has not yet been clearly established. It is important to note that myoblasts transplanted from the second or the third thoracic segment, associated with wing and haltere imaginal disks, respectively, can both fuse to and contribute to abdominal muscles (Roy and Raghavan, 1997). It is thus possible that Antp could be involved in the thoracic myoblast fusion process, a hypothesis that has never been directly addressed because of a lack of appropriate tools. Rivlin et al. (2001) demonstrated using allelic combinations of Ubx mutations, leading to different levels of ectodermal transformation, that the transformed T3 segment could contain IFM, normally present solely in T2 (Rivlin et al., 2001). This placed ectodermal inductive cues in the central position of adult IFM patterning. However, this observation has not been reproduced (Egger et al., 1990; Fernandes et al., 1994; Dutta et al., 2010), leaving the role for Hox proteins in adult muscle development not elucidated. We note that no Hox regulatory networks were identified for the development of adult abdominal and leg muscles, leaving the role of Hox in adult muscle development largely undetermined.

It is noteworthy that the well-characterized Hox PBC cofactors, Exd and Hth are involved in the adult muscle development, controlling the fate decision between fibrillar flight and tubular leg muscles, but appear to do so in a Hox-independent fashion (Bryantsev et al., 2012). Knowing that for the proper somatic musculature patterning, Hox proteins use other muscle-specific transcription factors as cofactors, it is thus likely that Hox control of somatic muscle development is independent of their canonical cofactors and reciprocally, Exd/Hth function is independent of Hox proteins in this context.

Hox Control of Vertebrate Skeletal Musculature

In vertebrates, Hox involvement in skeletal, myotome-derived musculature patterning remains controversial. In the limb musculoskeletal system, Hox genes have an essential role in skeletal and connective tissue patterning, but limb muscle precursors seem to depend completely on environmental cues for their proper development (reviewed in Pineault and Wellik, 2014). Early grafting experiments in the chick/quail embryos demonstrated that muscle precursors do not possess intrinsic patterning information and their development is influenced by the surrounding mesenchyme (Chevallier et al., 1977; Christ et al., 1977). This view has been later challenged by a study suggesting that the axial identity of the somite is important for the generation of non-limb skeletal muscles, arguing toward an autonomous role for Hox in the body wall muscle patterning (Nowicki and Burke, 2000; Alvares et al., 2003). Furthermore, microarray analyses on purified skeletal muscle myoblasts showed that the Hox code is present along the cranio-caudal axis, specifically in embryonic, but not fetal myoblasts, indicating that myoblasts carry some intrinsic positional information (Biressi et al., 2007).

Direct evidence for Hox role in skeletal muscle patterning came from the mouse forelimb zeugopod, where Hoxa11 and Hoxd11 are expressed in the muscle connective tissue and tendons, but not in differentiated muscle cells (Swinehart et al., 2013). In Hoxa11/d11 double mutant mice, several muscles and tendons of the forelimb zeugopod are absent or improperly patterned, and importantly, this is a direct consequence of loss of Hox function and not a secondary effect due to defects in the skeletal patterning. In addition to extending the well-established role of Hox in the patterning of the axial skeleton to muscles and tendons, this study also reinforces the initial view that muscle precursors are patterned by their environment, at least at the level of the limb. However, it is still unknown whether the Hox positional information is conveyed by the muscle mesenchyme

or is encoded in muscle precursors themselves, which could be dependent on the somite axial level.

Recently, much attention has been given to the role of Hox proteins in adult muscle stem cells (MuSCs), also known as satellite cells, able to regenerate adult muscles upon injury (Machado et al., 2017; Evano et al., 2020; Yoshioka et al., 2021). It is now accepted that depending on the anatomical position of the muscle, satellite cells display heterogeneity in their proliferative and regenerative properties (Ono et al., 2010). Searching for underlying molecular determinants, Hox-A and Hox-C clusters were found to have different methylation profiles and thus be differentially expressed in adult muscles and their satellite cells derived from somites, compared to the ones derived from the cranial mesoderm (Evano et al., 2020; Yoshioka et al., 2021). Knowing that the cranial-derived, Hox-free, satellite cells display higher regenerative capacity than the limb, Hox-expressing satellite cells, an appealing possibility is that Hox could be involved in this process (Evano et al., 2020). Indeed, it has been demonstrated that during the aging process, *Hoxa9* is up-regulated in the limb satellite cells, leading to their cell cycle entry default and senescence (Schwörer et al., 2016). It does so by regulating the targets of several developmental pathways, including those of the Wnt, TGFβ, and JAK/STAT pathways. Interestingly, *Hoxa9* deletion in satellite cells from aged adult mice was sufficient to improve their regenerative capacity, suggesting that Hox expression in these cells would have a negative effect on muscle regeneration. A conditional depletion of *Hoxa10* leads to a repair default of some somite, but not cranio-derived muscles, explained by a genomic instability and consequent proliferation arrest of adult satellite cells (Yoshioka et al., 2021). Different Hox would thus assume opposed functions in satellite cell proliferation and could account for different capacities of distinct muscle groups to regenerate, a hypothesis that needs to be investigated further. In *Drosophila*, a muscle satellite cell population has been identified in adult thoracic flight muscles (Chaturvedi et al., 2017), but it remains unknown whether Hox proteins are involved in their transcriptional regulation like in vertebrates.

Hox Control of *Drosophila* Neuromuscular Circuits

Hox involvement in central nervous system development is clearly established and is out of the scope of this review (Rogulja-Ortmann and Technau, 2008; Estacio-Gómez and Díaz-Benjumea, 2014; Meng and Heckscher, 2021). It is nevertheless noteworthy here that several parallels can be drawn between neuronal and somatic muscle lineage specification by Hox. In both tissues Hox act at the very early steps of progenitor specification. Like in the larval DA3 muscle where Hox allocate a correct number of progenitors and further control muscle size in a segment-specific manner, in the neuroectoderm their expression is required to specify NB1-1 derived thoracic and abdominal lineage comprising a different number of neurons (Prokop and Technau, 1994; Prokop et al., 1998; Enriquez et al., 2010). Hox can convey the proper tissue pattern via the regulation of specific TFs, shared across distinct tissues. For example, Antp regulates the expression of *collier* (*col*) in both muscle and neural clusters to generate muscle/neuronal diversity (Enriquez et al., 2010; Karlsson et al., 2010). Yet the transcriptional control mechanisms at play would appear different, since the *cis*-regulatory element used by Antp in the muscle lineage is distinct from the CNS regulatory sequence, which has to date not been identified. One important difference though, is that in the CNS, the Hox cofactors Exd and Hth are directly involved in the control of *col* expression. Like in the somatic and cardiac mesoderm, AbdB is able to suppress neuronal fate in the most posterior abdominal segments (Lovato et al., 2002; Ponzielli et al., 2002; Birkholz et al., 2013).

Crosstalk between neuronal and muscle lineages, more specifically between motor neurons (MNs) and somatic/skeletal muscles are at the basis of voluntary movements, crucial for locomotion, feeding, mating and interactions with the environment. For their establishment, MNs need to be correctly specified and differentiated and project their axons toward specific muscle groups. Neuromuscular circuitry defaults are associated with numerous neuromuscular diseases, thus the understanding of their proper development has a direct clinical significance (reviewed in Tripodi and Arber, 2012; Li et al., 2018).

In the CNS, Hox expression is shifted posteriorly compared to their mesodermal expression. Dfd is expressed in two subesophageal segments (S1 and S2), Scr is expressed in S3, Antp is expressed in the ventral cord in T1 to A9, Ubx from T3 to A9, Abd-A from A1 to A9 and Abd-B from A7 to A9. Pb is expressed in the thoracic leg motoneurons (Baek, 2011; **Figure 1B**). Hox involvement in neuromuscular circuitry establishment was first proposed in the context of larval crawling movements, following the observation that Ubx and AbdA are necessary for their generation, providing a genetic explanation for locally specialized locomotor circuit establishment (Dixit et al., 2008). This study suggested that Ubx and AbdA not only specify larval abdominal muscles required for these peristaltic movements, but also the neuronal circuitry allowing for the properly synchronized movements. Moreover, the single removal of either Ubx or AbdA did not compromise the peristaltic movement, demonstrating a genetic redundancy in Hox function since Ubx and AbdA can substitute for each other in this context. Ubx and Antp have been shown to be required for motoneuron segmental diversity in the embryo (lineages NB7-3 and NB2-4t), by regulating cell-death and cell-survival, respectively, in an antagonistic manner (Rogulja-Ortmann et al., 2008). Several Ubx and Antp putative binding sites were identified in the pro-apoptotic *reaper* (*rpr*) gene enhancer, suggesting that their competitive binding could account for their opposed regulatory mode of motoneuron survival.

An example of a molecular mechanism behind the role of Ubx/AbdA in locomotor circuit establishment has been elucidated in embryonic abdominal ventrolateral (VL) muscles. Ubx is expressed in both muscle cells and MNs and is required for the correct establishment of contacts between them (**Figure 1B**). In the VL2 muscle, it controls the activation of Wnt4 signaling as it does in the visceral mesoderm (Graba et al., 1995), and the expression of Sulf1, a sulfatase implicated in Wnt and BMP gradient establishment at neuromuscular junctions. In the VL1

MN, Ubx interacts with the components of the Wnt4 pathway. Signaling molecules regulated by Ubx in VL2 upon their secretion serve for the proper VL1 MN axonal extension toward the more dorsal VL1 muscle. Interestingly, it was suggested that Ubx, TCF, and Armadillo (Arm) can form a Wnt4-induced transcriptional complex (Hessinger et al., 2017). The requirement of Hox in both muscles and MNs was also demonstrated in larval hatching and feeding motor units, with the central Hox, *Dfd* (Friedrich et al., 2016). Here, Dfd assures correct innervation of muscles required for mouth hook movements by the maxillary nerve (**Figure 1B**). This demonstrates that Hox could provide a regulatory code for the correct muscle-motoneuron recognition in different *Drosophila* neuromuscular circuits. In line with this hypothesis, taking advantage of available ChIP-seq data, several potential Dfd target genes with functions in muscles, nerves and synaptic recognition were identified, such as *Connectin* (*Con*) and *capricious* (*caps*) (Friedrich et al., 2016).

Concerning adult muscle innervation, the role of Hox is well established in the case of leg muscles, where the second leg pair is required for flight take-off, and T1 and T3 legs are required for grooming the head and the abdomen, respectively (Baek et al., 2013; Enriquez et al., 2015). Focusing on a major subpopulation of leg MNs, arising from the LinA (also called Lin15B) lineage, it was shown that during the third larval stage, Antp is expressed in the newborn LinA MNs in all three thoracic segments, whereas Ubx was localized only in T3 subpopulation. During mid-pupal stage, the expression of Hox changes, with Antp expression being confined to all LinA MNs solely in T2 and Ubx solely in T3 and this segment-specific expression pattern is maintained until the adult stage. Hox cofactors Exd and Hth are ubiquitously present from the larval until the adult stage. In this context, Antp and Ubx with their cofactors are required for LinA MN survival, and Antp with Hth are further required for the correct axonal and dendritic morphology and axonal branching (**Figure 1B**). We note that Hox proteins do not specify leg motoneurons *per se*, but assure instead motoneuron survival and offer a unique code for their correct branching to distinct leg parts. Interestingly, this highlighted the importance of Antp protein levels, serving as a timing mechanism for correct proximal (early born) vs. distal (late-born) axon targeting (Baek et al., 2013). Antp protein levels seem to play an instructive role as well in T2 vs. T3 somatic muscle pattern establishment (as mentioned above), suggesting that several Hox functions could be dose-dependent.

While looking for a genetic explanation for distinct morphological characteristics of individual motoneurons, an important role was attributed to anterior Hox *proboscipedia* (*pb*). Concentrating on the LinB (Lin24) lineage of leg MNs, it was demonstrated that Pb, expressed in three of seven neurons composing this population, is required for their proper dendritic morphology and axonal patterning (**Figure 1B**). Clonal removal of *pb* in T2 LinB MNs affected the linearity of path/stability during high speed walking (Enriquez et al., 2015).

A role for Hox in adult muscle innervation was also shown with Ubx whose misexpression in T2 MNs compromised adult IFM development (Dutta et al., 2010). Collectively, muscle development and homeostasis require both Hox autonomous and non-autonomous function, in the *Drosophila* mesoderm and neuroectoderm, respectively.

Hox Control of Vertebrate Neuromuscular Circuits

In vertebrates, abundant evidence argues toward an important Hox function in motor circuit establishment (reviewed in Jung and Dasen, 2015). For the correct locomotor circuitry establishment, a large variety of different MN subtypes need to form precise connections with target muscles. A well-studied group of MNs, spinal MNs display different columnar, divisional and pool identities allowing them to contact more than 50 different limb muscles. Concentrating on chick brachial lateral motor columnar (LMC) neurons, it was shown that Hox3, 4, 5, 7, and 8 proteins divide LMC into subdomains along the rostrocaudal axis. Importantly, changing MN transcriptional identity by manipulation of Hox protein levels resulted in corresponding changes in muscle connectivity. It has been proposed that in this context, Hox proteins confer different pool identities by regulating downstream transcription factors in MNs, such as Nkx6 (NK6 Homeobox 1) (Dasen et al., 2003, 2005). Hox6 proteins do not have a role in the initial LMC specification but are required for further LMC pool identity establishment and proper limb innervation (Lacombe et al., 2013). In the search for the link between neuronal identity specification by Hox and muscle innervation in chick and mice, the transcription factor FoxP1 (Forkhead box protein P1) was identified as a Hox accessory factor, allowing to fine-tune Hox output (Dasen et al., 2008). The molecular mechanism behind this process also involves cell surface receptor encoding genes *Ret proto-oncogene* (*Ret*) and *Glial cell line-derived neurotrophic factor receptor alpha* (*Gfrα*). In this context, Hox proteins in collaboration with their cofactor Meis1 were shown to regulate the spatial pattern and expression levels of these genes in LMC neurons, required for proper MN differentiation and connectivity (Catela et al., 2016). Interestingly, digit-innervating MNs in chick and mice also employ a Hox code for their specification, that is however different from the one used in more proximal limb muscles. In this context, joint expression of Hoxc8 and Hoxc9 are required for correct digit innervation (Mendelsohn et al., 2017).

Hox proteins not only control LMC neuromuscular circuitry at the limb level, but also at the thoracic level, where MNs innervating hypaxial muscles are specified by Hoxc9, acting as a repressor of a limb-innervating MN fate (Jung et al., 2010). Non-limb innervating MNs at the cervical spinal cord level within the phrenic motor column (PMC) also require Hox for their correct development. Interestingly, mice lacking *Hox5* genes (*Hoxa5* and *Hoxc5*) in these neurons die of respiratory failure as a consequence of altered diaphragm innervation (Philippidou et al., 2012; Philippidou and Dasen, 2013). Furthermore, spinocerebellar tract neurons (SCTNs) that relay sensory/proprioceptive information to the CNS from muscles and tendons also use Hox-dependent transcriptional program for their diversification. Discrete populations of SCTNs along rostro-caudal axis display a combinatorial expression of several Hox genes (from *Hox4* to *Hox10*) and their genetic

manipulation leads to defaults in muscle-cerebellum connectivity (Baek et al., 2019). Collectively, both spinal MNs and sensory SCTNs use Hox-code for their proper specification, suggesting a general role for Hox proteins in the proper muscle-neuron connectivity establishment.

CARDIAC MUSCLE

Hox Control of *Drosophila* Cardiac and Alary Muscles

Drosophila cardiac musculature develops during embryogenesis from lateral mesodermal cells that migrate toward the dorsal midline during dorsal closure and form a tube named the cardiac tube, also known as the dorsal vessel (reviewed in Monier et al., 2007; Bataillé et al., 2015; Rotstein and Paululat, 2016). The cardiac tube is subdivided into two parts, with the anterior narrow portion (T1–A4) named aorta and the posterior larger one (A5–A8) named heart, with the hemolymph flowing from the posterior to anterior, assuring its distribution throughout the organism (**Figure 1C**). Aorta can be further subdivided into anterior (T1–T3) and posterior (A1–A4) parts. In the anterior/thoracic aorta, each segment contains four pairs of cardiomyoblasts (CMs) that express the homeodomain-containing transcription factor Tinman (*tin*) (Nkx2.5 in vertebrates). The posterior/abdominal part is constituted of six pairs of CMs, with two anterior ones expressing the orphan nuclear receptor Seven-up (*svp*) (ortholog of Nuclear Receptor Subfamily 2 members) and the four posterior CM pairs expressing *tin* (except A8 that contains only two *tin*-expressing CM pairs). *Svp*-expressing cardioblasts in the heart further differentiate into ostiae, inflow valves that allow for the hemolymph pumping, accounting for the partially open circulatory system in *Drosophila* (Molina and Cripps, 2001). The metamerically repeated expression of *svp*, *tin* and other genes suggested the cell-identity specification in a segment-specific fashion. During metamorphosis, the heart part of the cardiac tube is histolyzed, and the adult heart develops from the larval posterior aorta myocytes that undergo a transdifferentiation without cell proliferation (Monier et al., 2005).

Hox genes have a rather complex expression pattern in the dorsal vessel. The anterior aorta is considered as a Hox-free region. The posterior aorta expresses *Antp* (in part of the T3, in A1 and weakly in A2 CM) and *Ubx* (from A2 to A5 CM) and the heart expresses *abd-A* (from the fifth pair of A5 CM to the second pair in A8) and *AbdB* (in the two posterior CM pairs in A8) in both cardioblasts and at least some pericardial cells (Lo et al., 2002; Lovato et al., 2002; Ponzielli et al., 2002; Zmojdzian et al., 2018; **Figure 1C**). Hox were shown to be responsible for dorsal vessel patterning in a cell-autonomous fashion (reviewed in Monier et al., 2007). Independent studies demonstrated that Ubx and AbdA are responsible for heart cardioblast specification (**Figure 1C**). In *abd-A* null embryos the heart is transformed into the aorta and conversely, its mesodermal overexpression is sufficient to specify aorta as heart cardioblasts instead (Lo et al., 2002; Ponzielli et al., 2002). In *Ubx* or *Ubx/abd-A* double mutants, the anterior part of the aorta is affected (Ponzielli et al., 2002). Antp does not specify cardiac lineage *per se* but generates CM diversity by controlling *svp* expression in the A1 segment (Perrin et al., 2004). Interestingly, *AbdB* mesodermal overexpression suppresses cardiac fate while its loss-of-function leads to supernumerary cardioblasts but also somatic nuclei (Lovato et al., 2002; Ponzielli et al., 2002). The combined mutations of *Scr*, *Antp*, *Ubx*, *abd-A*, and *AbdB* transformed the whole dorsal vessel into the aorta, further showing that the aorta fate is a ground state of the cardiac tube (Perrin et al., 2004).

Even though Hox proteins clearly regulate the cell lineage choice between the anterior aorta and posterior aorta/heart only a few of their target genes involved in this process have so far been identified. Antp, Ubx, and AbdA control *svp* expression in their respective segments in the embryonic heart, a function that can be once again considered as generic (Perrin et al., 2004; Ryan et al., 2005). *Svp* itself was shown to be regulated by Hedgehog (Hh) signaling coming from the ectoderm. Interestingly, it was suggested that cardioblasts can respond to Hh signals only in Hox-expressing cells, explaining why the anterior aorta (a Hox-free region) does not express *svp* even though it receives Hh (Ryan et al., 2005). To explain heart-patterning control by Hox, it was suggested that Hox could regulate symmetric/asymmetric division of progenitors giving rise to *tin* and *svp*-expressing cardioblasts and pericardiac cells, although this hypothesis remains to be tested. AbdA regulates the expression of *Troponin C-akin-1* (*Tina-1*), a heart-specific marker, of unspecified function (Lovato et al., 2002). Some target genes were identified at later stages, where Hox further pattern the individual cardiomyoblasts. AbdA in collaboration with Svp was suggested to activate *Wg* expression in heart cardioblasts expressing *svp*. In the Tin-expressing cardioblasts, AbdA activates expression of a Na^+-driven anion exchanger (*ndae1*), involved in ionic homeostasis (Perrin et al., 2004). No gene level mechanisms explaining the Hox cardiac target gene regulation have however been identified, leaving the possibility of the existence of tissue-specific *cis*-regulatory modules used by Hox, as is the case of the somatic musculature.

Interestingly, ecdysone-dependent repression of Ubx in A1-A4 *tin*-expressing myocytes is required during the mid-pupal stage for adult heart formation (**Figure 1C**). Adult heart develops during metamorphosis by a remodeling of the larval posterior aorta. If Ubx expression is maintained during pupal stages in posterior aorta *tin*-expressing cells, they adopt A5 characteristics, resulting in the adult remodeling alteration (Monier et al., 2005). This argues in favor of a hypothesis stipulating that in the process of organogenesis, Hox input is necessary for the activation of downstream signaling networks but once these are activated, Hox presence is no longer needed and can be even detrimental for the rest of the development (Hombría and Lovegrove, 2003). The modification of AbdA is also required for heart metamorphosis, but occurring at the functional instead of expression level, yet also in an ecdysone-dependent manner (**Figure 1C**). Instead of conferring heart fate to the CMs like AbdA does during embryogenesis, here it regulates the reprogramming of A5 segment that becomes the terminal chamber in the adult. It is interesting to note that the switch in AbdA function occurs also at the transcriptional level, where early during development

AbdA regulates positively *Ih* (a voltage-gated ion channel) expression but represses it in the pupa (Monier et al., 2005). We highlight that Hth is expressed only in the anterior aorta which does not express any Hox (Perrin et al., 2004), and thus Hox function in the *Drosophila* cardiac tube, like in the somatic muscle development (Bryantsev et al., 2012) is independent of Exd/Hth cofactors.

Besides cardiac muscle, Hox also control the patterning of seven pairs of alary muscles (AMs), specialized skeletal muscle connecting the cardiac tube at the level of Svp-expressing pericardial cells, with the lateral exoskeleton (LaBeau et al., 2009; Bataillé et al., 2015; **Figure 1C**). AMs were recently shown to maintain the dorsal vessel at the tracheal trunk position (Bataillé et al., 2020). During metamorphosis, four AM pairs remain, originating from larval posterior AMs (Lehmacher et al., 2012). Interestingly, three anterior AM pairs undergo a process of dedifferentiation and give rise to adult ventral longitudinal muscles (VLMs) of unknown function (Molina and Cripps, 2001; Schaub et al., 2015). Two Hox genes are expressed in AMs at the embryonic stage, *Ubx* in the three most anterior pairs and *abd-A* in the four posterior AM pairs. Consistent with the Ubx expression pattern, in *Ubx* mutant embryos, 2–3 anterior AM pairs do not form. The absence of AbdA does not compromise posterior AM formation, probably because Ubx and AbdA functions are redundant in this context. Conversely, Ubx or AbdA overexpression leads to supernumerary AM (LaBeau et al., 2009). Furthermore, modulation of Hox expression is required for correct AM transdifferentiation, since overexpression of AbdA in anterior AMs, leading to Ubx suppression, prevents VLM formation (**Figure 1C**). Like in the case of cardiac tube remodeling, ecdysone signaling is also required for AM transdifferentiation (Schaub et al., 2015). However, it is not known whether in this case ecdysone pathway modulates Hox activity like during adult heart development. Concerning Hox target genes, *optomor-blind-related-gene-1* (*org-1*, ortholog of vertebrate T-box Transcription factor Tbx1) was proposed as being directly or indirectly regulated by Ubx during adult VLM formation (Schaub et al., 2015). Knowing that *org-1* is also required for embryonic AM development, it is possible that Ubx (and also AbdA) could regulate *org-1* also during this stage (Boukhatmi et al., 2014).

Hox Control of Vertebrate Cardiac Muscle

Many similarities can be found between *Drosophila* and vertebrate cardiac myogenesis even though at a first glance they seem very distinct. Both are developed from mesodermal precursors that converge toward the midline to give rise to a linear, contractile tube, that is further looped and developed into a multi-chambered organ in vertebrates (Zaffran and Frasch, 2002; Zaffran and Kelly, 2012; Lescroart and Zaffran, 2018). In birds and mammalians, cardiogenic precursors/first heart field progenitors (FHF) converging at the anterior midline express Mesp1 (Mesoderm Posterior bHLH Transcription Factor 1) and form a so-called "cardiac crescent." The cardiac crescent develops into a transient heart tube with an inner endocardial and outer myocardial layer that will mainly contribute to the left ventricle. The heart tube elongates by the addition of second heart field (SHF) progenitors, located in the pharyngeal mesoderm (itself formed by cells of both splanchnic and paraxial mesoderm). SHF segregates along the AP axis into posterior SHF, contributing to the atrial myocardium at the venous pole and anterior SHF (also called AHF) that contribute to the outflow tract (OFT) (connecting the ventricles to the future aorta) and the right ventricle at the arterial pole. The heart is finally shaped by rightward looping and myocardium expansion leading to the formation of four integrated cardiac chambers, two ventricles and two atria. A specialized population of neural crest cells (NCC) contribute to the development of large arteries and outflow septum. A large number of inductive signaling molecules have been linked with vertebrate cardiac development, such as NK homeodomain proteins (e.g., Nkx2.5, *Drosophila tinman* ortholog), GATA (*Drosophila pannier* ortholog) and T-box families.

In contrast to skeletal muscles, a role for Hox in vertebrate cardiac muscle development is very well established (reviewed in Lescroart and Zaffran, 2018). It was recently shown that anterior Hox genes (*Hoxa1*, *Hoxa2*, *Hoxb1*, and *Hoxb2*) are expressed in a subpopulation of Mesp1-expressing cardiovascular progenitors that seem to be the last to emerge from the primitive streak. Interestingly, progenitors that do not express Hox seem to be unipotent in contrast with Hox-expressing progenitors that are bipotent, contributing either to cardiomyocytes and smooth muscles or cardiomyocytes and endothelial cells (Lescroart et al., 2014, 2018). Hox genes (*Hoxa1*, *Hoxa3*, and *Hoxb1*) are also expressed in SHF progenitors and their expression patterns subdivide this cell population in distinct domains: *Hoxa1* and *Hoxb1* are expressed in AHF with different anterior limits and *Hoxa3* is expressed in posterior SHF. While these progenitors contribute to both poles of the heart, *Hoxb1*-expressing progenitors are found only in the proximal OFT and atria and cells expressing *Hoxa1* and *Hoxa3* only in the distal OFT and some regions of the atria. SHF cells are thus pre-patterned before their addition to the developing heart (Bertrand et al., 2011). The same Hox genes are differentially expressed along the rostro-caudal axis in cardiac NCCs (Bertrand et al., 2011; Lescroart and Zaffran, 2018).

Concerning the role of Hox in cardiac development, it has been suggested (although not directly demonstrated) that *Hoxb1* could play a role in cardiac progenitor migration from the primitive streak (Lescroart and Zaffran, 2018). Mice mutant for *Hoxa1* develop heart patterning defects, such as OFT malformations, that have been also observed in human patients, giving a direct role to *Hoxa1* in OFT patterning (Makki and Capecchi, 2012). Additionally, mice lacking *Hoxb1* display OFT and ventricular septum (wall separating the two ventricles) defects. In this case, *Hoxb1* mutation led on the one hand to the upregulation of *fgf8* levels and abnormal SHF proliferation and on the other hand the upregulation of the SHF differentiation markers, α-actinin and MF20 and thus premature SHF differentiation. Furthermore, in compound *Hoxa1*, *Hoxb1* mutant mice, the OFT and ventral septum deficits were exacerbated, suggesting a genetic interaction

between them (Roux et al., 2015). Interestingly, not only *Hoxb1* loss of function but also its overexpression leads to cardiac malformations (Zaffran et al., 2018). Extending this concept further, transcriptional profiling has shown that *Hoxb1* activates the posterior program of the SHF and inhibits the premature differentiation of progenitors by directly repressing *Natriuretic peptide precursor A* (*Nppa*) and *B* (*Nppb*) expression (Stefanovic et al., 2020).

Anterior Hox genes also play a role in cardiac NCC (Chisaka and Capecchi, 1991; Roux et al., 2017). *Hoxa3*-mutant mice display defects in carotid arteries as well as defaults in size and shape of heart compartments (Chisaka and Capecchi, 1991). *Hoxa1* and *Hoxb1* are required for cardiac NCC migration and their mutation leads to subsequent large artery patterning and outflow septum defects (Roux et al., 2017). Interestingly, an *in vitro* study using mouse embryonic stem cells has demonstrated a role for *Hoxa10* in the timing of cardiac cell differentiation, suggesting an unexpected role for posterior Hox in vertebrate cardiac development (Behrens et al., 2013). In contrast to *Drosophila* heart development, Hox PBC/MEIS cofactors were found to be associated with Hox function in vertebrates (Lescroart and Zaffran, 2018).

VISCERAL MUSCLE

Hox Control of *Drosophila* Visceral Musculature

The *Drosophila* gut is formed by the assembly of cells originating from all three germ layers; ectodermal, endodermal and mesodermal cells. The mesodermal visceral muscles are located in the external midgut layer surrounding the endoderm and are responsible for peristaltic movements. Five Hox genes (*Scr*, *Antp*, *Ubx*, *abd-A*, and *AbdB*) are expressed in the midgut visceral mesoderm and all except *Scr* and *AbdB* pattern the unsegmented gut together with ectodermal cues. Their expression in this tissue is parasegmental (*Scr* and *AbdB* excepted), and non-overlapping (a parasegment is a metameric unit composed of a posterior part of one segment and an anterior part of the adjacent segment). *Scr* mostly overlaps with PS4, *Antp* is expressed in PS5 and 6, *Ubx* in PS7, *abd-A* in PS8-12 and *AbdB* in the end of the midgut (**Figure 1D**). Curiously, there is a one-to-two cell gap between *Scr* and *Antp* expression domains. In the gut inner endoderm only *labial* (*lab*) expression can be detected (Diederich et al., 1989). Antp, Ubx, and AbdA are responsible for three midgut constrictions establishment, subdividing the midgut into four distinct chambers and seemingly helping to the proper gut elongation (**Figure 1D**). In homozygous *Antp* mutant embryos, the first constriction doesn't form (Tremml and Bienz, 1989; Reuter and Scott, 1990). The establishment of the second constriction is dependent on both Ubx and AbdA. The third constriction is fully specified by AbdA (Tremml and Bienz, 1989). Scr is not involved in midgut subdivision, but in *Scr* mutant conditions, four protrusions located in the anterior midgut called gastric caeca do not form (Reuter and Scott, 1990). Finally, AbdB is required for the gut left-right asymmetry establishment by controlling the activity of a gene encoding for the type ID unconventional myosin (*myosinID*), a function presumed to be independent of Hox patterning function (Coutelis et al., 2013; **Figure 1D**).

Concerning Hox target genes in the visceral mesoderm, it was first predicted that Hox could regulate cytoskeleton or genes able to drive mesodermal cell contraction around the underlying endoderm, explaining the constriction establishment (Reuter and Scott, 1990). One such gene was identified, *beta3-tubulin*, encoding a cytoskeleton-associated protein whose expression is regulated by Ubx (Hinz et al., 1992). Importantly, Ubx is required for *decapentaplegic* (*dpp*) expression in the visceral mesoderm and together with AbdA controls *wingless* (*wg*) expression (Immerglück et al., 1990). The *cis*-regulatory region in the *dpp* gene regulated directly by Ubx and AbdA has been successfully identified and constituted the first example of a Hox target gene enhancer (Capovilla et al., 1994). Later, the enhancer in *wg* gene bound by AbdA and Mad (Mothers against dpp, transcription factor and Dpp signaling target), driving its expression was also identified (Grienenberger et al., 2003). In this particular case, the sole AbdA binding without Dpp input does not allow *wg* activation, once again demonstrating that a cooperative binding between Hox and transcription factors is required to convey a proper cell fate. All the following visceral mesoderm target genes identified are activated by Wg or Dpp and are thus indirect targets: In the anterior midgut mesoderm, Antp regulates *teashirt* (*tsh*) expression and Ubx, AbdA, Wg, and Dpp regulate its expression in the central part (Mathies et al., 1994). Through Wg and Dpp signaling, Ubx and AbdA activate *pointed* (*pnt*) and *odd-paired* (*opa*) in the specific posterior midgut mesodermal regions (Bilder et al., 1998).

Hox Control of Vertebrate Visceral Musculature

In vertebrates as in *Drosophila*, the gut develops both from endoderm and mesoderm. More precisely, it develops from the splanchnic mesoderm, itself derived from the lateral plate mesoderm. The splanchnic mesoderm and the endoderm involute to form a primitive gut tube. The tube develops further into foregut, midgut comprising the small intestine, cecum and anterior portion of the large intestine and hindgut comprising the remainder of the large intestine and rectum. Hox genes are collinearly expressed along the lateral plate mesodermal component of the gut but also in the endoderm and the expression of many persists in the adult (Beck et al., 2000; Beck, 2002). The detailed, complex Hox expression patterns in different organs of the gut mesoderm has been extensively summarized previously (Beck et al., 2000; Choi et al., 2006).

Few Hox functions have been identified in the mouse gut, notably because of the co-expression in the same gut regions of two or more Hox from different paralogous groups, leading to high functional redundancy. Despite this, it has been shown that *Hoxd13* and *Hoxd12* are required for the generation of the anal sphincter (Kondo et al., 1996). When all *Hoxd* genes are deleted (except *Hoxd1* and *Hoxd3*) the ileocecal sphincter (separating small and large intestine) doesn't form and the ileocecal smooth-musculature is disorganized (Zákány and Duboule, 1999). Hox

genes are thus clearly required to pattern the unsubdivided gut mesoderm both in *Drosophila* and vertebrates. Similarly, deletion of the anterior part of the *HoxD* locus (from *Hoxd1* to *Hoxd10*) provokes agenesis of the caecum (at the junction of the ileum and large intestine). Interestingly, it has been demonstrated that it is not the combined deletion of these genes but instead the resulting strong ectopic *Hoxd12* expression that accounts for this phenotype (Zacchetti et al., 2007). Trying to find a common framework for the role of the different *HoxD* genes, it has been recently reported that *Hoxd3* deletion alone results in gut growth deficit, giving an essential role in the gut development to this Hox gene (Zakany et al., 2017).

A role has also been attributed to Hox genes from other paralogous groups, such as *Hoxc4*, whose deletion results in esophageal musculature disorganization and obstruction of the organ (Boulet and Capecchi, 1996). Mice carrying a *Hoxa4* transgene, resulting in its strong overexpression in the gut, develop a congenital megacolon phenotype characterized by a largely distended colon (Wolgemuth et al., 1989). Deletion of *Hoxa5* leads to stomach morphogenesis defaults, presumably by controlling the epithelio-mesenchymal signaling. Indeed, in the *Hoxa5*-deficient mice stomach epithelium, the expression pattern of *Ihh*, *Shh*, and *Fgf10* changed and the expression levels of *Ptc* and *Gli* increased (Aubin et al., 2002). Misexpression of *Hoxc8* under the control of *Hoxa4* regulatory elements, resulting in a shift in the anterior boundary of its expression, results in several hamartomatous lesions, where gastric epithelium was found embedded within the stomach musculature (Pollock et al., 1992).

CONCLUDING REMARKS

Hox genes are involved in the patterning of all muscle types in *Drosophila* and vertebrates. While how they achieve this is not completely resolved, plethora studies in different model organisms have gone a long way to determine their role in the numerous muscle types. One emerging concept would appear that Hox act at numerous stages of muscle development, where at early stages, Hox appear to play a specifying role by providing spatial cues along the anterior–posterior axis, and at later stages controlling basic cellular functions such as proliferation, cell survival, death etc. This follows the role Hox play in patterning of the ectoderm and so it is perhaps without surprise that Hox contribute similarly in muscle development through mesoderm patterning.

As in patterning of the ectoderm, in muscle development there is often requirement for the cooperation between Hox and other tissue-specific transcription factors, which are themselves Hox target genes together with signaling molecules. Examples about Hox functional conservation can be found mainly in vertebrate cardiac muscle, but conservation also exists in skeletal and gut muscles, further suggesting a universal role for Hox in mesodermal patterning. While how this achieved is not fully understood, knowledge gained in how Hox generate diversity in the CNS should help understand their role in generating different muscle types, which together allow for the development of more complex organisms.

Even though a large number of Hox-dependent functions across different mesodermal derivatives are now known, there is however only sparse evidence about the underlying molecular mechanisms. To fully understand how Hox orchestrate muscle development, it is essential to define the network of genes they regulate, in addition to the tissue-specific transcription factors such as the identity TFs in muscle and temporal TFs in the CNS. The discovery of novel *cis*-regulatory regions of Hox target genes was historically a difficult and laborious process, but now with the vast advances in genome-wide approaches, both spatially and temporally, at the level of whole tissue or single cell, it is soon possible to better define Hox regulatory regions and target genes. Such a genome-wide spatio-temporal approach will thus allow us to fully grasp the complex and intricate networks defining how Hox proteins regulate muscle development.

AUTHOR CONTRIBUTIONS

GP designed and drew illustrations, and wrote the manuscript with inputs from CM-Z, YG, and AS. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

The authors would like to thank the Fondation ARC (grant PJA 20181208014) and AFM-Téléthon (grant 23721) for financial support of their research.

REFERENCES

Alvares, L. E., Schubert, F. R., Thorpe, C., Mootoosamy, R. C., Cheng, L., Parkyn, G., et al. (2003). Intrinsic, Hox-dependent cues determine the fate of skeletal muscle precursors. *Dev. Cell* 5, 379–390. doi: 10.1016/S1534-5807(03)00263-6

Aubin, J., Déry, U., Lemieux, M., Chailler, P., and Jeannotte, L. (2002). Stomach regional specification requires Hoxa5-driven mesenchymal-epithelial signaling. *Development* 129, 4075–4087. doi: 10.1242/dev.129.17.4075

Baek, M. (2011). *Development of Leg Motor Neurons in Drosophila Melanogaster*. Ph.D. thesis. New York, NY: Columbia University. doi: 10.7916/D84J0N30

Baek, M., Enriquez, J., and Mann, R. S. (2013). Dual role for Hox genes and Hox co-factors in conferring leg motoneuron survival and identity in *Drosophila*. *Development* 140, 2027–2038. doi: 10.1242/dev.090902

Baek, M., Menon, V., Jessell, T. M., Hantman, A. W., and Dasen, J. S. (2019). Molecular logic of spinocerebellar tract neuron diversity and connectivity. *Cell Rep.* 27, 2620–2635.e4. doi: 10.1016/j.celrep.2019.04.113

Bataillé, L., Colombié, N., Pelletier, A., Paululat, A., Lebreton, G., Carrier, Y., et al. (2020). Alary muscles and thoracic alary-related muscles are atypical striated muscles involved in maintaining the position of internal organs. *Development* 147:dev185645. doi: 10.1242/dev.185645

Bataillé, L., Frendo, J.-L., and Vincent, A. (2015). Hox control of *Drosophila* larval anatomy; The Alary and Thoracic Alary-Related Muscles. *Mech. Dev.* 138(Pt 2), 170–176. doi: 10.1016/j.mod.2015.07.005

Baylies, M. K., and Bate, M. (1996). twist: a myogenic switch in *Drosophila*. *Science* 272, 1481–1484. doi: 10.1126/science.272.5267.1481

Beck, F. (2002). Homeobox genes in gut development. *Gut* 51, 450–454. doi: 10.1136/gut.51.3.450

Beck, F., Tata, F., and Chawengsaksophak, K. (2000). Homeobox genes and gut development. *BioEssays* 22, 431–441. doi: 10.1002/(SICI)1521-1878(200005)22:5<431::AID-BIES5>3.0.CO;2-X

Behrens, A. N., Iacovino, M., Lohr, J. L., Ren, Y., Zierold, C., Harvey, R. P., et al. (2013). Nkx2-5 mediates differential cardiac differentiation through interaction with Hoxa10. *Stem Cells Dev.* 22, 2211–2220. doi: 10.1089/scd.2012.0611

Bertrand, N., Roux, M., Ryckebüsch, L., Niederreither, K., Dollé, P., Moon, A., et al. (2011). Hox genes define distinct progenitor sub-domains within the second heart field. *Dev. Biol.* 353, 266–274. doi: 10.1016/j.ydbio.2011.02.029

Bilder, D., Graba, Y., and Scott, M. P. (1998). Wnt and TGFbeta signals subdivide the AbdA Hox domain during Drosophila mesoderm patterning. *Development* 125, 1781–1790. doi: 10.1242/dev.125.9.1781

Biressi, S., Tagliafico, E., Lamorte, G., Monteverde, S., Tenedini, E., Roncaglia, E., et al. (2007). Intrinsic phenotypic diversity of embryonic and fetal myoblasts is revealed by genome-wide gene expression analysis on purified cells. *Dev. Biol.* 304, 633–651. doi: 10.1016/j.ydbio.2007.01.016

Birkholz, O., Vef, O., Rogulja-Ortmann, A., Berger, C., and Technau, G. M. (2013). Abdominal-B and caudal inhibit the formation of specific neuroblasts in the *Drosophila* tail region. *Development* 140, 3552–3564. doi: 10.1242/dev.096099

Boukhatmi, H., Schaub, C., Bataillé, L., Reim, I., Frendo, J.-L., Frasch, M., et al. (2014). An Org-1–Tup transcriptional cascade reveals different types of alary muscles connecting internal organs in *Drosophila*. *Development* 141, 3761–3771. doi: 10.1242/dev.111005

Boulet, A. M., and Capecchi, M. R. (1996). Targeted disruption of hoxc-4 causes esophageal defects and vertebral transformations. *Dev. Biol.* 177, 232–249. doi: 10.1006/dbio.1996.0159

Bryantsev, A. L., Duong, S., Brunetti, T. M., Chechenova, M. B., Lovato, T. L., Nelson, C., et al. (2012). Extradenticle and homothorax control adult muscle fiber identity in *Drosophila*. *Dev. Cell* 23, 664–673. doi: 10.1016/j.devcel.2012.08.004

Capovilla, M., Brandt, M., and Botas, J. (1994). Direct regulation of decapentaplegic by *Ultrabithorax* and its role in Drosophila midgut morphogenesis. *Cell* 76, 461–475. doi: 10.1016/0092-8674(94)90111-2

Capovilla, M., Kambris, Z., and Botas, J. (2001). Direct regulation of the muscle-identity gene apterous by a Hox protein in the somatic mesoderm. *Development* 128, 1221–1230. doi: 10.1242/dev.128.8.1221

Carmena, A., Buff, E., Halfon, M. S., Gisselbrecht, S., Jiménez, F., Baylies, M. K., et al. (2002). Reciprocal regulatory interactions between the Notch and Ras signaling pathways in the *Drosophila* embryonic mesoderm. *Dev. Biol.* 244, 226–242. doi: 10.1006/dbio.2002.0606

Carmena, A., Gisselbrecht, S., Harrison, J., Jiménez, F., and Michelson, A. M. (1998). Combinatorial signaling codes for the progressive determination of cell fates in the *Drosophila* embryonic mesoderm. *Genes Dev.* 12, 3910–3922. doi: 10.1101/gad.12.24.3910

Catela, C., Shin, M. M., Lee, D. H., Liu, J.-P., and Dasen, J. S. (2016). Hox proteins coordinate motor neuron differentiation and connectivity programs through Ret/Gfrα genes. *Cell Rep.* 14, 1901–1915. doi: 10.1016/j.celrep.2016.01.067

Chal, J., and Pourquié, O. (2017). Making muscle: skeletal myogenesis in vivo and in vitro. *Development* 144, 2104–2122. doi: 10.1242/dev.151035

Chaturvedi, D., Reichert, H., Gunage, R. D., and VijayRaghavan, K. (2017). Identification and functional characterization of muscle satellite cells in *Drosophila*. *ELife* 6:e30107. doi: 10.7554/eLife.30107.021

Chevallier, A., Kieny, M., and Mauger, A. (1977). Limb-somite relationship: origin of the limb musculature. *Development* 41, 245–258. doi: 10.1242/dev.41.1.245

Chisaka, O., and Capecchi, M. R. (1991). Regionally restricted developmental defects resulting from targeted disruption of the mouse homeobox gene hox-1.5. *Nature* 350, 473–479. doi: 10.1038/350473a0

Choi, M. Y., Romer, A. I., Hu, M., Lepourcelet, M., Mechoor, A., Yesilaltay, A., et al. (2006). A dynamic expression survey identifies transcription factors relevant in mouse digestive tract development. *Development* 133, 4119–4129. doi: 10.1242/dev.02537

Christ, B., Jacob, H. J., and Jacob, M. (1977). Experimental analysis of the origin of the wing musculature in avian embryos. *Anat. Embryol.* 150, 171–186. doi: 10.1007/BF00316649

Coutelis, J.-B., Géminard, C., Spéder, P., Suzanne, M., Petzoldt, A. G., and Noselli, S. (2013). Drosophila left/right asymmetry establishment is controlled by the Hox gene abdominal-B. *Dev. Cell* 24, 89–97. doi: 10.1016/j.devcel.2012.11.013

Dasen, J. S., De Camilli, A., Wang, B., Tucker, P. W., and Jessell, T. M. (2008). Hox repertoires for motor neuron diversity and connectivity gated by a single accessory factor, *FoxP1*. *Cell* 134, 304–316. doi: 10.1016/j.cell.2008.06.019

Dasen, J. S., Liu, J.-P., and Jessell, T. M. (2003). Motor neuron columnar fate imposed by sequential phases of Hox-c activity. *Nature* 425, 926–933. doi: 10.1038/nature02051

Dasen, J. S., Tice, B. C., Brenner-Morton, S., and Jessell, T. M. (2005). A Hox regulatory network establishes motor neuron pool identity and target-muscle connectivity. *Cell* 123, 477–491. doi: 10.1016/j.cell.2005.09.009

Diederich, R. J., Merrill, V. K., Pultz, M. A., and Kaufman, T. C. (1989). Isolation, structure, and expression of labial, a homeotic gene of the Antennapedia Complex involved in *Drosophila* head development. *Genes Dev.* 3, 399–414. doi: 10.1101/gad.3.3.399

Dixit, R., Vijayraghavan, K., and Bate, M. (2008). Hox genes and the regulation of movement in *Drosophila*. *Dev. Neurobiol.* 68, 309–316. doi: 10.1002/dneu.20589

Dobi, K. C., Schulman, V. K., and Baylies, M. K. (2015). Specification of the somatic musculature in *Drosophila*: Specification of the somatic musculature in *Drosophila*. *Wiley Interdiscip. Rev. Dev. Biol.* 4, 357–375. doi: 10.1002/wdev.182

Domsch, K., Schröder, J., Janeschik, M., Schaub, C., and Lohmann, I. (2021). The Hox transcription factor Ubx ensures somatic myogenesis by suppressing the mesodermal master regulator twist. *Cell Rep.* 34:108577. doi: 10.1016/j.celrep.2020.108577

Dutta, D., Umashankar, M., LewisEdward, B., Rodrigues, V., and VijayRaghavan, K. (2010). Hox genes regulate muscle founder cell pattern autonomously and regulate morphogenesis through motor neurons. *J. Neurogenet.* 24, 95–108. doi: 10.3109/01677063.2010.494317

Egger, M. D., Harris, S., Peng, B., Schneiderman, A. M., and Wyman, R. J. (1990). Morphometric analysis of thoracic muscles in wildtype and in bithorax *Drosophila*. *Anat. Rec.* 226, 373–382. doi: 10.1002/ar.1092260315

Endo, T. (2015). Molecular mechanisms of skeletal muscle development, regeneration, and osteogenic conversion. *Bone* 80, 2–13. doi: 10.1016/j.bone.2015.02.028

Enriquez, J., and Vincent, A. (2010). Segmental variations in the patterns of somatic muscles: what roles for Hox? *Fly* 4, 249–252. doi: 10.4161/fly.4.3.12281

Enriquez, J., Boukhatmi, H., Dubois, L., Philippakis, A. A., Bulyk, M. L., Michelson, A. M., et al. (2010). Multi-step control of muscle diversity by Hox proteins in the *Drosophila* embryo. *Development* 137, 457–466. doi: 10.1242/dev.045286

Enriquez, J., Venkatasubramanian, L., Baek, M., Peterson, M., Aghayeva, U., and Mann, R. S. (2015). Specification of individual adult motor neuron morphologies by combinatorial transcription factor codes. *Neuron* 86, 955–970. doi: 10.1016/j.neuron.2015.04.011

Estacio-Gómez, A., and Díaz-Benjumea, F. J. (2014). Roles of Hox genes in the patterning of the central nervous system of *Drosophila*. *Fly* 8, 26–32. doi: 10.4161/fly.27424

Evano, B., Gill, D., Hernando-Herraez, I., Comai, G., Stubbs, T. M., Commere, P.-H., et al. (2020). Transcriptome and epigenome diversity and plasticity of muscle stem cells following transplantation. *PLoS Genet.* 16:e1009022. doi: 10.1371/journal.pgen.1009022

Fernandes, J., Bate, M., and Vijayraghavan, K. (1991). Development of the indirect flight muscles of *Drosophila*. *Development* 113, 67–77. doi: 10.1242/dev.113.1.67

Fernandes, J., Celniker, S. E., Lewis, E. B., and VijayRaghavan, K. (1994). Muscle development in the four-winged *Drosophila* and the role of the *Ultrabithorax* gene. *Curr. Biol.* 4, 957–964. doi: 10.1016/S0960-9822(00)00219-0

Friedrich, J., Sorge, S., Bujupi, F., Eichenlaub, M. P., Schulz, N. G., Wittbrodt, J., et al. (2016). Hox function is required for the development and maintenance of the *Drosophila* feeding motor unit. *Cell Rep.* 14, 850–860. doi: 10.1016/j.celrep.2015.12.077

Graba, Y., Gieseler, K., Aragnol, D., Laurenti, P., Mariol, M. C., Berenger, H., et al. (1995). DWnt-4, a novel *Drosophila* Wnt gene acts downstream of homeotic complex genes in the visceral mesoderm. *Development* 121, 209–218. doi: 10.1242/dev.121.1.209

Greig, S., and Akam, M. (1993). Homeotic genes autonomously specify one aspect of pattern in the *Drosophilla* mesoderm. *Nature* 362, 630–632. doi: 10.1038/362630a0

Grienenberger, A., Merabet, S., Manak, J., Iltis, I., Fabre, A., Bérenger, H., et al. (2003). Tgfβ signaling acts on a Hox response element to confer specificity and diversity to Hox protein function. *Development* 130, 5445–5455. doi: 10.1242/dev.00760

Gunage, R. D., Dhanyasi, N., Reichert, H., and VijayRaghavan, K. (2017). Drosophila adult muscle development and regeneration. *Semin. Cell Dev. Biol.* 72, 56–66. doi: 10.1016/j.semcdb.2017.11.017

Hessinger, C., Technau, G. M., and Rogulja-Ortmann, A. (2017). The *Drosophila* Hox gene *Ultrabithorax* acts in both muscles and motoneurons to orchestrate formation of specific neuromuscular connections. *Development* 144, 139–150. doi: 10.1242/dev.143875

Hinz, U., Wolk, A., and Renkawitz-Pohl, R. (1992). Ultrabithorax is a regulator of beta 3 tubulin expression in the Drosophila visceral mesoderm. *Development* 116, 543–554. doi: 10.1242/dev.116.3.543

Hombría, J. C.-G., and Lovegrove, B. (2003). Beyond homeosis—HOX function in morphogenesis and organogenesis. *Differentiation* 71, 461–476. doi: 10.1046/j.1432-0436.2003.7108004.x

Hooper, J. E. (1986). Homeotic gene function in the muscles of *Drosophila* larvae. *EMBO J.* 5, 2321–2329. doi: 10.1002/j.1460-2075.1986.tb04500.x

Immerglück, K., Lawrence, P. A., and Bienz, M. (1990). Induction across germ layers in *Drosophila* mediated by a genetic cascade. *Cell* 62, 261–268. doi: 10.1016/0092-8674(90)90364-K

Jorgensen, E. M., and Garber, R. L. (1987). Function and misfunction of the two promoters of the *Drosophila Antennapedia* gene. *Genes Dev.* 1, 544–555. doi: 10.1101/gad.1.6.544

Jung, H., and Dasen, J. S. (2015). Evolution of patterning systems and circuit elements for locomotion. *Dev. Cell* 32, 408–422. doi: 10.1016/j.devcel.2015.01.008

Jung, H., Lacombe, J., Mazzoni, E. O., Liem, K. F., Grinstein, J., Mahony, S., et al. (2010). Global control of motor neuron topography mediated by the repressive actions of a single hox gene. *Neuron* 67, 781–796. doi: 10.1016/j.neuron.2010.08.008

Karlsson, D., Baumgardt, M., and Thor, S. (2010). Segment-specific neuronal subtype specification by the integration of anteroposterior and temporal cues. *PLoS Biol.* 8:e1000368. doi: 10.1371/journal.pbio.1000368

Kondo, T., Dolle, P., Zakany, J., and Duboule, D. (1996). Function of posterior HoxD genes in the morphogenesis of the anal sphincter. *Development* 122, 2651–2659. doi: 10.1242/dev.122.9.2651

LaBeau, E. M., Trujillo, D. L., and Cripps, R. M. (2009). Bithorax complex genes control alary muscle patterning along the cardiac tube of *Drosophila*. *Mech. Dev.* 126, 478–486. doi: 10.1016/j.mod.2009.01.001

Lacombe, J., Hanley, O., Jung, H., Philippidou, P., Surmeli, G., Grinstein, J., et al. (2013). Genetic and functional modularity of Hox activities in the specification of limb-innervating motor neurons. *PLoS Genet.* 9:e1003184. doi: 10.1371/journal.pgen.1003184

Lawrence, P. A., and Brower, D. L. (1982). Myoblasts from *Drosophila* wing disks can contribute to developing muscles throughout the fly. *Nature* 295, 55–57. doi: 10.1038/295055a0

Lawrence, P. A., and Johnston, P. (1984). The genetic specification of pattern in a *Drosophila* muscle. *Cell* 36, 775–782. doi: 10.1016/0092-8674(84)90357-X

Lehmacher, C., Abeln, B., and Paululat, A. (2012). The ultrastructure of *Drosophila* heart cells. *Arthropod Struct. Dev.* 41, 459–474. doi: 10.1016/j.asd.2012.02.002

Lescroart, F., and Zaffran, S. (2018). Hox and Tale transcription factors in heart development and disease. *Int. J. Dev. Biol.* 62, 837–846. doi: 10.1387/ijdb.180192sz

Lescroart, F., Chabab, S., Lin, X., Rulands, S., Paulissen, C., Rodolosse, A., et al. (2014). Early lineage restriction in temporally distinct populations of Mesp1 progenitors during mammalian heart development. *Nat. Cell Biol.* 16, 829–840. doi: 10.1038/ncb3024

Lescroart, F., Wang, X., Lin, X., Swedlund, B., Gargouri, S., Sànchez-Dànes, A., et al. (2018). Defining the earliest step of cardiovascular lineage segregation by single-cell RNA-seq. *Science* 359, 1177–1181. doi: 10.1126/science.aao4174

Li, L., Xiong, W.-C., and Mei, L. (2018). Neuromuscular junction formation, aging, and disorders. *Annu. Rev. Physiol.* 80, 159–188. doi: 10.1146/annurev-physiol-022516-034255

Lo, P. C. H., Skeath, J. B., Gajewski, K., Schulz, R. A., and Frasch, M. (2002). Homeotic genes autonomously specify the anteroposterior subdivision of the *Drosophila* dorsal vessel into aorta and heart. *Dev. Biol.* 251, 307–319. doi: 10.1006/dbio.2002.0839

Lovato, T. L., Nguyen, T. P., Molina, M. R., and Cripps, R. M. (2002). The Hox gene abdominal-a specifies heart cell fate in the *Drosophila* dorsal vessel. *Development* 129, 5019–5027. doi: 10.1242/dev.129.21.5019

Machado, L., Esteves de Lima, J., Fabre, O., Proux, C., Legendre, R., Szegedi, A., et al. (2017). In situ fixation redefines quiescence and early activation of skeletal muscle stem cells. *Cell Rep.* 21, 1982–1993. doi: 10.1016/j.celrep.2017.10.080

Makki, N., and Capecchi, M. R. (2012). Cardiovascular defects in a mouse model of HOXA1 syndrome. *Hum. Mol. Genet.* 21, 26–31. doi: 10.1093/hmg/ddr434

Mallo, M., and Alonso, C. R. (2013). The regulation of Hox gene expression during animal development. *Development* 140, 3951–3963. doi: 10.1242/dev.068346

Mathies, L. D., Kerridge, S., and Scott, M. P. (1994). Role of the teashirt gene in *Drosophila* midgut morphogenesis: secreted proteins mediate the action of homeotic genes. *Development* 120, 2799–2809. doi: 10.1242/dev.120.10.2799

Mendelsohn, A. I., Dasen, J. S., and Jessell, T. M. (2017). Divergent Hox coding and evasion of retinoid signaling specifies motor neurons innervating digit muscles. *Neuron* 93, 792–805.e4. doi: 10.1016/j.neuron.2017.01.017

Meng, J. L., and Heckscher, E. S. (2021). Development of motor circuits: from neuronal stem cells and neuronal diversity to motor circuit assembly. *Curr. Top. Dev. Biol.* 142, 409–442. doi: 10.1016/bs.ctdb.2020.11.010

Michelson, A. M. (1994). Muscle pattern diversification in *Drosophila* is determined by the autonomous function of homeotic genes in the embryonic mesoderm. *Development* 120, 755–768. doi: 10.1242/dev.120.4.755

Molina, M. R., and Cripps, R. M. (2001). Ostia, the inflow tracts of the *Drosophila* heart, develop from a genetically distinct subset of cardial cells. *Mech. Dev.* 109, 51–59. doi: 10.1016/S0925-4773(01)00509-3

Monier, B., Astier, M., Sémériva, M., and Perrin, L. (2005). Steroid-dependent modification of Hox function drives myocyte reprogramming in the *Drosophila* heart. *Development* 132, 5283–5293. doi: 10.1242/dev.02091

Monier, B., Tevy, F., Perrin, L., Capovilla, M., and Semeriva, M. (2007). Downstream of homeotic genes: in the heart of Hox function. *Fly* 1, 59–67. doi: 10.4161/fly.3993

Musumeci, G., Castrogiovanni, P., Coleman, R., Szychlinska, M. A., Salvatorelli, L., Parenti, R., et al. (2015). Somitogenesis: from somite to skeletal muscle. *Acta Histochem.* 117, 313–328. doi: 10.1016/j.acthis.2015.02.011

Nowicki, J. L., and Burke, A. C. (2000). Hox genes and morphological identity: axial versus lateral patterning in the vertebrate mesoderm. *Development* 127, 4265–4275. doi: 10.1242/dev.127.19.4265

Ono, Y., Boldrin, L., Knopp, P., Morgan, J. E., and Zammit, P. S. (2010). Muscle satellite cells are a functionally heterogeneous population in both somite-derived and branchiomeric muscles. *Dev. Biol.* 337, 29–41. doi: 10.1016/j.ydbio.2009.10.005

Paul, R., Giraud, G., Domsch, K., Duffraisse, M., Marmigère, F., Khan, S., et al. (2021). Hox dosage contributes to flight appendage morphology in *Drosophila*. *Nat. Commun.* 12:2892. doi: 10.1038/s41467-021-23293-8

Pearson, J. C., Lemons, D., and McGinnis, W. (2005). Modulating Hox gene functions during animal body patterning. *Nat. Rev. Genet.* 6, 893–904. doi: 10.1038/nrg1726

Perrin, L., Monier, B., Ponzielli, R., Astier, M., and Semeriva, M. (2004). Drosophila cardiac tube organogenesis requires multiple phases of Hox activity. *Dev. Biol.* 272, 419–431. doi: 10.1016/j.ydbio.2004.04.036

Philippidou, P., and Dasen, J. S. (2013). Hox genes: choreographers in neural development, architects of circuit organization. *Neuron* 80, 12–34. doi: 10.1016/j.neuron.2013.09.020

Philippidou, P., Walsh, C. M., Aubin, J., Jeannotte, L., and Dasen, J. S. (2012). Sustained Hox5 gene activity is required for respiratory motor neuron development. *Nat. Neurosci.* 15, 1636–1644. doi: 10.1038/nn.3242

Pineault, K. M., and Wellik, D. M. (2014). Hox genes and limb musculoskeletal development. *Curr. Osteoporos. Rep.* 12, 420–427. doi: 10.1007/s11914-014-0241-0

Pollock, R. A., Jay, G., and Bieberich, C. J. (1992). Altering the boundaries of Hox3.1 expression: evidence for antipodal gene regulation. *Cell* 71, 911–923.

Ponzielli, R., Astier, M., Chartier, A., Gallet, A., Thérond, P., and Sémériva, M. (2002). Heart tube patterning in *Drosophila* requires integration of axial and segmental information provided by the bithorax complex genes and hedgehog signaling. *Development* 129, 4509–4521.

Poovathumkadavil, P., and Jagla, K. (2020). Genetic control of muscle diversification and homeostasis: insights from *Drosophila*. *Cells* 9:1543.

Prokop, A., and Technau, G. (1994). Early tagma-specific commitment of *Drosophila* CNS progenitor NB1-1. *Development* 120, 2567–2578.

Prokop, A., Bray, S., Harrison, E., and Technau, G. M. (1998). Homeotic regulation of segment-specific differences in neuroblast numbers and proliferation in the *Drosophila* central nervous system. *Mech. Dev.* 74, 99–110.

Reuter, R., and Scott, M. P. (1990). Expression and function of the homoeotic genes Antennapedia and sex combs reduced in the embryonic midgut of *Drosophila*. *Development* 109, 289–303.

Rezsohazy, R., Saurin, A. J., Maurel-Zaffran, C., and Graba, Y. (2015). Cellular and molecular insights into Hox protein action. *Development* 142, 1212–1227.

Rivlin, P. K., Gong, A., Schneiderman, A. M., and Booker, R. (2001). The role of Ultrabithorax in the patterning of adult thoracic muscles in *Drosophila* melanogaster. *Dev. Genes Evol.* 211, 55–66. doi: 10.1007/s004270000126

Rogulja-Ortmann, A., and Technau, G. M. (2008). Multiple roles for Hox genes in segment-specific shaping of CNS lineages. *Fly* 2, 316–319. doi: 10.4161/fly.7464

Rogulja-Ortmann, A., Renner, S., and Technau, G. M. (2008). Antagonistic roles for Ultrabithorax and Antennapedia in regulating segment-specific apoptosis of differentiated motoneurons in the *Drosophila* embryonic central nervous system. *Development* 135, 3435–3445. doi: 10.1242/dev.023986

Rotstein, B., and Paululat, A. (2016). On the morphology of the *Drosophila* Heart. *J. Cardiovasc. Dev. Dis.* 3:15. doi: 10.3390/jcdd3020015

Roux, M., Laforest, B., Capecchi, M., Bertrand, N., and Zaffran, S. (2015). Hoxb1 regulates proliferation and differentiation of second heart field progenitors in pharyngeal mesoderm and genetically interacts with Hoxa1 during cardiac outflow tract development. *Dev. Biol.* 406, 247–258. doi: 10.1016/j.ydbio.2015.08.015

Roux, M., Laforest, B., Eudes, N., Bertrand, N., Stefanovic, S., and Zaffran, S. (2017). Hoxa1 and Hoxb1 are required for pharyngeal arch artery development. *Mech. Dev.* 143, 1–8. doi: 10.1016/j.mod.2016.11.006

Roy, S., and Raghavan, K. V. (1997). Homeotic genes and the regulation of myoblast migration, fusion, and fibre-specific gene expression during adult myogenesis in *Drosophila*. *Development* 124, 3333–3341. doi: 10.1242/dev.124.17.3333

Roy, S., Shashidhara, L. S., and VijayRaghavan, K. (1997). Muscles in the *Drosophila* second thoracic segment are patterned independently of autonomous homeotic gene function. *Curr. Biol.* 7, 222–227. doi: 10.1016/S0960-9822(06)00117-5

Ruiz Gómez, M., and Bate, M. (1997). Segregation of myogenic lineages in *Drosophila* requires numb. *Development* 124, 4857–4866. doi: 10.1242/dev.124.23.4857

Ryan, K., Dk, H., and Rm, C. (2005). Homeotic selector genes control the patterning of seven-up expressing cells in the *Drosophila* dorsal vessel. *Mech. Dev.* 122, 1023–1033. doi: 10.1016/j.mod.2005.04.007

Sánchez-Herrero, E. (2013). Hox targets and cellular functions. *Scientifica* 2013:738257. doi: 10.1155/2013/738257

Saurin, A. J., Delfini, M. C., Maurel-Zaffran, C., and Graba, Y. (2018). The generic facet of Hox protein function. *Trends Genet.* 34, 941–953. doi: 10.1016/j.tig.2018.08.006

Schaub, C., März, J., Reim, I., and Frasch, M. (2015). Org-1-dependent lineage reprogramming generates the ventral longitudinal musculature of the *Drosophila* heart. *Curr. Biol.* 25, 488–494. doi: 10.1016/j.cub.2014.12.029

Schnorrer, F., and Dickson, B. J. (2004). Muscle building: mechanisms of myotube guidance and attachment site selection. *Dev. Cell* 7, 9–20. doi: 10.1016/j.devcel.2004.06.010

Schulman, V. K., Dobi, K. C., and Baylies, M. K. (2015). Morphogenesis of the somatic musculature in *Drosophila melanogaster*: morphogenesis of the *Drosophila* musculature. *Wiley Interdiscip. Rev. Dev. Biol.* 4, 313–334. doi: 10.1002/wdev.180

Schwörer, S., Becker, F., Feller, C., Baig, A. H., Köber, U., Henze, H., et al. (2016). Epigenetic stress responses induce muscle stem-cell ageing by Hoxa9 developmental signals. *Nature* 540, 428–432. doi: 10.1038/nature20603

Stefanovic, S., Laforest, B., Desvignes, J.-P., Lescroart, F., Argiro, L., Maurel-Zaffran, C., et al. (2020). Hox-dependent coordination of mouse cardiac progenitor cell patterning and differentiation. *ELife* 9:e55124. doi: 10.7554/eLife.55124.sa2

Swinehart, I. T., Schlientz, A. J., Quintanilla, C. A., Mortlock, D. P., and Wellik, D. M. (2013). Hox11 genes are required for regional patterning and integration of muscle, tendon and bone. *Development* 140, 4574–4582. doi: 10.1242/dev.096693

Taylor, M. V. (2006). "Comparison of muscle development in *Drosophila* and vertebrates," in *Muscle Development in Drosophila*, ed. H. Sink (New York, NY: Springer New York), 169–203. doi: 10.1007/0-387-32963-3_14

Tremml, G., and Bienz, M. (1989). Homeotic gene expression in the visceral mesoderm of *Drosophila* embryos. *EMBO J.* 8, 2677–2685. doi: 10.1002/j.1460-2075.1989.tb08408.x

Tripodi, M., and Arber, S. (2012). Regulation of motor circuit assembly by spatial and temporal mechanisms. *Curr. Opin. Neurobiol.* 22, 615–623. doi: 0.1016/j.conb.2012.02.011

Wolgemuth, D. J., Behringer, R. R., Mostoller, M. P., Brinster, R. L., and Palmiter, R. D. (1989). Transgenic mice overexpressing the mouse homoeobox-containing gene Hox-1.4 exhibit abnormal gut development. *Nature* 337, 464–467. doi: 10.1038/337464a0

Yoshioka, K., Nagahisa, H., Miura, F., Araki, H., Kamei, Y., Kitajima, Y., et al. (2021). Hoxa10 mediates positional memory to govern stem cell function in adult skeletal muscle. *Sci. Adv.* 7:eabd7924. doi: 10.1126/sciadv.abd7924

Zacchetti, G., Duboule, D., and Zakany, J. (2007). Hox gene function in vertebrate gut morphogenesis: the case of the caecum. *Development* 134, 3967–3973. doi: 10.1242/dev.010991

Zaffran, S., and Frasch, M. (2002). Early signals in cardiac development. *Circ. Res.* 91, 457–469. doi: 10.1161/01.RES.0000034152.74523.A8

Zaffran, S., and Kelly, R. G. (2012). New developments in the second heart field. *Differ. Res. Biol. Divers.* 84, 17–24. doi: 10.1016/j.diff.2012.03.003

Zaffran, S., Odelin, G., Stefanovic, S., Lescroart, F., and Etchevers, H. C. (2018). Ectopic expression of Hoxb1 induces cardiac and craniofacial malformations. *Genesis* 56:e23221. doi: 10.1002/dvg.23221

Zákány, J., and Duboule, D. (1999). Hox genes and the making of sphincters. *Nature* 401, 761–762. doi: 10.1038/44511

Zakany, J., Darbellay, F., Mascrez, B., Necsulea, A., and Duboule, D. (2017). Control of growth and gut maturation by HoxD genes and the associated lncRNA Haglr. *Proc. Natl. Acad. Sci. U.S.A.* 114, E9290–E9299. doi: 10.1073/pnas.1712511114

Zmojdzian, M., de Joussineau, S., Da Ponte, J. P., and Jagla, K. (2018). Distinct subsets of Eve-positive pericardial cells stabilise cardiac outflow and contribute to Hox gene-triggered heart morphogenesis in *Drosophila*. *Development* 145:dev158717. doi: 10.1242/dev.158717

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Cell and Developmental Biology

REVIEW
published: 16 November 2021
doi: 10.3389/fcell.2021.787339

Mechanisms Underlying Hox-Mediated Transcriptional Outcomes

Brittany Cain[1] *and Brian Gebelein*[1,2]*

[1]Division of Developmental Biology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, United States, [2]Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, OH, United States

Metazoans differentially express multiple Hox transcription factors to specify diverse cell fates along the developing anterior-posterior axis. Two challenges arise when trying to understand how the Hox transcription factors regulate the required target genes for morphogenesis: First, how does each Hox factor differ from one another to accurately activate and repress target genes required for the formation of distinct segment and regional identities? Second, how can a Hox factor that is broadly expressed in many tissues within a segment impact the development of specific organs by regulating target genes in a cell type-specific manner? In this review, we highlight how recent genomic, interactome, and *cis*-regulatory studies are providing new insights into answering these two questions. Collectively, these studies suggest that Hox factors may differentially modify the chromatin of gene targets as well as utilize numerous interactions with additional co-activators, co-repressors, and sequence-specific transcription factors to achieve accurate segment and cell type-specific transcriptional outcomes.

Keywords: Hox, transcription factor, chromatin accessibility, *cis*-regulatory modules (CRMs), protein-protein interaction

OPEN ACCESS

Edited by:
Edwina McGlinn,
Monash University, Australia

Reviewed by:
Stefan Thor,
The University of Queensland, Australia
Montserrat Corominas,
University of Barcelona, Spain

***Correspondence:**
Brian Gebelein
Brian.Gebelein@cchmc.org

Specialty section:
This article was submitted to Morphogenesis and Patterning, a section of the journal Frontiers in Cell and Developmental Biology

Received: *30 September 2021*
Accepted: *28 October 2021*
Published: *16 November 2021*

Citation:
Cain B and Gebelein B (2021) Mechanisms Underlying Hox-Mediated Transcriptional Outcomes. Front. Cell Dev. Biol. 9:787339. doi: 10.3389/fcell.2021.787339

INTRODUCTION

Hox genes have long fascinated developmental biologists for the essential roles that they play in specifying different segment and regional identities along the developing anterior-posterior (A-P) axis of metazoans. Classic genetic studies first revealed that Hox gene mutations can result in homeotic transformations, and thereby cause one part of the organism to be transformed into the likeness of another region. As an example, *Drosophila* with Hox mutations can have obvious developmental abnormalities that include the misspecification of appendages as evidenced by the transformation of antennae into legs (Kaufman et al., 1980; Abbott and Kaufman, 1986; Schneuwly et al., 1987; Casares and Mann, 1998) or the conversion of the haltere into an extra set of wings (Lewis, 1978; Bender et al., 1983; Carroll et al., 1995). Subsequent studies in other organisms including a variety of vertebrate animals revealed that mutations within the highly conserved Hox gene family can cause a wide variety of homeotic transformations across metazoans as reviewed by Mark et al. (1997) and Quinonez and Innis (2014).

Hox genes were originally discovered in *Drosophila melanogaster*. In total, *Drosophila* has eight Hox genes that are separated into two distinct chromosomal clusters: The *Antennapedia* cluster consists of five Hox genes [*labial* (*lab*), *proboscipedia* (*pb*), *Deformed* (*Dfd*), *Sex combs reduced* (*Scr*), and *Antennapedia* (*Antp*)] that collectively regulate head and anterior thoracic development, whereas the three Hox genes in the *Bithorax* cluster [*Ultrabithorax* (*Ubx*), *abdominal-A* (*abd-A*),

and *Abdominal-B* (*Abd-B*)] specify cell fates within the third thoracic segment and the abdominal segments (Morata et al., 1990; Maeda and Karch, 2009). In general, the order of the Hox genes on the chromosome correspond with the location along the A-P axis that the Hox genes act in the embryo (Lewis, 1978; Mann, 1997; Noordermeer and Duboule, 2013; Luo et al., 2019; Hajirnis and Mishra, 2021). For example, genes at the 3′ end of the Hox gene cluster tend to mediate anterior development whereas the 5′ genes tend to control posterior structures. In contrast to the single set of eight Hox genes in *Drosophila*, vertebrates have undergone genome duplication events such that humans have four distinct Hox clusters (labeled HOXA, HOXB, HOXC, and HOXD, respectively) encoding 39 Hox genes that have been categorized into 13 paralogs (HOX1-13). Importantly, the mammalian Hox genes exhibit the same spatial collinearity along the A-P axis as in *Drosophila* (Duboule and Dolle, 1989; Graham et al., 1989). For example, HOX1 genes on the 3′ end of each cluster regulate anterior structures including the hindbrain (Singh et al., 2020), while HOX13 genes on the 5' end of each cluster control posterior and distal structures including digit development (Desanlis et al., 2020). Based on sequence conservation, the relative positions of each Hox gene within a cluster, and their roles in A-P patterning, the Hox genes have been broadly categorized into anterior (*lab*, *pb*, *Dfd*, and *Scr* in *Drosophila* and *Hox1-5* in vertebrates), central (*Antp*, *Ubx*, and *abd-A* in *Drosophila* and *Hox6-8* in vertebrates), and posterior groups (*Abd-B* in *Drosophila* and *Hox9-13* in vertebrates) (Hueber et al., 2010). It is important to note that not all Hox paralogs remain in each of the duplicated vertebrate Hox clusters. For example, cluster HOXB does not have posterior factors HOXB10-B12, and cluster HOXC lacks paralogs of HOXC1-C3 in humans (Mark et al., 1997). In short, metazoans encode variable numbers of Hox genes that are typically found clustered along the chromosome to specify the different cell fates that form along the A-P axis body plan.

The mysteries underlying how Hox genes control distinct body regions only grew upon the discovery that each encodes a homeodomain transcription factor (TF) capable of binding highly similar AT-rich DNA sequences (McGinnis et al., 1984a; McGinnis et al., 1984b). In fact, Hox genes are members of a much larger homeodomain TF family that consists of over 200 members in mammals, and many of these genes control distinct developmental processes and fates despite encoding TFs that bind highly similar DNA sequences (Berger et al., 2008; Jolma et al., 2013; Bürglin and Affolter, 2016). Taken together, these conflicting genetic and biochemical findings raise a fundamental paradox: How can a family of homeodomain TFs capable of binding highly similar DNA sequences *in vitro*, regulate distinct and diverse cell fates *in vivo*?

During the past two decades, many molecular, genetic, and genomic approaches have begun to reveal that numerous mechanisms likely underlie the ability of Hox TFs to specify different cell fates along the A-P body axis. In total, these studies have made considerable progress in defining mechanisms that enhance Hox DNA target specificity, especially by the formation of larger DNA binding complexes with other TFs. For example, the Extradenticle (Exd, *Drosophila*)/Pre-B cell leukemia homeobox (Pbx, vertebrate) and/or Homothorax (Hth, *Drosophila*)/Myeloid ecotropic viral integration site (Meis, vertebrate) TFs have been shown to form cooperative DNA binding complexes with Hox TFs and thereby enhance Hox DNA binding specificity (Mann and Affolter, 1998; Moens and Selleri, 2006; Merabet and Mann, 2016). Through a combination of structural, biochemical, and genetic studies, the formation of Hox/Exd and Hox/Pbx complexes have uncovered several key concepts that underlie how Hox TFs gain DNA binding specificity including the critical role of not just nucleotide identity but DNA shape (Zeiske et al., 2018), the concept of latent specificity (Slattery et al., 2011), and the importance of low affinity versus high affinity binding sites (Crocker et al., 2015; Zandvakili et al., 2019). These mechanisms, which by and large are used to increase Hox target gene specificity, have been reviewed in several excellent articles (Merabet and Mann, 2016; Kribelbauer et al., 2019; De Kumar and Darland, 2021).

In this review, we focus on how large-scale genomic and interactome data have uncovered numerous potential Hox regulatory elements and protein interactors that present both new opportunities and challenges. Genomic DNA binding studies from tissues and cells have revealed that Hox TFs, like most sequence-specific TFs, bind thousands of potential *cis*-regulatory elements but only a subset of these binding events are likely to be associated with significant changes in the expression of nearby genes (Walter et al., 1994; Farnham, 2009; Biggin, 2011; Choo and Russell, 2011; Walhout, 2011; Fisher et al., 2012). In addition, comparative studies between Hox TFs have revealed differences in their ability to bind inaccessible (i.e., closed chromatin) DNA elements. Such differences in ability to bind DNA wrapped in nucleosomes may indicate that Hox TFs have the potential to elicit pioneer-like activities that promote the opening of closed chromatin, thereby expanding the already large number of possible genomic binding sites. However, since Hox TFs are capable of mediating both transcriptional activation and repression, simply detecting Hox TF binding to an element cannot easily be used to predict transcriptional outcome. Intriguingly, protein-protein interaction assays have uncovered that Hox TFs can interact with many different proteins including other sequence-specific TFs as well as factors involved in mediating gene activation and repression. Integrating these large-scale findings with existing *cis*-regulatory logic studies of confirmed Hox target genes suggests that Hox TFs are likely to require numerous protein-protein interactions with other TFs to gain the required regulatory specificity to ensure accurate gene activation or repression outcomes occur in a reproducible and robust manner.

HOX TRANSCRIPTION FACTOR BINDING AND CHROMATIN ACCESSIBILITY

Hox factors, like all TFs, must bind specific DNA regulatory elements to mediate accurate transcriptional responses. Since all the cells within an organism have the same genomic material,

differences in the chromatin landscape of a cell can play a large role in dictating which DNA elements are available for transcription factor binding. Thus, chromatin accessibility helps to define which genes can be activated during the specification of distinct cell fates along the body plan. Intriguingly, comparative genomic accessibility studies using Formaldehyde-Assisted Isolation of Regulatory Elements sequencing (FAIRE-seq) on *Drosophila* imaginal discs revealed that the wing, haltere, and metathoracic leg imaginal discs have very similar chromatin profiles (McKay and Lieb, 2013). For example, comparison between the wing and haltere imaginal discs showed that except for genomic regions flanking the *Ultrabithorax* (*Ubx*) Hox gene these two tissues have largely identical accessible *cis* regulatory elements (McKay and Lieb, 2013). Similar results were obtained using the Assay for Transposase-Accessible Chromatin sequencing (ATAC-seq) methods with ~98% of the accessible DNA sequences being the same between age-matched wing and haltere discs (Loker et al., 2021).

The above findings suggest that the Ubx Hox factor, which is differentially expressed in the *Drosophila* imaginal discs, directs the formation of different cell and tissue fates by regulating distinct target genes within highly similar chromatin landscapes. But does the expression of this Hox TF alter the chromatin landscape during the process of cell fate specification and morphogenesis? A recent elegant study addressed this question to better define the role of the Ubx TF in regulating haltere development by focusing on the relatively small percentage of loci (~2% of accessible regions) that were differentially accessible in haltere versus wing discs (Loker et al., 2021). Importantly, Loker et al. combined chromatin accessibility data with Ubx Chromatin Immunoprecipitation sequencing (ChIP-seq) assays and transcriptomics (RNA-seq) to show that Ubx genomic binding correlates with the opening and closing of specific loci to mediate distinct transcriptional outputs during *Drosophila* haltere development (Loker et al., 2021). In particular, they found that Ubx could modify the chromatin landscape to both reduce chromatin accessibility to repress gene transcription in the capitulum and proximal hinge of the haltere and increase chromatin accessibility to activate gene transcription in the distal hinge with the aid of the Hth and Exd Hox co-factor proteins (Loker et al., 2021). Since *Ubx* is required for haltere fate and the loss of *Ubx* function results in the transformation of haltere tissue into wing tissue (Lewis, 1978; Bender et al., 1983; Carroll et al., 1995), these data are congruent with the idea that the primary difference between these two serially homologous appendages is the expression of *Ubx* and that once expressed, the Ubx TF directs haltere development by modulating chromatin accessibility and target gene expression within an initial chromatin landscape capable of forming either a wing or a haltere (McKay and Lieb, 2013; Loker et al., 2021). Thus, while many of the accessible genomic regions across imaginal disc tissues are the same, Hox TFs are likely to modify this landscape to activate and/or repress key target genes during cell fate specification and morphogenesis.

The finding that Hox TF binding can increase genomic accessibility raises the possibility that Hox TFs have pioneer-like activities. By definition, pioneer transcription factors can both bind DNA that is wrapped around a nucleosome and promote chromatin remodeling to make DNA elements accessible for other TFs (Iwafuchi-Doi and Zaret, 2014; Zaret, 2020). To assess the ability of Hox TFs to bind inaccessible DNA and promote chromatin opening, recent comparative genomic binding and accessibility studies have been performed for Hox TFs in both a *Drosophila* cell line (Kc167 cells) (Beh et al., 2016; Porcelli et al., 2019) and in a mouse motor neuron progenitor culture system (Bulajić et al., 2020). Intriguingly, these data indicate that some, but not all, Hox factors can readily bind inaccessible chromatin. By intersecting ATAC-seq and ChIP-seq profiles of the eight *Drosophila* Hox TFs, Porcelli et al. showed that this ability to bind inaccessible chromatin is shared by the anterior factors, Lab, Dfd, and Pb, as well as the posterior Hox factor, Abd-B (**Figure 1A**; Porcelli et al., 2019). Further, by comparing ATAC-seq profiles before and after inducing Dfd and Abd-B expression in respective Kc167 cell lines, Porcelli et al. found that Dfd and Abd-B can increase chromatin accessibility of their targets (Porcelli et al., 2019). These findings are consistent with a previous finding that 42% of Abd-B specific peaks were bound outside of the cells DNaseI accessible regions in Kc167 cells (Beh et al., 2016). The enhanced ability of Abd-B to bind inaccessible chromatin was also supported by studies of the mammalian Abd-B orthologs in neural progenitors and undifferentiated motor neurons (Bulajić et al., 2020). In particular, Bulajić et al. found that the HOXC9 and HOXC13 posterior Hox TFs bound significantly more inaccessible genomic regions than the HOXC6 and HOXC8, which are classified as central Hox TFs (**Figure 1B**; Bulajić et al., 2020). Consistent with these findings, the HOXD13 TF also demonstrated pioneer factor-like activity by increasing chromatin accessibility of targets to guide proximal to distal limb development (**Figure 1B**; Desanlis et al., 2020), supporting a mechanism in which select Hox factors can bind inaccessible chromatin and increase chromatin accessibility of its targets (**Figure 1**; **Figure 2A**).

While the above findings are congruent with the idea of the posterior Abd-B-like Hox factors being able to readily bind inaccessible DNA, additional studies revealed that not all posterior Hox TFs may equally share such properties. For example, comparative studies between several posterior HOX TFs in the motor neuron progenitor assay revealed clear differences with HOXC9 and HOXC13 binding many more inaccessible regions than HOXC10, HOXA9, or HOXD9 (**Figure 1B**; Bulajić et al., 2020). Moreover, the ability of the human HOXC13 factor to bind inaccessible DNA was predominantly influenced by the DNA binding domain and C-terminus (Bulajić et al., 2020). Thus, while it has been argued based on structural studies that posterior Hox TFs may have enhanced binding to inaccessible DNA due to high affinity electrostatic interactions between the narrow groove of DNA and the Hox N-terminal arm of the homeodomain (LaRonde-LeBlanc and Wolberger, 2003; Beh et al., 2016), we currently lack a molecular understanding of why only a subset of posterior HOX TFs readily bind inaccessible chromatin.

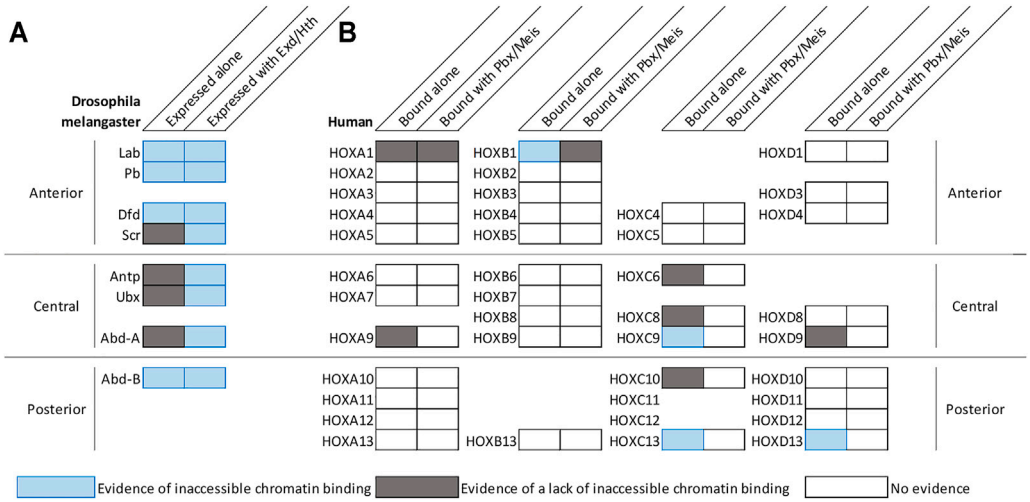


FIGURE 1 | Select Hox factors can bind inaccessible chromatin with and/or without the help of common co-factors. **(A)** Diagram summarizing the genomic DNA binding activities of Hox TFs in *Drosophila* Kc167 cells (data from Beh et al., 2016; Porcelli et al., 2019). By comparing the chromatin accessibility profiles of cells prior to Hox factor transfection and genome binding profiles after Hox factor transient transfection in Kc167 cells, Beh et al. and Porcelli et al. demonstrated that anterior factors and posterior *Drosophila* factors tend to have the ability to bind inaccessible chromatin (Beh et al., 2016; Porcelli et al., 2019). Furthermore, Exd and Hth expression tend to enhance a factor's ability to bind to inaccessible chromatin (Beh et al., 2016; Porcelli et al., 2019). It is important to note that the ability to bind inaccessible chromatin of Abd-B was not enhanced and the ability of Scr was only slightly enhanced by Exd/Hth. **(B)** Diagram summarizing the genomic DNA binding activities of human Hox factors in motor neuron cells (Bulajić et al., 2020), mouse embryonic stem cells (Singh et al., 2021), and mouse limb buds (Desanlis et al., 2020). The genomic binding and accessibility profiles were intersected to assess inaccessible chromatin binding. Nearby PBX and MEIS motifs were used to determine co-binding. *Drosophila* and human Hox factors follow a similar trend that posterior factors can bind inaccessible chromatin more so than central factors.

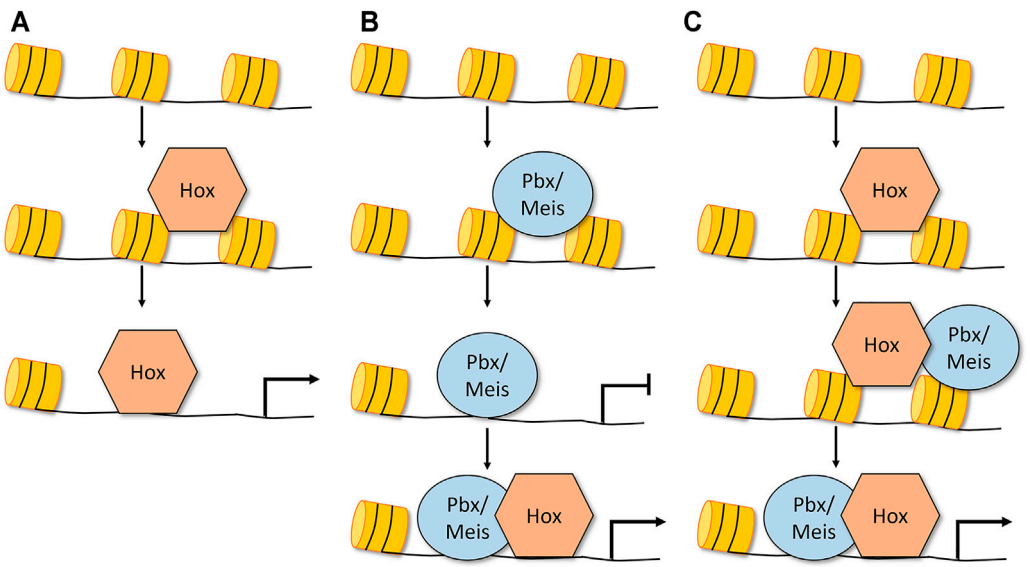


FIGURE 2 | Mechanisms by which Hox factors could exhibit pioneer-like activity and alter chromatin accessibility. **(A)** Dfd and Abd-B Hox factors alone can bind to inaccessible chromatin and increase the chromatin accessibility of its targets without co-factor expression (Porcelli et al., 2019). **(B)** Pbx/Meis in vertebrates have been shown to bind inaccessible chromatin and promote chromatin opening. In this model, the Hox factor gains access to accessible DNA, forms a complex with Pbx/Meis, and this larger TF complex is required for accurate target regulation (Sagerstrom, 2004; Choe et al., 2014; Mariani et al., 2021). **(C)** Another potential model is that Hox TFs, a subset of which are capable of binding inaccessible DNA, recruit Exd/Pbx and Hth/Meis and together these complexes promote chromatin opening (Porcelli et al., 2019). Note that in each of these models, the role of the Hox factor in chromatin opening has not yet been confirmed.

THE IMPACT OF PBX/EXD AND MEIS/HTH ON ALTERING HOX BINDING IN CHROMATIN

Hox factors are well known to interact with the Pbx/Exd and Meis/Hth TFs to enhance DNA binding specificity on naked DNA *in vitro* (Mann and Affolter, 1998; Uhl et al., 2010; Merabet and Mann, 2016). More recently, these factors have also been shown to influence the binding of Hox TFs as well as other TFs to genomic DNA elements embedded in chromatin. The first study describing such an activity for Pbx and Meis was in association with MyoD, a non-Hox basic-Helix-Loop-Helix transcription factor that promotes muscle cell development (Berkes et al., 2004). Pbx was shown to bind the inactive *myogenin* promoter in undifferentiated C2C12 myoblast cells at a time-point that preceded MyoD binding by 6 h (Bergstrom et al., 2002; Berkes et al., 2004). The subsequent binding of MyoD with Pbx during differentiation correlated with *myogenin* promoter activation, consistent with previous studies that found MyoD is able to remodel chromatin and activate target genes (Gerber et al., 1997). Together, these findings suggest that Pbx is binding to and marking the inaccessible chromatin for activation upon recruitment of MyoD, a mechanism that might extend to Pbx's interactions with Hox factors (Sagerstrom, 2004). In fact, Choe et al. found that Pbx/Meis bound numerous loci as early as the zebrafish blastula and that later in development the Hoxb1a TF was required for these loci to become fully active (Choe et al., 2014). More recently, Mariani et al. used a combination of DNA accessibility assays, Pbx ChIP-seq assays, and transcriptomics on wild type and Pbx knockout cells undergoing paraxial mesoderm differentiation to show that Pbx factors are required to bind and open essential chromatin regions during the maturation of paraxial mesoderm cells (Mariani et al., 2021). Importantly, the authors used genome editing to show that a Pbx binding site in a regulatory element of the *msgn1* gene that specifies paraxial mesoderm cell fate is required for its chromatin accessibility. Thus, either the loss of the Pbx protein or the disruption of the Pbx binding site resulted in the loss of *msgn1* enhancer DNA accessibility and *msgn1* gene activation (Mariani et al., 2021). These data support a model in which Pbx marks and opens the inaccessible chromatin for subsequent gene regulation by Hox factors (**Figure 2B**).

In *Drosophila*, Porcelli et al. systematically assessed how the Exd and Hth co-factors impact genomic accessibility and Hox DNA binding profiles by taking advantage of the fact that Kc167 cells lack Hth expression, which thereby restricts Exd to the cytoplasm (Porcelli et al., 2019). These studies revealed that the expression of Hth, which concomitantly localizes Exd to the nucleus, generally increased the genomic binding of all the *Drosophila* Hox factors but Abd-B to inaccessible chromatin (**Figure 1A**; Porcelli et al., 2019). This was previously shown for Ubx in which ~30% of the Ubx and Hth specific binding sites did not intersect with the cell line's DNase1 profile prior to Hox gene expression, whereas in the absence of Hth and nuclear Exd only ~5% of Ubx bound regions intersected with this DNaseI inaccessible chromatin profile (Beh et al., 2016). Moreover, by comparing chromatin profiles before and after Ubx and Hth induction, Ubx was shown to open the surrounding chromatin of its targets with the help of Hth (Porcelli et al., 2019). These data support the idea that the formation of Ubx/Hth/Exd complexes can promote chromatin remodeling and DNA accessibility, which is consistent with the findings that Ubx increases chromatin accessibility to activate gene transcription in the Hth and Exd expressing cells of the distal hinge in the haltere disc (Loker et al., 2021).

In agreement with these *Drosophila* findings, a recent study in mouse embryonic stem cells studies found that like Lab (Porcelli et al., 2019), the HOXB1 homologue is capable of binding to both inaccessible and accessible DNA (Singh et al., 2021). Intriguingly, by also performing ChIP-seq assays for PBX1 and various chromatin marks, the authors found that those HOXB1 peaks found in inaccessible DNA were not bound by PBX1 and were predominantly located in gene deserts of nucleosome-bound chromatin. In contrast, the HOXB1 regions that were also bound by PBX1 tended to be in more accessible chromatin that correlated with open chromatin marks such as H2K27ac, H3Kme1, and H3Kme3 (**Figure 1B**; Singh et al., 2021). These findings suggest that while HOXB1 has the capacity to bind inaccessible DNA, it may have a limited ability to convert that binding event into accessible chromatin unless co-bound with the PBX1 factor.

Collectively, the above genomic data in both *Drosophila* and vertebrates support the idea that the Pbx/Exd and Meis/Hth factors have some degree of pioneer TF activity. Consistent with this model, a recent nucleosome consecutive affinity purification-systematic evolution of ligands by exponential enrichment assay provided evidence that MEIS3 is capable of binding nucleosome bound DNA *in vitro* (Zhu et al., 2018), and a comparative study of pioneer TFs highlighted that PBX contains a truncated alpha recognition helix that mimics the structure that allows the FOXA3, OCT4, PU1, and ASCL1 pioneer TFs to bind nucleosome bound DNA (Fernandez Garcia et al., 2019). In total, these studies provide support for the following model: the Pbx/Exd and Meis/Hth factors can bind inaccessible DNA, promote chromatin opening, and ultimately regulate target gene expression *via* the subsequent recruitment of Hox TFs as well as other non-Hox TFs such as MyoD (**Figure 2B**). What remains less clear is if the Hox TFs are only involved in the final step of target gene activation or if the Hox TFs also participate with Pbx/Exd and Meis/Hth in the process of chromatin remodeling. Moreover, since at least a subset of Hox TFs also bind inaccessible DNA, it is possible that at some regulatory elements Hox TFs can use a pioneer-like activity to bind inaccessible DNA and subsequently recruit the Pbx/Exd and/or Meis/Hth factors to open chromatin and regulate target gene expression (**Figure 2C**). Thus, the differential ability of Hox TFs and the Pbx/Exd and Meis/Hth TFs to bind accessible versus inaccessible DNA provide an additional potential regulatory mechanism that may underlie how the anterior, central, and posterior Hox TFs accurately control target gene expression during animal development.

HOX FACTORS AS MULTI-FUNCTIONAL TRANSCRIPTIONAL ACTIVATORS AND REPRESSORS

Once bound to DNA, Hox factors ultimately function by altering the expression of downstream target genes. Unlike some TFs that are thought to function predominantly as transcriptional activators or repressors, the Hox TFs are capable of mediating both transcriptional outcomes (Pearson et al., 2005; Zandvakili and Gebelein, 2016). In the past, considerable work has been done to map activation and repression domains of the Hox factors as well as to determine how mutating these regions impacts transcriptional output. For example, a structure function study of Abd-A in *Drosophila* revealed how a Hox protein can utilize multiple distinct Exd interaction domains to differentially regulate target genes and morphological outcomes (Merabet et al., 2011). Further, a combination of mutational analyses and transcriptional output assays using Gal4 drivers showed that Dfd in *Drosophila* (Li et al., 1999) and HoxA7 in NIH3T3 cells (Schnabel and Abate-Shen, 1996) as well as HoxD4 in P19 embryonal carcinoma cells (Rambaldi et al., 1994) possessed a proline alanine rich region in the N-terminus that can activate transcriptional output. However, this activity was masked by the homeodomain and C-terminus in the context of the full proteins. In fact, there is increasing evidence that the homeodomain itself can be a large driver of transcriptional repression, and that the extent of this repression is paralog specific. A recent study quantitively measured protein domain transcriptional activity using a novel high-throughput sequencing technique, HT-recruit (Tycko et al., 2020). This study fused a large library of TF protein domains to the rTetR DNA binding domain within a lentivirus and assessed their ability to alter a citrine reporter gene under the control of TetO binding sites. After subjecting infected cells to doxycycline, cells were sorted for citrine-ON versus citrine-OFF and the read count ratio between the off and on cells was used to quantify the repression capability of each protein domain. Through this technique, they discovered that the repression capability of the Hox homeodomains was colinear and correlated with paralog such that posterior Hox factors had a more potent repression activity than the anterior Hox factors (Tycko et al., 2020). The authors then connected the enhanced repression of posterior factors to a more positively charged N-terminal arm in the homeodomain, specifically a RKKR motif (Tycko et al., 2020). This connection is consistent with a previous mutational study that found that mutating a similar region of HoxA7 to the amino acids of HoxB4 resulted in reduced repression activity (Schnabel and Abate-Shen, 1996). Altogether, these findings highlight the importance of the homeodomain in repression as well as exemplifies how transcriptional outputs across Hox proteins can vary. Moreover, these data provide further evidence that the same Hox TFs, such as HoxA7, HoxD4, and Abd-A, can have both functional activation and repression domains.

Given that Hox TFs have the capacity to both activate and repress transcription, it is not surprising that a wide variety of co-activator and co-repressor proteins have been shown to physically interact with Hox TFs as reviewed in (Mann et al., 2009; Zandvakili and Gebelein, 2016; De Kumar and Darland, 2021). Many large scale interactome analyses have been performed with Hox factors, and each of these have identified a substantial number of potential protein-protein interactions that could modify the ability of the Hox TFs to mediate gene activation and/or gene repression (Giot et al., 2003; Stanyon et al., 2004; Rual et al., 2005; Yu et al., 2011; Lambert et al., 2012; Rolland et al., 2014; Bischof et al., 2018; Shokri et al., 2019; Carnesecchi et al., 2020; Luck et al., 2020). For example, the Ubx and Abd-A Hox factors were screened for interactions against 260 different TFs in the *Drosophila* embryo using a split-fluorescence assay coupled with ectopic expression using the Gal4-UAS system (Bischof et al., 2018). Unexpectedly, both of these Hox TFs interacted with a large number of the tested TFs, as Ubx interacted with 163 of the 260 TFs (62%), and Abd-A interacted with 149 of the TFs (57%) (Bischof et al., 2018). However, it should be noted that an additional large-scale TF-TF interaction screen tested a number of different Hox TFs using a yeast 2-hybrid assay and found that the Hox TFs, including Ubx and Abd-A, interact with relatively few tested TFs (Shokri et al., 2019). These conflicting results are likely to be attributed to differences in sensitivity between the two assays as well as the fact that the fluorescence complementation assay was performed in *Drosophila* cells that express additional co-factor proteins that may allow large scale complex formation whereas the two-hybrid approach was performed in yeast. More recently, a proximity-dependent Biotin IDentification (BioID) assay in multiple cell types of the *Drosophila* embryo revealed that Ubx interacts with many proteins involved in processes from chromatin modification to mRNA processing (Carnesecchi et al., 2020). Surprisingly, however, while most of the BioID identified Ubx interactors were found to occur in a tissue-specific manner, the vast majority of the proteins that do interact with Ubx are broadly expressed across many tissues (Carnesecchi et al., 2020). This finding raises the possibility that the ability of Ubx to interact with ubiquitous regulatory proteins can be modified in a tissue-specific manner, although the mechanisms regulating such tissue-specific interactions are currently unknown. Nevertheless, these data raise the possibility that the Hox TFs gain in DNA binding specificity by forming complexes with numerous additional TFs, many of which are expressed in a tissue-restricted manner, and gain in regulatory specificity (i.e., activate versus repress) by interacting with many different co-activator and co-repressor proteins that are widely expressed in numerous cell types.

CASE STUDIES ON THE *CIS*-REGULATORY LOGIC OF HOX TRANSCRIPTION FACTORS

To better understand how Hox TFs regulate target genes in specific tissues, a select number of *cis*-regulatory modules (CRMs) have been extensively characterized using a combination of DNA binding assays, transcriptional reporter assays, and loss- and gain-of-function genetics. In this review, we are going to focus on our current knowledge of the *cis*-regulatory logic of two well-characterized *Drosophila* CRMs, one of which is specifically regulated by Abd-A and the other

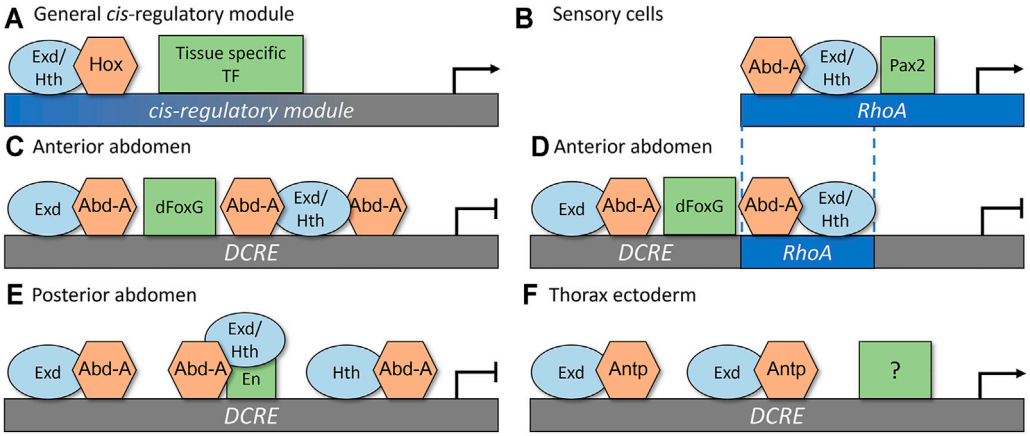


FIGURE 3 | *Rhomboid-A* (*RhoA*) and *Distal-less conserved regulatory element* (*DCRE*) are regulated through tissue specific co-activators and co-repressors to mediate distinct transcriptional outputs. **(A)** A simplistic general model for how a Hox CRM yields cell-specific outputs *via* the direct integration of a Hox factor with the common co-factors Exd and Hth, and a tissue specific TF to mediate either activation or repression **(B)** In sensory cells, Abd-A, Exd, Hth, and Pax2 interact together to activate gene expression by forming complexes on the *RhoA* CRM. **(C)** In the anterior abdomen, Abd-A/Exd/Hth complexes bind together with dFoxG to repress the *DCRE* CRM. **(D)** Replacing two of the Hox binding sites from the *DCRE* with the Abd-A/Hth/Exd sites that mediate activation *via* the *RhoA* CRM did not alter the transcriptional response of the *DCRE*. These data show that the configuration of Hox binding sites does not dictate gene activation versus repression. Instead, it is the presence of nearby dFoxG sites [repression in **(D)**] or Pax2 sites [activation in **(B)**] that dictates transcriptional outcomes. **(E)** In the posterior abdomen, En interacts with Abd-A and Exd/Hth to repress the *DCRE* CRM. **(F)** In the thorax ectoderm, the Antp Hox factor cooperatively binds with Exd to stimulate gene expression. This activation activity does not require the dFoxG nor the En sites. In this model, we propose that an unknown tissue specific TF mediates activation of *DCRE* CRM to result in ectoderm specific gene activation.

is regulated by the Abd-A, Ubx, and Antp Hox factors. Intriguingly, Abd-A regulates these two CRMs in different cell types and in opposing ways. In the developing peripheral nervous system, Abd-A triggers the secretion of epidermal growth factor ligands from a specific subset of abdominal sensory organ precursor cells by activating the expression of the *rhomboid* (*rho*) serine protease gene *via* a highly conserved CRM called *RhoA* (Brodu et al., 2002; Li-Kroeger et al., 2008). In contrast, Abd-A, as well as Ubx, suppresses leg development in abdominal segments by repressing the expression of the *Distal-less* (*Dll*) homeodomain protein in ectodermal cells *via* the *Dll* conserved regulatory element (*DCRE*) (Vachon et al., 1992; Gebelein et al., 2002, Gebelein et al., 2004). In addition to being repressed by both Abd-A and Ubx in the abdomen, the *DCRE* CRM can also be activated by the Antennapedia (Antp) Hox factor in thoracic segments (Uhl et al., 2016).

To determine how these CRMs recruit specific Hox factors to mediate distinct cell type- and segment-specific outputs, comparative studies on the TF binding sites (TFBSs), the TF complexes, and the genetic requirements of each TF have revealed several insights into the principals underlying Hox *cis*-regulatory logic (**Figure 3**). First, the same Hox, Exd, and Hth binding sites are capable of mediating either activation or repression. The *RhoA* CRM contains a single set of adjacent Exd/Hth/Hox binding sites (**Figure 3B**; Brodu et al., 2002; Li-Kroeger et al., 2008), whereas the *DCRE* CRM contains three Hox sites, each of which is directly adjacent to a Exd or Hth site (**Figures 3C,E**; Gebelein et al., 2002, Gebelein et al., 2004; Uhl et al., 2016). Each configuration of binding sites is capable of cooperatively binding Abd-A/Hth/Exd complexes. Further, mutations within these binding sites disrupt the ability of Abd-A to either activate gene expression in sensory cells (**Figure 3B**; Li-Kroeger et al., 2012) or repress gene expression in the abdominal ectoderm (**Figures 3C,E**; Gebelein et al., 2002, 2004; Uhl et al., 2016). Moreover, swapping the "activating" Exd/Hth/Hox sites from the *RhoA* CRM into the *DCRE* demonstrated that the Abd-A Hox factor can also use this same configuration of binding sites to mediate transcriptional repression (**Figure 3D**; Zandvakili et al., 2019). These data suggest that differences in the conformation of Exd, Hth, and Hox TFBSs do not reveal how the Abd-A Hox complex mediates distinct outcomes in different cell types.

Second, accurate Hox-dependent transcriptional outcomes by the *RhoA* and *DCRE* CRMs require nearby TFBSs for additional tissue-restricted TFs (**Figure 3A**). For example, the *RhoA* CRM encodes a binding site for the Pax2 TF (**Figure 3B**, the *Drosophila* Pax2 gene name is *shaven*, *sv*, but for simplicity we will call it Pax2) and mutations within the *RhoA* CRM that disrupt Pax2 binding compromise Abd-A mediated activation (Li-Kroeger et al., 2012; Zandvakili et al., 2018). Moreover, Pax2, Abd-A, Exd, and Hth could utilize these TFBSs to form specific TF complexes on the *RhoA* CRM (Li-Kroeger et al., 2012) and altering the spacing and orientation between the Pax2 and Exd/Hth/Hox site disrupted *RhoA* activity in abdominal sensory organ cells (Zandvakili et al., 2019). Given that the expression of the *Drosophila Pax2* gene is predominately restricted to sensory organ cells in the embryo (Li-Kroeger et al., 2012), the direct regulation of the *RhoA* CRM by Abd-A and Pax2 provides insight into both the abdominal and sensory specific activity of this enhancer.

The *DCRE* CRM similarly requires additional TFBSs to mediate abdominal-specific repression by the Ubx and Abd-A Hox factors (**Figures 3C,E**; Gebelein et al., 2004). In fact, these

two Hox factors were found to require different TFs in the anterior versus the posterior compartments of the abdominal segments. In the posterior compartment cells, the Ubx and Abd-A Hox factors form cooperative complexes with the Engrailed (En) TFs on adjacent binding sites within the *DCRE* (**Figure 3E**). In the anterior compartment, the Ubx and Abd-A Hox factors cooperate with the *Drosophila* FoxG factors, which are encoded by the largely redundant *sloppy-paired 1* (*slp1*) and *sloppy-paired 2* (*slp2*) genes *via* nearby Hox and dFoxG binding sites within the *DCRE* (**Figure 3C**; Gebelein et al., 2004). Moreover, like in the *RhoA* CRM, the spacing between the Hox and dFoxG TFBSs contributed to optimal activity as adding a 5 bp sequence disrupted repression (Zandvakili et al., 2019). However, adding longer space sequences between the Hox and dFoxG sites (+10, +15, and +20 bps) resulted in strong transcriptional repression, suggesting that unlike the Pax2 and Hox sites in *RhoA*, the configurations of dFoxG and Hox sites in the *DCRE* are not rigidly fixed to mediate transcriptional repression (Zandvakili et al., 2019). Since the dFoxG factors are specifically expressed in the anterior compartment cells, whereas En is specifically expressed in the posterior compartment cells, these findings again highlight how a CRM can integrate both Hox TFs and tissue-restricted TFs to yield accurate abdominal-specific outcomes.

Third, the *DCRE* CRM can contribute to both Hox-mediated transcriptional repression in the abdomen and Hox-mediated transcriptional activation in the thorax. In addition to repressing *Dll* expression in the abdomen, the *DCRE* can also use a subset of the Hox TFBSs to stimulate gene expression in the thoracic leg primordia cells *via* the Antp Hox TF (Uhl et al., 2016). In particular, Antp can utilize the two Hox/Exd sites, but not the adjacent Hth/Hox site, to stimulate *DCRE*-mediated activation in thoracic cells (**Figure 3F**). However, unlike *DCRE* mediated repression, the dFoxG and the En binding sites are not required for this Hox-dependent activity, suggesting that Antp is likely to cooperate with other TFs to stimulate the *DCRE* (**Figure 3F**; Uhl et al., 2016). Additional *cis*-regulatory studies on the *six2* target gene in mammals have also revealed that the same Hox binding sites can be used to activate *six2 via* Hox11 factors in the kidney and repress *six2 via* Hoxa2 in the branchial arch and facial mesenchyme (Yallowitz et al., 2009). Thus, the same Hox binding sites can contribute to both activation and repression, but the appropriate transcriptional response will likely depend upon integrating distinct combinations of additional TFs.

While the thorough characterization of the *RhoA* and *DCRE* CRMs provide new insight into the *cis*-regulatory logic of the segment- and cell type-specific transcriptional responses, does this *cis*-regulatory logic provide insight into how Abd-A represses the *DCRE* and activates the *RhoA* CRMs? Currently, it is unclear how Abd-A/Pax2 complexes activate the *RhoA* CRM in sensory cells, but it is interesting to note that vertebrate Pax2 and Hox11 factors are thought to collaborate to activate the *six2* target gene and Hox11 contains an activation domain required for this function (Gong et al., 2007; Yallowitz et al., 2009). Moreover, the integration of Abd-A with either the Slp1/2 dFoxG factors or the En homeodomain factor provides a likely mechanism of repression as both dFoxG and En have been shown to recruit the well-established Groucho co-repressor protein (Jiménez et al., 1997; Andrioli et al., 2004). In addition, Agelopoulos et al. used a novel cell- and gene-specific ChIP strategy to demonstrate that while the regulatory element containing the *DCRE* loops and contacts the *Dll* promoter in thoracic segments, consistent with gene activation, the *DCRE* region does not contact the *Dll* promoter in the abdominal segments (Agelopoulos et al., 2012). Using whole embryo ChIP, the authors then found that the *Dll* enhancer region containing the *DCRE* was not only bound by Ubx and Abd-A but also was highly correlated with histone variant, H2Av. These data suggest that Ubx and/or Abd-A may recruit H2Av and result in decreased interactions between the *DCRE* and the *Dll* promoter (Agelopoulos et al., 2012). These data also further highlight how some Hox factors, most notably Ubx, can activate and repress gene expression by both increasing and decreasing chromatin accessibility in a gene-specific manner (Agelopoulos et al., 2012; Loker et al., 2021).

DISCUSSION

Several properties of the Hox factors make it particularly challenging to develop a general model that predicts the outcome of a Hox binding event to a regulatory element. First, Hox factors are expressed within many cell types of a segment, and yet most Hox target genes are regulated in only a small subset of the cells that express the Hox TF. Second, as mentioned throughout this review, all the members of the Hox TF family share a homeodomain that binds highly similar AT-rich sequences, and yet Hox TFs specify different segment identities and regional cell fates. Thus, we need to determine both the mechanisms that underlie how the same broadly expressed Hox TF can regulate target genes in a cell-type specific manner, and the mechanisms that underlie what makes each Hox TF different from each other to regulate the distinct combinations of target genes needed to specify different cell fates along the A-P axis.

This review summarized several studies that suggest select Hox factors can bind to inaccessible chromatin by intersecting genome binding and chromatin accessibility profiles (Beh et al., 2016; Porcelli et al., 2019; Bulajić et al., 2020). Moreover, at least in the case of Dfd and Abd-B, Hox TFs have the potential to increase chromatin accessibility even in the absence of nuclear Exd and Hth (Porcelli et al., 2019). However, intersecting genome binding and chromatin accessibility profiles only provides correlative evidence of pioneer-like activity, as it is possible that other TFs regulated by Dfd and Abd-B are ultimately the ones that open the chromatin. Thus, additional studies that combine the use of ATAC-seq, ChIP-seq, and Hox mutational analysis with genomic editing of known Hox binding sites will be required to confirm or refute Hox pioneer-like activity, much like the studies of Mariani et al. for PBX established pioneer activity in mouse epiblast stem cells (Mariani et al., 2021).

Through a combination of large-scale genomic, bioinformatics, and protein interactome approaches, the scientific field has recently identified numerous Hox-bound genomic elements as well as many Hox protein interactors

that are likely to contribute to the diverse regulatory potential of Hox factors. Of particular note is that Hox TFs were found to form complexes with a surprisingly large number of other sequence specific TFs (Bischof et al., 2018; Carnesecchi et al., 2020). Taken together with the finding that several well-characterized Hox regulatory elements require TFBSs for additional TFs that are expressed in tissue- and cell type-restricted patterns, these results suggest that Hox-regulated CRMs function by integrating specific Hox TFs with numerous other TFs to yield accurate segment-, cell-, and gene-specific regulatory outcomes. The question that arises from these studies is do Hox TFs regulate each target gene by interactions with distinct combinations of TFs, and thus each Hox-regulated CRM will contain a relatively unique combination of TFBSs? Or do the Hox TFs regulate many different target genes *via* interactions with a common group of TFs such that potential *cis*-regulatory codes can be identified and used to predict Hox-regulated CRM output?

To answer these questions, future experiments will be needed to generate additional genomic binding data and transcriptomics data for Hox TFs as well as their potential partner proteins in defined cell types. For example, intersecting ChIP-seq for Hox TFs, Pbx/Exd, Hth/Meis, and other TFs from the same cell types would allow one to segregate Hox genomic binding events into distinct bins that are associated with the binding or lack thereof of additional TFs. Moreover, the use of higher resolution binding assays such as Cleavage Under Targets & Release Using Nuclease (CUT&RUN) or ChIP-Exonuclease can provide near bp resolution binding that reveals if adjacent sites are also occupied near the Hox TF binding site. Such an approach was recently utilized for the Gsx2 homeodomain TF to reveal distinct monomer versus dimer binding events using CUT&RUN assays and nucleotide footprinting analysis (Salomone et al., 2021). By combining high-resolution genomic binding data with transcriptomic studies using wild type and specific mutant cells (i.e., Hox mutant, Pbx/Exd mutant, etc), we will be better positioned to define which binding events are associated with gene expression changes. Lastly, bioinformatics can be used to perform unbiased searches for additional TF motifs as well as to search for potential constraints on the relationships between Hox TF sites and other TFBSs. Such an approach has already identified that many Hox genomic binding events enrich for coupled Pbx/Hox (vertebrates) or Exd/Hox (*Drosophila*) motifs even when the genomic binding assay was performed using a complex tissue composed of many cell types (Loker et al., 2021; Singh et al., 2021). These findings are consistent with the Pbx/Exd TFs serving as widespread Hox co-factor proteins in many tissues. Moreover, a recent study for HoxB1 combined genomic binding assays with transcriptomics and unbiased motif enrichment analysis to show that HoxB1 genomic binding events associated with gene repression, but not gene activation, are enriched for TFBSs for the REST transcriptional repressor (Singh et al., 2021). Hence, future studies focused on genomic binding assays for many Hox and other TFs in specific tissues will be needed to determine which TFs are likely to collaborate with specific Hox factors. Armed with the sequences of these potential regulatory elements, bioinformatics approaches will help to reveal if specific *cis*-regulatory codes underlie how Hox TFs are integrated with each different TF to regulate cell specific gene expression.

AUTHOR CONTRIBUTIONS

The topic of this review was jointly conceived by BC and BG. The initial draft of the manuscript was written by BC and BG. The figures were initially generated by BC and edited by BG.

FUNDING

This work was supported by a National Institutes of Health grant #GM079428 to BG.

ACKNOWLEDGMENTS

Many of the ideas presented in this review were stimulated by present and past members of the Gebelein laboratory.

REFERENCES

Abbott, M. K., and Kaufman, T. C. (1986). The Relationship between the Functional Complexity and the Molecular Organization of the Antennapedia Locus of *Drosophila melanogaster*. *Genetics* 114 (3), 919–942. doi:10.1093/genetics/114.3.919

Agelopoulos, M., McKay, D. J., and Mann, R. S. (2012). Developmental Regulation of Chromatin Conformation by Hox Proteins in Drosophila. *Cel Rep.* 1 (4), 350–359. doi:10.1016/j.celrep.2012.03.003

Andrioli, L. P., Oberstein, A. L., Corado, M. S. G., Yu, D., and Small, S. (2004). Groucho-dependent Repression by Sloppy-Paired 1 Differentially Positions Anterior Pair-Rule Stripes in the Drosophila Embryo. *Develop. Biol.* 276 (2), 541–551. doi:10.1016/j.ydbio.2004.09.025

Beh, C. Y., El-Sharnouby, S., Chatzipli, A., Russell, S., Choo, S. W., and White, R. (2016). Roles of Cofactors and Chromatin Accessibility in Hox Protein Target Specificity. *Epigenetics Chromatin* 9 (1), 1–17. doi:10.1186/s13072-015-0049-x

Bender, W., Akam, M., Karch, F., Beachy, P. A., Peifer, M., Spierer, P., et al. (1983). Molecular Genetics of the Bithorax Complex in *Drosophila melanogaster*. *Science* 221 (4605), 23–29. doi:10.1126/science.221.4605.23

Berger, M. F., Badis, G., Gehrke, A. R., Talukder, S., Philippakis, A. A., Peña-Castillo, L., et al. (2008). Variation in Homeodomain DNA Binding Revealed by High-Resolution Analysis of Sequence Preferences. *Cell* 133 (7), 1266–1276. doi:10.1016/j.cell.2008.05.024

Bergstrom, D. A., Penn, B. H., Strand, A., Perry, R. L. S., Rudnicki, M. A., and Tapscott, S. J. (2002). Promoter-specific Regulation of MyoD Binding and Signal Transduction Cooperate to Pattern Gene Expression. *Mol. Cel* 9 (3), 587–600. doi:10.1016/S1097-2765(02)00481-1

Berkes, C. A., Bergstrom, D. A., Penn, B. H., Seaver, K. J., Knoepfler, P. S., and Tapscott, S. J. (2004). Pbx marks Genes for Activation by MyoD Indicating a Role for a Homeodomain Protein in Establishing Myogenic Potential. *Mol. Cel* 14 (4), 465–477. doi:10.1016/S1097-2765(04)00260-6

Biggin, M. D. (2011). Animal Transcription Networks as Highly Connected, Quantitative Continua. *Develop. Cel* 21 (4), 611–626. doi:10.1016/j.devcel.2011.09.008

Bischof, J., Duffraisse, M., Furger, E., Ajuria, L., Giraud, G., Vanderperre, S., et al. (2018). Generation of a Versatile BiFC ORFeome Library for Analyzing Protein-Protein Interactions in Live Drosophila. *ELIFE* 7, 1–24. doi:10.1101/34348310.7554/eLife.38853

Brodu, V., Elstob, P. R., and Gould, A. P. (2002). Abdominal a Specifies One Cell Type in Drosophila by Regulating One Principal Target Gene. *Development* 129 (12), 2957–2963. doi:10.1242/dev.129.12.2957

Bulajić, M., Srivastava, D., Dasen, J. S., Wichterle, H., Mahony, S., and Mazzoni, E. O. (2020). Differential Abilities to Engage Inaccessible Chromatin Diversify Vertebrate Hox Binding Patterns. *Development* 147, dev194761. doi:10.1242/dev.194761

Bürglin, T. R., and Affolter, M. (2016). Homeodomain Proteins: an Update. *Chromosoma* 125 (3), 497–521. doi:10.1007/s00412-015-0543-8

Carnesecchi, J., Sigismondo, G., Domsch, K., Baader, C. E. P., Rafiee, M. R., Krijgsveld, J., et al. (2020). Multi-level and Lineage-specific Interactomes of the Hox Transcription Factor Ubx Contribute to its Functional Specificity. *Nat. Commun.* 11, 1388. doi:10.1038/s41467-020-15223-x

Carroll, S. B., Weatherbee, S. D., and Langeland, J. A. (1995). Homeotic Genes and the Regulation and Evolution of Insect wing Number. *Nature* 375 (6526), 58–61. doi:10.1038/375058a0

Casares, F., and Mann, R. S. (1998). Control of Antennal versus Leg Development in Drosophila. *Nature* 392 (6677), 723–726. doi:10.1038/33706

Choe, S.-K., Ladam, F., and Sagerström, C. G. (2014). TALE Factors Poise Promoters for Activation by Hox Proteins. *Develop. Cel* 28 (2), 203–211. doi:10.1016/j.devcel.2013.12.011

Choo, S. W., and Russell, S. (2011). Genomic Approaches to Understanding Hox Gene Function. *Adv. Genet.* 76, 55. doi:10.1016/B978-0-12-386481-9.00003-1

Crocker, J., Abe, N., Rinaldi, L., McGregor, A. P., Frankel, N., Wang, S., et al. (2015). Low Affinity Binding Site Clusters Confer HOX Specificity and Regulatory Robustness. *Cell* 160 (1–2), 191–203. doi:10.1016/j.cell.2014.11.041

De Kumar, B., and Darland, D. C. (2021). The Hox Protein Conundrum: The "Specifics" of DNA Binding for Hox Proteins and Their Partners. *Develop. Biol.* 477 (June), 284–292. doi:10.1016/j.ydbio.2021.06.002

Desanlis, I., Kherdjemil, Y., Mayran, A., Bouklouch, Y., Gentile, C., Sheth, R., et al. (2020). HOX13-dependent Chromatin Accessibility Underlies the Transition towards the Digit Development Program. *Nat. Commun.* 11, 2491. doi:10.1038/s41467-020-16317-2

Duboule, D., and Dollé, P. (1989). The Structural and Functional Organization of the Murine HOX Gene Family Resembles that of Drosophila Homeotic Genes. *EMBO J.* 8 (5), 1497–1505. doi:10.1002/j.1460-2075.1989.tb03534.x

Farnham, P. J. (2009). Insights from Genomic Profiling of Transcription Factors. *Nat. Rev. Genet.* 10 (9), 605–616. doi:10.1038/nrg2636

Fernandez Garcia, M., Moore, C. D., Schulz, K. N., Alberto, O., Donague, G., Harrison, M. M., et al. (2019). Structural Features of Transcription Factors Associating with Nucleosome Binding. *Mol. Cel* 75 (5), 921–932. doi:10.1016/j.molcel.2019.06.009

Fisher, W. W., Li, J. J., Hammonds, A. S., Brown, J. B., Pfeiffer, B. D., Weiszmann, R., et al. (2012). DNA Regions Bound at Low Occupancy by Transcription Factors Do Not Drive Patterned Reporter Gene Expression in Drosophila. *Proc. Natl. Acad. Sci.* 109 (52), 21330–21335. doi:10.1073/pnas.1209589110

Gebelein, B., Culi, J., Ryoo, H. D., Zhang, W., and Mann, R. S. (2002). Specificity of Distalless Repression and Limb Primordia Development by Abdominal Hox Proteins. *Develop. Cel* 3 (4), 487–498. doi:10.1016/S1534-5807(02)00257-5

Gebelein, B., McKay, D. J., and Mann, R. S. (2004). Direct Integration of Hox and Segmentation Gene Inputs during Drosophila Development. *Nature* 431 (7009), 653–659. doi:10.1038/nature02946

Gerber, A. N., Klesert, T. R., Bergstrom, D. A., and Tapscott, S. J. (1997). Two Domains of MyoD Mediate Transcriptional Activation of Genes in Repressive Chromatin: A Mechanism for Lineage Determination in Myogenesis. *Genes Develop.* 11 (4), 436–450. doi:10.1101/gad.11.4.436

Giot, L., Bader, J. S., Brouwer, C., Chaudhuri, A., Kuang, B., Li, Y., et al. (2003). A Protein Interaction Map of *Drosophila melanogaster*. *Science* 302 (5651), 1727–1736. doi:10.1126/science.1090289

Gong, K.-Q., Yallowitz, A. R., Sun, H., Dressler, G. R., and Wellik, D. M. (2007). A Hox-Eya-Pax Complex Regulates Early Kidney Developmental Gene Expression. *Mol. Cel Biol.* 27 (21), 7661–7668. doi:10.1128/mcb.00465-07

Graham, A., Papalopulu, N., and Krumlauf, R. (1989). The Murine and Drosophila Homeobox Gene Complexes Have Common Features of Organization and Expression. *Cell* 57 (3), 367–378. doi:10.1016/0092-8674(89)90912-4

Hajirnis, N., and Mishra, R. K. (2021). Homeotic Genes: Clustering, Modularity, and Diversity. *Front. Cel Develop. Biol.* 9 (August), 1–20. doi:10.3389/fcell.2021.718308

Hueber, S. D., Weiller, G. F., Djordjevic, M. A., and Frickey, T. (2010). Improving Hox Protein Classification across the Major Model Organisms. *PLoS ONE* 5, e10820. doi:10.1371/journal.pone.0010820

Iwafuchi-Doi, M., and Zaret, K. S. (2014). Pioneer Transcription Factors in Cell Reprogramming. *Genes Dev.* 28 (24), 2679–2692. doi:10.1101/gad.253443.114

Jiménez, G., Paroush, Z. e., and Ish-Horowicz, D. (1997). Groucho Acts as a Corepressor for a Subset of Negative Regulators, Including Hairy and Engrailed. *Genes Develop.* 11 (22), 3072–3082. doi:10.1101/gad.11.22.3072

Jolma, A., Yan, J., Whitington, T., Toivonen, J., Nitta, K. R., Rastas, P., et al. (2013). DNA-binding Specificities of Human Transcription Factors. *Cell* 152 (1–2), 327–339. doi:10.1016/j.cell.2012.12.009

Kaufman, T. C., Lewis, R., and Wakimoto, B. (1980). Cytogenetic Analysis of Chromosome 3 in *Drosophila melanogaster*: The Homoeotic Gene Complex in Polytene Chromosome Interval 84A-B. *Genetics* 94 (1), 115–133. doi:10.1093/genetics/94.1.115

Kribelbauer, J. F., Rastogi, C., Bussemaker, H. J., and Mann, R. S. (2019). Low-affinity Binding Sites and the Transcription Factor Specificity Paradox in Eukaryotes. *Annu. Rev. Cel Dev. Biol.* 35, 357–379. doi:10.1146/annurev-cellbio-100617-062719

Lambert, B., Vandeputte, J., Remacle, S., Bergiers, I., Simonis, N., Twizere, J. C., et al. (2012). Protein Interactions of the Transcription Factor Hoxa1. *BMC Dev. Biol.* 12, 29. doi:10.1186/1471-213X-12-29

LaRonde-LeBlanc, N. A., and Wolberger, C. (2003). Structure of HoxA9 and Pbx1 Bound to DNA: Hox Hexapeptide and DNA Recognition Anterior to Posterior. *Genes Develop.* 17 (16), 2060–2072. doi:10.1101/gad.1103303

Lewis, E. B. (1978). A Gene Complex Controlling Segmentation in Drosophila. *Nature* 276 (5688), 565–570. doi:10.1038/276565a0

Li, X., Murre, C., and McGinnis, W. (1999). Activity Regulation of a Hox Protein and a Role for the Homeodomain in Inhibiting Transcriptional Activation. *EMBO J.* 18 (1), 198–211. doi:10.1093/emboj/18.1.198

Li-Kroeger, D., Cook, T. A., and Gebelein, B. (2012). Integration of an Abdominal Hox Complex with PAX2 Yields Cell-specific EGF Secretion from Drosophila Sensory Precursor Cells. *Development* 139 (9), 1611–1619. doi:10.1242/dev.077842

Li-Kroeger, D., Witt, L. M., Grimes, H. L., Cook, T. A., and Gebelein, B. (2008). Hox and Senseless Antagonism Functions as a Molecular Switch to Regulate EGF Secretion in the Drosophila PNS. *Develop. Cel* 15 (2), 298–308. doi:10.1016/j.devcel.2008.06.001

Loker, R., Sanner, J. E., and Mann, R. S. (2021). Cell-type-specific Hox Regulatory Strategies Orchestrate Tissue Identity. *Curr. Biol.* 31 (19), 4246–4255. doi:10.1016/j.cub.2021.07.030

Luck, K., Kim, D. K., Lambourne, L., Spirohn, K., Begg, B. E., Bian, W., et al. (2020). A Reference Map of the Human Binary Protein Interactome. *Nature* 580 (7803), 402–408. doi:10.1038/s41586-020-2188-x

Luo, Z., Rhie, S. K., and Farnham, P. J. (2019). The Enigmatic Hox Genes: Can We Crack Their Code? *Cancers* 11, 323. doi:10.3390/cancers11030323

Maeda, R. K., and Karch, F. (2009). The Bithorax Complex of Drosophila. *Curr. Top. Develop. Biol.* 88 (09), 1–33. doi:10.1016/S0070-2153(09)88001-0

Mann, R. S., and Affolter, M. (1998). Hox Proteins Meet More Partners. *Curr. Opin. Genet. Develop.* 8 (4), 423–429. doi:10.1016/S0959-437X(98)80113-5

Mann, R. S., Lelli, K. M., and Joshi, R. (2009). Hox Specificity Unique Roles for Cofactors and Collaborators. *Curr. Top. Develop. Biol.* 88, 63. doi:10.1016/S0070-2153(09)88003-4

Mann, R. S. (1997). Why areHox Genes Clustered? *BioEssays* 19 (8), 661–664. doi:10.1002/bies.950190804

Mariani, L., Guo, X., Menezes, N. A., Drozd, A. M., Çakal, S. D., Wang, Q., et al. (2021). A TALE/HOX Code Unlocks WNT Signalling Response towards Paraxial Mesoderm. *Nat. Commun.* 12 (1), 5136–5219. doi:10.1038/s41467-021-25370-4

Mark, M., Rijli, F. M., and Chambon, P. (1997). Homeobox Genes in Embryogenesis and Pathogenesis. *Pediatr. Res.* 42 (4), 421–429. doi:10.1203/00006450-199710000-00001

McGinnis, W., Garber, R. L., Wirz, J., Kuroiwa, A., and Gehring, W. J. (1984a). A Homologous Protein-Coding Sequence in drosophila Homeotic Genes and its Conservation in Other Metazoans. *Cell* 37 (2), 403–408. doi:10.1016/0092-8674(84)90370-2

McGinnis, W., Levine, M. S., Hafen, E., Kuroiwa, A., and Gehring, W. J. (1984b). A Conserved DNA Sequence in Homoeotic Genes of the Drosophila Antennapedia and Bithorax Complexes. *Nature* 308 (5958), 428–433. doi:10.1038/308428a0

McKay, D. J., and Lieb, J. D. (2013). A Common Set of DNA Regulatory Elements Shapes Drosophila Appendages. *Develop. Cel* 27 (3), 306–318. doi:10.1016/j.devcel.2013.10.009

Merabet, S., Litim-Mecheri, I., Karlsson, D., Dixit, R., Saadaoui, M., Monier, B., et al. (2011). Insights into Hox Protein Function from a Large Scale Combinatorial Analysis of Protein Domains. *Plos Genet.* 7, e1002302. doi:10.1371/journal.pgen.1002302

Merabet, S., and Mann, R. S. (2016). To Be Specific or Not: The Critical Relationship between Hox and TALE Proteins. *Trends Genet.* 32 (6), 334–347. doi:10.1016/j.tig.2016.03.004

Moens, C. B., and Selleri, L. (2006). Hox Cofactors in Vertebrate Development. *Develop. Biol.* 291 (2), 193–206. doi:10.1016/j.ydbio.2005.10.032

Morata, G., Macías, A., Urquía, N., and González-Reyes, A. (1990). Homoeotic Genes. *Semin. Cel Biol.* 1 (3), 219–227.

Noordermeer, D., and Duboule, D. (2013). Chromatin Architectures and Hox Gene Collinearity. *Curr. Top. Develop. Biol.* 104, 113. doi:10.1016/B978-0-12-416027-9.00004-8

Pearson, J. C., Lemons, D., and McGinnis, W. (2005). Modulating Hox Gene Functions during Animal Body Patterning. *Nat. Rev. Genet.* 6 (12), 893–904. doi:10.1038/nrg1726

Porcelli, D., Fischer, B., Russell, S., and White, R. (2019). Chromatin Accessibility Plays a Key Role in Selective Targeting of Hox Proteins. *Genome Biol.* 20 (1), 115–119. doi:10.1186/s13059-019-1721-4

Quinonez, S. C., and Innis, J. W. (2014). Human HOX Gene Disorders. *Mol. Genet. Metab.* 111 (1), 4–15. doi:10.1016/j.ymgme.2013.10.012

Rambaldi, I., Kovàcs, E. N., and Featherstone, M. S. (1994). A Proline-Rich Transcriptional Activation Domain in Murine HOXD-4 (HOX-4.2). *Nucl. Acids Res.* 22 (3), 376–382. doi:10.1093/nar/22.3.376

Rolland, T., Taşan, M., Charloteaux, B., Pevzner, S. J., Zhong, Q., Sahni, N., et al. (2014). A Proteome-Scale Map of the Human Interactome Network. *Cell* 159 (5), 1212–1226. doi:10.1016/j.cell.2014.10.050

Rual, J.-F., Venkatesan, K., Hao, T., Hirozane-Kishikawa, T., Dricot, A., Li, N., et al. (2005). Towards a Proteome-Scale Map of the Human Protein-Protein Interaction Network. *Nature* 437 (7062), 1173–1178. doi:10.1038/nature04209

Sagerström, C. G. (2004). pbX Marks the Spot. *Develop. Cel* 6, 737–738. doi:10.1016/j.devcel.2004.05.015

Salomone, J., Qin, S., Fufa, T. D., Cain, B., Farrow, E., Guan, B., et al. (2021). Conserved Gsx2/Ind Homeodomain Monomer versus Homodimer DNA Binding Defines Regulatory Outcomes in Flies and Mice. *Genes Dev.* 35 (1), 157–174. doi:10.1101/GAD.343053.120

Schnabel, C. A., and Abate-Shen, C. (1996). Repression by HoxA7 Is Mediated by the Homeodomain and the Modulatory Action of its N-Terminal-Arm Residues. *Mol. Cel Biol.* 16 (6), 2678–2688. doi:10.1128/mcb.16.6.2678

Schneuwly, S., Klemenz, R., and Gehring, W. J. (1987). Redesigning the Body Plan of Drosophila by Ectopic Expression of the Homoeotic Gene Antennapedia. *Nature* 325 (6107), 816–818. doi:10.1038/325816a0

Shokri, L., Inukai, S., Hafner, A., Weinand, K., Hens, K., Vedenko, A., et al. (2019). A Comprehensive *Drosophila melanogaster* Transcription Factor Interactome. *Cel Rep.* 27 (3), 955–970. doi:10.1016/j.celrep.2019.03.071

Singh, N. P., De Kumar, B., Paulson, A., Parrish, M. E., Scott, C., Zhang, Y., et al. (2021). Genome-wide Binding Analyses of HOXB1 Revealed a Novel DNA Binding Motif Associated with Gene Repression. *J. Dev. Biol.* 9 (1), 1–20. doi:10.3390/JDB9010006

Singh, N. P., de Kumar, B., Paulson, A., Parrish, M. E., Zhang, Y., Florens, L., et al. (2020). A Six-Amino-Acid Motif Is a Major Determinant in Functional Evolution of HOX1 Proteins. *Genes Dev.* 34 (23–24), 1680–1696. doi:10.1101/gad.342329.120

Slattery, M., Riley, T., Liu, P., Abe, N., Gomez-Alcala, P., Dror, I., et al. (2011). Cofactor Binding Evokes Latent Differences in DNA Binding Specificity between Hox Proteins. *Cell* 147 (6), 1270–1282. doi:10.1016/j.cell.2011.10.053

Stanyon, C. A., Liu, G., Mangiola, B. A., Patel, N., Giot, L., Kuang, B., et al. (2004). A Drosophila Protein-Interaction Map Centered on Cell-Cycle Regulators. *Genome Biol.* 5, R96. doi:10.1186/gb-2004-5-12-r96

Tycko, J., DelRosso, N., Hess, G. T., Banerjee, A., Van, M. V., Ego, B. K., et al. (2020). High-Throughput Discovery and Characterization of Human Transcriptional Effectors. *Cell* 183 (7), 2020–2035. doi:10.1016/j.cell.2020.11.024

Uhl, J. D., Cook, T. A., and Gebelein, B. (2010). Comparing Anterior and Posterior Hox Complex Formation Reveals Guidelines for Predicting Cis-Regulatory Elements. *Dev. Biol.* 343 (1–2), 154–166. doi:10.1016/j.ydbio.2010.04.004

Uhl, J. D., Zandvakili, A., and Gebelein, B. (2016). A Hox Transcription Factor Collective Binds a Highly Conserved Distal-Less Cis-Regulatory Module to Generate Robust Transcriptional Outcomes. *Plos Genet.* 12 (4), e1005981. doi:10.1371/journal.pgen.1005981

Vachon, G., Cohen, B., Pfeifle, C., McGuffin, M. E., Botas, J., and Cohen, S. M. (1992). Homeotic Genes of the Bithorax Complex Repress Limb Development in the Abdomen of the Drosophila Embryo through the Target Gene Distal-Less. *Cell* 71 (3), 437–450. doi:10.1016/0092-8674(92)90513-C

Walhout, A. J. (2011). What Does Biologically Meaningful Mean? A Perspective on Gene Regulatory Network Validation. *Genome Biol.* 12 (4), 109–117. doi:10.1186/gb-2011-12-4-109

Walter, J., Dever, C. A., and Biggin, M. D. (1994). Two Homeo Domain Proteins Bind with Similar Specificity to a Wide Range of DNA Sites in Drosophila Embryos. *Genes Develop.* 8 (14), 1678–1692. doi:10.1101/gad.8.14.1678

Yallowitz, A. R., Gong, K.-Q., Swinehart, I. T., Nelson, L. T., and Wellik, D. M. (2009). Non-homeodomain Regions of Hox Proteins Mediate Activation versus Repression of Six2 via a Single Enhancer Site *In Vivo*. *Develop. Biol.* 335 (1), 156–165. doi:10.1016/j.ydbio.2009.08.020

Yu, H., Tardivo, L., Tam, S., Weiner, E., Gebreab, F., Fan, C., et al. (2011). Next-generation Sequencing to Generate Interactome Datasets. *Nat. Methods* 8 (6), 478–480. doi:10.1038/nmeth.1597

Zandvakili, A., Campbell, I., Gutzwiller, L. M., Weirauch, M. T., and Gebelein, B. (2018). Degenerate Pax2 and Senseless Binding Motifs Improve Detection of Low-Affinity Sites Required for Enhancer Specificity. *Plos Genet.* 14 (4), e1007289–25. doi:10.1371/journal.pgen.1007289

Zandvakili, A., and Gebelein, B. (2016). Mechanisms of Specificity for Hox Factor Activity. *J. Dev. Biol.* 4, 16. doi:10.3390/jdb4020016

Zandvakili, A., Uhl, J. D., Campbell, I., Salomone, J., Song, Y. C., and Gebelein, B. (2019). The Cis-Regulatory Logic Underlying Abdominal Hox-Mediated Repression versus Activation of Regulatory Elements in Drosophila. *Develop. Biol.* 445 (2), 226–236. doi:10.1016/j.ydbio.2018.11.006

Zaret, K. S. (2020). Pioneer Transcription Factors Initiating Gene Network Changes. *Annu. Rev. Genet.* 54, 367–385. doi:10.1146/annurev-genet-030220-015007

Zeiske, T., Baburajendran, N., Kaczynska, A., Brasch, J., Palmer, A. G., Shapiro, L., et al. (2018). Intrinsic DNA Shape Accounts for Affinity Differences between Hox-Cofactor Binding Sites. *Cel Rep.* 24 (9), 2221–2230. doi:10.1016/j.celrep.2018.07.100

Zhu, F., Farnung, L., Kaasinen, E., Sahu, B., Yin, Y., Wei, B., et al. (2018). The Interaction Landscape between Transcription Factors and the Nucleosome. *Nature* 562 (7725), 76–81. doi:10.1038/s41586-018-0549-5

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Cell and Developmental Biology

ORIGINAL RESEARCH
published: 26 November 2021
doi: 10.3389/fcell.2021.767454

The Lung Elastin Matrix Undergoes Rapid Degradation Upon Adult Loss of *Hox5* Function

*Mu-Hang Li[1], Leilani M. Marty-Santos[2], Paul R. van Ginkel[2], Aubrey E. McDermott[2], Andrew J. Rasky[3], Nicholas W. Lukacs[3] and Deneen M. Wellik[2]**

[1] Genetics Training Program, University of Wisconsin-Madison, Madison, WI, United States, [2] Department of Cell and Regenerative Biology, University of Wisconsin School of Medicine and Public Health, Madison, WI, United States, [3] Department of Pathology, University of Michigan, Ann Arbor, MI, United States

Hox genes encode transcription factors that are critical for embryonic skeletal patterning and organogenesis. The *Hoxa5*, *Hoxb5,* and *Hoxc5* paralogs are expressed in the lung mesenchyme and function redundantly during embryonic lung development. Conditional loss-of-function of these genes during postnatal stages leads to severe defects in alveologenesis, specifically in the generation of the elastin network, and animals display bronchopulmonary dysplasia (BPD) or BPD-like phenotype. Here we show the surprising results that mesenchyme-specific loss of *Hox5* function at adult stages leads to rapid disruption of the mature elastin matrix, alveolar enlargement, and an emphysema-like phenotype. As the elastin matrix of the lung is considered highly stable, adult disruption of the matrix was not predicted. Just 2 weeks after deletion, adult *Hox5* mutant animals show significant increases in alveolar space and changes in pulmonary function, including reduced elastance and increased compliance. Examination of the extracellular matrix (ECM) of adult *Tbx4rtTA; TetOCre; Hox5a^{f}a^{f}bbcc* lungs demonstrates a disruption of the elastin network although the underlying fibronectin, interstitial collagen and basement membrane appear unaffected. An influx of macrophages and increased matrix metalloproteinase 12 (MMP12) are observed in the distal lung 3 days after *Hox5* deletion. In culture, fibroblasts from *Hox5* mutant lungs exhibit reduced adhesion. These findings establish a novel role for *Hox5* transcription factors as critical regulators of lung fibroblasts at adult homeostasis.

Keywords: *Hox* genes, lung homeostasis, extracellular matrix, distal lung fibroblasts, lung macrophages, brochopulmonary dysplasia, emphysema

OPEN ACCESS

Edited by:
Marie Kmita, Montreal Clinical Research Institute (IRCM), Canada

Reviewed by:
Rongbo Li, University of California, San Diego, United States
Stephen Gaunt, Independent Researcher, Cambridge, United Kingdom

****Correspondence:***
Deneen M. Wellik
wellik@wisc.edu

Specialty section:
This article was submitted to Morphogenesis and Patterning, a section of the journal Frontiers in Cell and Developmental Biology

Received: *30 August 2021*
Accepted: *11 October 2021*
Published: *26 November 2021*

Citation:
Li M-H, Marty-Santos LM, van Ginkel PR, McDermott AE, Rasky AJ, Lukacs NW and Wellik DM (2021) The Lung Elastin Matrix Undergoes Rapid Degradation Upon Adult Loss of Hox5 Function. Front. Cell Dev. Biol. 9:767454. doi: 10.3389/fcell.2021.767454

INTRODUCTION

Hox genes encode transcription factors that are well known for their role in patterning the anterior-posterior (AP) body axis during embryogenesis. All mammals have a total of 39 *Hox* genes located in four, tightly linked chromosomal clusters, subdivided into 13 related paralogous groups based on their expression and shared function. In addition to their roles in patterning the skeleton, *Hox* paralog groups also function redundantly in the proper formation of many organs, including the thymus, thyroid, lungs, pancreas, kidney and reproductive tract (Jeannotte et al., 1993; Aubin et al., 1997; Taylor et al., 1997; Manley and Capecchi, 1998; Wellik et al., 2002; Yallowitz et al., 2011). Multiple studies have demonstrated the functional redundancy exhibited by the members of

paralog groups and loss-of-function of multiple paralogous genes results in more severe phenotypes than loss of a single *Hox* gene (Chisaka and Capecchi, 1991; Horan et al., 1995; Chen and Capecchi, 1997, 1999; Manley and Capecchi, 1998; Wellik et al., 2002, p. 11; Wellik and Capecchi, 2003; McIntyre et al., 2007; Boucherat et al., 2013; Hrycaj et al., 2015; Larsen et al., 2015).

Developmentally, critical roles for *Hox5* genes have been demonstrated in skeleton patterning, central nervous system formation, and lung organogenesis (Tuggle et al., 1990; Jeannotte et al., 1993; Aubin et al., 1997; Mandeville et al., 2006; McIntyre et al., 2007; Hrycaj et al., 2015, 2018a). During lung development, *Hox5* genes (*Hoxa5*, *Hoxb5*, and *Hoxc5*) are exclusively expressed in the mesenchyme of the lung (Aubin et al., 1997; Boucherat et al., 2013; Hrycaj et al., 2015). *Hoxa5* single homozygous mutants (*Hox5 aaBBCC*) exhibit a reduction in *Ttf-1* and *Hnf-3* expression and defects in surfactant production. There is also a high rate of perinatal lethality associated with improper tracheal morphogenesis and occlusion of the proximal airways. *Hoxb5* and *Hoxc5* single-mutant mice exhibit no overt embryonic lung phenotypes and are viable, as are *Hoxb5/Hoxc5* double mutant animals (Boucherat et al., 2013). The extent of functional redundancy of all three *Hox5* alleles was demonstrated by generating *Hoxa5*; *Hoxb5*; *Hoxc5* triple mutant embryos *(Hox5 aabbcc)*. Lungs from these embryos undergo only a few early branches and newborn animals die with severely hypoplastic lungs due to the loss of Wnt2/2b signaling in the early lung mesoderm. Compound, 4-allele *Hox5* mutant (*Hox5 AabbCc*) lungs show no observable defect at embryonic stages, but display expanded, simplified alveoli at postnatal stages compared to controls (Hrycaj et al., 2015). Expression of all three *Hox5* genes in the lung decreases from mid to late embryogenesis, then peaks to its highest level at postnatal stages, but is maintained throughout adult life. Postnatal deletion of *Hoxa5* in the background of *Hoxb5/Hoxc5* nulls leads to BPD (Hrycaj et al., 2018a).

In this study, we demonstrate that *Hox5* function remains important at adult stages for proper lung homeostasis. When *Hoxa5* deletion is induced at adult stages, the elastin matrix is disrupted within days after deletion. The distal airways expand and pulmonary function tests demonstrate that mutant lungs become significantly more compliant and less elastic just 2-weeks after deletion. The matrix disruption appears to be specific to elastin as laminin, interstitial collagen and fibronectin scaffolds appear unaffected. Examination of the distal lung just 3 days post-deletion shows increased expression of matrix metalloproteinase 12 (MMP12), also known as neutrophil elastase, and an influx of F4/80 + and CD68 + macrophage populations. Similar to what we previously reported at postnatal stages, fibroblasts from the *Hox5* triple, adult conditional mutant lung exhibit reduced adhesion and decreased integrin α5 protein expression. Our results are consistent with a model in which induced loss of fibroblast cell adhesion leads to elastin matrix instability. This work demonstrates that lung maintenance requires continued *Hox5* function in lung fibroblasts. Our work provides insight into the pathophysiological process and putative targets for molecular and cellular therapies for lung diseases.

MATERIALS AND METHODS

Mice and Tissue Isolation

All mice used in this study have been previously reported (Zhang et al., 2013; Hrycaj et al., 2018a). Mice were treated with 2 mg/ml Doxycycline (DOT Sci., #DSD43020) (in water with 2.5 mg sucrose added per ml) at the age of 8 weeks for 3 days or 2 weeks. Mice were euthanized and perfused with phosphate buffered saline (PBS) via the right ventricle. Lungs were isolated, inflated and fixed as previously reported (Hrycaj et al., 2018b). The left lung lobes were vacuum embedded in paraffin; the right superior lung lobes were embedded in OCT (Fisher Sci., #23730571); the right middle and inferior lobes were digested for fibroblast isolation; the right accessory lobes were used for protein or RNA extraction. All experiments were performed following protocols approved by the Institutional Animal Care and Use Committee (IACUC) guidelines at the University of Michigan or the University of Wisconsin-Madison.

Lung Whole Mount Imaging

Adult left lung lobes were fixed in 4% paraformaldehyde (PFA) in PBS overnight at 4°C then transferred to absolute MeOH through MeOH/PBS dilution series: 25, 50, 75 and 100% MeOH. Tissues then were incubated in Dent's bleach (MeOH:DMSO:30% H_2O_2 = 4:1:1) for 2 h at room temperature to remove any coloration, and transferred to absolute MeOH for imaging on a Leica MZ125 dissecting microscope. PFA (Sigma-Aldrich, #P6148), MeOH (Sigma-Aldrich, #179337), DMSO (Sigma-Aldrich, #D2650), H_2O_2 (Sigma-Aldrich, #323381).

Chord Length Analyses

Hematoxylin (Fisher Sci., #SH30500D) and Eosin (Fisher Sci., #SE23) staining was performed on 7 μm paraffin lung sections. Images were captured on a Nikon Ds-Fi3 camera. Mean alveolar chord length (MACL) measurements were taken using the grid function on ImageJ 2.0, as previously described (Sajjan et al., 2009).

Pulmonary Function Tests

Analyses were performed as previously described (Hrycaj et al., 2018b). Briefly, 10 week old control and *Tbx4rtTA; TetOCre; Hox5a^f a^f bbcc* mice (treated with Dox from 8 to 10 weeks) were anesthetized prior to the insertion of a tracheal tube. Mechanical breathing measurements were performed at baseline to examine changes in lung function.

Immunohistochemistry/ Immunofluorescence

Paraffin sections were deparaffinized in xylenes and rehydrated in an ethanol series prior to antigen retrieval in 10 mM Sodium Citrate buffer. Cryosections were washed in PBS to remove excess OCT. Sections were blocked in 5% Normal Donkey Serum (Sigma-Aldrich, #566460) and incubated in primary antibodies in 4^{o}C overnight. A complete list of primary antibodies and the dilution used is provided in **Supplementary Table 1**. Sections were rinsed, incubated in secondary antibodies at room

temperature for 2 h, with a 10 min 1 μg/ml DAPI (Thermo Sci., #62248) incubation at room temperature and mounted using ProLong Gold mountant (Fisher Sci., #P36930). Images were captured on a Nikon Eclipse Ti-U camera or on a Keyence BZ-X810 fluorescence microscope.

RNA Extraction and Quantitative Reverse Transcription PCR

Total RNA was extracted from the right accessory lobe of wildtype and mutant mice using the RNeasy Mini Kit (Qiagen, #74104) and dissolved in 32 μl of DNase/RNase-free deionized water. cDNA was synthesized from 1 μg RNA using the iScript Reverse Transcription Supermix (Bio-Rad, #1708841). Quantitative real-time PCR was conducted using 2 × SYBR Green qPCR Master Mix (Fisher Sci., #4309155) on a StepOnePlus^{TM} Real-Time PCR System Machine (Fisher Sci., #4376600). Threshold cycles (Ct) in target gene expression were calculated and compared to Ct values of house-keeping gene β-*actin*. Primers for quantitative reverse transcription PCR (RT-qPCR) are listed in **Supplementary Table 2**.

Western Blot Analysis

Lung right accessory lobes were lysed in radioimmunoprecipitation assay buffer (50 mM Tris·HCl, pH 7.2, 150 mM NaCl, 0.1% Triton X-100, 1% sodium deoxycholate, 5 mM EDTA) containing Complete Mini Protease Inhibitor Mixture (Sigma-Aldrich, #11873580001), and extracts were cleared by centrifugation at 20,000× *g* for 30 min at 4°C. Total protein content was assessed using the Pierce BCA protein assay kit (Thermo Sci. #23227) and analyzed by SDS/PAGE after boiling in Laemmli sample buffer (Bio-Rad, #1610737). Proteins were transferred to low-fluorescence polyvinylidene fluoride (Cytiva, #10600022), blocked in 3% BSA with sodium azide, and probed with primary antibodies as indicated in **Supplementary Table 1**. Fluorescence signals were detected on an Azure imaging system or a LI-COR Odyssey Fc imaging system. Western blot densitometry analysis was quantified using ImageJ 2.0.

Lung Fibroblast Isolation and Primary Culture

Following perfusion with PBS and lung isolation, the trachea and proximal airway were removed, and the right middle and right inferior lung lobes were minced and digested in 2 mg/ml Collagenase I (Gibco, #17100017) and 3 mg/ml Dispase (Gibco, #17105041) for 2 h at 37°C. This cell suspension was incubated with red blood cell lysis buffer on ice and filtered through a 100 μm nylon cell strainer (Fisherbrand, #22363549). Cells were then washed twice in fresh DMEM/F12 media (Gibco, #11320033). Cells were plated in 6-well tissue culture plates using DMEM/F12 media supplemented with 10% FBS (Gibco, #10437010) and 1% 10,000 U/ml Penicillin-Streptomycin (Gibco, #15140122). Media was changed every third day.

Adhesion Assay

Assays were performed as previously described (Hrycaj et al., 2018a). Non-tissue culture treated polystyrene 96-well flat bottom microtiter plates (Denville, # T1097) were coated and incubated with bovine plasma fibronectin (Sigma-Aldrich, #F1141) at 20 μg/ml for 1 h at 37°C. Plates were then blocked with 100 μl/well of 1% BSA (Sigma-Aldrich, #A7906) in serum-free DMEM/F12 (Gibco, #11320033) for 30 min at 37°C. Fibroblasts were seeded at 10,000 cells per individual well. Plates were centrifuged (top side up) at 10 × g for 5 min to reduce the variability inherent in the settling of cells onto the plate surface and were incubated for 1 h at 37°C with 5% CO_2. Non-adherent cells were removed by centrifugation (top side down) at 48 × g for 5 min. Adherent cells remaining on the plate were fixed and stained with 1% formaldehyde (Fisher Sci., #BP228-100), 0.5% crystal violet (Sigma-Aldrich, #C6158), 20% MeOH followed by PBS washes. Individual wells were imaged on a Leica MZ125 dissecting microscope and manually quantified using ImageJ 2.0 cell counter function.

Statistical Analyses

GraphPad Prism software (version 8.4.3 for macOS) was used to perform unpaired Student's *t*-test analysis, with *P*-values less than 0.05 considered significant and *P*-values greater than or equal to 0.05 considered not significant.

RESULTS

Conditional *Hox5* Triple Mutant Adult Mice Exhibit Expanded Distal Airspaces and Altered Lung Mechanics

To investigate whether *Hox5* plays a role in the adult lung, we used our previously described *Hoxa5* conditional allele combined with the lung mesenchyme-specific *Tbx4rtTA*; *TetOCre* driver and enacted deletion with Doxycycline (Dox) beginning at 8 weeks of age in the presence of *Hoxb5/Hoxc5* loss-of-function (*Hoxa5$^{LoxP/LoxP}$; Hoxb5$^{-/-}$; Hoxc5$^{-/-}$; Tbx4rtTA; TetOCre* +) (referred to as *Hox5* adult conditional mutants or as *Tbx4rtTA; TetOCre; Hox5a^{f} a^{f} bbcc* throughout) (Hrycaj et al., 2018a). Examination of the lungs 2 weeks after the initiation of *Hox5* deletion (at 10 weeks of age) resulted in significantly expanded distal airspaces that are clearly visible in the whole mount lung tissue after inflation (**Supplementary Figure 1**). This is supported by histological examination of sections and morphological measurements of mean alveolar chord length (MACL). *Tbx4rtTA; TetOCre; Hox5a^{f} a^{f} bbcc* Dox-treated mutant animals demonstrate a ∼45% increase in MACL 2-weeks post-deletion and thinned septal walls compared to wild-type animals (**Figures 1A,A',C,C',D**). Additionally, adult mice carrying only the *Hox5bbcc* alleles in the absence of Cre only developed a moderate phenotype (∼20% increase in MACL) (**Figures 1B,B',D**), demonstrating that conditional, loss of *Hoxa5* function contributes directly to the adult phenotype.

Measurements of pulmonary function revealed changes in the *Hox5* adult 10-week-old conditional mutants (Dox treated from 8 to 10 weeks) consistent with an emphysema-like phenotype. Lung compliance was increased as measured by increased chord compliance (Cchord), compliance at zero pressure (Cp0)

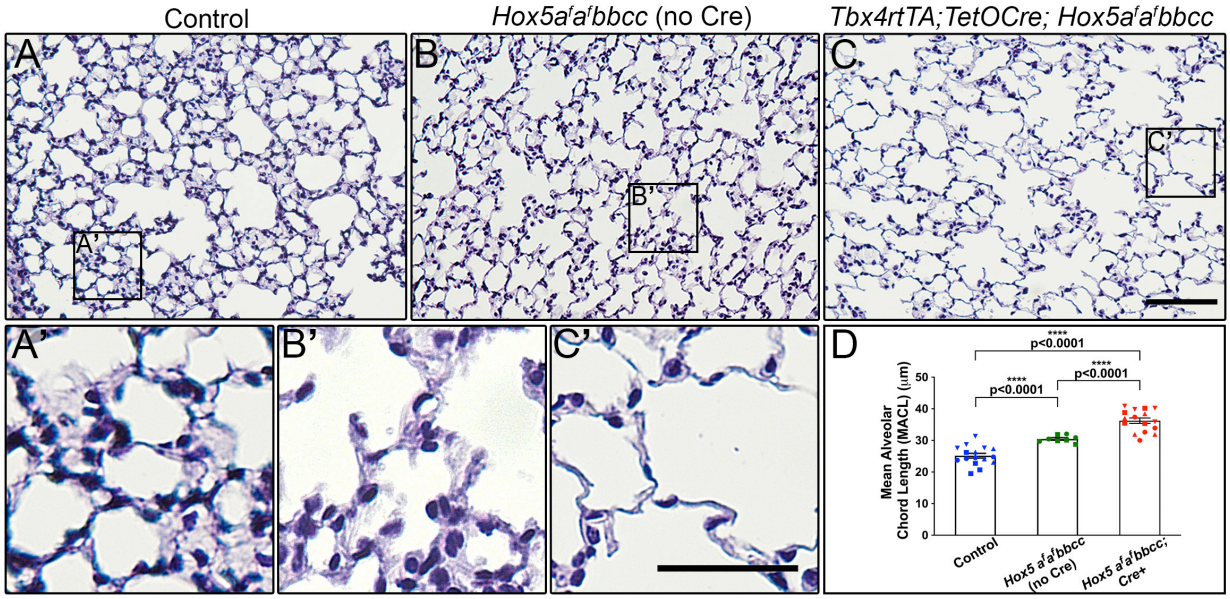


FIGURE 1 | Deletion of *Hox5* function at adult stages leads to distal lung airway expansion. H&E sections through 10-week-old wild-type animals **(A,A')**, *Hox5a^{f}a^{f}bbcc* mutants (no Cre) **(B,B')** and *Hoxa5$^{LoxP/LoxP}$; Hoxb5$^{-/-}$; Hoxc5$^{-/-}$; Tbx4rtTA; TetOCre* + mutants (*Hox5a^{f}a^{f}bbcc; Cre* +, **C,C')**, treated with Doxycycline from 8 to 10 weeks of age. Measured mean alveolar chord length (MACL) values **(D)**. *Hox5a^{f}a^{f}bbcc* double mutant animals (*Hoxa5* floxed alleles with no Cre) exhibit enlarged, simplified alveoli with a ∼20% increase in alveolar chord length compared to control lungs. The distal airway expansion phenotype with adult, conditional *Hox5* triple mutants increases by ∼45% compared to control lungs. Each shape represents an individual animal. Control, *Hox5AABBCC*. Scale bars: 50 μm **(C)**; 25 μm **(C')**. *P*-values and statistical significance (****$P < 0.0001$) were determined by an unpaired Student's *t*-test.

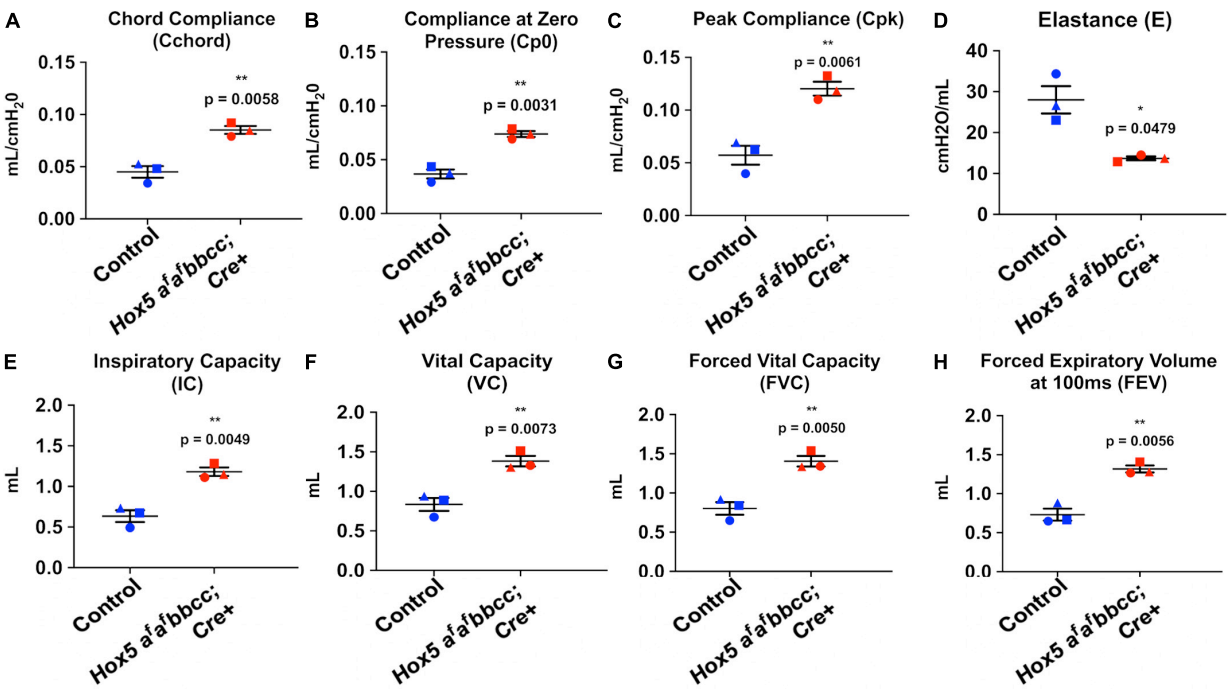


FIGURE 2 | Adult conditional *Hox5* mutant lungs exhibit abnormal pulmonary function. Lung compliance, elastance and volume were measured in mice by fast flow maneuvers using orotracheal intubation and tracheostomy. Increased lung compliance of Dox-treated (8–10 weeks) *Hox5* conditional mutants is indicated by significant increases in chord compliance (Cchord) **(A)**, compliance at zero pressure (Cp0) **(B)**, and peak compliance (Cpk) **(C)**. *Hox5* adult conditional mutants also exhibit a significant decrease in elastance (E) **(D)**, and a significant increased lung volume indicated in inspiratory capacity (IC) **(E)**, vital capacity (VC) **(F)**, forced vital capacity (FVC) **(G)**, and forced expiratory volume (FEV) **(H)**. Each shape represents an individual animal. *P*-values and statistical significance (*$0.01 \leq P < 0.05$; **$0.001 \leq P < 0.01$) were determined by an unpaired Student's *t*-test.

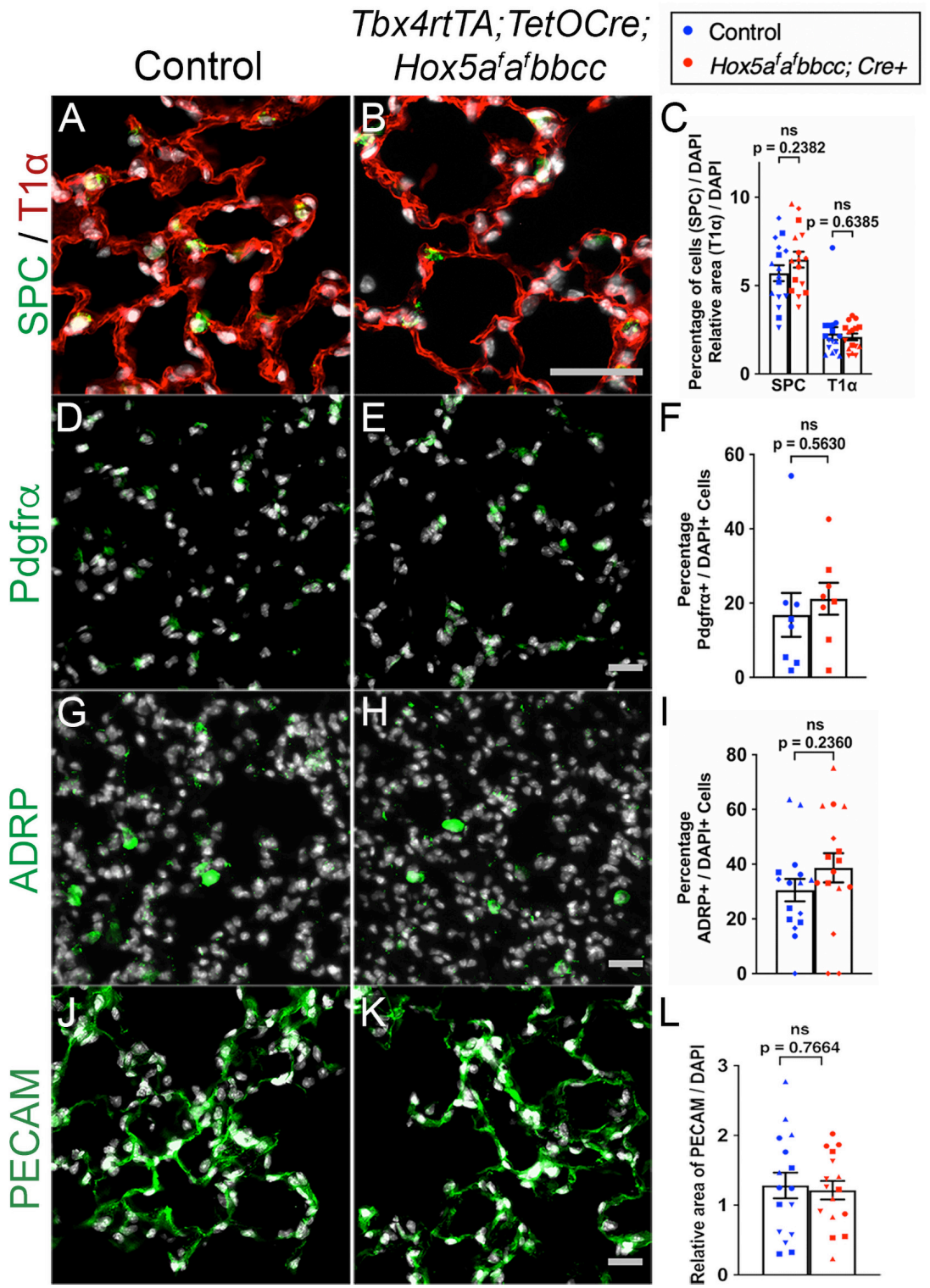


FIGURE 3 | Distal lung cell types are present and appear similar in controls and *Hox5* conditional mutants. Lung paraffin sections **(A,B,G,H,J,K)** and cryosections **(D,E)** of control and *Hox5* adult, conditional mutant mice show similar expression of SPC (green, AECII cells), T1α (red, AECI cells), PDGFRα (green, fibroblasts), ADRP (green, lipofibroblasts) and PECAM (green, endothelial cells); DAPI in gray. Quantification of pixel intensity of T1α **(C)**, and PECAM **(L)** were normalized to pixel intensity of DAPI per field image. Quantifications of SPC **(C)**, PDGFRα **(F)** and ADRP **(I)** cell numbers were normalized to DAPI-positive cell numbers in each panel quantified. Each shape represents an individual animal (ns, not significant). Scale bars: 50 μm. *P*-values were determined by an unpaired Student's *t*-test.

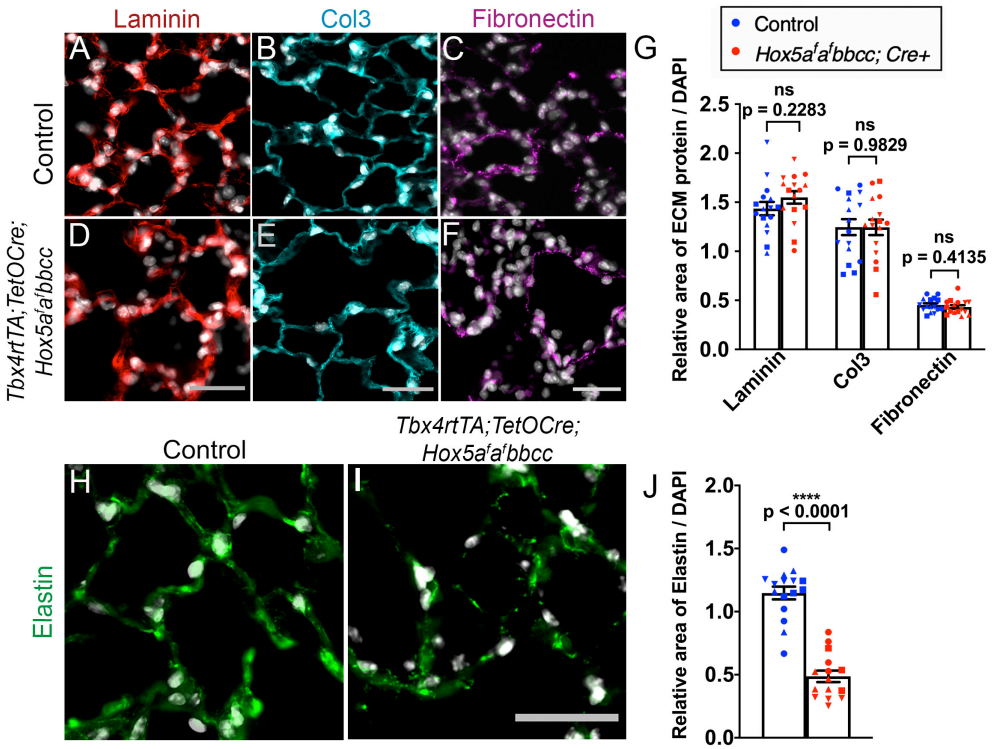


FIGURE 4 | The elastin network is disrupted in the *Hox5* adult, conditional mutant lungs. Lung paraffin sections from control and *Hox5* adult, conditional mutant mice stained with laminin (red, **A,D**), collagen3 (cyan, **B,E**), fibronectin (magenta, **C,F**) and elastin (green, **H,I**) with DAPI in gray in all. Quantifications of pixel intensity of laminin, collagen3, and fibronectin was normalized to pixel intensity of DAPI, and show no differences between controls and mutants **(G)**. The elastin network is disrupted, and the total elastin pixel intensity normalized to pixel intensity of DAPI per field image is significantly decreased in mutant lungs compared to control lungs **(J)**. Each shape represents an individual animal (ns, not significant). Scale bars: 25 μm. *P*-values and statistical significance (****$P < 0.0001$) were determined by an unpaired Student's *t*-test.

and peak compliance (Cpk), with changes accompanied by a significant decrease in elastance (**Figures 2A–D**). Concomitantly, *Hox5* adult conditional mutants exhibit increased lung volumes, demonstrated by significant increases in inspiratory capacity (IC), vital capacity (VC), forced vital capacity (FVC), and forced expiratory volume (FEV) (**Figures 2E–H**).

The Elastin Network Is Disrupted in *Hox5* Mutant Lungs, but Cell Types and Other Extracellular Matrix Components Appear Normal

To examine whether changes occurred in distal lung cell composition, we examined epithelial, fibroblast, and endothelial cell populations in the distal lung after mesenchymal deletion of *Hoxa5*. We observed no significant changes in the morphology or distribution in T1α + alveolar epithelial type I cells (AECI) or Surfactant Protein C (SPC) + alveolar epithelial type II cells (AECII) (**Figures 3A–C**). We also observed no changes in the number or relative distribution of Platelet-derived growth factor receptor alpha (PDGFRα) + fibroblasts (**Figures 3D–F**), Adipocyte differentiation-related protein (ADRP) + lipofibroblasts (**Figures 3G–I**), or Platelet endothelial cell adhesion molecule (PECAM) + endothelial cells (**Figures 3J,K**) in the distal lung of the *Hox5* mutants. Additionally, no changes in mRNA levels of these cell markers was observed using RT-qPCR (**Supplementary Figures 2A–E**).

We analyzed several key ECM components in *Hox5* adult, conditional mutant distal lungs administered Dox from 8 to 10 weeks. Immunohistochemical immunofluorescence (IHC-IF) analyses with subsequent ImageJ quantification of pixel intensity demonstrated ECM phenotypes strikingly similar to those previously described with postnatal deletion (Hrycaj et al., 2018a). We observed no changes in the basement membrane component laminin, interstitial collagen 3 or fibronectin in the *Hox5* adult conditional mutant lungs compared to wild-type lungs (**Figures 4A–G**). However, we observed a substantial decrease in elastin staining (∼42% of wild-type levels) and disruption of the integrity of the elastin network in *Hox5* adult conditional mutant lungs (**Figures 4H–J**). This phenotype was particularly surprising as it indicates rapid loss and/or destruction of the elastin matrix that was fully established normally at 8-weeks of age and is considered very stable (Shapiro et al., 1991). RT-qPCR shows no differences of elastin mRNA level between controls and mutants, which indicates the changes in elastin structural integrity are likely an indirect consequence of other disruptions in the distal lung (**Supplementary Figure 2F**).

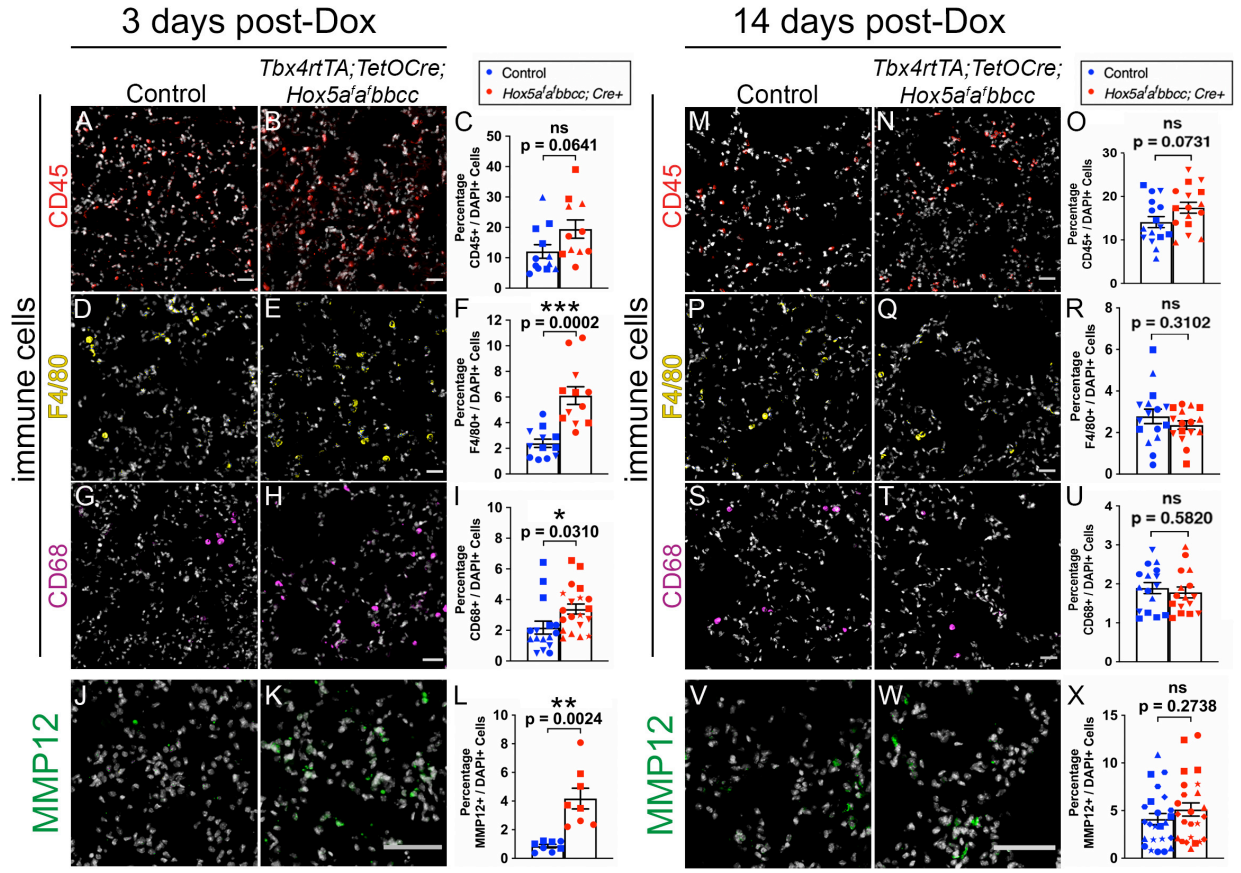


FIGURE 5 | Increased inflammatory response of macrophages and MMP12 expression levels 3 days after *Hox5* deletion. Cryosections of control and mutant mice lungs (3 days post-Dox treatment initiated at 8 weeks of age) were stained with CD45 (red, **A,B**), F4/80 (yellow, **D,E**), CD68 (magenta, **G,H**), and MMP12 (green, **J,K**) with DAPI in gray. Quantifications of CD45- **(C)**, F4/80- **(F),** CD68- **(I)**, and MMP12- **(L)** positive cell numbers were normalized to DAPI-positive cell numbers and demonstrate significantly increased infiltration of macrophages and MMP12 staining levels in *Hox5* mutant distal lungs. No increase of lung leukocyte, macrophage populations or MMP12 expression levels was observed 14 days post-Dox treatment **(M–X)**. Each shape represents an individual animal (ns, not significant). Scale bars: 50 μm. *P*-values and statistical significance (*$0.01 \leq P < 0.05$; **$0.001 \leq P < 0.01$; ***$0.0001 \leq P < 0.001$) were determined by an unpaired Student's *t*-test.

Increased Inflammatory Response and Elastase Expression in the *Hox5* Triple Conditional Mutant Lung 3 Days After Dox Induction

Previous studies have demonstrated that leukocytes and macrophages secrete matrix metalloproteinases (MMPs), including neutrophil elastase (MMP12), that are capable of degrading the extracellular matrix during lung development and homeostasis (Gibbs et al., 1999; Grumelli et al., 2004; Zeng et al., 2014; Morrell et al., 2020). Further, in a study using *Hox5* compound null alleles (*Hoxa5*$^{+/-}$; *Hoxb5*$^{-/-}$; *Hoxc5*$^{+/-}$), mice present with an increased Th2 cells response and exacerbated lung tissue pathology in asthma models (Ptaschinski et al., 2017). In order to determine whether an increase in either leukocytes or macrophages could account for the alveolar enlargement and the decrease in elastin in our *Hox5* conditional mutant lungs, we performed IHC-IF and quantifications for Cluster of differentiation 45 (CD45, also known as the leukocyte antigen) (**Figures 5A–C** and **Supplementary Figures 3A,B**), the macrophage marker F4/80 (**Figures 5D–F** and **Supplementary Figures 3C,D**), and Cluster of differentiation 68 (CD68), which labels tissue-resident macrophages (**Figures 5G–I** and **Supplementary Figures 3E,F**) immediately following *Hox5* deletion (3 days post-Dox treatment initiated at 8 weeks of age). We observed a significant increase in F4/80+ and CD68+ macrophages number in *Hox5* mutant distal lungs (**Figures 5F,I**), consistent with an increased inflammatory response following *Hox5* mesenchymal deletion.

Consistent with the influx of macrophages, we observed a significant increase of MMP12, a major elastase reported to be secreted by activated lung macrophages (McGarry Houghton, 2015), in *Hox5* conditional mutant lungs compared to controls 3 days after Doxycycline induction (**Figures 5J–L**). MMPs are shown to target both ECM components and adhesion receptors to alter cell behaviors in lung and other cancer cells (Yu and Stamenkovic, 2000; Rolli et al., 2003; Zeng et al., 2014). A previous study also suggests enhanced production of MMPs in *Hoxa5*$^{-/-}$ juvenile mice lungs (Mandeville et al., 2006). However, by 14 days after the initiation of deletion, there was

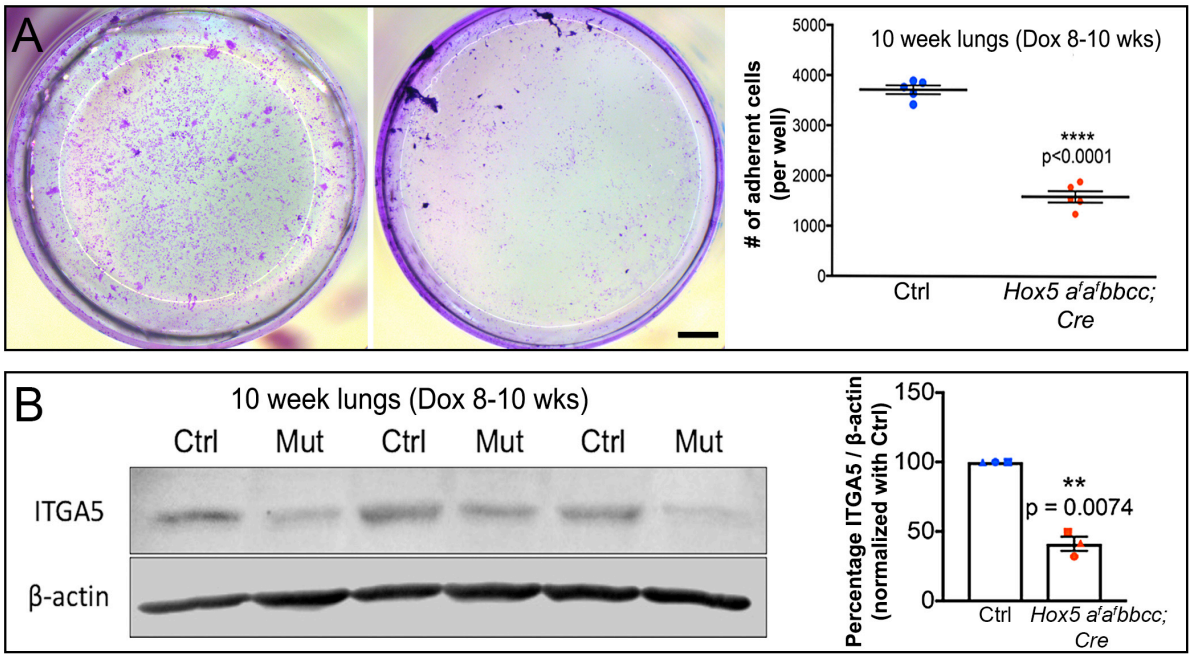


FIGURE 6 | *Hox5* adult, mutant lung fibroblasts are less adherent in culture compared to controls. Lung fibroblasts were isolated, seeded on fibronectin-coated plates and incubated for 1 h prior to low-gravity centrifugation then fixation. Cells were stained with crystal violet, imaged, then counted. *Hox5* adult, conditional mutant stromal cells are significantly less adherent in this assay **(A)**. Western blotting of protein from right accessory lobes from *Hox5* adult, conditional mutant lungs show ∼50% lower ITGA5 levels compared to control lung tissues **(B)**. Each shape represents an individual animal. Scale bar: 10 mm. *P*-values and statistical significance (**$0.001 \leq P < 0.01$; ****$P < 0.0001$) were determined by an unpaired Student's *t*-test.

no observable differences either in the number or location of leukocytes, macrophages (**Figures 5M–U**) or in the expression of MMP12 (**Figures 5V–X** and **Supplementary Figure 4**) between *Hox5* mutants and control lungs.

Together, these data are consistent with an acute inflammatory response in *Hox5 adult* mutant distal lungs after conditional *Hox5* deletion in the lung fibroblasts.

Hox5 Conditional Triple Mutant Fibroblasts Display Reduced Adhesion to Fibronectin *in vitro*

Our previous work examining lung fibroblasts from postnatal *Hox5* conditional mutants demonstrated that these cells exhibit reduced adhesion to fibronectin compared to controls (Hrycaj et al., 2018a). To examine whether *Hox5* plays a similar role in regulating adhesion of lung fibroblasts in response to deletion at adult stages (8–10 weeks of Dox-induced deletion), we cultured primary lung cells harvested from adult control and *Hox5* conditional mutant animals on fibronectin-coated plates. *In vitro* adhesion assays demonstrated that *Hox5* adult mutant fibroblasts exhibit a ∼60% decrease in their ability to adhere to fibronectin (**Figure 6A**). Consistent with what we reported after postnatal loss of *Hox5* function, there is a dramatic decrease in the protein expression of integrin α5, an important component of fibroblast adhesion, in mutants compared to controls (**Figure 6B**).

DISCUSSION

Although *Hox* genes were initially reported to be embryonic patterning factors, an increasing body of work has demonstrated that these genes play a role in the homeostasis, maintenance and repair in postnatal and adult tissues (Pineault et al., 2015; Rux and Wellik, 2017; Rux et al., 2017; Nova-Lampeti et al., 2018; Bradaschia-Correa et al., 2019; Song et al., 2020). Prior studies from our laboratory have demonstrated a requirement for *Hox5* in lung fibroblast adhesion and establishment of the elastin network during postnatal alveologenesis (Hrycaj et al., 2018a). In our present work, we extend these findings and show that loss of *Hox5* in the adult lung mesenchyme leads to a rapid expansion of the distal airspaces, apparently resulting from the degradation of the previously established elastin matrix responsible for maintaining alveolar structures in the distal lung. While other extracellular matrix components and distal lung cell types do not appear affected, we see a loss of adhesion in the fibroblasts of the distal lung and report an increase in macrophages and elastase (MMP12), which might contribute to the rapid change in alveolar structures.

The lung ECM, composed primarily of elastin, collagens, and proteoglycans, determines its mechanical properties. The elastin network specifically provides the elastic recoil necessary for exhalation and is reported to be stable, with a reported ^{14}C half-life of ∼74 years or longer in humans (Shapiro et al., 1991). However, perturbations of the elastin network of the adult lung can cause destruction of the alveolar walls that result in emphysema, and mutations in components of the elastogenesis pathway, such as tropoelastin, fibrillins and fibulins are also associated with a predisposition to emphysema (Rodriguez-Revenga et al., 2004; Urban et al., 2005; Hersh et al., 2006; Hucthagowder et al., 2006). Additionally, elastin quantity is a marker of susceptibility to emphysema when the lung is challenged, as mice heterozygous for the elastin gene ($Eln^{+/-}$) are more prone to develop emphysema after prolonged exposure to cigarette smoke (Shifren et al., 2007). This predisposition to emphysema was proposed to result from a lower availability of cross-linked elastin leading to degradation of the larger elastin fibers, and changes in the adult ECM leading to an increase in collagen that may ultimately prevent matrix remodeling by restricting cell movement (Shifren et al., 2007). In the adult *Hox5* conditional triple mutant lung, there is a rapid destruction of an already established elastin network. This could be mediated by an observed increase in hematopoietic-derived cell numbers, which leads to increased activated elastase (MMP12) in response to exposure of the matrix when fibroblasts lose adherence. The degradation of elastin would lead to an expansion of the distal airspaces reminiscent of that seen in emphysema (Grumelli et al., 2004; Shifren et al., 2007).

We previously showed that postnatal loss of *Hox5* in the lung mesenchyme leads to a decrease in the protein levels of the integrin α5β1 heterodimer (Hrycaj et al., 2018a), a phenotype which is recapitulated in the adult *Hox5* conditional triple mutant. Integrin α5β1 mediates the binding of lung fibroblasts to fibronectin (Watt and Hodivala, 1994; Liu et al., 2010, p. 1; Epstein Shochet et al., 2017), and is in turn regulated by the formation of focal adhesions (Cai et al., 2009). Overexpression of HoxA5 in EOMA cell lines has been shown to stabilize focal adhesions by increasing Akt expression (Arderiu et al., 2007), and other *Hox* genes also play a role in the regulation of adhesions *in vitro* (Jones et al., 1992). Intriguingly, increased MMP12 has been associated with a reduction of focal adhesions in patients with anti-alpha1 trypsin deficiency, resulting in an enhanced severity of emphysema (Baraldo et al., 2015). Overall, these data are consistent with an important role for *Hox5* genes in the regulation of lung fibroblast adhesion. Elucidation of the factors that regulate maintenance of the adult lung elastin matrix, including a better understanding of the role played by fibroblasts and by *Hox5* genes, will ultimately lead to better treatments for lung disease diseases related to lung mesenchymal behavior such as BPD, emphysema, and idiopathic pulmonary fibrosis (IPF).

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

The animal study was reviewed and approved by Institutional Animal Care and Use Committee (IACUC) at the University of Michigan or Institutional Animal Care and Use Committee (IACUC) at the University of Wisconsin-Madison.

AUTHOR CONTRIBUTIONS

LM-S, DW, and M-HL designed the research. LM-S, M-HL, PvG, AR, and NL performed the research. M-HL, LM-S, AM, PvG, and DW analyzed the data. M-HL, LM-S, and DW wrote the manuscript. M-HL, PvG, AM, and DW revised the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by the National Heart, Lung, and Blood Institute (NHLBI) R01-HL137364 to DW and F32-HL140969 to LM-S.

ACKNOWLEDGMENTS

The authors would like to acknowledge Emma Snyder-White and Xiaohu Wang for their assistance with data analysis, and Alex Hurley for his assistance with mouse genotyping.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2021.767454/full#supplementary-material

Supplementary Figure 1 | Whole mount lung picture of *Hox5* conditional triple mutant lungs at 10 weeks of age (after Dox deletion from 8 to 10 weeks) show significantly expanded distal airspaces compared to controls (arrowheads, **A,B**). Scale bar: 100 μm.

Supplementary Figure 2 | mRNA levels of *Sftpc* **(A)**, *T1α* **(B)**, *Pdgfrα* **(C)**, *Adrp* **(D)**, *Pecam1* **(E)** and *Elastin* **(F)** were normalized to β-*actin* and show no significant differences in control and *Hox5* conditional triple mutant lungs at 10 weeks of age (after Dox deletion from 8 to 10 weeks). Each shape represents an individual animal (ns, not significant). *P*-values were determined by an unpaired Student's *t*-test.

Supplementary Figure 3 | Low magnification depicting IHC-IF staining of CD45 (red, **A,B**), F4/80 (yellow, **C,D**) and CD68 (magenta, **E,F**) in 3 days post-Dox control and *Hox5* adult conditional mutant lungs. Scale bar: 1 mm.

Supplementary Figure 4 | Protein levels of MMP12 appear low overall, but unchanged in 3-day **(A)** or 14-day **(B)** post-Dox treatment mutant lungs compared to controls demonstrated by western blots. The abundance of MMP12 protein was normalized to the total amount of protein in each lane. Each shape represents an individual animal (ns, not significant). *P*-values were determined by an unpaired Student's *t*-test.

REFERENCES

Arderiu, G., Cuevas, I., Chen, A., Carrio, M., East, L., and Boudreau, N. J. (2007). HoxA5 stabilizes adherens junctions via increased Akt1. *Cell Adh. Migr.* 1, 185–195.

Aubin, J., Lemieux, M., Tremblay, M., Bérard, J., and Jeannotte, L. (1997). Early postnatal lethality in Hoxa-5 mutant mice is attributable to respiratory tract defects. *Dev. Biol.* 192, 432–445. doi: 10.1006/dbio.1997.8746

Baraldo, S., Turato, G., Lunardi, F., Bazzan, E., Schiavon, M., Ferrarotti, I., et al. (2015). Immune activation in α1-antitrypsin-deficiency emphysema. beyond the protease-antiprotease paradigm. *Am. J. Respir. Crit. Care Med.* 191, 402–409. doi: 10.1164/rccm.201403-0529OC

Boucherat, O., Montaron, S., Bérubé-Simard, F.-A., Aubin, J., Philippidou, P., Wellik, D. M., et al. (2013). Partial functional redundancy between Hoxa5 and Hoxb5 paralog genes during lung morphogenesis. *Am. J. Physiol. Lung Cell. Mol. Physiol.* 304, L817–L830. doi: 10.1152/ajplung.00006.2013

Bradaschia-Correa, V., Leclerc, K., Josephson, A. M., Lee, S., Palma, L., Litwa, H. P., et al. (2019). Hox gene expression determines cell fate of adult periosteal stem/progenitor cells. *Sci. Rep.* 9:5043. doi: 10.1038/s41598-019-41639-7

Cai, W.-J., Li, M. B., Wu, X., Wu, S., Zhu, W., Chen, D., et al. (2009). Activation of the integrins alpha 5beta 1 and alpha v beta 3 and focal adhesion kinase (FAK) during arteriogenesis. *Mol. Cell. Biochem.* 322, 161–169. doi: 10.1007/s11010-008-9953-8

Chen, F., and Capecchi, M. R. (1997). Targeted mutations in Hoxa-9 and Hoxb-9 reveal synergistic interactions. *Dev. Biol.* 181, 186–196. doi: 10.1006/dbio.1996.8440

Chen, F., and Capecchi, M. R. (1999). Paralogous mouse Hox genes, Hoxa9, Hoxb9, and Hoxd9, function together to control development of the mammary gland in response to pregnancy. *Proc. Natl. Acad. Sci. U.S.A.* 96, 541–546. doi: 10.1073/pnas.96.2.541

Chisaka, O., and Capecchi, M. R. (1991). Regionally restricted developmental defects resulting from targeted disruption of the mouse homeobox gene hox-1.5. *Nature* 350, 473–479. doi: 10.1038/350473a0

Epstein Shochet, G., Brook, E., Israeli-Shani, L., Edelstein, E., and Shitrit, D. (2017). Fibroblast paracrine TNF-α signaling elevates integrin A5 expression in idiopathic pulmonary fibrosis (IPF). *Respir. Res.* 18:122. doi: 10.1186/s12931-017-0606-x

Gibbs, D. F., Warner, R. L., Weiss, S. J., Johnson, K. J., and Varani, J. (1999). Characterization of matrix metalloproteinases produced by rat alveolar macrophages. *Am. J. Respir. Cell Mol. Biol.* 20, 1136–1144. doi: 10.1165/ajrcmb.20.6.3483

Grumelli, S., Corry, D. B., Song, L.-Z., Song, L., Green, L., Huh, J., et al. (2004). An immune basis for lung parenchymal destruction in chronic obstructive pulmonary disease and emphysema. *PLoS Med.* 1:e8. doi: 10.1371/journal.pmed.0010008

Hersh, C. P., DeMeo, D. L., Lazarus, R., Celedón, J. C., Raby, B. A., Benditt, J. O., et al. (2006). Genetic association analysis of functional impairment in chronic obstructive pulmonary disease. *Am. J. Respir. Crit. Care Med.* 173, 977–984. doi: 10.1164/rccm.200509-1452OC

Horan, G. S., Kovàcs, E. N., Behringer, R. R., and Featherstone, M. S. (1995). Mutations in paralogous Hox genes result in overlapping homeotic transformations of the axial skeleton: evidence for unique and redundant function. *Dev. Biol.* 169, 359–372. doi: 10.1006/dbio.1995.1150

Hrycaj, S. M., Dye, B. R., Baker, N. C., Larsen, B. M., Burke, A. C., Spence, J. R., et al. (2015). Hox5 genes regulate the Wnt2/2b-Bmp4-signaling axis during lung development. *Cell Rep.* 12, 903–912. doi: 10.1016/j.celrep.2015.07.020

Hrycaj, S. M., Marty-Santos, L., Cebrian, C., Rasky, A. J., Ptaschinski, C., Lukacs, N. W., et al. (2018a). Hox5 genes direct elastin network formation during alveologenesis by regulating myofibroblast adhesion. *Proc. Natl. Acad. Sci. U.S.A.* 115, E10605–E10614. doi: 10.1073/pnas.1807067115

Hrycaj, S. M., Marty-Santos, L., Rasky, A. J., Lukacs, N. W., and Wellik, D. M. (2018b). Loss of Hox5 function results in myofibroblast mislocalization and distal lung matrix defects during postnatal development. *Sci. China Life Sci.* 61, 1030–1038. doi: 10.1007/s11427-017-9290-1

Hucthagowder, V., Sausgruber, N., Kim, K. H., Angle, B., Marmorstein, L. Y., and Urban, Z. (2006). Fibulin-4: a novel gene for an autosomal recessive cutis laxa syndrome. *Am. J. Hum. Genet.* 78, 1075–1080. doi: 10.1086/504304

Jeannotte, L., Lemieux, M., Charron, J., Poirier, F., and Robertson, E. J. (1993). Specification of axial identity in the mouse: role of the Hoxa-5 (Hox1.3) gene. *Genes Dev.* 7, 2085–2096. doi: 10.1101/gad.7.11.2085

Jones, F. S., Prediger, E. A., Bittner, D. A., De Robertis, E. M., and Edelman, G. M. (1992). Cell adhesion molecules as targets for Hox genes: neural cell adhesion molecule promoter activity is modulated by cotransfection with Hox-2.5 and -2.4. *Proc. Natl. Acad. Sci. U.S.A.* 89, 2086–2090. doi: 10.1073/pnas.89.6.2086

Larsen, B. M., Hrycaj, S. M., Newman, M., Li, Y., and Wellik, D. M. (2015). Mesenchymal Hox6 function is required for mouse pancreatic endocrine cell differentiation. *Development* 142, 3859–3868. doi: 10.1242/dev.126888

Liu, S., Xu, S., Blumbach, K., Eastwood, M., Denton, C. P., Eckes, B., et al. (2010). Expression of integrin beta1 by fibroblasts is required for tissue repair in vivo. *J. Cell Sci.* 123, 3674–3682. doi: 10.1242/jcs.070672

Mandeville, I., Aubin, J., LeBlanc, M., Lalancette-Hébert, M., Janelle, M.-F., Tremblay, G. M., et al. (2006). Impact of the Loss of Hoxa5 Function on Lung Alveogenesis. *Am. J. Pathol.* 169, 1312–1327. doi: 10.2353/ajpath.2006.051333

Manley, N. R., and Capecchi, M. R. (1998). Hox group 3 paralogs regulate the development and migration of the thymus, thyroid, and parathyroid glands. *Dev. Biol.* 195, 1–15. doi: 10.1006/dbio.1997.8827

McGarry Houghton, A. (2015). Matrix metalloproteinases in destructive lung disease. *Matrix Biol.* 4, 167–174. doi: 10.1016/j.matbio.2015.02.002

McIntyre, D. C., Rakshit, S., Yallowitz, A. R., Loken, L., Jeannotte, L., Capecchi, M. R., et al. (2007). Hox patterning of the vertebrate rib cage. *Development* 134, 2981–2989. doi: 10.1242/dev.007567

Morrell, E. D., Mikacenic, C., Gong, K.-Q., Kosamo, S., Wurfel, M. M., and Manicone, A. M. (2020). Alveolar MMP28 is associated with clinical outcomes and measures of lung injury in acute respiratory distress syndrome. *Crit. Care* 24:141. doi: 10.1186/s13054-020-02847-0

Nova-Lampeti, E., Aguilera, V., Oporto, K., Guzmán, P., Ormazábal, V., Zúñiga, F., et al. (2018). *Hox Genes in Adult Tissues and Their Role in Endothelial Cell Differentiation and Angiogenesis*. London: IntechOpen.

Pineault, K. M., Swinehart, I. T., Garthus, K. N., Ho, E., Yao, Q., Schipani, E., et al. (2015). Hox11 genes regulate postnatal longitudinal bone growth and growth plate proliferation. *Biol. Open* 4, 1538–1548. doi: 10.1242/bio.012500

Ptaschinski, C., Hrycaj, S. M., Schaller, M. A., Wellik, D. M., and Lukacs, N. W. (2017). Hox5 paralogous genes modulate Th2 cell function during chronic allergic inflammation via regulation of Gata3. *J. Immunol.* 199, 501–509. doi: 10.4049/jimmunol.1601826

Rodriguez-Revenga, L., Iranzo, P., Badenas, C., Puig, S., Carrió, A., and Milà, M. (2004). A novel elastin gene mutation resulting in an autosomal dominant form of cutis laxa. *Arch. Dermatol.* 140, 1135–1139. doi: 10.1001/archderm.140.9.1135

Rolli, M., Fransvea, E., Pilch, J., Saven, A., and Felding-Habermann, B. (2003). Activated integrin αvβ3 cooperates with metalloproteinase MMP-9 in regulating migration of metastatic breast cancer cells. *Proc. Natl. Acad. Sci. U.S.A.* 100, 9482–9487. doi: 10.1073/pnas.1633689100

Rux, D. R., and Wellik, D. M. (2017). Hox genes in the adult skeleton: novel functions beyond embryonic development. *Dev. Dyn.* 246, 310–317. doi: 10.1002/dvdy.24482

Rux, D. R., Song, J. Y., Pineault, K. M., Mandair, G. S., Swinehart, I. T., Schlientz, A. J., et al. (2017). Hox11 function is required for region-specific fracture repair. *J. Bone Miner. Res.* 32, 1750–1760. doi: 10.1002/jbmr.3166

Sajjan, U., Ganesan, S., Comstock, A. T., Shim, J., Wang, Q., Nagarkar, D. R., et al. (2009). Elastase- and LPS-exposed mice display altered responses to rhinovirus infection. *Am. J. Physiol. Lung Cell. Mol. Physiol.* 297, L931–L944. doi: 10.1152/ajplung.00150.2009

Shapiro, S. D., Endicott, S. K., Province, M. A., Pierce, J. A., and Campbell, E. J. (1991). Marked longevity of human lung parenchymal elastic fibers deduced from prevalence of D-aspartate and nuclear weapons-related radiocarbon. *J. Clin. Invest.* 87, 1828–1834. doi: 10.1172/JCI115204

Shifren, A., Durmowicz, A. G., Knutsen, R. H., Hirano, E., and Mecham, R. P. (2007). Elastin protein levels are a vital modifier affecting normal lung development and susceptibility to emphysema. *Am. J. Physiol. Lung Cell. Mol. Physiol.* 292, L778–L787. doi: 10.1152/ajplung.00352.2006

Song, J. Y., Pineault, K. M., Dones, J. M., Raines, R. T., and Wellik, D. M. (2020). Hox genes maintain critical roles in the adult skeleton. *Proc. Natl. Acad. Sci. U.S.A.* 117, 7296–7304. doi: 10.1073/pnas.1920860117

Taylor, H. S., Vanden Heuvel, G. B., and Igarashi, P. (1997). A conserved hox axis in the mouse and human female reproductive system: late establishment and persistent adult expression of the hoxa cluster genes. *Biol. Reprod.* 57, 1338–1345. doi: 10.1095/biolreprod57.6.1338

Tuggle, C. K., Zakany, J., Cianetti, L., Peschle, C., and Nguyen-Huu, M. C. (1990). Region-specific enhancers near two mammalian homeo box genes define adjacent rostrocaudal domains in the central nervous system. *Genes Dev.* 4, 180–189. doi: 10.1101/gad.4.2.180

Urban, Z., Gao, J., Pope, F. M., and Davis, E. C. (2005). Autosomal dominant cutis laxa with severe lung disease: synthesis and matrix deposition of mutant tropoelastin. *J. Invest. Dermatol.* 124, 1193–1199. doi: 10.1111/j.0022-202X.2005.23758.x

Watt, F. M., and Hodivala, K. J. (1994). Cell adhesion. Fibronectin and integrin knockouts come unstuck. *Curr. Biol.* 4, 270–272. doi: 10.1016/s0960-9822(00)00063-4

Wellik, D. M., and Capecchi, M. R. (2003). Hox10 and Hox11 genes are required to globally pattern the mammalian skeleton. *Science* 301, 363–367. doi: 10.1126/science.1085672

Wellik, D. M., Hawkes, P. J., and Capecchi, M. R. (2002). Hox11 paralogous genes are essential for metanephric kidney induction. *Genes Dev.* 16, 1423–1432. doi: 10.1101/gad.993302

Yallowitz, A. R., Hrycaj, S. M., Short, K. M., Smyth, I. M., and Wellik, D. M. (2011). Hox10 genes function in kidney development in the differentiation and integration of the cortical stroma. *PLoS One* 6:e23410. doi: 10.1371/journal.pone.0023410

Yu, Q., and Stamenkovic, I. (2000). Cell surface-localized matrix metalloproteinase-9 proteolytically activates TGF-beta and promotes tumor invasion and angiogenesis. *Genes Dev.* 14, 163–176.

Zeng, Q., Li, S., Zhou, Y., Ou, W., Cai, X., Zhang, L., et al. (2014). Interleukin-32 contributes to invasion and metastasis of primary lung adenocarcinoma via NF-kappaB induced matrix metalloproteinases 2 and 9 expression. *Cytokine* 65, 24–32. doi: 10.1016/j.cyto.2013.09.017

Zhang, W., Menke, D. B., Jiang, M., Chen, H., Warburton, D., Turcatel, G., et al. (2013). Spatial-temporal targeting of lung-specific mesenchyme by a Tbx4 enhancer. *BMC Biol.* 11:111. doi: 10.1186/1741-7007-11-111

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Cell and Developmental Biology

REVIEW
published: 07 January 2022
doi: 10.3389/fcell.2021.786993

Roles of *Drosophila* Hox Genes in the Assembly of Neuromuscular Networks and Behavior

Rohit Joshi[1]*, *Rashmi Sipani*[1,2†] and *Asif Bakshi*[1,2†]

[1]Laboratory of Drosophila Neural Development, Centre for DNA Fingerprinting and Diagnostics (CDFD), Hyderabad, India, [2]Graduate Studies, Manipal Academy of Higher Education, Manipal, India

Hox genes have been known for specifying the anterior-posterior axis (AP) in bilaterian body plans. Studies in vertebrates have shown their importance in developing region-specific neural circuitry and diversifying motor neuron pools. In *Drosophila*, they are instrumental for segment-specific neurogenesis and myogenesis early in development. Their robust expression in differentiated neurons implied their role in assembling region-specific neuromuscular networks. In the last decade, studies in *Drosophila* have unequivocally established that Hox genes go beyond their conventional functions of generating cellular diversity along the AP axis of the developing central nervous system. These roles range from establishing and maintaining the neuromuscular networks to controlling their function by regulating the motor neuron morphology and neurophysiology, thereby directly impacting the behavior. Here we summarize the limited knowledge on the role of *Drosophila* Hox genes in the assembly of region-specific neuromuscular networks and their effect on associated behavior.

Keywords: Hox, *Drosophila*, motor neuron (MN), behavior, feeding, locomotion, self righting behavior, neuromuscular network

OPEN ACCESS

Edited by:
Ernesto Sánchez-Herrero, Centre for Molecular Biology Severo Ochoa (CSIC), Spain

Reviewed by:
Claudio Alonso, University of Sussex, United Kingdom
Jimena Berni, University of Sussex, United Kingdom

***Correspondence:**
Rohit Joshi
rohit@cdfd.org.in
rj2152@gmail.com

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Morphogenesis and Patterning, a section of the journal Frontiers in Cell and Developmental Biology

Received: *30 September 2021*
Accepted: *14 December 2021*
Published: *07 January 2022*

Citation:
Joshi R, Sipani R and Bakshi A (2022) Roles of Drosophila *Hox Genes in the Assembly of Neuromuscular Networks and Behavior. Front. Cell Dev. Biol. 9:786993. doi: 10.3389/fcell.2021.786993*

INTRODUCTION

Feeding, locomotion, and reproduction are some of the most fundamental behaviors exhibited by bilaterians. Regional specialization of the muscles, as well as the central nervous system (CNS) along the anterior-posterior (AP) axis, is a prerequisite for the successful and reproducible execution of these behaviors (Philippidou and Dasen, 2013). Role of Hox genes in assembling region-specific neural circuitry and MN diversification has been examined in vertebrates (Dasen et al., 2005; Dasen et al., 2008; di Sanguinetto et al., 2008; Philippidou and Dasen, 2013) and to some extent in insects (Dixit et al., 2008; Dutta et al., 2010; Baek et al., 2013; Garaulet et al., 2014; Enriquez et al., 2015; Picao-Osorio et al., 2015; Friedrich et al., 2016; Issa et al., 2019; Garaulet et al., 2020). Hox genes are also involved in the specification, survival, and functioning of other neuronal cell types (van den Akker et al., 1999; Pattyn et al., 2003; Gaufo et al., 2004; Holstege et al., 2008; Miguez et al., 2012; Huber et al., 2012; Bussell et al., 2014; Baek et al., 2019); however, such reports (for specific cell types) are limited. The MNs are central players in the functioning of neuromuscular networks. Their role in fine tuning muscle control is important for behavioral execution. The loss of MNs or perturbation of their function owing to a disease affects the behavior and leads to progressive muscle wasting. Therefore, studying their specification and functioning will give insights into the molecular basis of complex behaviors and disease.

In *Drosophila,* Hox genes are known to establish segment-specific patterns of myogenesis and neurogenesis (Michelson, 1994; Technau et al., 2014). However, the molecular basis of how Hox genes play a role in the specification and regional adjustment of the motor neuron (MN) networks is just beginning to be understood. Therefore much remains to be learned about their role in the assembly, maintenance, and functioning of segment-specific neuromuscular networks. In this regard, *Drosophila* as a model organism offers many unique advantages over other models (Hales et al., 2015; Schlegel et al., 2017; Yamaguchi and Yoshida, 2018). These advantages are a short life cycle, a fully sequenced genome, less redundancy than vertebrates, and a wide array of molecular genetic tools. In its short life cycle, *Drosophila* undergoes remarkable morphological and behavioral changes with different modes of feeding and locomotion for different stages. In just 10 days, it progresses from a static, non-feeding embryonic stage to a crawling and feeding larva, followed by an immobile and non-feeding pupal stage, eventually eclosing as a sexually active adult with an entirely different mode of navigation, locomotion, and foraging. It has a wide repertoire of simple, well-established behaviors (Nichols et al., 2012; Neckameyer and Bhatt, 2016), and many of the neuromuscular modules executing these behaviors are simple and well investigated. Compared to the vertebrates, *Drosophila* has a relatively less complex nervous system and musculature, and a fantastic array of molecular tools for reproducibly making subtle genetic manipulations in a cell-specific manner. The effect of these manipulations can be assayed in live and behaving animals (Korona et al., 2017; Martín and Alcorta, 2017), which is a tremendous advantage in correlating a gene to behavior.

In this review, we summarize existing *Drosophila* literature elucidating the role of Hox genes in the assembly and functioning of region-specific muscle-MN connections and their contribution in executing associated behaviors.

ROLE OF HOX GENES IN THE SPECIFICATION OF ANTERIOR-POSTERIOR AXIS IN *DROSOPHILA* CENTRAL NERVOUS SYSTEM

Hox genes are a family of homeodomain (HD) containing transcription factors (TFs), which play an important role in determining the anterior-posterior (AP) axis of bilaterian organisms (Hart et al., 1985; Regulski et al., 1985; Akam, 1989; Carroll, 1995). They are known to specify the AP axis by differentially regulating their downstream target genes with the help of TALE-HD containing cofactors Pbx/Exdradenticle (Exd) and Meis/Homothorax (Hth) (Mann and Chan, 1996; Mann and Affolter, 1998; Moens and Selleri, 2006; Merabet et al., 2007; Lelli et al., 2011; Saadaoui et al., 2011; Hudry et al., 2012). Hox genes execute these functions by giving the segments where they are expressed a very distinct identity, translating into divergent morphologies/properties along the AP axis of the body (including epidermal structure, CNS, and musculature). In *Drosophila,* there are eight Hox genes (compared to 39 Hox genes in vertebrates) which are organized into two complexes-the Antennapedia Complex (Antp-C) (Kaufman et al., 1990) [comprising the genes *labial (lab)*, *proboscipedia (pb)*, *Deformed (Dfd)*, *Sex combs reduced (Scr)*, *Antennapedia* (*Antp*)], and the Bithorax Complex (BX-C) (Sánchez-Herrero et al., 1985; Tiong et al., 1985; Maeda and Karch, 2006) [consisting of the genes *Ultrabithorax (Ubx)*, *abdominal-A* (*abd-A)* and *Abdominal-B (Abd-B)*]. *Drosophila* CNS consists of the brain and segmented ventral nerve cord (VNC). Hox genes pattern the VNC, which is the *Drosophila* equivalent of the vertebrate hindbrain and spinal cord. The embryonic VNC specified by Hox factors can be broadly divided into five regions, namely: supra-esophageal ganglia (SPG) expressing *pb* and *labial*; sub-esophageal ganglia (SEG) [composed of maxillary (Mx), mandibular (Mn), and labial segments (Lb)] expressing Hox genes *Dfd, Scr* and *Antp*; thoracic ganglia (T1-T3 segments) expressing *Antp* and *Ubx*; abdominal ganglia (A1-A7 segments) expressing *Ubx*, *abd-A,* and *Abd-B*; and the terminal region (A8-A10 segments) expressing *Abd-B* (Hirth et al., 1998; Kuert et al., 2014) (**Figure 1**). The expression of *pb* has also been reported in other segments (SEG to A9) of VNC (Baek et al., 2013; Enriquez et al., 2015; Hirth et al., 1998).

The neurogenesis in *Drosophila* happens in two phases, embryonic and larval, separated by a period of mitotic quiescence for the neural stem cells (called neuroblast-NB), which are the progenitors and generate all the neurons and glial cells of the CNS (Homem and Knoblich, 2012). In embryonic stages (stages 9–11), NBs delaminate from the neuroectoderm in each segment (Hartenstein and Wodarz, 2013). Five such successive waves of delamination generate 30 NBs per hemisegment of the embryo (Truman and Bate, 1988; Doe, 1992; Hartenstein and Wodarz, 2013). This blueprint of the CNS, when superimposed with the spatial genes [responsible for determining the AP and DV (Dorso-Ventral) (Skeath, 1999) axis and segment polarity genes (Bhat, 1999)], gives the NBs their specific positional identity (Schmid et al., 1999; Truman et al., 2004). This spatial identity of the NBs, in collaboration with the sequentially expressing temporal series TFs (Doe, 2017) expressed during embryogenesis, results in the generation of a segment-specific variety of cell types and cell numbers in the embryo. The transition of these temporal series TFs (Hunchback>Kruppel>Pdm>Castor >Grh) is intricately coupled to the NB cell cycle, which precisely times the expression of these factors and further contributes to specific cell type generation (Isshiki et al., 2001). The embryonic phase generates neurons required for larval CNS and eventually contributes to 10% of the adult neurons, while postembryonic neurogenesis contributes to the remaining 90% of the adult neurons (Homem and Knoblich, 2012). Hox genes contribute to the generation of the cellular variety along the AP axis in both embryonic and post-embryonic phase of neurogenesis by regulating fate specification, quiescence, proliferation, differentiation and apoptosis of NBs and their progeny (Prokop and Technau, 1994; Prokop et al., 1998; Bello et al., 2003; Miguel-Aliaga and Thor, 2004; Berger et al., 2005a; Berger

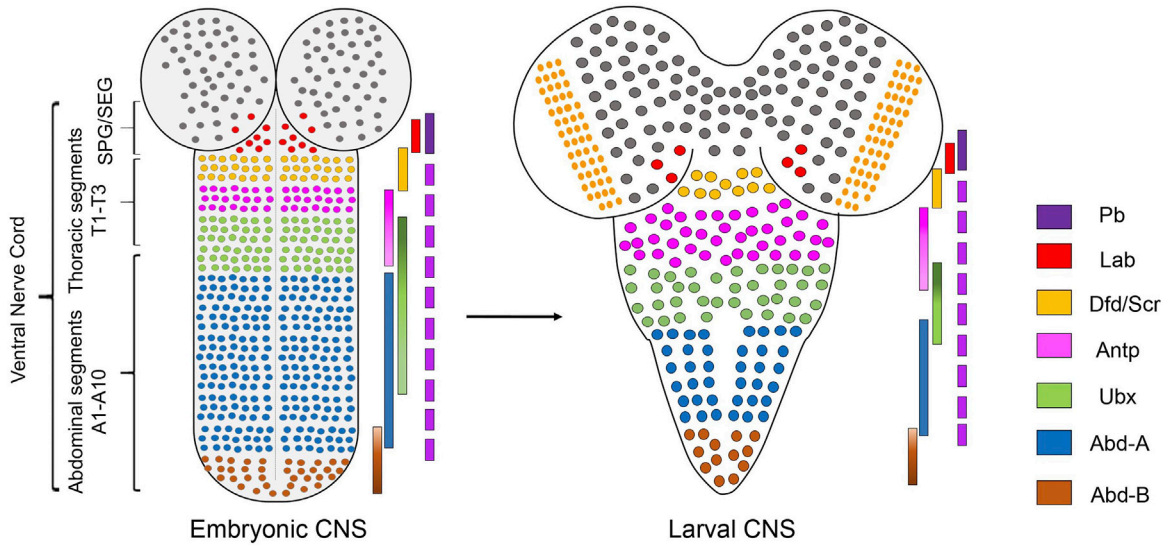


FIGURE 1 | Expression of Hox genes in *Drosophila* CNS. Schematics of embryonic (stage 14) and second instar larval CNS show Hox genes' expression pattern in different regions along the AP-axis. *Drosophila* CNS has a brain and ventral nerve cord (VNC). VNC is divided into SPG/SEG region, thoracic segments (T1-T3), abdominal (A1-A7) segments, and terminal (A8-A10) segments. The precise extent of Hox gene expression in these regions is shown by overlapping color-coded bars. Pb is expressed in all the segments from SEG to A9.

TABLE 1 | Role of Hox genes in establishing neuromuscular networks and behaviour in *Drosophila*.

Function	Hox gene involved	Location of action of Hox	Specific roles of Hox	References
Peristaltic movement in larval locomotion (**Figure 2A**)	Ubx/AbdA	Muscles and neurons	Establishing region specific neuromuscular networks. Region-specific matching of MN and Muscle (suggested)	Dixit et al. (2008)
Establishment of neuromuscular network for adult legs (**Figures 2B,C**)	Scr/ Antp/Ubx	Thoracic LinA MNs	LinA MN survival Axonal targeting of LinA MNs and innervation of leg muscles Axonal arborisation on leg muscle Antp level dependent axonal targeting to proximal and distal leg regions	Baek et al. (2013)
Adult locomotion	Pb	Thoracic LinB MNs	Regulation of axonal and dendritic morphology with the help of mTFs Targeting of 3 LinB MNs that innervate leg muscles Controlling the walking stance of the adults at high speed	Enriquez et al. (2015)
Larval feeding (**Figure 3**)	Dfd	MHE muscles and maxillary nerve motor neurons	Regulation of axonal outgrowth of MNs from the SEG that innervate the MHE muscles Formation and maintenance of synapses at the NMJ in the MHE by regulation of molecules controlling synaptic specificity	Friedrich et al. (2016)
Establishment of embryonic muscle innervation pattern (**Figure 4**)	Ubx	VL2 muscles and VL1 MNs	Regulation of Wnt4 and Sulf-1 in VL2 muscles that signal and repel away approaching growth cones of VL1 MNs Controlling expression of target genes in VL1 MNs to repel VL1 MNs from VL2 muscles	Hessinger et al. (2017)
Female egg-laying	Ubx	Fru^+ neurons	Proper oviduct innervation by Fru^+ ILP7 MNs Maintenance of MN synapses on oviduct and radial muscles	Garaulet et al. (2014)
Larval Self-righting behavior (**Figures 5A,B**)	Ubx	Larval SR node MNs	Regulation of neural Ca^{2+} activity of the SR node MNs	Picao-Osorio et al. (2015)
Adult Self-righting behavior (**Figures 5C,D**)	Ubx	Adult SR node MNs (these are distinct from larval SR MNs)	Regulation of neural Ca^{2+} activity of the SR node MNs Maintenance of synaptic structures on the adult leg muscles	Issa et al. (2019)

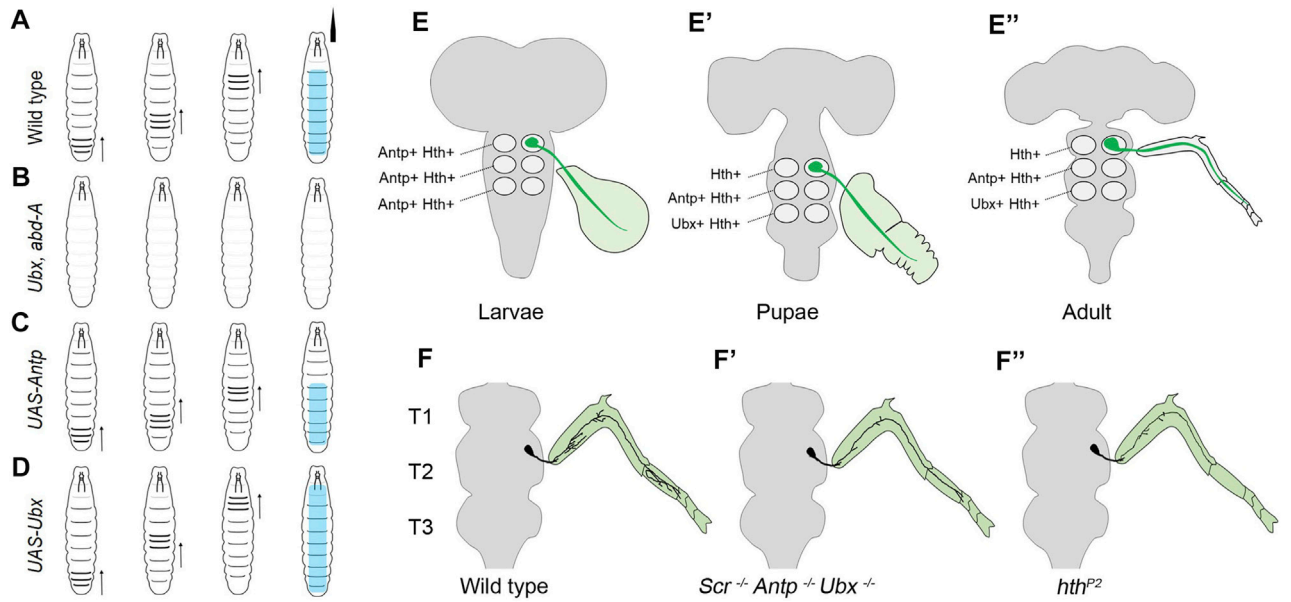


FIGURE 2 | Summary of the role of Hox genes in larval peristalsis and leg innervating MNs. **(A)** Schematic showing abdominal peristaltic movement in wild-type larva. **(B)** Shows that abdominal peristaltic movements are lost in *Ubx, abd-A* double mutant. **(C)** Shows that Antp overexpression transforms anterior segments but transformed segments do not show peristaltic movements. The abdominal peristaltic movements are unaffected. **(D)** Shows that in the case of *Ubx* overexpression, thoracic and more anterior segments get transformed and gain abdominal peristaltic movements. The direction of peristalsis is shown with an arrow, and its extent is shown in cyan color. **(E–E")** Shows that thoracic MNs (in green) innervate primordial leg tissue in larval and pupal stages and adult leg muscles. Also shown is the change in the expression code of Antp, Ubx, and Hth in LinA MNs of different thoracic segments (T1-T3) across different developmental stages. **(F–F")** Shows the wild-type arborization pattern of thoracic MNs innervating to the adult leg muscles. This axonal arborization is affected in triple Hox triple mutants (Scr^{-}, $Antp^{-}$, Ubx^{-}) and hth^{P2} mutants. This suggests that Hox/Hth is required for the survival, targeting and morphology of MNs innervating to the adult leg.

et al., 2005b; Rogulja-Ortmann et al., 2008; Tsuji et al., 2008; Kannan et al., 2010; Karlsson et al., 2010; Suska et al., 2011; Kuert et al., 2012; Baek et al., 2013; Birkholz et al., 2013a; Birkholz et al., 2013b; Estacio-Gómez et al., 2013; Baumgardt et al., 2014; Kuert et al., 2014; Arya et al., 2015; Becker et al., 2016; Khandelwal et al., 2017; Monedero Cobeta et al., 2017; Yaghmaeian Salmani et al., 2018; Bahrampour et al., 2019; Ghosh et al., 2019; Bakshi et al., 2020).

In the CNS, Hox genes are expressed in NBs in early embryonic stages, but their expression from the NBs is largely excluded thereafter. However, Hox genes continue to express in the neurons as they differentiate, project axons/dendrites, and form synaptic connections (Hirth et al., 1998) in embryonic and postembryonic stages. This led to the suggestion that Hox genes may have a role in the assembly and functioning of neuromuscular networks (Hirth et al., 1998; Dixit et al., 2008). More so, since the regionally distinct muscle patterns are known to be established by Hox genes early in development (Michelson, 1994). However, barring their role in regulating neuronal differentiation and apoptosis (Hirth et al., 1998; Miguel-Aliaga and Thor, 2004; Rogulja-Ortmann et al., 2008; Karlsson et al., 2010; Suska et al., 2011; Baek et al., 2013; Estacio-Gómez et al., 2013), the utility of the sustained expression of Hox genes in neurons had not been entirely clear. Therefore, the role of Hox genes in the regional specialization of the MNs, and their contribution to the assembly of functional neuromuscular networks (along the AP axis) remained unaddressed for a long time. Here, we focus on the reported roles of *Drosophila* Hox genes in the establishment of functional neuromuscular networks and behavior (summarized in **Table 1**). We also attempt to identify some common themes in the context of neuromuscular network assembly and functioning, which are independent of the conventional role of Hox genes in AP axis determination.

ROLE OF HOX GENES IN LOCOMOTION

One of the first reports implicating the role of Hox genes in assembling the segment-specific neuromuscular networks in *Drosophila* was by Dixit et al. (2008) from Bate and Vijay Raghavan groups in Cambridge and Bangalore. This report established the role of Hox genes in regulating segmental peristaltic movements in larval locomotion (Dixit et al., 2008). This work, in many ways, laid the foundation for exploring the molecular basis of the genetic control that results in equivalent cells of CNS (along the AP axis) to form regionally specialized neuromuscular networks.

This study showed that the thoracic and abdominal segments of the larval body have distinct movement patterns during larval peristalsis (**Figure 2A**). The experiments suggested that the abdominal peristaltic movements critically relied on BX-C, specifically *Ubx* and *abd-A* (**Figure 2B**). The mutants for *Ubx* or *abd-A* were used to show that either of these genes was

necessary for developing the neuromuscular networks coordinating these movements. Interestingly, the ubiquitous expression of either of the two genes was also sufficient for transforming the anterior segments to posterior identity (**Figure 2D**). This resulted in morphological transformation of the epidermal denticle belts (which had been known from earlier studies) and the anterior segments gaining the peristaltic activity like abdominal ones. However, this did not happen in the case of Antp overexpression (**Figure 2C**). In the absence of anatomical data, the study speculated that these movements rely on region-specific muscle architecture and their precise innervation by cognate MNs. In agreement with this, the expression of *Ubx* or *AbdA* in muscles was not sufficient for anterior segments to show a peristaltic pattern similar to posterior segments. This observation supported the idea that in the case of Hox-dependent segmental transformation, the MNs and the entire neural circuitry are reorganized to match the transformed pattern of muscle. This study also suggested that a one-to-one match in the identity of the muscle with that of the underlying neural circuitry is required for the proper execution of abdominal peristaltic movements.

Subsequently, two studies comprehensively addressed the developmental role of Hox genes in survival, targeting, and morphology of thoracic MNs, which innervate the leg muscles responsible for adult locomotion (Baek et al., 2013; Enriquez et al., 2015). These studies were built upon previous work which had shown that 50 MNs innervating the *Drosophila* adult leg muscles are generated by 11 NBs located in each thoracic hemisegment (Baek and Mann, 2009; Brierley et al., 2012). Two-third of these 50 MNs are generated by two NB lineages, LinA (or Lin 15), which generate 28 MNs, and LinB (or Lin 24), which generate seven MNs (Truman et al., 2004; Baek and Mann, 2009). These studies had characterized stereotypic axonal and dendritic morphologies of all the 50 MNs (generated by LinA and LinB) at the single-cell level, down to their synaptic innervations of the 14 leg muscles on the four segments of the adult leg (Baek and Mann, 2009; Brierley et al., 2012).

The first study by Baek et al. (2013), from Mann's group in New York focussed on LinA MNs and showed that Hox genes (*Scr*, *Antp*, and *Ubx*) and their TALE-HD containing cofactor Hth are required for the survival of the MNs in all the three thoracic segments. They found that all newly born thoracic MNs express Antp and Hth in all the three thoracic segments in larval stages (**Figures 2E–E'**). As the development progressed, this expression code transformed from being $Antp^{+}/Hth^{+}$ in all the segments to Hth^{+} in T1, $Antp^{+}/Hth^{+}$ in T2, and Ubx^{+}/Hth^{+} in the T3 segment in the late pupal stage and adults (**Figures 2E–E'**). This change in TF code is suggested to specialize the MNs innervating the adult legs to execute their segment-specific functions (Szebenyi, 1969; Kaplan and Trout, 1974; Dawkins and Dawkins, 1976; Trimarchi and Schneiderman, 1993; Dickinson et al., 2000). Exd was found to express in all the segments in all the stages. Subsequent clonal analysis (Wu and Luo, 2006) done with *Hox triple* (Scr^{-}, $Antp^{-}$, Ubx^{-}) and *hth* mutants indicated that expression of both *Hox and hth* genes is required autonomously within the thoracic lineages for survival, targeting, and morphology of the adult MNs (**Figures 2F–F'**) (Baek et al., 2013). Since Antp was the only Hox factor to be expressed initially in NBs, *Antp* single mutant was tested. Curiously, this mutant recapitulated most of the phenotype exhibited by *Hox triple* mutant, except the death of MNs in the T3 segment, which was attributed to *Ubx* and *Antp* redundancy in this segment. Subsequently, the death of the MNs was blocked by the expression of p35, a viral protein commonly used to block apoptosis (LaCount et al., 2000). In this case, the surviving MNs in *Hox triple* (Scr^{-}, $Antp^{-}$, Ubx^{-}) or *hth* mutant backgrounds targeted roughly to the same region along the proximal-distal (PD) axis of the adult leg segments, with terminal arborization defects. This suggested that *Hox* (and *hth*) genes are not needed by LinA progeny to assume the thorax-specific lineage identity or the MN fate. However, they are required for the appropriate specification of the finer morphological features of these MNs necessary for the functional muscle innervation (**Figures 2F–F'**). This was similar to what is known for the role of Hox genes in vertebrates MN specification, wherein MN identity is established independent of Hox genes (Jessell, 2000; Dasen et al., 2005).

The study also provided a novel alternative mechanism to diversify cell fate within a given lineage by modulating the expression level of Hox factor Antp. Usually, NB progeny rely on temporal series TFs for fate diversification (Doe, 2017). Baek et al. (2013) observed that within the same lineage, Antp is expressed at high levels in late-born MNs and low levels in early-born MNs. This variation in the Antp gene expression levels in MNs was found to have an instructional role in their axonal targeting. It was observed that high Antp expressing late-born MNs targeted the distal region, and low Antp expressing early-born MNs targeted the proximal regions of the adult leg. Expectedly, this pattern could be reversed by overexpression or the knockdown of *Antp*. Though, in the null allele of *Antp,* both distal and proximal targeting of MN was affected with no specific bias, indicating that low level gave a distinct phenotype from the absence of *Antp*. This variation in the expression levels of *Antp* had cell-autonomous consequences in MN innervation and did not show any defect in the leg muscles of the adult fly. This was in contrast to an earlier work by Dutta et al. from VijayRaghavan and Rodriguez groups at Bangalore, where Hox dysregulation in MNs resulted in muscle development defects (Dutta et al., 2010). The experiment in this study shows that knockdown of *Ubx* in adult MNs resulted in modest reduction and developmental deformity in adult leg muscles (Dutta et al., 2010). On the other hand, *Ubx* overexpression in the MNs innervating thoracic dorsolateral muscle (DLM) of adults caused a dramatic reduction in the number of DLM fibers (Dutta et al., 2010). These observations implied an active communication between the adult thoracic MNs and their muscle targets. Dutta et al. (2010) also suggested that Hox expression needs to be tightly controlled within a narrow range for the assembly of functional neuromuscular network. It is to be noted that this study relied on a chronic knockdown/

overexpression of *Ubx* in the MNs, while Baek et al. (2013) used temporally controlled overexpression or knockdown.

The requirement of Hox genes in determining the morphology of thoracic MNs was followed up by Enriquez et al. from Mann's group in New York (Enriquez et al., 2015). This work identified the role of Hox gene *proboscipedia (pb)* in determining the morphological characteristics of three thoracic MNs (Enriquez et al., 2015). The *proboscipedia (pb)* expresses from the supraesophageal region to the A9 segment in embryonic and larval CNS (Hirth et al., 1998; Baek et al., 2013), but its role in neurogenesis had not been tested. Previous studies had suggested that the morphology of MNs arising out of thoracic LinA and LinB lineages were under precise genetic control, which had a bearing on their function (Baek and Mann, 2009; Brierley et al., 2012). However, the genetic determinants regulating the individual neuronal morphology for these (or any neuron) were not known. LinB (or Lin 24), which produces only seven MNs in each hemisegment (of 3 thoracic segments) were an attractive system to address this problem owing to few but well-characterized MNs in this lineage (Baek and Mann, 2009; Brierley et al., 2012). Enriquez et al. screened 230 antibodies against different TFs for their expression in larval LinB lineage [marked by GFP using MARCM (Wu and Luo, 2006)]. They identified six TFs whose combinatorial expression was sufficient to uniquely classify the seven MNs of LinB lineage. These factors were Empty spiracle (Ems), the Zinc finger homeodomain factors 1 and 2 (Zfh1 and Zfh2), the Hox TF Proboscipedia (Pb), the Pax6 ortholog Twin of Eyeless (Toy), and Prospero (Pros). Further, they observed that the TFs combinations observed in each of the seven MNs were not observed in any other neuron of the CNS. After that, they used lineage tracing experiments to correlate larval LinB MNs (with unique TF code) to their adult counterparts, each of which corresponds to distinct morphology and muscle innervation. This supported the idea that the characteristic expression of these six TFs probably results in distinct axonal and dendritic morphologies of these MNs. Consequently, these factors were called morphology TFs (mTFs).

Next, a clever combination of MARCM (Wu and Luo, 2006) with the Flybow technique (Hadjieconomou et al., 2011) was used to mark adult MNs where they removed or overexpressed *pb* to analyze its effect in MNs of LinB. The Hox gene *pb* was shown to be essential for the morphological identity of 3 out 7 MNs of LinB. Interestingly, loss or overexpression of *pb* did not affect the expression of the other five mTFs, which was in contrast to what is known for the temporal series TFs (Doe, 2017), which play a crucial role in generating neuronal diversity. When *pb* mutant LinB NB was analyzed, it was observed that the number of MNs generated in the lineage was unaffected. These MNs did not lose or change their identity; they remained glutaminergic, and their axons targeted the leg muscles. However, there was a reduction in the area covered by dendrites of MNs, and specific axon targeting defects were observed on adult leg muscles. Conversely, misexpression of *pb* in LinA MNs resulted in the relocalization of their dendrites to an area on the neuropil, where typically LinB dendrites were located. However, LinA retained many of its features and did not gain all the characteristics of LinB. Since *pb* mutant MNs show defective leg muscle innervation, the adults with *pb* mutant MNs were tested for walking behavior. Most walking parameters were normal, except that at high speed, the flies with *pb* mutant LinB MNs showed more wobble in walking than the control adults. This indicated a role of Pb expressing LinB MNs in stable walking at high speed. In order to establish that mTF code (of Pb with other factors) was instructive for the MN morphology, the TF code of specific LinB MNs was changed to other MNs in the LinB lineage by simultaneous knockdown and overexpression of the mTFs. It was observed that altering the mTF code resulted in the predictable transformation of the morphology, which supported the idea that different combinations of mTFs determined the MN's morphological identity and led to the suggestion that role of *pb* in morphology had a critical bearing on the fly walking behavior. These results also established a genetic basis for the morphology of the MNs. They also suggested that MNs rely on a unique combination of different mTFs, which collectively give them their distinct signature morphologies. An idea proposed in the study is that temporal TFs most likely direct a stepwise change in the mTFs code for successive MNs (generated in LinB) and thus progressively change their morphology. To test this idea, it will be an important (though tedious) task to delineate the role of individual mTFs in determining the final MN morphology in the context of LinB. The results also raise the question of whether Pb plays a similar role in determining the morphology of other thoracic MNs (working with a different set of mTFs). Alternatively, considering Pb expression in other segments; it is a possibility that Pb may contribute to determining the morphologies of MNs found in other segments of VNC as well.

Moreover, since different levels of Antp have already been shown to play a role in regulating the morphology of MNs (Baek et al., 2013), one wonders if there is any interaction between *pb* (or other mTFs) and resident Hox gene in determining the final MN morphology. It is to be noted here that Antp expressing LinA MNs did not express Pb (Baek et al., 2013). Experimental testing of this idea will also determine whether the identity of the NB has any consequence on the choice of the mTFs employed. However, the existence of a lineage-specific combination of the mTFs has already been ruled out by Enriquez et al. (2015).

HOX CODE FOR NEUROMUSCULAR ASSEMBLY IN EMBRYOGENESIS AND LARVAL FEEDING CIRCUIT

Feeding is a fundamental behavior necessary for the survival of an animal. In *Drosophila,* the feeding behavior has been investigated in larval and adult stages (Pool and Scott, 2014; Miroschnikow et al., 2020). The neurons responsible for feeding behavior and taste perception reside in the maxillary (Mx) and mandibular (Mn) neuromeres of larval SEG, which express Hox gene *Dfd* (McGinnis and Krumlauf, 1992; Hirth et al., 1998; Kuert et al., 2014). The *Dfd* loss of function mutants die during embryogenesis due to their inability to hatch. The hypomorphic alleles that survive until adulthood starve to death, owing to their inability to extend proboscis and ingest

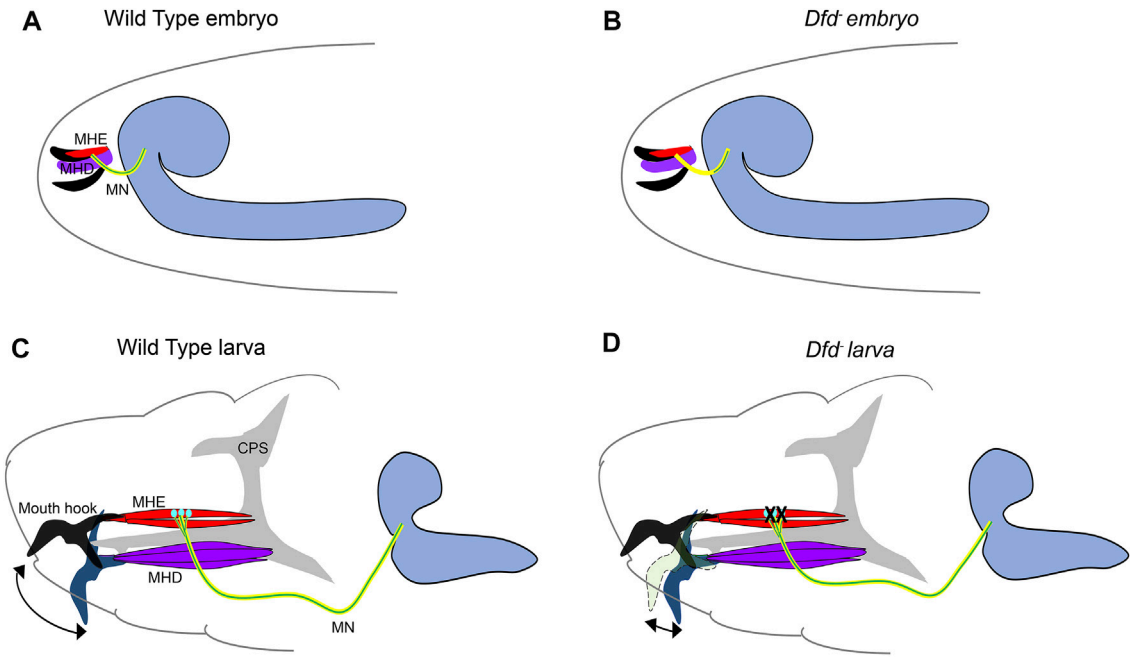


FIGURE 3 | Role of *Dfd* in embryonic axonogenesis and larval feeding. Schematics show the major components responsible for embryonic mouth hook movement and larval feeding circuitry. The maxillary nerve (shown in yellow, originates from the SEG region of embryonic/larval CNS and innervates both MHD (shown in purple) and MHE (shown in red) muscles. The *Dfd* positive motorneurons (shown in green) from SEG synapse with the MHE muscles only and are crucial for the mouth hook elevation during embryonic hatching and larval feeding. The grey shaded region represents the Cephalopharyngeal Skeleton (CPS), extension and retraction of which is integral to the feeding process. The CNS is shown in blue on the right. The mouth hooks are shown in black in the embryos. In larvae, black mouth hook shows the extent of its elevation in wild-type, blue mouth hook shows the extent of its depression in wild type, and light green mouth hooks show the extent of its elevation in a Dfd^3 mutant. **(A,B)** Shows the schematic of wild type and *Dfd* mutant embryos, latter show a severe restriction in axon outgrowth for the Dfd positive motor neurons resulting in failure of these neurons to project to the MHE. These mutants show hatching defects. **(C,D)** Shows the schematic of the wild type and *Dfd* mutant larvae. In the absence of *Dfd*, the mouth hook elevation is drastically reduced (double-sided arrow indicates the extent of mouth hook elevation and depression in wild type and *Dfd* mutant larvae). *Dfd* is required in both MHE and MNs to regulate assembly and maintenance of the feeding motor unit to execute optimum mouth hook movement.

food (Merrill et al., 1987; Regulski et al., 1987; Restifo and Merrill, 1994).

Building on these observations, a study by Friedrich et al. from Lohmann's group at Heidelberg investigated the role of Dfd in larval feeding behavior (Friedrich et al., 2016) and exploited the fact that both hatching and feeding rely on the same motor circuit responsible for the up and down movement of the larval mouth hooks (Pereanu et al., 2007; Schoofs et al., 2010; Hückesfeld et al., 2015) (**Figure 3**). These movements are executed by mouth hook elevator (MHE) and depressor (MHD) muscles in the larval head, which receive synaptic input from neurons in Mx, Mn, and Lb neuromeres and contribute to the larval feeding circuit (Pereanu et al., 2007; Schoofs et al., 2010; Hückesfeld et al., 2015). The authors show that the MNs from SEG expressed Dfd and innervate the elevator but not the depressor muscles (**Figure 3A**). Congruent to these observations, embryos mutant for *Dfd* were found to exhibit axonogenesis defects (**Figure 3B**) and consequent failure to hatch into larvae.

Blocking the synaptic transmission in Dfd expressing MNs using tetanus toxin also compromised embryonic hatching. Next, using a temperature-sensitive allele (Dfd^3) it was shown that Dfd is chronically required in the assembly, maintenance, and functionality of the feeding circuit. It was observed that Dfd^3 embryos exhibited mouth hook movement and hatching defects when raised to non-permissive temperature in late embryogenesis (which is much after the formation of synapses). In corroboration to this, Dfd^3 larvae, when shifted to non-permissive temperature as late as in the third instar stage, showed head-mouth hook movement defects, further establishing the chronic requirement for the gene (**Figure 3D**). Interestingly in both these cases, the innervation of the elevator muscle was found to be normal. Similar knockdown (KD) of Dfd in neurons by RNA interference or Dfd^3 allele exhibited a significant change in synaptic morphology coupled with the reduction in the expression of a synaptic gene, Ankyrin-2 extra-large (Ank2-XL). However, unlike in the case of Ubx KD in MNs (Dutta et al., 2010) reported earlier by Dutta et al., muscles in the larval feeding circuit were normal in the case of Dfd KD in MNs. Since Dfd was found to be expressed both in the elevator muscles and the MNs from SEG, it was proposed that Hox expression in both these cells types provides them with a molecular code to identify each other during synaptic assembly. In agreement with this idea, a synaptic target recognition molecule "Connectin" (*Con*) (Nose et al., 1994) was found to be amongst the direct transcriptional targets of Dfd in CNS. Interestingly, this homophilic cell adhesion molecule Con was expressed in MNs and muscle devoid of Dfd protein, and its expression was regained in the MNs mutant for *Dfd*. This suggested that Dfd functioned to bring together MNs

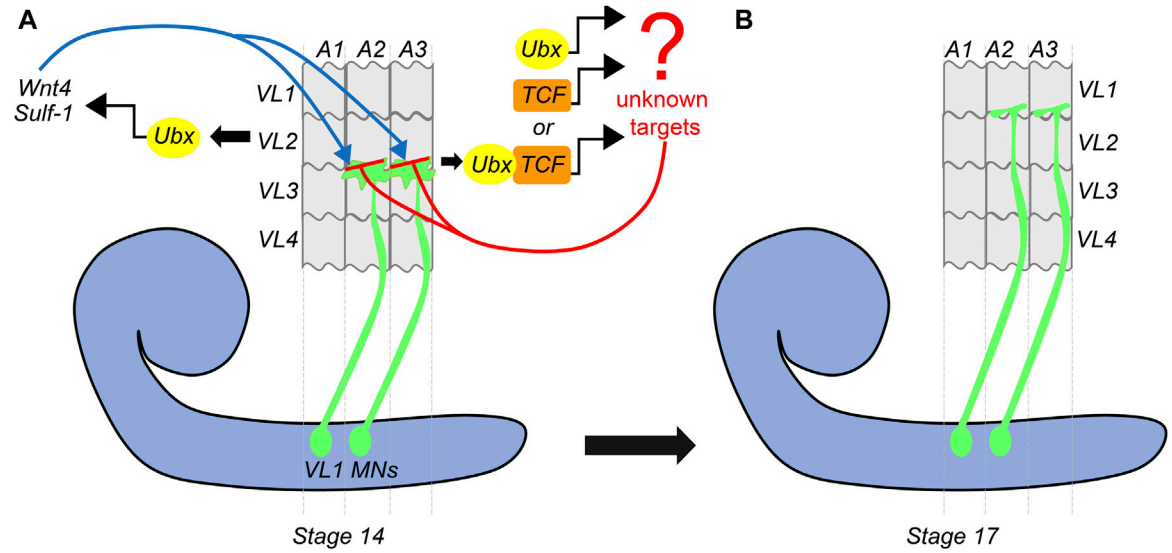


FIGURE 4 | Dual role of Ubx in embryonic muscle innervation. Schematic of stage-14 and 17 embryos show the mechanism of innervation of ventrolateral muscles (VL1-4) of the embryonic body wall (shown in grey) by VL1-MNs (shown in green) coming from the embryonic CNS (shown in blue). **(A)** At stage-14, the approaching growth cone of the VL1 MNs is repelled by Ubx expressing VL2 muscles. Ubx mediates activation of Wnt4 and Sulf-1 in the VL2 muscles, which then interact with the Wnt receptors on the growth cone of the MNs. This leads to the activation of Wnt signaling (armadillo/TCF) in the VL1 MNs. Ubx and TCF in these MNs act together or in parallel to regulate the expression of unknown target genes, resulting in the repulsion of the VL1-MNs by VL2 muscles, thereby pushing them to their final target (VL1 muscles) by stage 17. **(B)** This suggests that Ubx expression in both VL2 muscles and VL1-MNs is required for establishing precise neuromuscular connections in the embryo.

and cognate muscles by actively repressing *Con* in the cells of the feeding circuit. However, the identity of cell adhesion molecule(s) positively required by these cells for assembly of the neuromuscular feeding unit is yet to be determined.

The work established the role of Dfd as a critical coordinator for the formation, maintenance, and functioning of the neuromuscular network in the larval feeding circuit. The results also showed that synaptic stability and plasticity are determined by the half-life of synaptic proteins as well as the transcriptional program, which sustains the supply of synaptic components that maintains the neuromuscular junction. Lastly, it was proposed that Hox genes provide the molecular code for matching the MNs and muscles during developmental synaptic assembly through their transcriptional targets. However, even though Dfd was shown to have a role in the functioning of the feeding circuit, it remains to be investigated whether Dfd played a role in regulating the neural activity of the motor neurons to regulate the feeding behavior.

Continuing on the theme of Hox gene providing a molecular basis for matching the MNs and muscles, a subsequent study by Hessinger et al. (2017) from Rogulja-Ortmann and Technau groups at Mainz established a similar role for Hox gene *Ubx* in the assembly of the embryonic neuromuscular junction. This study unraveled the mechanism of how *Ubx* plays a role in determining the target specificity of the MN and its cognate muscle during embryogenesis. In the abdominal segments (A2-A7 segments) of embryonic CNS, ventrally projecting RP MNs innervate ventrolateral (VL) muscles on the embryonic body wall. The RP MNs 1, 3, 4, and 5 are some of the MNs known to innervate four VL muscles (VL1-4) in the abdominal segments (Bossing et al., 1996; Landgraf et al., 1997). Hessinger et al. focussed on the innervation of RP5 and V MNs (referred to as VL1-MNs) onto the VL1 muscles of abdominal segments (**Figure 4**). Through meticulous genetics, the study established that precise innervation of VL1 muscle by its cognate MNs (VL1-MNs) relies on Ubx mediated activation of Wnt4 signaling in VL2 muscle (**Figure 4A**). The authors found that AbdA had no role in this innervation, which entirely relied on Ubx dependent expression of the *Wnt4* and *sulfatase 1* gene (*sulf1*-known to be necessary for axon guidance) in the VL2 muscle. Wnt4 and Sulf1 expression in VL2 muscle played an instrumental role in repelling the axons of the MNs facilitating them to innervate their correct target, which was the VL1 muscle (**Figure 4B**). *Wnt4* is a member of the Wnt family of signaling molecules while Sulf1 is a sulfatase implicated in regulating Wnt and BMP gradient in neuromuscular junction (Nose et al., 1994; Inaki et al., 2007). The secretion of Wnt4 and Sulf1 by VL2 was paralleled with the activation of canonical Wnt4 signaling in VL1-MNs. This facilitated the repulsion of MNs away from VL2 muscles, thereby establishing a precise neuromuscular connection (between VL1 muscle and VL1-MNs). Congruent to this, the knockdown of the canonical Wnt4 signaling pathway in the VL1-MNs resulted in their targeting defects. On the expected lines in *Ubx* mutants, *Wnt4* and *sufl1* genes were downregulated in VL2 muscles. Consequently, VL1-MNs could not go past VL2 muscles, and the innervation of VL1 muscles by these MNs was lost. Finally, as was observed in the case of larval peristalsis (Dixit et al., 2008) and feeding circuitry (Friedrich et al., 2016), it was the simultaneous expression of Ubx in both MNs and the muscles which rescued the *Ubx* mutant phenotype.

Collectively, these studies highlight the importance of Hox genes in establishing a complementary molecular code between MN and muscles for the functional assembly of the neuromuscular networks.

ROLE OF BX-C miRNA MEDIATED HOX REGULATION IN BEHAVIOR

BX-C has a bidirectionally transcribed microRNA (miRNA) locus with two overlapping miRNA's on the opposite strand (*iab4/8*). This miRNA locus lies between *abd-A* and *Abd-B* and has been shown to target neighboring homeotic genes and results in homeotic transformation on overexpression (Ronshaugen et al., 2005; Bender, 2008; Stark et al., 2008; Tyler et al., 2008). Phenotypically, the homozygous deletion for the miRNA shows sterility and no other significant phenotype (Bender, 2008; Lemons et al., 2012). Recent studies have focussed on the role of this locus in CNS development, sterility, and adult behavior.

ROLE OF UBX IN EGG-LAYING BEHAVIOR

One of these studies by Garaulet et al. (2014) from Lai's group in New York investigated the role of BX-C miRNA in CNS patterning and female sterility. Garaulet et al. demonstrated that in contrast to the embryonic epidermis where AbdA and AbdB repress anterior Hox gene *Ubx*, in larval CNS, it is the BX-C miRNA, that represses the BX-C genes outside their normal domain of expression. The deletion of this locus results in deregulation of Hox genes *Ubx* and *abd-A* and their cofactor *exd* and *hth* in posterior VNC of larval CNS. This was in agreement with what had been reported for this miRNA previously in the embryonic CNS as well (Bender, 2008; Thomsen et al., 2010; Gummalla et al., 2012). The subsequent genetic analysis shows that the sterility phenotype reported in miRNA-deleted females was substantially rescued by heterozygosity for BX-C genes (*Ubx*, *abdA*) and their cofactor *hth*. This effect was recapitulated by the targeted knockdown of *Ubx* in neurons, thereby establishing that deregulation of *Hox* and *hth* genes in neurons is critical for the sterility phenotype. Phenotypically, the ovary morphology in the mutant female flies was normal, and flies were capable of mating. Since the defect seemed to be in the egg-laying, therefore the focus shifted to the oviduct. The oviduct has two kinds of inputs, *Insulin-like peptide 7* (ILP7) expressing excitatory glutaminergic MNs and inhibitory octopaminergic neurons terminating on radial muscles and epithelial linings (Rodríguez-Valentín et al., 2006; Yang et al., 2008; Castellanos et al., 2013). Significantly, the BX-C miRNA deletion did not alter the number or the transmitter identity for the ILP7+ MNs or the octopaminergic neurons. However, there was a reduction in the innervation of ILP7+ MNs on the oviduct and synaptic bouton count of the MNs on the radial muscles. These defects were substantially rescued by heterozygosity of *Ubx* and *abd-A,* but not by *hth*. However, the overexpression of *Ubx* or *hth* specifically in ILP7+ MNs did not recapitulate the sterility. This suggested that the broad de-repression of these genes in CNS was the cause of adult sterility. A search for functional neuronal domain responsible for the sterility was narrowed down to the Fruitless (Fru) expressing neurons (Stockinger et al., 2005) [which include ILP7+ MNs of oviduct as well (Castellanos et al., 2013)]. The Ubx and Hth overexpression in Fru+ neurons resulted in significant female sterility (90% in Ubx and 22% in Hth), suggesting that these neurons contribute to the female egg-laying program. However, other neuronal lineages from Fru expressing domain relevant for fertility and egg-laying behavior were not identified. Quite surprisingly, a subsequent study by the same group with a new deletion allele for BX-C miRNA showed that female flies were normal in their egg-laying behavior and had a functional neuromuscular control at the genital tract (Garaulet et al., 2020). Instead, this study suggested that the miRNA-deleted female had a behavioral shift from a virgin state to a post-mated state. This shift was attributed to the misregulation of *hth* in CNS. However, whether the misregulation for Hox genes (*Ubx* and *abd-A*) play a role in the behavioral shift was not reported. Also, this study did not investigate the innervation of ILP7+ MNs in females homozygous for new miRNA deletion. This suggests that either Hox genes of BX-C have no role in this behavioral shift for the female flies or the same is yet to be investigated.

ROLE OF *UBX* IN SELF RIGHTING BEHAVIOR

Continuing on the theme of BX-C miRNA mediated repression of homeotic genes, two elegant studies from the Alonso Lab at Sussex in the UK have uncovered a role of the homeotic gene *Ubx* and the BX-C miRNA in self-righting (SR) motor behavior in *Drosophila* larvae and adults (Picao-Osorio et al., 2015; Issa et al., 2019). These studies show for the first time a post-developmental role of Hox gene and the importance of maintaining a very fine control over Hox expression in CNS to regulate neural physiology and behavior. SR is an innate reflex that corrects the body orientation when it is out of its normal upright position (**Figures 5A,C**). This response is evolutionarily conserved amongst all the bilaterians (Penn and Jane Brockmann, 1995; Faisal and Matheson, 2001; Jusufi et al., 2011).

The first study by Picao-Osorio et al. (2015) established a role of BX-C miRNA *iab4* in the regulation of *Ubx* in a defined group of MNs required to execute the SR behavior in larvae. Initially, the larvae for the deletion of BX-C miRNA were tested for different behavior assays of abdominal peristaltic waves, turning, and SR. All the behaviors were normal except for the SR behavior where miRNA-deleted larvae took a long time to turn themselves over after being put on their dorsal side (**Figure 5A**). Since *Ubx* was a known target of BX-C miRNA in VNC (Bender, 2008; Tyler et al., 2008; Thomsen et al., 2010; Garaulet et al., 2014), it was tested by targeted overexpression in its native transcriptional domain, and its role was confirmed in SR defects. Next, the cellular basis of aberrant SR behavior was narrowed down to *Ubx* regulation by *iab4* to two metameric MNs in larval VNC (SR node neurons or SRN). The SRN innervated the lateral transverse (LT) muscles of the larval body wall, the LT1/2 (**Figure 5B**) (Picao-Osorio et al.,

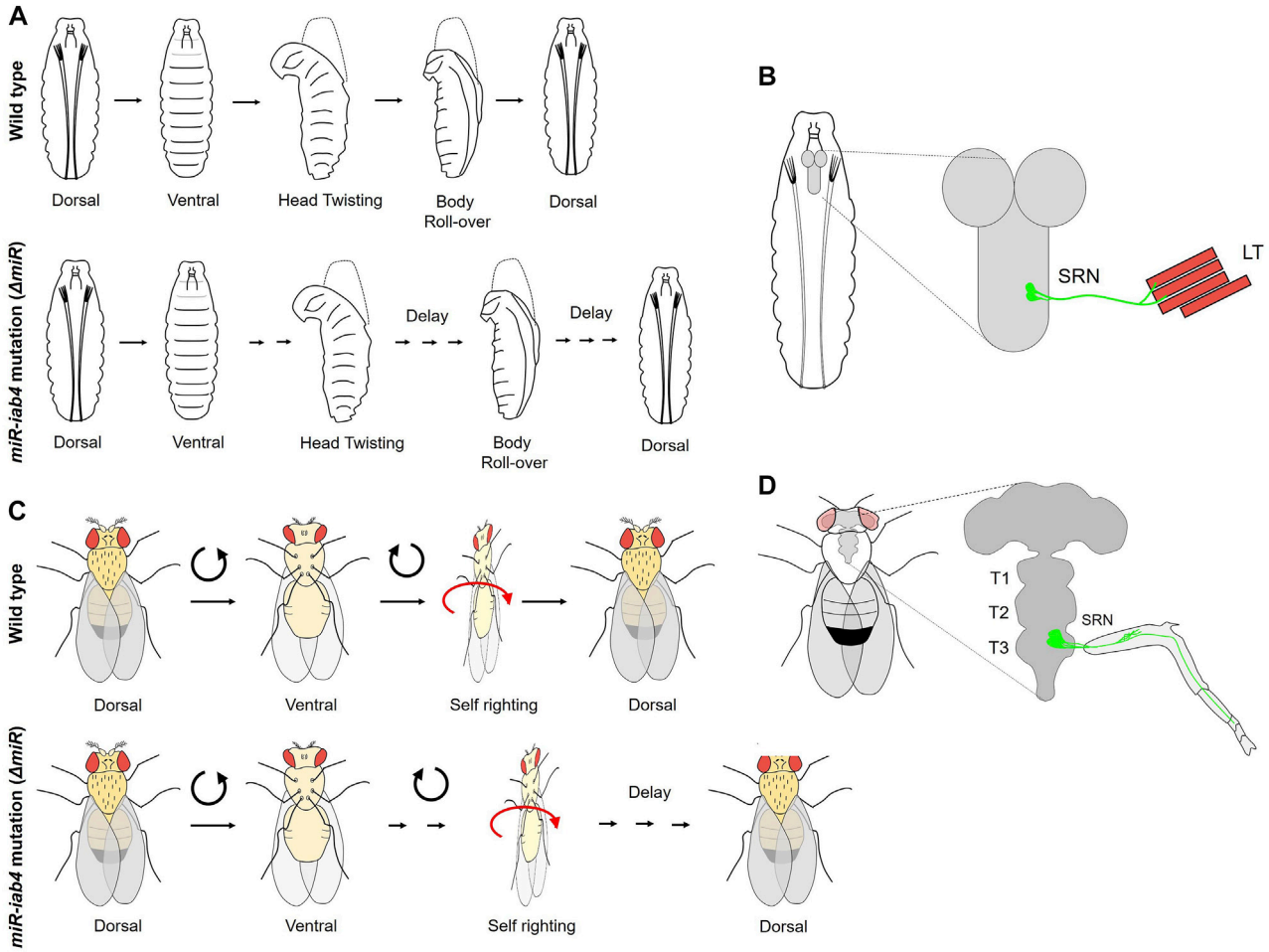


FIGURE 5 | Role of BX-C miRNA in self-righting behavior. **(A)** The larval SR behavioral response in wild-type and *miR-iab4* mutants is shown. Head twisting and body roll-over get significantly delayed in miR-iab4 mutants compared to wild-type larvae. **(B)** Shows that SR node (SRN) neurons originating from abdominal segments innervate to the lateral transverse (LT1/2) muscles of the larval body wall. **(C)** Shows adult self-righting behavioral response in wild type and *miR-iab4* mutants, latter shows delayed SR behavior in adults as well. **(D)** Shows the innervation of the adult leg muscles by SRN neurons originating from T3 segments.

2015). Interestingly, the authors did not find any developmental consequence of Ubx dysregulation in larval CNS, and conditional expression of Ubx in the SRN in larval stages could recapitulate the SR behavior defects. This suggested that tampering with the levels of Ubx in these neurons specifically affected physiology and behavior. Similarly, a specific MN GAL4 line, which innervated LT muscles, was used for misexpressing *Ubx* and was shown to delay the SR in larvae. These observations were further corroborated by the differences in calcium activity traces of the SR MNs [measured using *in vivo* calcium sensor GCaMP (Chen et al., 2013)] in miRNA-deleted larvae compared to the wild-type controls. Artificial thermogenic activation (Hamada et al., 2008) or silencing (Kitamoto, 2001) of SR MNs also resulted in SR behavior defects, which was also reflected by the difference in calcium activity traces (Chen et al., 2013) in the test and the wild-type controls. However, it was not clear from this study whether similar SR movements in morphologically distinct organisms like larvae or adults relied on common or different genetic modules. To address this, the same group investigated and found a role of miRNA-mediated *Ubx* regulation in adult SR behavior (**Figure 5C**) (Issa et al., 2019). In this case, as well, overexpressing *Ubx* in its native domain could recapitulate these defects. Subsequently, *Ubx* was upregulated in two different subsets of adult leg MNs. However, the SR defect was reported in only in one case, further restricting the MNs responsible for SR defects in adults. These MNs were different from those required for executing SR behavior in larvae (Picao-Osorio et al., 2015). The downregulation of *Ubx* in adult-specific SR MNs was sufficient to rescue the behavioral defects reported in miRNA deletion. This knockdown also increased the number of synaptic varicosities on the femur muscles of the adult leg and rescued the neural activity in MNs back to the wild-type levels. These results supported a previously suggested idea that Hox genes have a role in assembling and maintaining the synaptic structures (Friedrich et al., 2016). The *Drosophila* larva and adult are divergent in lifestyle, behavioral properties, muscle structure, and neural constitution. Therefore, this study suggests that similar movements performed by organisms with distinct

biomechanical, morphological, and neural structures could rely on the same miRNA/Hox genetic module, which can be redeployed in different developmental stages for equivalent behavior.

Importantly, these studies show that the miRNA-dependent post-transcriptional regulation of Hox gene *Ubx* can control the neural activity of MN to regulate the behavior of an animal. This function of Hox genes in neural physiology is independent of their role in development. The authors also suggest that other behavioral modules (like postural adjustment and locomotion) could also be controlled by miRNA. Furthermore, it is also possible that other adult movement-associated behaviors (like flight, walking, and jumping) (Szebenyi, 1969; Kaplan and Trout, 1974; Dawkins and Dawkins, 1976; Trimarchi and Schneiderman, 1993; Dickinson et al., 2000) may also be regulated by miRNA-mediated regulation of Ubx or Antp. For instance, Baek et al. show that Antp and Hth are the primary factors expressed in all thoracic leg MNs in larval stages. However, in the late pupal and adult stage, the T1 MNs express Hth, T2 express Antp and Hth, and T3 express Ubx and Hth (Baek et al., 2013). Therefore, one possibility worth considering is whether BX-C miRNA-mediated regulation of Antp and Hth also contributes to adult SR or other movement-associated behaviors. This is plausible considering that Hth has already been shown to be a target of BX-C miRNA in CNS (Bender, 2008; Tyler et al., 2008; Thomsen et al., 2010; Garaulet et al., 2014; Garaulet et al., 2020). Since 40% of the miRNA in the *Drosophila* genome were shown to affect the larval SR behavior (Picao-Osorio et al., 2017), therefore it may also be worthwhile to check whether any of these miRNA's contribute to the regulation of SR or other movement associated behaviors through the regulation of Hox (*Ubx* or *Antp*) or Hth.

Lastly, a tempting question is whether the miRNA/Hox genetic module could also function in MNs of other behavioral circuits like feeding, mating, courtship, grooming, and virgin/mated behavioral shift. Moreover, if such control exists, it needs to be investigated whether it is executed primarily through Hox genes or other miRNA targets.

CONCLUSIONS

The survival of an organism depends on its ability to successfully and reproducibly execute a multitude of essential behaviors. This critically relies on Hox-dependent region-specific neuromuscular networks established along the AP axis of the body. Hox genes have been extensively investigated for their role in MN specification and motor circuit assembly in the hindbrain and the spinal cord of vertebrate CNS (Jessell, 2000; Philippidou and Dasen, 2013). The MNs in the hindbrain have a clustered organization, while in the spinal cord MNs are organized into longitudinal columns. At lower cervical (brachial) and lumbar levels of the spinal cord, MNs of the lateral motor column (LMC) project axons toward the forelimbs and hindlimbs (Landmesser, 2001). These columnar identities are regulated by the action of one or multiple Hox genes. Hox genes also diversify the MNs within LMC to generate approximately 50 MN pools targeting different limb muscles (Dasen et al., 2005). The cross-repressive interactions between different Hox genes set up a distinct transcriptional profile for each pool, which contributes to their clustering and peripheral muscle innervation (Dasen et al., 2008). Expectedly, individual Hox mutants in vertebrates affect the pool sizes, their position, and MN arborization. For example, in the case of HoxC6 mutants, brachial LMC size is reduced (Tiret et al., 1998; Vermot et al., 2005; Lacombe et al., 2013). Similarly, in the lumbar region where Hox10 is a major determinant of LMC identity, different mutant combinations for Hox10 result in defects in hindlimb innervation and compromise MN survival (Wahba et al., 2001; Lin and Carpenter, 2003; Shah et al., 2004; Wu et al., 2008). It has also been shown in the spinal cord MNs that acquisition of their basic MN identity and features is Hox independent (Jessell, 2000). These observations are reminiscent of the thoracic LinA/LinB lineage in *Drosophila*, which generate MNs innervating the adult leg muscles. In the case of LinB lineage, the Hox gene Pb and other mTFs play an instructional role in giving unique axonal and dendritic arborization to three MNs of the lineage, thereby regulating the morphological diversity of the MNs (Enriquez et al., 2015). Interestingly, Pb was not required for the survival of these MNs. This underlines the importance of Hox in determining the uniqueness of neuronal morphology. This genetic control of the morphological diversification of MNs was also shown to be critical in their functional capability for flawless walking at high speed (Baek and Mann, 2009; Baek et al., 2013; Enriquez et al., 2015). The role of Hox genes in determining the morphology of the vertebrate MNs has been reported. However, in our limited knowledge, no similar functional correlation between MN morphology and behavior has been established so far in the vertebrates. The observations made in thoracic LinA MNs are closer to what is reported in the vertebrates. In the case of *Antp* or the Hox triple (Scr^{-}, $Antp^{-}$, Ubx^{-}) mutants LinA NBs, MNs were reported to undergo apoptosis. When the cell death was blocked, the surviving neurons took their fate as thoracic MNs. These MNs innervate the right target muscles and exhibit subtle arborization defects (Baek et al., 2013). This was similar to what was reported in the case of vertebrate. However, unlike vertebrates, the majority of the *Drosophila* LinA MNs do not show expression of more than one Hox factor, or Hox gene cross-regulation playing a central role in determining MN identity (Baek et al., 2013). Only in the case of the T3 segment, LinA MNs express Antp in larval stages and Ubx in pupal and adult stages (**Figures 2E–E'**) (Baek et al., 2013). The mutant analysis for these MNs suggested that Ubx expression represses Antp, and these two genes function redundantly in these cells of the T3 segment (Baek et al., 2013). None of the thoracic LinA MNs expressed Hox gene Pb (Baek et al., 2013). The apparent differences in the role of Hox genes in *Drosophila* compared to their elaborate role in specifying MN pool identity might be due to the complex limb musculature found in the vertebrates, which need a very refined control from MNs. It is reported that 11 Hox genes are required to diversify the MN pools, which innervate the muscles of anterior limbs alone (Dasen et al., 2005). On the other hand, *Drosophila* leg musculature is not as complex and therefore may

not require such complex transcriptional code to generate a large diversity of MNs. However, all these conclusions in *Drosophila* and their comparisons with vertebrates are based on studies done in LinA and LinB lineages, which constitutes only two-third of the leg MNs. It is possible that detailed analysis of other leg innervating MNs in *Drosophila* may give some additional insights (Truman et al., 2004; Baek and Mann, 2009).

The other *Drosophila* studies discussed here (summarized in **Table 1**) highlight the importance of Hox genes in setting up a molecular code for the functional assembly of neuromuscular networks (Friedrich et al., 2016; Hessinger et al., 2017). These studies also established that the requirement of Hox genes in the cells is not transient and restricted to the formation of the networks, but is chronic and is required for the maintenance and functioning of the networks much after they are established (Friedrich et al., 2016). At the cellular level, Hox genes have been shown to play a role in the survival of the MNs (Baek et al., 2013), their muscle innervation, and in determining their axonal and dendritic morphology (discussed above) (Baek et al., 2013; Enriquez et al., 2015). Notably, studies (with Antp) also established that the level of Hox genes in the adult MNs could regulate their axonal targeting and innervation onto the muscles, with low expressing MNs targeting proximal leg muscles and vice versa (Baek et al., 2013). The studies with BX-C miRNA emphasized the importance of maintaining a fine control over Hox expression in the MNs to establish a functional neuromuscular network and its role in executing the behavior (Garaulet et al., 2014; Picao-Osorio et al., 2015; Issa et al., 2019). More specifically, the miRNA-mediated control of Ubx expression on the neural activity of the SR MNs was the first instance where fine control over the levels of Hox gene has been shown to impact both neurophysiology and behavior. How exactly is this effect executed in MNs, and whether the miRNA-mediated regulation of Hox levels impacts other adult behaviors remains to be addressed.

Many roles discussed here go beyond the conventional developmental roles played by Hox genes in AP axis determination. These studies establish that in addition to giving the neurons their positional identity and the capacity to form the region-specific neural circuitry, Hox genes have a functional requirement in adult stages in regulating, at the very least, the morphology and neural activity of the MNs and their functions. Therefore these functions, to some extent, explain the sustained and robust neuronal expression of these genes post differentiation (Hirth et al., 1998; Technau et al., 2006). In order to further understand the role of Hox genes in the assembly of neuromuscular networks along the AP axis as well as their function beyond, there is a need to identify their targets in MNs. For instance, both Hox and Hth were similarly required for thoracic MNs to survive (Baek et al., 2013), but phenotypes like axonal and dendritic morphology differed when either Hox or Hth were individually removed (Baek et al., 2013; Enriquez et al., 2015). This supported the idea that distinct genetic programs downstream of Hox and Hth control axonal and dendritic morphology independent of each other. Therefore, identifying Hox and Exd/Hth targets specifically in MNs will be useful to understand their role in neuromuscular circuit assembly and their morphological diversification. Hox target genes have been identified in past using various approaches (Leemans et al., 2001; Barmina et al., 2005; Mohit et al., 2006; Hersh et al., 2007; Hueber et al., 2007; Agrawal et al., 2011; Choo et al., 2011; Pavlopoulos and Akam, 2011; Slattery et al., 2011; Sorge et al., 2012; Beh et al., 2016; Prasad et al., 2016; Shlyueva et al., 2016). However, none of these approaches were geared towards identifying the targets within CNS or its specific cell types. Identifying Hox targets in MNs may have been technically difficult so far, but a finer refinement of targeted DamID (TaDa) (Southall et al., 2013) to an elegant nano DAM technique (https://www.biorxiv.org/content/10.1101/2021.06.07.447332v2) may provide a useful mean for identifying MN specific targets genes downstream to these factors. Lastly, the distinct morphological phenotypes observed in MNs in *Antp* and *hth* mutants also suggest that Antp may be using cofactors other than canonical Hox cofactors (like Exd and Hth) (Baek et al., 2013). This is not unusual as Hox genes have been shown to use cooperative co-factors other than Exd/Hth (Gebelein et al., 2004; Ghosh et al., 2019), as well as novel collaborative co-factors in both neural and non-neural cell types (Gebelein et al., 2004; Mann et al., 2009; Baëza et al., 2015; Khandelwal et al., 2017; Bischof et al., 2018; Ghosh et al., 2019; Bakshi et al., 2020; Carnesecchi et al., 2020). However, the question remains whether any of these non-canonical cofactors facilitate Hox genes to carry out their conventional and newly discovered roles in MNs.

AUTHOR CONTRIBUTIONS

RJ conceptualized, researched and wrote the manuscript. RS researched and helped in manuscript writing and made the figures. AB researched and helped in manuscript writing and made the figures.

FUNDING

RJ acknowledges funding support from the Department of Science and Technology, India (CRG/2021/ 003275); Department of Biotechnology, India (BT/PR26385/MED/122/110/2017, BT/PR27455/BRB/10/1647/2018 and BT/PR41306/MED/122/259/2020); CDFD core funds and UGC, India (Fellowship award to AB) (UGC Ref No. 22/06/2014(i)EU-V, 2061430472), and ICMR, India (Fellowship award to RS) (ICMR Ref. No. 3/1/3/JRF-2012/HRD-63 (40260)).

REFERENCES

Agrawal, P., Habib, F., Yelagandula, R., and Shashidhara, L. S. (2011). Genome-level Identification of Targets of Hox Protein Ultrabithorax in Drosophila: Novel Mechanisms for Target Selection. *Sci. Rep.* 1, 205. doi:10.1038/srep00205

Akam, M. (1989). Hox and HOM: Homologous Gene Clusters in Insects and Vertebrates. *Cell* 57, 347–349. doi:10.1016/0092-8674(89)90909-4

Arya, R., Sarkissian, T., Tan, Y., and White, K. (2015). Neural Stem Cell Progeny Regulate Stem Cell Death in a Notch and Hox Dependent Manner. *Cell Death Differ* 22, 1378–1387. doi:10.1038/cdd.2014.235

Baek, M., Enriquez, J., and Mann, R. S. (2013). Dual Role for Hox Genes and Hox Co-factors in Conferring Leg Motoneuron Survival and Identity in Drosophila. *Development* 140, 2027–2038. doi:10.1242/dev.090902

Baek, M., and Mann, R. S. (2009). Lineage and Birth Date Specify Motor Neuron Targeting and Dendritic Architecture in Adult Drosophila. *J. Neurosci.* 29, 6904–6916. doi:10.1523/jneurosci.1585-09.2009

Baek, M., Menon, V., Jessell, T. M., Hantman, A. W., and Dasen, J. S. (2019). Molecular Logic of Spinocerebellar Tract Neuron Diversity and Connectivity. *Cel Rep.* 27, 2620–2635. doi:10.1016/j.celrep.2019.04.113

Baëza, M., Viala, S., Heim, M., Dard, A., Hudry, B., Duffraisse, M., et al. (2015). Inhibitory Activities of Short Linear Motifs Underlie Hox Interactome Specificity In Vivo. *eLife* 4, e06034. doi:10.7554/eLife.06034

Bahrampour, S., Jonsson, C., and Thor, S. (2019). Brain Expansion Promoted by Polycomb-Mediated Anterior Enhancement of a Neural Stem Cell Proliferation Program. *Plos Biol.* 17, e3000163. doi:10.1371/journal.pbio.3000163

Bakshi, A., Sipani, R., Ghosh, N., and Joshi, R. (2020). Sequential Activation of Notch and Grainyhead Gives Apoptotic Competence to Abdominal-B Expressing Larval Neuroblasts in Drosophila Central Nervous System. *Plos Genet.* 16, e1008976. doi:10.1371/journal.pgen.1008976

Barmina, O., Gonzalo, M., McIntyre, L. M., and Kopp, A. (2005). Sex- and Segment-specific Modulation of Gene Expression Profiles in Drosophila. *Develop. Biol.* 288, 528–544. doi:10.1016/j.ydbio.2005.09.052

Baumgardt, M., Karlsson, D., Salmani, B. Y., Bivik, C., MacDonald, R. B., Gunnar, E., et al. (2014). Global Programmed Switch in Neural Daughter Cell Proliferation Mode Triggered by a Temporal Gene cascade. *Develop. Cel.* 30, 192–208. doi:10.1016/j.devcel.2014.06.021

Becker, H., Renner, S., Technau, G. M., and Berger, C. (2016). Cell-Autonomous and Non-cell-autonomous Function of Hox Genes Specify Segmental Neuroblast Identity in the Gnathal Region of the Embryonic CNS in Drosophila. *Plos Genet.* 12, e1005961. doi:10.1371/journal.pgen.1005961

Beh, C. Y., El-Sharnouby, S., Chatzipli, A., Russell, S., Choo, S. W., and White, R. (2016). Roles of Cofactors and Chromatin Accessibility in Hox Protein Target Specificity. *Epigenetics & Chromatin* 9, 1. doi:10.1186/s13072-015-0049-x

Bello, B. C., Hirth, F., and Gould, A. P. (2003). A Pulse of the Drosophila Hox Protein Abdominal-A Schedules the End of Neural Proliferation via Neuroblast Apoptosis. *Neuron* 37, 209–219. doi:10.1016/s0896-6273(02)01181-9

Bender, W. (2008). MicroRNAs in the Drosophila Bithorax Complex. *Genes Dev.* 22, 14–19. doi:10.1101/gad.1614208

Berger, C., Pallavi, S. K., Prasad, M., Shashidhara, L. S., and Technau, G. M. (2005a). A Critical Role for Cyclin E in Cell Fate Determination in the central Nervous System of Drosophila melanogaster. *Nat. Cel Biol* 7, 56–62. doi:10.1038/ncb1203

Berger, C., Pallavi, S. K., Prasad, M., Shashidhara, L. S., and Technau, G. M. (2005b). Cyclin E Acts under the Control of Hox-Genes as a Cell Fate Determinant in the Developing central Nervous System. *Cell Cycle* 4, 422–425. doi:10.4161/cc.4.3.1524

Bhat, K. M. (1999). Segment Polarity Genes in Neuroblast Formation and Identity Specification during Drosophila Neurogenesis. *Bioessays* 21, 472–485. doi:10.1002/(sici)1521-1878(199906)21:6<472::aid-bies4>3.0.co;2-w

Birkholz, O., Rickert, C., Berger, C., Urbach, R., and Technau, G. M. (2013). Neuroblast Pattern and Identity in the Drosophila Tail Region and Role of Doublesex in the Survival of Sex-specific Precursors. *Development* 140, 1830–1842. doi:10.1242/dev.090043

Birkholz, O., Vef, O., Rogulja-Ortmann, A., Berger, C., and Technau, G. M. (2013). Abdominal-B and Caudal Inhibit the Formation of Specific Neuroblasts in the Drosophila Tail Region. *Development* 140, 3552–3564. doi:10.1242/dev.096099

Bischof, J., Duffraisse, M., Furger, E., Ajuria, L., Giraud, G., Vanderperre, S., et al. (2018). Generation of a Versatile BiFC ORFeome Library for Analyzing Protein-Protein Interactions in Live Drosophila. *eLife* 7, e38853. doi:10.7554/eLife.38853

Bossing, T., Udolph, G., Doe, C. Q., and Technau, G. M. (1996). The Embryonic Central Nervous System Lineages of Drosophila Melanogaster. *Develop. Biol.* 179, 41–64. doi:10.1006/dbio.1996.0240

Brierley, D. J., Rathore, K., VijayRaghavan, K., and Williams, D. W. (2012). Developmental Origins and Architecture of Drosophila Leg Motoneurons. *J. Comp. Neurol.* 520, 1629–1649. doi:10.1002/cne.23003

Bussell, J. J., Yapici, N., Zhang, S. X., Dickson, B. J., and Vosshall, L. B. (2014). Abdominal-B Neurons Control Drosophila virgin Female Receptivity. *Curr. Biol.* 24, 1584–1595. doi:10.1016/j.cub.2014.06.011

Carnesecchi, J., Sigismondo, G., Domsch, K., Baader, C. E. P., Rafiee, M.-R., Krijgsveld, J., et al. (2020). Multi-level and Lineage-specific Interactomes of the Hox Transcription Factor Ubx Contribute to its Functional Specificity. *Nat. Commun.* 11, 1388. doi:10.1038/s41467-020-15223-x

Carroll, S. B. (1995). Homeotic Genes and the Evolution of Arthropods and Chordates. *Nature* 376, 479–485. doi:10.1038/376479a0

Castellanos, M. C., Tang, J. C. Y., and Allan, D. W. (2013). Female-biased Dimorphism Underlies a Female-specific Role for post-embryonic Ilp7 Neurons in Drosophila Fertility. *Development* 140, 3915–3926. doi:10.1242/dev.094714

Chen, T.-W., Wardill, T. J., Sun, Y., Pulver, S. R., Renninger, S. L., Baohan, A., et al. (2013). Ultrasensitive Fluorescent Proteins for Imaging Neuronal Activity. *Nature* 499, 295–300. doi:10.1038/nature12354

Choo, S. W., White, R., and Russell, S. (2011). Genome-wide Analysis of the Binding of the Hox Protein Ultrabithorax and the Hox Cofactor Homothorax in Drosophila. *PloS one* 6, e14778. doi:10.1371/journal.pone.0014778

Dasen, J. S., De Camilli, A., Wang, B., Tucker, P. W., and Jessell, T. M. (2008). Hox Repertoires for Motor Neuron Diversity and Connectivity Gated by a Single Accessory Factor, FoxP1. *Cell* 134, 304–316. doi:10.1016/j.cell.2008.06.019

Dasen, J. S., Tice, B. C., Brenner-Morton, S., and Jessell, T. M. (2005). A Hox Regulatory Network Establishes Motor Neuron Pool Identity and Target-Muscle Connectivity. *Cell* 123, 477–491. doi:10.1016/j.cell.2005.09.009

Dawkins, R., and Dawkins, M. (1976). Hierachical Organization and Postural Facilitation: Rules for Grooming in Flies. *Anim. Behav.* 24, 739–755.

di Sanguinetto, S. A. D. T., Dasen, J. S., and Arber, S. (2008). Transcriptional Mechanisms Controlling Motor Neuron Diversity and Connectivity. *Curr. Opin. Neurobiol.* 18, 36–43. doi:10.1016/j.conb.2008.04.002

Dickinson, M. H., Farley, C. T., Full, R. J., Koehl, M. A. R., Kram, R., and Lehman, S. (2000). How Animals Move: an Integrative View. *Science* 288, 100–106. doi:10.1126/science.288.5463.100

Dixit, R., Vijayraghavan, K., and Bate, M. (2008). Hox Genes and the Regulation of Movement inDrosophila. *Devel Neurobio* 68, 309–316. doi:10.1002/dneu.20589

Doe, C. Q. (1992). Molecular Markers for Identified Neuroblasts and Ganglion Mother Cells in the Drosophila central Nervous System. *Development* 116, 855–863. doi:10.1242/dev.116.4.855

Doe, C. Q. (2017). Temporal Patterning in the Drosophila CNS. *Annu. Rev. Cel Dev. Biol.* 33, 219–240. doi:10.1146/annurev-cellbio-111315-125210

Dutta, D., Umashankar, M., Lewis, E. B., Rodrigues, V., and Vijayraghavan, K. (2010). Hox Genes Regulate Muscle Founder Cell Pattern Autonomously and Regulate Morphogenesis through Motor Neurons. *J. Neurogenet.* 24, 95–108. doi:10.3109/01677063.2010.494317

Enriquez, J., Venkatasubramanian, L., Baek, M., Peterson, M., Aghayeva, U., and Mann, R. S. (2015). Specification of Individual Adult Motor Neuron Morphologies by Combinatorial Transcription Factor Codes. *Neuron* 86, 955–970. doi:10.1016/j.neuron.2015.04.011

Estacio-Gómez, A., Moris-Sanz, M., Schäfer, A.-K., Perea, D., Herrero, P., and Díaz-Benjumea, F. J. (2013). Bithorax-complex Genes Sculpt the Pattern of Leucokinergic Neurons in the Drosophila central Nervous System. *Development* 140, 2139–2148. doi:10.1242/dev.090423

Faisal, A. A., and Matheson, T. (2001). Coordinated Righting Behaviour in Locusts. *J. Exp. Biol.* 204, 637–648. doi:10.1242/jeb.204.4.637

Friedrich, J., Sorge, S., Bujupi, F., Eichenlaub, M. P., Schulz, N. G., Wittbrodt, J., et al. (2016). Hox Function Is Required for the Development and Maintenance of the Drosophila Feeding Motor Unit. *Cel Rep.* 14, 850–860. doi:10.1016/j.celrep.2015.12.077

Garaulet, D. L., Castellanos, M. C., Bejarano, F., Sanfilippo, P., Tyler, D. M., Allan, D. W., et al. (2014). Homeotic Function of Drosophila Bithorax-Complex miRNAs Mediates Fertility by Restricting Multiple Hox Genes and TALE Cofactors in the CNS. *Develop. Cel.* 29, 635–648. doi:10.1016/j.devcel.2014.04.023

Garaulet, D. L., Zhang, B., Wei, L., Li, E., and Lai, E. C. (2020). miRNAs and Neural Alternative Polyadenylation Specify the Virgin Behavioral State. *Develop. Cel.* 54, 410–423. doi:10.1016/j.devcel.2020.06.004

Gaufo, G. O., Wu, S., and Capecchi, M. R. (2004). Contribution of Hox Genes to the Diversity of the Hindbrain Sensory System. *Development* 131, 1259–1266. doi:10.1242/dev.01029

Gebelein, B., McKay, D. J., and Mann, R. S. (2004). Direct Integration of Hox and Segmentation Gene Inputs during Drosophila Development. *Nature* 431, 653–659. doi:10.1038/nature02946

Ghosh, N., Bakshi, A., Khandelwal, R., Rajan, S. G., and Joshi, R. (2019). The Hox Gene Abdominal-B Uses DoublesexF as a Cofactor to Promote Neuroblast Apoptosis in the Drosophila central Nervous System. *Development* 146, dev175158. doi:10.1242/dev.175158

Gummalla, M., Maeda, R. K., Castro Alvarez, J. J., Gyurkovics, H., Singari, S., Edwards, K. A., et al. (2012). Abd-A Regulation by the Iab-8 Noncoding RNA. *Plos Genet.* 8, e1002720. doi:10.1371/journal.pgen.1002720

Hadjieconomou, D., Rotkopf, S., Alexandre, C., Bell, D. M., Dickson, B. J., and Salecker, I. (2011). Flybow: Genetic Multicolor Cell Labeling for Neural Circuit Analysis in Drosophila melanogaster. *Nat. Methods* 8, 260–266. doi:10.1038/nmeth.1567

Hales, K. G., Korey, C. A., Larracuente, A. M., and Roberts, D. M. (2015). Genetics on the Fly: A Primer on the Drosophila Model System. *Genetics* 201, 815–842. doi:10.1534/genetics.115.183392

Hamada, F. N., Rosenzweig, M., Kang, K., Pulver, S. R., Ghezzi, A., Jegla, T. J., et al. (2008). An Internal thermal Sensor Controlling Temperature Preference in Drosophila. *Nature* 454, 217–220. doi:10.1038/nature07001

Hart, C., Awgulewitsch, A., Fainsod, A., McGinnis, W., and Ruddle, F. H. (1985). Homeo Box Gene Complex on Mouse Chromosome 11: Molecular Cloning, Expression in Embryogenesis, and Homology to a Human Homeo Box Locus. *Cell* 43, 9–18. doi:10.1016/0092-8674(85)90007-8

Hartenstein, V., and Wodarz, A. (2013). Initial Neurogenesis in Drosophila. *Wiley Interdiscip. Rev. Dev. Biol.* 2, 701–721. doi:10.1002/wdev.111

Hersh, B. M., Nelson, C. E., Stoll, S. J., Norton, J. E., Albert, T. J., and Carroll, S. B. (2007). The UBX-Regulated Network in the Haltere Imaginal Disc of D. melanogaster. *Develop. Biol.* 302, 717–727. doi:10.1016/j.ydbio.2006.11.011

Hessinger, C., Technau, G. M., and Rogulja-Ortmann, A. (2017). The Drosophila Hox Gene Ultrabithorax Acts in Both Muscles and Motoneurons to Orchestrate Formation of Specific Neuromuscular Connections. *Development* 144, 139–150. doi:10.1242/dev.143875

Hirth, F., Hartmann, B., and Reichert, H. (1998). Homeotic Gene Action in Embryonic Brain Development of Drosophila. *Development* 125, 1579–1589. doi:10.1242/dev.125.9.1579

Holstege, J. C., de Graaff, W., Hossaini, M., Cano, S. C., Jaarsma, D., van den Akker, E., et al. (2008). Loss of Hoxb8 Alters Spinal Dorsal Laminae and Sensory Responses in Mice. *Proc. Natl. Acad. Sci.* 105, 6338–6343. doi:10.1073/pnas.0802176105

Homem, C. C. F., and Knoblich, J. A. (2012). Drosophila Neuroblasts: a Model for Stem Cell Biology. *Development* 139, 4297–4310. doi:10.1242/dev.080515

Huber, L., Ferdin, M., Holzmann, J., Stubbusch, J., and Rohrer, H. (2012). HoxB8 in Noradrenergic Specification and Differentiation of the Autonomic Nervous System. *Develop. Biol.* 363, 219–233. doi:10.1016/j.ydbio.2011.12.026

Hückesfeld, S., Schoofs, A., Schlegel, P., Miroschnikow, A., and Pankratz, M. J. (2015). Localization of Motor Neurons and Central Pattern Generators for Motor Patterns Underlying Feeding Behavior in Drosophila Larvae. *PloS one* 10, e0135011. doi:10.1371/journal.pone.0135011

Hudry, B., Remacle, S., Delfini, M.-C., Rezsohazy, R., Graba, Y., and Merabet, S. (2012). Hox Proteins Display a Common and Ancestral Ability to Diversify Their Interaction Mode with the PBC Class Cofactors. *Plos Biol.* 10, e1001351. doi:10.1371/journal.pbio.1001351

Hueber, S. D., Bezdan, D., Henz, S. R., Blank, M., Wu, H., and Lohmann, I. (2007). Comparative Analysis of Hox Downstream Genes inDrosophila. *Development* 134, 381–392. doi:10.1242/dev.02746

Inaki, M., Yoshikawa, S., Thomas, J. B., Aburatani, H., and Nose, A. (2007). Wnt4 Is a Local Repulsive Cue that Determines Synaptic Target Specificity. *Curr. Biol.* 17, 1574–1579. doi:10.1016/j.cub.2007.08.013

Issa, A. R., Picao-Osorio, J., Rito, N., Chiappe, M. E., and Alonso, C. R. (2019). A Single MicroRNA-Hox Gene Module Controls Equivalent Movements in Biomechanically Distinct Forms of Drosophila. *Curr. Biol.* 29, 2665–2675. doi:10.1016/j.cub.2019.06.082

Isshiki, T., Pearson, B., Holbrook, S., and Doe, C. Q. (2001). Drosophila Neuroblasts Sequentially Express Transcription Factors Which Specify the Temporal Identity of Their Neuronal Progeny. *Cell* 106, 511–521. doi:10.1016/s0092-8674(01)00465-2

Jessell, T. M. (2000). Neuronal Specification in the Spinal Cord: Inductive Signals and Transcriptional Codes. *Nat. Rev. Genet.* 1, 20–29. doi:10.1038/35049541

Jusufi, A., Zeng, Y., Full, R. J., and Dudley, R. (2011). Aerial Righting Reflexes in Flightless Animals. *Integr. Comp. Biol.* 51, 937–943. doi:10.1093/icb/icr114

Kannan, R., Berger, C., Myneni, S., Technau, G. M., and Shashidhara, L. S. (2010). Abdominal-A Mediated Repression of Cyclin E Expression during Cell-Fate Specification in the Drosophila central Nervous System. *Mech. Dev.* 127, 137–145. doi:10.1016/j.mod.2009.09.008

Kaplan, W. D., and Trout, W. E. (1974). Genetic Manipulation of an Abnormal Jump Response in Drosophila. *Genetics* 77, 721–739. doi:10.1093/genetics/77.4.721

Karlsson, D., Baumgardt, M., and Thor, S. (2010). Segment-specific Neuronal Subtype Specification by the Integration of Anteroposterior and Temporal Cues. *Plos Biol.* 8, e1000368. doi:10.1371/journal.pbio.1000368

Kaufman, T. C., Seeger, M. A., and Olsen, G. (1990). Molecular and Genetic Organization of the Antennapedia Gene Complex of Drosophila melanogaster. *Adv. Genet.* 27, 309–362. doi:10.1016/s0065-2660(08)60029-2

Khandelwal, R., Sipani, R., Govinda Rajan, S., Kumar, R., and Joshi, R. (2017). Combinatorial Action of Grainyhead, Extradenticle and Notch in Regulating Hox Mediated Apoptosis in Drosophila Larval CNS. *Plos Genet.* 13, e1007043. doi:10.1371/journal.pgen.1007043

Kitamoto, T. (2001). Conditional Modification of Behavior inDrosophila by Targeted Expression of a Temperature-Sensitiveshibire Allele in Defined Neurons. *J. Neurobiol.* 47, 81–92. doi:10.1002/neu.1018

Korona, D., Koestler, S. A., and Russell, S. (2017). Engineering the Drosophila Genome for Developmental Biology. *J. Dev. Biol.* 5, 16. doi:10.3390/jdb5040016

Kuert, P. A., Bello, B. C., and Reichert, H. (2012). The Labial Gene Is Required to Terminate Proliferation of Identified Neuroblasts in Postembryonic Development of the Drosophila Brain. *Biol. Open* 1, 1006–1015. doi:10.1242/bio.20121966

Kuert, P. A., Hartenstein, V., Bello, B. C., Lovick, J. K., and Reichert, H. (2014). Neuroblast Lineage Identification and Lineage-specific Hox Gene Action during Postembryonic Development of the Subesophageal Ganglion in the Drosophila central Brain. *Develop. Biol.* 390, 102–115. doi:10.1016/j.ydbio.2014.03.021

Lacombe, J., Hanley, O., Jung, H., Philippidou, P., Surmeli, G., Grinstein, J., et al. (2013). Genetic and Functional Modularity of Hox Activities in the Specification of Limb-Innervating Motor Neurons. *Plos Genet.* 9, e1003184. doi:10.1371/journal.pgen.1003184

LaCount, D. J., Hanson, S. F., Schneider, C. L., and Friesen, P. D. (2000). Caspase Inhibitor P35 and Inhibitor of Apoptosis Op-IAP Block In Vivo Proteolytic Activation of an Effector Caspase at Different Steps. *J. Biol. Chem.* 275, 15657–15664. doi:10.1074/jbc.m000791200

Landgraf, M., Bossing, T., Technau, G. M., and Bate, M. (1997). The Origin, Location, and Projections of the Embryonic Abdominal Motorneurons ofDrosophila. *J. Neurosci.* 17, 9642–9655. doi:10.1523/jneurosci.17-24-09642.1997

Landmesser, L. T. (2001). The Acquisition of Motoneuron Subtype Identity and Motor Circuit Formation. *Int. J. Dev. Neurosci.* 19, 175–182. doi:10.1016/s0736-5748(00)00090-3

Leemans, R., Loop, T., Egger, B., He, H., Kammermeier, L., Hartmann, B., et al. (2001). Identification of Candidate Downstream Genes for the Homeodomain Transcription Factor Labial in Drosophila through Oligonucleotide-Array Transcript Imaging. *Genome Biol.* 2, RESEARCH0015. doi:10.1186/gb-2001-2-5-research0015

Lelli, K. M., Noro, B., and Mann, R. S. (2011). Variable Motif Utilization in Homeotic Selector (Hox)-Cofactor Complex Formation Controls Specificity. *Proc. Natl. Acad. Sci.* 108, 21122–21127. doi:10.1073/pnas.1114118109

Lemons, D., Paré, A., and McGinnis, W. (2012). Three Drosophila Hox Complex microRNAs Do Not Have Major Effects on Expression of Evolutionarily Conserved Hox Gene Targets during Embryogenesis. *PloS one* 7, e31365. doi:10.1371/journal.pone.0031365

Lin, A. W., and Carpenter, E. M. (2003). Hoxa10 andHoxd10 Coordinately Regulate Lumbar Motor Neuron Patterning. *J. Neurobiol.* 56, 328–337. doi:10.1002/neu.10239

Maeda, R. K., and Karch, F. (2006). The ABC of the BX-C: the Bithorax Complex Explained. *Development* 133, 1413–1422. doi:10.1242/dev.02323

Mann, R. S., and Affolter, M. (1998). Hox Proteins Meet More Partners. *Curr. Opin. Genet. Develop.* 8, 423–429. doi:10.1016/s0959-437x(98)80113-5

Mann, R. S., and Chan, S.-K. (1996). Extra Specificity from Extradenticle: the Partnership between HOX and PBX/EXD Homeodomain Proteins. *Trends Genet.* 12, 258–262. doi:10.1016/0168-9525(96)10026-3

Mann, R. S., Lelli, K. M., and Joshi, R. (2009). Hox Specificity Unique Roles for Cofactors and Collaborators. *Curr. Top. Dev. Biol.* 88, 63–101. doi:10.1016/s0070-2153(09)88003-4

Martín, F., and Alcorta, E. (2017). Novel Genetic Approaches to Behavior in Drosophila. *J. Neurogenet.* 31, 288–299. doi:10.1080/01677063.2017.1395875

McGinnis, W., and Krumlauf, R. (1992). Homeobox Genes and Axial Patterning. *Cell* 68, 283–302. doi:10.1016/0092-8674(92)90471-n

Merabet, S., Saadaoui, M., Sambrani, N., Hudry, B., Pradel, J., Affolter, M., et al. (2007). A Unique Extradenticle Recruitment Mode in the Drosophila Hox Protein Ultrabithorax. *Proc. Natl. Acad. Sci.* 104, 16946–16951. doi:10.1073/pnas.0705832104

Merrill, V. K. L., Turner, F. R., and Kaufman, T. C. (1987). A Genetic and Developmental Analysis of Mutations in the Deformed Locus in Drosophila melanogaster. *Develop. Biol.* 122, 379–395. doi:10.1016/0012-1606(87)90303-4

Michelson, A. M. (1994). Muscle Pattern Diversification in Drosophila Is Determined by the Autonomous Function of Homeotic Genes in the Embryonic Mesoderm. *Development* 120, 755–768. doi:10.1242/dev.120.4.755

Miguel-Aliaga, I., and Thor, S. (2004). Segment-specific Prevention of pioneer Neuron Apoptosis by Cell-Autonomous, Postmitotic Hox Gene Activity. *Development* 131, 6093–6105. doi:10.1242/dev.01521

Miguez, A., Ducret, S., Di Meglio, T., Parras, C., Hmidan, H., Haton, C., et al. (2012). Opposing Roles for Hoxa2 and Hoxb2 in Hindbrain Oligodendrocyte Patterning. *J. Neurosci.* 32, 17172–17185. doi:10.1523/JNEUROSCI.0885-12.2012

Miroschnikow, A., Schlegel, P., and Pankratz, M. J. (2020). Making Feeding Decisions in the Drosophila Nervous System. *Curr. Biol.* 30, R831–R840. doi:10.1016/j.cub.2020.06.036

Moens, C. B., and Selleri, L. (2006). Hox Cofactors in Vertebrate Development. *Develop. Biol.* 291, 193–206. doi:10.1016/j.ydbio.2005.10.032

Mohit, P., Makhijani, K., Madhavi, M. B., Bharathi, V., Lal, A., Sirdesai, G., et al. (2006). Modulation of AP and DV Signaling Pathways by the Homeotic Gene Ultrabithorax during Haltere Development in Drosophila. *Develop. Biol.* 291, 356–367. doi:10.1016/j.ydbio.2005.12.022

Monedero Cobeta, I., Salmani, B. Y., and Thor, S. (2017). Anterior-Posterior Gradient in Neural Stem and Daughter Cell Proliferation Governed by Spatial and Temporal Hox Control. *Curr. Biol.* 27, 1161–1172. doi:10.1016/j.cub.2017.03.023

Neckameyer, W. S., and Bhatt, P. (2016). Protocols to Study Behavior in Drosophila. *Methods Mol. Biol.* 1478, 303–320. doi:10.1007/978-1-4939-6371-3_19

Nichols, C. D., Becnel, J., and Pandey, U. B. (2012). Methods to Assay Drosophila Behavior. *J. visualized experiments : JoVE* 61, 3795. doi:10.3791/3795

Nose, A., Takeichi, M., and Goodman, C. S. (1994). Ectopic Expression of Connectin Reveals a Repulsive Function during Growth Cone Guidance and Synapse Formation. *Neuron* 13, 525–539. doi:10.1016/0896-6273(94)90023-x

Pattyn, A., Vallstedt, A., Dias, J. M., Samad, O. A., Krumlauf, R., Rijli, F. M., et al. (2003). Coordinated Temporal and Spatial Control of Motor Neuron and Serotonergic Neuron Generation from a Common Pool of CNS Progenitors. *Genes Dev.* 17, 729–737. doi:10.1101/gad.255803

Pavlopoulos, A., and Akam, M. (2011). Hox Gene Ultrabithorax Regulates Distinct Sets of Target Genes at Successive Stages of Drosophila Haltere Morphogenesis. *Proc. Natl. Acad. Sci.* 108, 2855–2860. doi:10.1073/pnas.1015077108

Penn, D., and Jane Brockmann, H. (1995). Age-biased Stranding and Righting in Male Horseshoe Crabs, Limulus polyphemus. *Anim. Behav.* 49, 1531–1539. doi:10.1016/0003-3472(95)90074-8

Pereanu, W., Spindler, S., Im, E., Buu, N., and Hartenstein, V. (2007). The Emergence of Patterned Movement during Late Embryogenesis ofDrosophila. *Devel Neurobio.* 67, 1669–1685. doi:10.1002/dneu.20538

Philippidou, P., and Dasen, J. S. (2013). Hox Genes: Choreographers in Neural Development, Architects of Circuit Organization. *Neuron* 80, 12–34. doi:10.1016/j.neuron.2013.09.020

Picao-Osorio, J., Johnston, J., Landgraf, M., Berni, J., and Alonso, C. R. (2015). MicroRNA-encoded Behavior in Drosophila. *Science* 350, 815–820. doi:10.1126/science.aad0217

Picao-Osorio, J., Lago-Baldaia, I., Patraquim, P., and Alonso, C. R. (2017). Pervasive Behavioral Effects of MicroRNA Regulation in Drosophila. *Genetics* 206, 1535–1548. doi:10.1534/genetics.116.195776

Pool, A.-H., and Scott, K. (2014). Feeding Regulation in Drosophila. *Curr. Opin. Neurobiol.* 29, 57–63. doi:10.1016/j.conb.2014.05.008

Prasad, N., Tarikere, S., Khanale, D., Habib, F., and Shashidhara, L. S. (2016). A Comparative Genomic Analysis of Targets of Hox Protein Ultrabithorax Amongst Distant Insect Species. *Sci. Rep.* 6, 27885. doi:10.1038/srep27885

Prokop, A., Bray, S., Harrison, E., and Technau, G. M. (1998). Homeotic Regulation of Segment-specific Differences in Neuroblast Numbers and Proliferation in the Drosophila central Nervous System. *Mech. Dev.* 74, 99–110. doi:10.1016/s0925-4773(98)00068-9

Prokop, A., and Technau, G. M. (1994). Early Tagma-specific Commitment of Drosophila CNS Progenitor NB1-1. *Development* 120, 2567–2578. doi:10.1242/dev.120.9.2567

Regulski, M., Harding, K., Kostriken, R., Karch, F., Levine, M., and McGinnis, W. (1985). Homeo Box Genes of the Antennapedia and Bithorax Complexes of Drosophila. *Cell* 43, 71–80. doi:10.1016/0092-8674(85)90013-3

Regulski, M., McGinnis, N., Chadwick, R., and McGinnis, W. (1987). Developmental and Molecular Analysis of Deformed ; a Homeotic Gene Controlling Drosophila Head Development. *EMBO J.* 6, 767–777. doi:10.1002/j.1460-2075.1987.tb04819.x

Restifo, L. L., and Merrill, V. K. L. (1994). Two Drosophila Regulatory Genes, Deformed and the Broad-Complex, Share Common Functions in Development of Adult CNS, Head, and Salivary Glands. *Develop. Biol.* 162, 465–485. doi:10.1006/dbio.1994.1102

Rodríguez-Valentín, R., López-González, I., Jorquera, R., Labarca, P., Zurita, M., and Reynaud, E. (2006). Oviduct Contraction inDrosophilais Modulated by a Neural Network that Is Both, Octopaminergic and Glutamatergic. *J. Cel. Physiol.* 209, 183–198. doi:10.1002/jcp.20722

Rogulja-Ortmann, A., Renner, S., and Technau, G. M. (2008). Antagonistic Roles for Ultrabithorax and Antennapediain Regulating Segment-specific Apoptosis of Differentiated Motoneurons in the Drosophila Embryonic central Nervous System. *Development* 135, 3435–3445. doi:10.1242/dev.023986

Ronshaugen, M., Biemar, F., Piel, J., Levine, M., and Lai, E. C. (2005). The Drosophila microRNA Iab-4 Causes a Dominant Homeotic Transformation of Halteres to Wings. *Genes Dev.* 19, 2947–2952. doi:10.1101/gad.1372505

Saadaoui, M., Merabet, S., Litim-Mecheri, I., Arbeille, E., Sambrani, N., Damen, W., et al. (2011). Selection of Distinct Hox-Extradenticle Interaction Modes fine-tunes Hox Protein Activity. *Proc. Natl. Acad. Sci.* 108, 2276–2281. doi:10.1073/pnas.1006964108

Sánchez-Herrero, E., Vernós, I., Marco, R., and Morata, G. (1985). Genetic Organization of Drosophila Bithorax Complex. *Nature* 313, 108–113.

Schlegel, P., Costa, M., and Jefferis, G. S. (2017). Learning from Connectomics on the Fly. *Curr. Opin. Insect Sci.* 24, 96–105. doi:10.1016/j.cois.2017.09.011

Schmid, A., Chiba, A., and Doe, C. Q. (1999). Clonal Analysis of Drosophila Embryonic Neuroblasts: Neural Cell Types, Axon Projections and Muscle Targets. *Development* 126, 4653–4689. doi:10.1242/dev.126.21.4653

Schoofs, A., Niederegger, S., van Ooyen, A., Heinzel, H.-G., and Spiess, R. (2010). The Brain Can Eat: Establishing the Existence of a central Pattern Generator for Feeding in Third Instar Larvae of Drosophila virilis and Drosophila melanogaster. *J. Insect Physiol.* 56, 695–705. doi:10.1016/j.jinsphys.2009.12.008

Shah, V., Drill, E., and Lance-Jones, C. (2004). Ectopic Expression of Hoxd10 in Thoracic Spinal Segments Induces Motoneurons with a Lumbosacral Molecular Profile and Axon Projections to the Limb. *Dev. Dyn.* 231, 43–56. doi:10.1002/dvdy.20103

Shlyueva, D., Meireles-Filho, A. C. A., Pagani, M., and Stark, A. (2016). Genome-Wide Ultrabithorax Binding Analysis Reveals Highly Targeted Genomic Loci at Developmental Regulators and a Potential Connection to Polycomb-Mediated Regulation. *PloS one* 11, e0161997. doi:10.1371/journal.pone.0161997

Skeath, J. B. (1999). At the Nexus between Pattern Formation and Cell-type Specification: the Generation of Individual Neuroblast Fates in the Drosophila Embryonic central Nervous System. *Bioessays* 21, 922–931. doi:10.1002/(sici)1521-1878(199911)21:11<922::aid-bies4>3.0.co;2-t

Slattery, M., Ma, L., Négre, N., White, K. P., and Mann, R. S. (2011). Genome-wide Tissue-specific Occupancy of the Hox Protein Ultrabithorax and Hox Cofactor Homothorax in Drosophila. *PloS one* 6, e14686. doi:10.1371/journal.pone.0014686

Sorge, S., Ha, N., Polychronidou, M., Friedrich, J., Bezdan, D., Kaspar, P., et al. (2012). Thecis-regulatory Code of Hox Function inDrosophila. *EMBO J.* 31, 3323–3333. doi:10.1038/emboj.2012.179

Southall, T. D., Gold, K. S., Egger, B., Davidson, C. M., Caygill, E. E., Marshall, O. J., et al. (2013). Cell-type-specific Profiling of Gene Expression and Chromatin Binding without Cell Isolation: Assaying RNA Pol II Occupancy in Neural Stem Cells. *Develop. Cel.* 26, 101–112. doi:10.1016/j.devcel.2013.05.020

Stark, A., Bushati, N., Jan, C. H., Kheradpour, P., Hodges, E., Brennecke, J., et al. (2008). A Single Hox Locus in Drosophila Produces Functional microRNAs from Opposite DNA Strands. *Genes Dev.* 22, 8–13. doi:10.1101/gad.1613108

Stockinger, P., Kvitsiani, D., Rotkopf, S., Tirián, L., and Dickson, B. J. (2005). Neural Circuitry that Governs Drosophila Male Courtship Behavior. *Cell* 121, 795–807. doi:10.1016/j.cell.2005.04.026

Suska, A., Miguel-Aliaga, I., and Thor, S. (2011). Segment-specific Generation of Drosophila Capability Neuropeptide Neurons by Multi-Faceted Hox Cues. *Develop. Biol.* 353, 72–80. doi:10.1016/j.ydbio.2011.02.015

Szebenyi, A. L. (1969). Cleaning Behaviour in Drosophila melanogaster. *Anim. Behav.* 17, 641–651. doi:10.1016/s0003-3472(69)80006-0

Technau, G. M., Berger, C., and Urbach, R. (2006). Generation of Cell Diversity and Segmental Pattern in the Embryonic central Nervous System ofDrosophila. *Dev. Dyn.* 235, 861–869. doi:10.1002/dvdy.20566

Technau, G. M., Rogulja-Ortmann, A., Berger, C., Birkholz, O., and Rickert, C. (2014). Composition of a Neuromere and its Segmental Diversification under the Control ofHoxGenes in the Embryonic CNS ofDrosophila. *J. Neurogenet.* 28, 171–180. doi:10.3109/01677063.2013.868459

Thomsen, S., Azzam, G., Kaschula, R., Williams, L. S., and Alonso, C. R. (2010). Developmental RNA Processing of 3′UTRs in Hox mRNAs as a Context-dependent Mechanism Modulating Visibility to microRNAs. *Development* 137, 2951–2960. doi:10.1242/dev.047324

Tiong, S., Bone, L. M., and Whittle, J. R. S. (1985). Recessive Lethal Mutations within the Bithorax-Complex in Drosophila. *Mol. Gen. Genet.* 200, 335–342. doi:10.1007/bf00425445

Tiret, L., Le Mouellic, H., Maury, M., and Brulet, P. (1998). Increased Apoptosis of Motoneurons and Altered Somatotopic Maps in the Brachial Spinal Cord of Hoxc-8-Deficient Mice. *Development* 125, 279–291. doi:10.1242/dev.125.2.279

Trimarchi, J. R., and Schneiderman, A. M. (1993). Giant Fiber Activation of an Intrinsic Muscle in the Mesothoracic Leg of Drosophila melanogaster. *J. Exp. Biol.* 177, 149–167. doi:10.1242/jeb.177.1.149

Truman, J. W., and Bate, M. (1988). Spatial and Temporal Patterns of Neurogenesis in the central Nervous System of Drosophila melanogaster. *Develop. Biol.* 125, 145–157. doi:10.1016/0012-1606(88)90067-x

Truman, J. W., Schuppe, H., Shepherd, D., and Williams, D. W. (2004). Developmental Architecture of Adult-specific Lineages in the Ventral CNS of Drosophila. *Development* 131, 5167–5184. doi:10.1242/dev.01371

Tsuji, T., Hasegawa, E., and Isshiki, T. (2008). Neuroblast Entry into Quiescence Is Regulated Intrinsically by the Combined Action of Spatial Hox Proteins and Temporal Identity Factors. *Development* 135, 3859–3869. doi:10.1242/dev.025189

Tyler, D. M., Okamura, K., Chung, W.-J., Hagen, J. W., Berezikov, E., Hannon, G. J., et al. (2008). Functionally Distinct Regulatory RNAs Generated by Bidirectional Transcription and Processing of microRNA Loci. *Genes Dev.* 22, 26–36. doi:10.1101/gad.1615208

van den Akker, E., Reijnen, M., Korving, J., Brouwer, A., Meijlink, F., and Deschamps, J. (1999). Targeted Inactivation of Hoxb8 Affects Survival of a Spinal Ganglion and Causes Aberrant Limb Reflexes. *Mech. Dev.* 89, 103–114. doi:10.1016/s0925-4773(99)00212-9

Vermot, J., Schuhbaur, B., Mouellic, H. L., McCaffery, P., Garnier, J.-M., Hentsch, D., et al. (2005). Retinaldehyde Dehydrogenase 2 and Hoxc8 Are Required in the Murine Brachial Spinal Cord for the Specification of Lim1+ Motoneurons and the Correct Distribution of Islet1+ Motoneurons. *Development* 132, 1611–1621. doi:10.1242/dev.01718

Wahba, G. M., Hostikka, S. L., and Carpenter, E. M. (2001). The Paralogous Hox Genes Hoxa10 and Hoxd10 Interact to Pattern the Mouse Hindlimb Peripheral Nervous System and Skeleton. *Develop. Biol.* 231, 87–102. doi:10.1006/dbio.2000.0130

Wu, J. S., and Luo, L. (2006). A Protocol for Mosaic Analysis with a Repressible Cell Marker (MARCM) in Drosophila. *Nat. Protoc.* 1, 2583–2589. doi:10.1038/nprot.2006.320

Wu, Y., Wang, G., Scott, S. A., and Capecchi, M. R. (2008). Hoxc10 and Hoxd10 Regulate Mouse Columnar, Divisional and Motor Pool Identity of Lumbar Motoneurons. *Development* 135, 171–182. doi:10.1242/dev.009225

Yaghmaeian Salmani, B., Monedero Cobeta, I., Rakar, J., Bauer, S., Curt, J. R., Starkenberg, A., et al. (2018). Evolutionarily Conserved Anterior Expansion of the central Nervous System Promoted by a Common PcG-Hox Program. *Development* 145, dev160747. doi:10.1242/dev.160747

Yamaguchi, M., and Yoshida, H. (2018). Drosophila as a Model Organism. *Adv. Exp. Med. Biol.* 1076, 1–10. doi:10.1007/978-981-13-0529-0_1

Yang, C.-h., Belawat, P., Hafen, E., Jan, L. Y., and Jan, Y.-N. (2008). Drosophila Egg-Laying Site Selection as a System to Study Simple Decision-Making Processes. *Science* 319, 1679–1683. doi:10.1126/science.1151842

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Cell and Developmental Biology

ORIGINAL RESEARCH
published: 31 January 2022
doi: 10.3389/fcell.2021.803370

Elevated Hoxb5b Expands Vagal Neural Crest Pool and Blocks Enteric Neuronal Development in Zebrafish

*Aubrey G. A. Howard IV[1], Aaron C. Nguyen[1], Joshua Tworig[2], Priya Ravisankar[3,4], Eileen W. Singleton[1], Can Li[2], Grayson Kotzur[1], Joshua S. Waxman[3] and Rosa A. Uribe[1]**

[1]BioSciences Department, Rice University, Houston, TX, United States, [2]Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA, United States, [3]Molecular Cardiovascular Biology Division, Cincinnati Children's Hospital Medical Center and Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, OH, United States, [4]Allen Institute of Immunology, Seattle, WA, United States

Neural crest cells (NCCs) are a migratory, transient, and multipotent stem cell population essential to vertebrate embryonic development, contributing to numerous cell lineages in the adult organism. While great strides have been made in elucidating molecular and cellular events that drive NCC specification, comprehensive knowledge of the genetic factors that orchestrate NCC developmental programs is still far from complete. We discovered that elevated Hoxb5b levels promoted an expansion of zebrafish NCCs, which persisted throughout multiple stages of development. Correspondingly, elevated Hoxb5b also specifically expanded expression domains of the vagal NCC markers *foxd3* and *phox2bb*. Increases in NCCs were most apparent after pulsed ectopic Hoxb5b expression at early developmental stages, rather than later during differentiation stages, as determined using a novel transgenic zebrafish line. The increase in vagal NCCs early in development led to supernumerary Phox2b$^+$ enteric neural progenitors, while leaving many other NCC-derived tissues without an overt phenotype. Surprisingly, these NCC-derived enteric progenitors failed to expand properly into sufficient quantities of enterically fated neurons and stalled in the gut tissue. These results suggest that while Hoxb5b participates in vagal NCC development as a driver of progenitor expansion, the supernumerary, ectopically localized NCC fail to initiate expansion programs in timely fashion in the gut. All together, these data point to a model in which Hoxb5b regulates NCCs both in a tissue specific and temporally restricted manner.

Keywords: neural crest, hox, zebrafish, enteric neuron, differentiation

OPEN ACCESS

Edited by:
Ernesto Sánchez-Herrero, Spanish National Research Council (CSIC), Spain

Reviewed by:
Sally Ann Moody, George Washington University, United States
Maria Elena De Bellard, California State University, Northridge, United States

****Correspondence:***
Rosa A. Uribe
rosa.uribe@rice.edu

Specialty section:
This article was submitted to Morphogenesis and Patterning, a section of the journal Frontiers in Cell and Developmental Biology

Received: *27 October 2021*
Accepted: *13 December 2021*
Published: *31 January 2022*

Citation:
Howard AGA, Nguyen AC, Tworig J, Ravisankar P, Singleton EW, Li C, Kotzur G, Waxman JS and Uribe RA (2022) Elevated Hoxb5b Expands Vagal Neural Crest Pool and Blocks Enteric Neuronal Development in Zebrafish. Front. Cell Dev. Biol. 9:803370. doi: 10.3389/fcell.2021.803370

INTRODUCTION

As an embryonic stem cell population in vertebrates, neural crest cells (NCCs) are renowned for their remarkable migratory capacity, as well as their multipotency. Born from the dorsal neural tube, NCCs migrate along stereotypic routes throughout the early embryo and give rise to a wide range of diverse tissue lineages, such as craniofacial skeleton, portions of the peripheral nervous system, and pigment cells (Rocha et al., 2020). NCCs exhibit regional potential along the anteroposterior (AP) neuraxis such that they may be divided into four general populations: cranial, vagal, trunk, and sacral (Le Douarin, 1982; Le Douarin and Kalcheim, 1999). Each of these populations give rise to numerous discrete lineages, for example, cranial NCC largely give rise to cell lineages in the head. Particularly of

interest are vagal NCCs, which contribute to several tissues, such as the cardiac outflow tract and nearly all of the enteric nervous system (ENS) (Tang et al., 2021) within the gut, and have been less well characterized than other populations (Hutchins et al., 2018). While the driving genetic factors which regulate the general pattern of NCC developmental trajectories have been well described (Simoes-Costa and Bronner, 2016; Martik and Bronner, 2017), we still have an incomplete understanding of what genes function in context of vagal NCC development and their subsequent differentiation.

Coincident with the anterior to posterior rise of NCC is the expression of Hox genes, a strongly conserved family of genes encoding for transcription factors most notable for their canonical role in body axis patterning (Amores et al., 1998; Pearson et al., 2005). Among their many roles, Hox transcription factors are known to play essential roles in establishing discrete partitions within the hindbrain, directing limb formation, regulating cardiac cell number, and guiding neural circuit formation within a variety of tissue contexts (Rancourt et al., 1995; Waxman et al., 2008; Minguillon et al., 2012; Breau et al., 2013; Di Bonito et al., 2013; Barsh et al., 2017). Organized in tight clusters in the genome, known as paralogy groups, Hox genes are labeled A–D to designate a particular chromosomal cluster, as well as by number, which represents the gene's chromosomal position within a particular cluster, ranging from 1–13 (Kudoh et al., 2001). Furthermore, not only are the Hox peptide sequences conserved between species, synteny of the Hox clusters is highly conserved (Santini et al., 2003). The role of Hox genes among vertebrates is also strongly conserved in their function, even among teleost fishes who have undergone a genome duplication during their evolutionary history (Amores et al., 1998; Parker et al., 2019). Each Hox gene is expressed along the anterior-posterior axis in nested domains collinear with their position in the chromosome, where they undergo complex regulatory interaction to establish discrete expression domains (Zhu et al., 2017). As such, earlier numbered Hox factors are commonly expressed within the cephalic tissues, while later numbered Hox factors are expressed more distally.

Within the context of cranial NCCs, Hox transcription factors have been shown to drive a number of NCC phenotypes. Overexpression of Hox factors in chicken, such as *Hoxa2, Hoxa3*, and *Hoxb4*, produces a variety of overlapping ablations of NCC-derived craniofacial skeleton (Creuzet et al., 2002). Similarly, the NCC-derived hyoid bones and presumptive thymic mesenchyme were greatly reduced or ablated in mice homozygous for single knockout for *Hoxa-3,* or double knockout of *Hoxa-3* and *Hoxd-3*, while also causing homeotic transformations of other structures throughout the animal (Condie and Capecchi, 1993, 1994). In addition to affecting formation of terminally differentiated craniofacial structures, earlier phases of NCC migration into pharyngeal arches and onset emigration of NCC from the neural tube are also acutely sensitive to changes in anterior Hox expression (Gouti et al., 2011; Parker et al., 2018). While the role of Hox transcription factors is less well characterized in posterior populations of NCC, vagal NCC in mice fail to colonize the gut following overexpression of *Hoxa-4*, which is endemic to the gut nervous network (Wolgemuth et al., 1989; Tennyson et al., 1993). The failure to form a complete enteric nervous system results in megacolon, a phenotype associated with human disease (Nagy and Goldstein, 2017). Overall, while we have learned much regarding Hox genes in the cranial NCC, the roles Hox transcription factors play within vagal and other posterior NCC populations requires further investigation.

One posterior Hox factor that has been implicated in vagal NCC development is Hoxb5. In mice, dominant negative abrogation of embryonic Hoxb5 activity alters vagal and trunk NCC development, with reduced NCCs observed en route to and along gut tissue, as well as decreased numbers of melanoblasts throughout the body (Lui et al., 2008; Kam and Lui, 2015). Additionally, Hoxb5 may regulate expression of key genes active in NCC development, notably *Foxd3* (Kam et al., 2014a) and *Phox2b* (Kam and Lui, 2015). The orthologous gene in zebrafish, *hoxb5b*, which is the primary ortholog in teleost fishes (Jarinova et al., 2008), was detected in differentiating NCC lineages at both 48 and 68 h post fertilization (hpf) (Howard et al., 2021), suggesting it may also play a role during zebrafish NCC development. While characterization of zebrafish *hoxb5b* mRNA expression *in situ* has been pervasively characterized in several early embryonic contexts, such as the mesoderm and limb buds (Kudoh et al., 2001; Jarinova et al., 2008; Waxman et al., 2008; Hortopan and Baraban, 2011; van der Velden et al., 2012), the functional role of *hoxb5b* with respect to NCC development had not yet been examined.

Here, we postulated that *hoxb5b* functions as a potential driver of vagal NCC development. We provide evidence that overexpression of *hoxb5b* is sufficient to grossly expand NCC populations throughout the zebrafish embryo, in addition to ectopic expansion of vagal domains marked by *foxd3* and *phox2bb*. The functional window of Hoxb5b activity was also restricted to a narrow developmental span, early during embryogenesis, rather than during NCC differentiation stages. The early expansion of NCC, however, did not lead to corresponding pan increases in NCC-derived tissues. Rather, elevated Hoxb5b activity expanded enteric neural progenitor cell pools along the gut, yet suppressed their subsequent expansion as they differentiated into enteric neurons, leading to overall fewer neurons along the gut. These data cumulatively support a model in which *hoxb5b* is a potent regulator of NCC expansion and cell number in zebrafish.

RESULTS

hoxb5b is Expressed in Post-Otic Vagal NCCs During Zebrafish Development

We examined *hoxb5b* expression in zebrafish embryos (**Figure 1A**) using whole mount *in situ* hybridization (ISH). At 32 hpf, *hoxb5b* was expressed bilaterally immediately posterior to the otic vesicle (post-otic), along

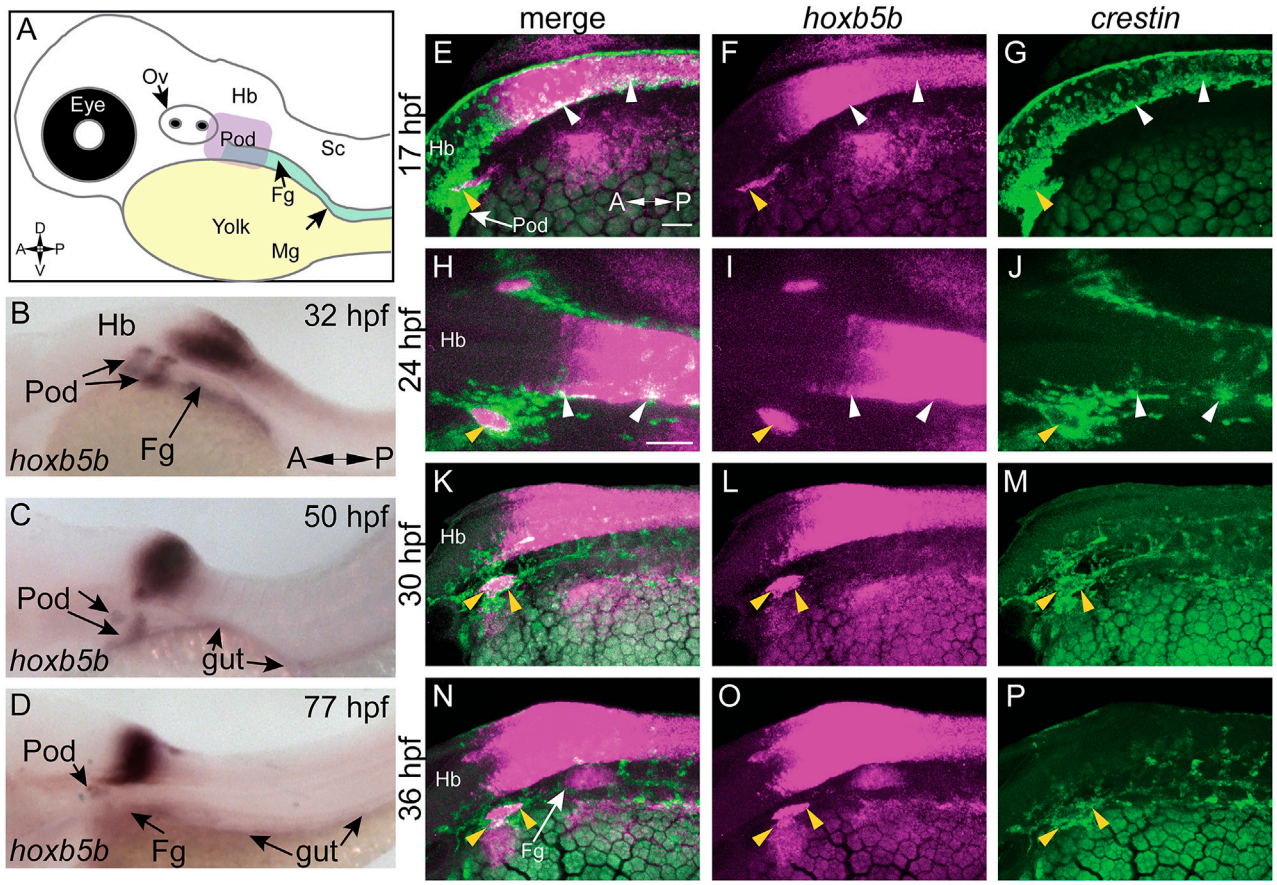


FIGURE 1 | *hoxb5b* is expressed during early NCC development. **(A)** Schematized embryo illustrating approximate locations of major relevant anatomical features, namely the post-otic domain (POD), eye, yolk, Otic Vesicle (Ov), Hindbrain (Hb), presumptive spinal cord (Sc), foregut mesenchyme (Fg), and Midgut mesenchyme (Mg). **(B–D)** *in situ* hybridization demonstrating *hoxb5b* expression in the posterior Hb, POD, and Fg during the second through third day of development. **(E–P)** Hybridization Chain Reaction probes against *crestin* and *hoxb5b* highlight their overlapping domains in the POD (yellow arrowheads) and along the dorsal length (white arrowheads). Scale bars E,H: 50 μM.

the foregut, in the hindbrain, and anterior spinal cord (**Figure 1B**), as previously described (Jarinova et al., 2008). *hoxb5b* persisted in all three domains at 50 hpf (**Figure 1C**), though the post-otic domains (POD) were slightly restricted and the hindbrain/spinal cord expression gained a more defined posterior boundary. By 77 hpf, *hoxb5b* expression remained largely in the hindbrain and foregut, with diminished yet persistent expression in the POD (**Figure 1D**). Together, these ISH data reveal changing post-otic spatiotemporal expression patterns of *hoxb5b* during the first 4 days of development.

We next examined the relationship of *hoxb5b* expression to the vagal NCC population during NCC specification and migration phases. NCCs were assayed by using the zebrafish pan-NCC marker *crestin* (Luo et al., 2001), in combination with *hoxb5b, via* hybridization chain reaction (HCR) (Choi et al., 2010, 2016, 2018). At 17 hpf, *crestin*$^{+}$/*hoxb5b*$^{+}$ domains were present dorsally (**Figures 1E–G**, white arrowheads), as well as ventral-laterally, along a post-otic stripe (**Figures 1E–G**, yellow arrowhead), revealing that *hoxb5b* is expressed within the POD vagal NCC population. *crestin*$^{+}$/*hoxb5b*$^{+}$ regions persisted by 24 hpf dorsally, in posterior hindbrain/anterior spinal cord axial levels (**Figures 1H–J**, white arrowheads). Concurrently at this stage, *hoxb5b* expression within the stripe became internalized within the POD vagal NCC population, marking the central area of this region, and highlighting several adjacent *crestin*$^{+}$/*hoxb5b*$^{+}$ domains (**Figures 1H–J**, yellow arrowhead). Between 30–36 hpf, the co-positive *hoxb5b*$^{+}$/*crestin*$^{+}$ stripe of POD vagal NCC persisted (Figures 1K–P, yellow arrowheads), and *crestin*$^{+}$/*hoxb5b*$^{+}$ regions were still observed along the dorsal neural tube. These data are consistent with our prior findings in which *hoxb5b* mRNA was present in posterior NCC at 48–50 hpf and 68–70 hpf (Howard et al., 2021). Collectively, the ISH and HCR data indicate that *hoxb5b* mRNA expression is coincident within the vagal NCCs area, persisting throughout the developmental window spanning NCC specification and well into their migration phase, highlighting *hoxb5b*'s potential role as a driver of vagal NCC development.

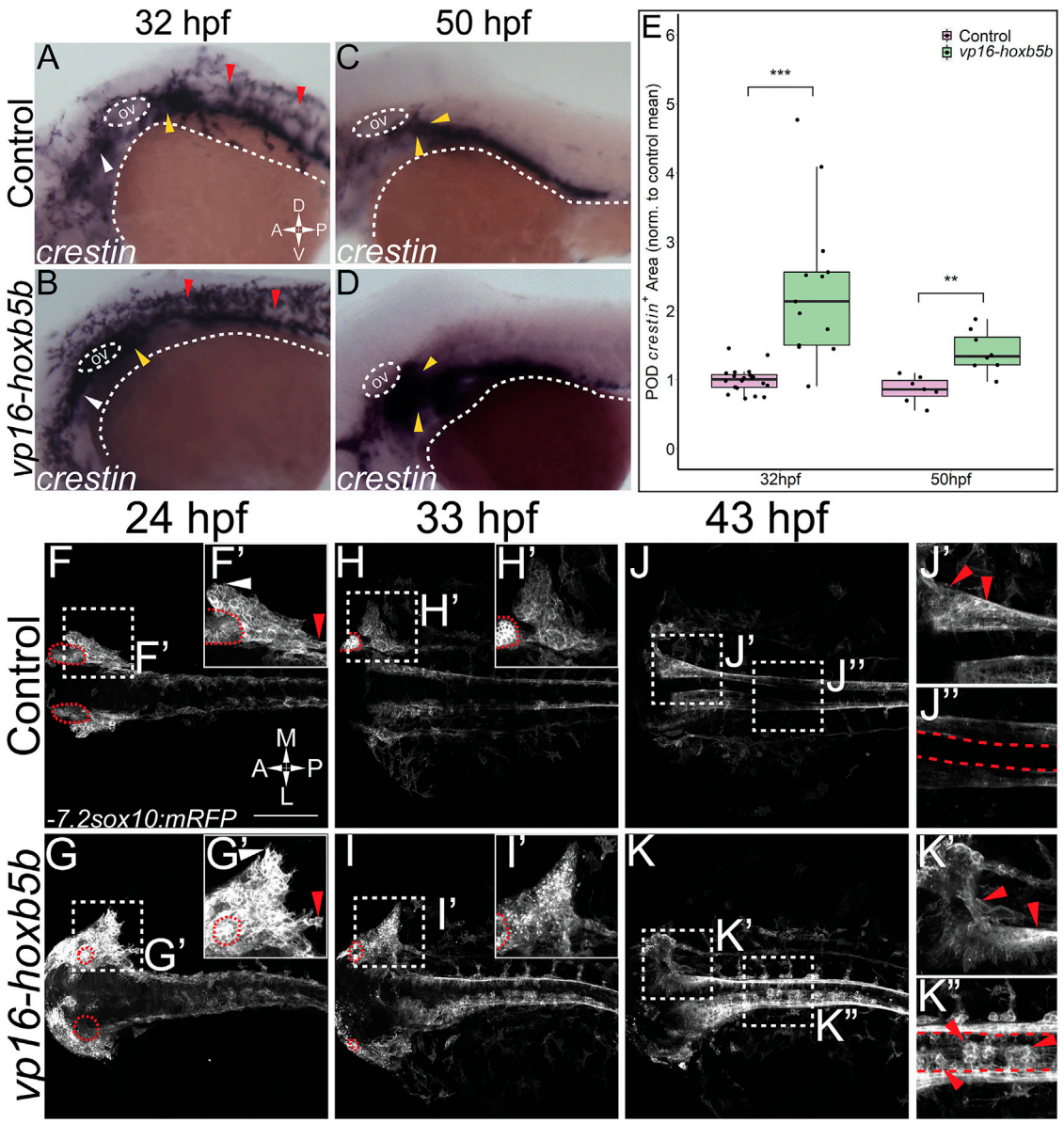


FIGURE 2 | Elevated Hoxb5b activity globally increases both number and localization of neural crest cells. **(A–D)** *in situ* hybridization for NCC using a *crestin* probe at both 32 hpf (A,B) and 50 hpf **(C,D)**. *crestin*$^{+}$ domains for embryos injected with 15 pg of *vp16-hoxb5b* mRNA **(B,D)** were expanded in the post-otic (yellow arrowheads), cranial (white arrowheads), and spinal cord (red arrowheads), compared to uninjected embryos **(A,C)**. **(E)** Quantification of expanded vagal *crestin*$^{+}$ domains shows significant expansion at both 32 hpf (control $n = 27$; *vp16-hox5b* $n = 21$; $p = 3.03 \times 10^{-5}$) and at 50 hpf (control $n = 7$; *vp16-hox5b* $n = 8$; $p = 0.00114$). **(F–K)** Maximum intensity projection stills taken from confocal time lapse movies of *sox10:mRFP* embryos. Controls ($n = 2$) were compared to 30 pg *vp16-hoxb5b* injected embryos ($n = 4$), examined from 24 hpf to 43 hpf and serially imaged along the dorsal aspect of the vagal domain. mRFP^{+} NCCs are grossly expanded in the vagal domain (**G,G'** arrowheads) over controls (**F,F'**, arrowheads). This expansion persists through the course of development, resulting in ectopically localized cells along the dorsal aspect of the embryo (**K"**) and in the post-otic pool (**K'**). Scale Bars in **(F)**: 100 μM. Anterior: Left.

Elevated Hoxb5b Activity Globally Promotes Expanded Localization and Increased Number of NCC *in vivo*

While other work has focused exclusively on the loss of function of Hoxb5 related genes (Lui et al., 2008; Dalgin and Prince, 2021), we have sought to understand the gain of function role of *hoxb5b* in the context of NCC development. To examine the possible role of *hoxb5b* in NCC, we employed a hyperactive *vp16-hoxb5b* fusion construct (Waxman et al., 2008; Waxman and Yelon, 2009). Injection of *vp16-hoxb5b* mRNA resulted in expansion of NCC, when compared to control embryos at 32 hpf, as assayed with an ISH probe against *crestin* (**Figures 2A,B**). The expansion in NCC territory was prominent in the pre-otic region (**Figure 2B**; white arrowheads), the POD (**Figure 2B**; yellow arrowheads)

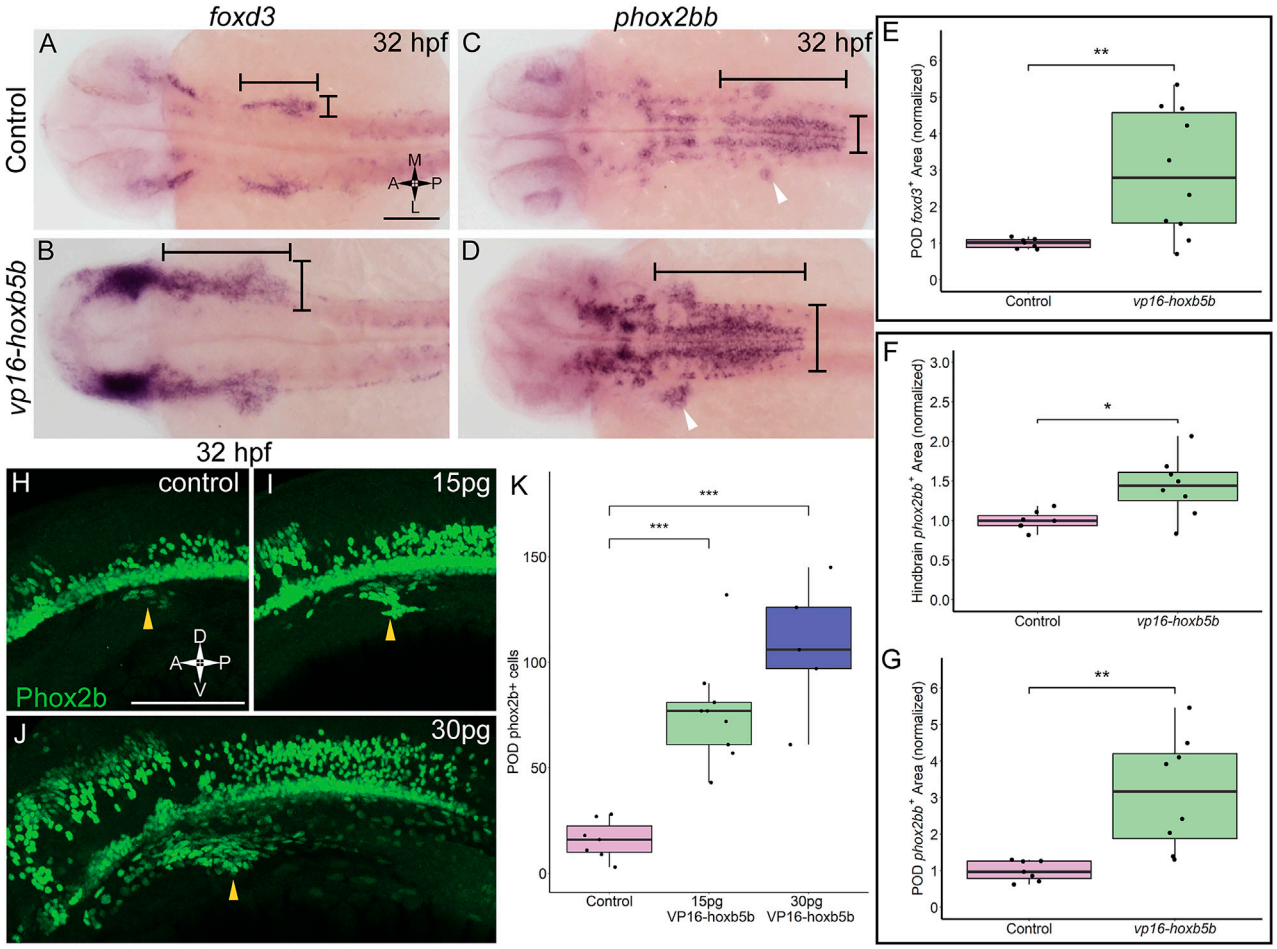


FIGURE 3 | Elevated Hoxb5b activity expands the expression domains of vagal markers *foxd3* and *phox2bb* during the first day in development. **(A,B)** *in situ* hybridization against *foxd3*, a marker for multipotent NCCs at this stage, in 32 hpf embryos injected with 30 pg *vp16-hoxb5b,* were compared to WT controls. **(C,D)** *in situ* hybridization against *phox2bb*, an autonomic NCC marker, in 32 hpf embryos injected with 30 pg *vp16-hoxb5b*, compared to WT controls. The POD is denoted with a white arrow head. **(E–G)** Quantified POD *foxd3* expression area (E, control $n = 7$; *vp16-hox5b* $n = 10$; $p = 0.00571$) or *phox2bb* expression area (F, control $n = 7$; *vp16-hox5b* $n = 8$; $p = 0.138$), as noted in representative images by black bars. The discrete POD expression domain is quantified from these same embryos as shown in **(G)** ($p = 0.00563$). Areas are normalized to control mean. **(H–K)** IHC detection of Phox2b$^+$ cells viewed along the lateral axis of 32 hpf embryos reveal that WT controls ($n = 7$) already have a nascent population [**(H)**, white arrowheads]. *vp16-hoxb5b* overexpression, at either 15 pg [**(I)**, $n = 9$] or 30 pg [**(J)**, $n = 5$] of mRNA injected, expands the Phox2b$^+$ vagal NCCs (white arrowheads). **(K)** Counted Phox2b$^+$ cells from 3 dimensional micrographs reveal increasing cells with the amount of *vp16-hoxb5b* mRNA (15 pg: $p = 3.03 \times 10^{-5}$; 30 pg: $p = 2.61 \times 10^{-5}$). Scale bar **(A,H)**: 100 μM.

and along the spinal cord-level of the trunk (**Figure 2B**; red arrowheads). Strikingly, post-otic expansion persisted along the dorsal-ventral and anterior-posterior axes well through NCC migration phases at 50 hpf (**Figures 2C,D**; arrowheads). Furthermore, quantification of the area occupied by POD NCCs corroborated the observed expansion of NCC localization (**Figure 2E**).

To better understand the spatiotemporal distribution of the increased NCC following *vp16-hoxb5b* expression, we utilized confocal microscopy to image *-7.2sox10:mRFP* transgenic embryos (referred to here as *sox10:mRFP*) (Kucenas et al., 2008), where NCC are labeled using a membrane bound RFP. Congruent with our prior findings (**Figures 2A,B**), at 24 and 33 hpf, confocal projections revealed *vp16-hoxb5b* expressing embryos exhibited broadened POD vagal NCC domains along the anterior-posterior and medio-lateral axes (**Figures 2G,G',I,I'**; arrowheads), when compared to control embryos (**Figures 2F,F',H,H'**; arrowheads). This expansion was coupled with an increase in the number of POD NCC. By 43 hpf, *vp16-hoxb5b* expressing embryos displayed a striking expansion of the POD NCC, as well as a disruption of the overall architecture of the domain, which extended further laterally from the dorsal midline (**Figure 2K'**), than in control (**Figure 2J'**). Moreover, *vp16-hoxb5b* promoted ectopic accumulation of cells along the dorsal midline of the spinal cord (**Figure 2K'**), which were not observed in control embryos (**Figure 2J"**). Considered together, these data (**Figure 2**) indicate that elevated Hoxb5b activity alters NCC localization along the embryo as well as expanding their cell number.

Elevated Hoxb5b Activity is Sufficient to Expand Vagal NCC Marker Expression

That *vp16-hoxb5b* expressing embryos contained an overabundance of NCC, and that the overproduced NCC were prominently enriched in the vagal axial levels suggests that excess Hoxb5b activity influences vagal NCC development in zebrafish. To examine if vagal NCC specification was altered following excess Hoxb5b activity, we assayed the expression of canonical marker genes of the vagal NCC, *foxd3* (Lister et al., 2006) and *phox2bb* (Elworthy et al., 2005). At 32 hpf, *foxd3* expression serves as an indicator of multipotent NCC while *phox2bb* indicates NCCs which are now specified to an autonomic neural lineage, particularly the ENS (Pattyn et al., 1999; Uribe and Bronner, 2015). *vp16-hoxb5b* expression was sufficient to widen *foxd3* expression domains at 32 hpf, principally along the anterior-posterior and mediolateral directions in the vagal region, when compared to control expression patterns (**Figures 3A,B**; black bars). We found that *vp16-hoxb5b* expression expanded the POD *foxd3*$^+$ area, leading to a 1.95 fold increase in the mean domain size compared to controls (**Figure 3E**). Additionally, *phox2bb* was also greatly expanded in response to increased Hoxb5b activity along the hindbrain (**Figures 3C,D**; black bars), with expansion uniformly in both the anterior-posterior and mediolateral axis, similar to the expansion of the *foxd3*. Measuring the hindbrain *phox2bb*$^+$ domain, we observed a 43% increase in the *phox2bb* expressed area throughout the hindbrain following elevated Hoxb5b (**Figure 3F**). Lastly, POD *phox2bb* expression was dramatically altered by elevated Hoxb5b (**Figures 3C,D**; white arrow heads), with a 2.14 times mean increase in domain size compared to wild-type controls (**Figure 3G**). In all, these data indicate that elevated Hoxb5b activity drastically expands vagal NCC marker expression along the embryo.

In support of the specific expansion of POD localized *phox2bb* expression, we also observed a corresponding increase in the number of Phox2b$^+$ cells, *via* whole mount fluorescent Immunohistochemistry (IHC), using an antibody against Phox2b (**Supplementary Figures S1A–C**; **Figure 3**). At 32 hpf, Phox2b$^+$ cells were observed in the POD, with 16 cells on average (**Figure 3L**; arrowhead, **Figure 3K**). Overexpression of Hoxb5b stimulated a dramatic expansion of POD Phox2b$^+$ cells (**Figures 3I,J**; arrowheads). The increase in POD Phox2b$^+$ cells was also concordant with an increase in *vp16-hoxb5b* dosage, with 77 and 107 mean POD Phox2b$^+$ cells per animal detected following injection with either 15 pg or 30 pg of mRNA, respectively (**Figure 3K**). The quantifiable increase in POD Phox2b$^+$ cells is confirmatory of our prior qualitative observations regarding increased cell number (**Figure 2**) and positions Hoxb5b as a potent driver of NCC number and localization.

Hoxb5b Overexpression Increases NCC Production During Early NCC Development

While we observed supernumerary NCCs ectopically localized, and expanded expression of vagal NCC specification factors following global expression of *vp16-hoxb5b* mRNA, exactly when during NCC development Hoxb5b may exert its influence was still unclear. To investigate the potential temporal role(s) of Hoxb5b during NCC development, we created and utilized a novel transgenic line, Tg (*hsp70l:EGFP-hoxb5b;cryaa:dsRed*)^{ci1014} (hereafter referred to as *hsp70l:GFP-hoxb5b*), which enables pan ectopic expression of a GFP-Hoxb5b protein fusion under the thermally inducible *hsp70l* promoter (Kwan et al., 2007) (**Figure 4A**). The transgene can then be activated by rapidly transferring embryos to warm 37°C culture conditions, which drives strong global expression of the EGFP-Hoxb5b fusion protein throughout the embryonic tissues (**Supplementary Figure S2A**). EGFP-Hoxb5b demonstrated robust and distinctive nuclear localization, which was still detectable over 24 h after embryos were returned to 28°C (**Supplementary Figures S2B,C**).

Two early phases in NCC development were tested, with heat shocks conducted starting either at 14 hpf, during NCC specification, or 22 hpf, early during the migratory span of posterior NCCs (**Figure 4B**). After heat shock at 14 hpf, whole mount ISH at 24 hpf showed enlargement of *crestin*$^+$ NCC domains in GFP-Hoxb5b$^+$ embryos, over the GFP-Hoxb5b$^-$ sibling controls (**Figures 4C,D**), particularly prevalent in the POD NCC (**Figure 4D**; yellow arrowheads). This increase in area was also accompanied by a qualitative increase in the number of *crestin*$^+$ cells along the posterior dorsal length of the embryo, similar to the phenotype observed in the previous Hoxb5b mRNA overexpression assays (**Figures 2A–E**). Quantified area of *crestin*$^+$ PODs confirmed the expansion of vagal NCCs (**Figure 4I**). Additionally, subtle expansion of cranial NCCs (**Figures 4C,D**, white arrowheads) and pre-otic NCCs (**Figures 4C,D**, red arrowheads) was also observed, though far less striking than that of the vagal population at this stage. These data indicate that NCC localization during early specification phase of NCC development is receptive to Hoxb5b activity.

GFP-Hoxb5b induction at 22 hpf also increased *crestin* staining by 26–28 hpf throughout the POD (**Figures 4E–H**, yellow arrowheads), which was especially prominent in the NCCs most proximal to the otic vesicles (**Figures 4G,H**, red arrowheads). As in the heat shock at 14 hpf, increased NCC localization was also observed in GFP-Hoxb5b$^+$ embryos across the cranial NCC populations (**Figures 4E–H**; white arrowheads). This induction of GFP-Hoxb5b also significantly expanded measurable vagal NCC area (**Figure 4J**), with a 26% mean increase in area compared to GFP$^-$ siblings. The expansion in area coupled with the increase in *crestin* staining replicates the results obtained via the microinjection assays presented in **Figure 2**, albeit more subtly, as well as indicates that 14–22 hpf is a critical period during which Hoxb5b is sufficient to alter NCC localization in the developing vertebrate body.

Increased Hoxb5b Alters Specific Vagal NCC-Derived Tissues

Due to the multipotent nature of NCC, it is possible a wide diversity of tissues can be affected by even a small perturbation in NCC development. Because we observed an overproduction of NCC following increases in Hoxb5b activity, we wondered if

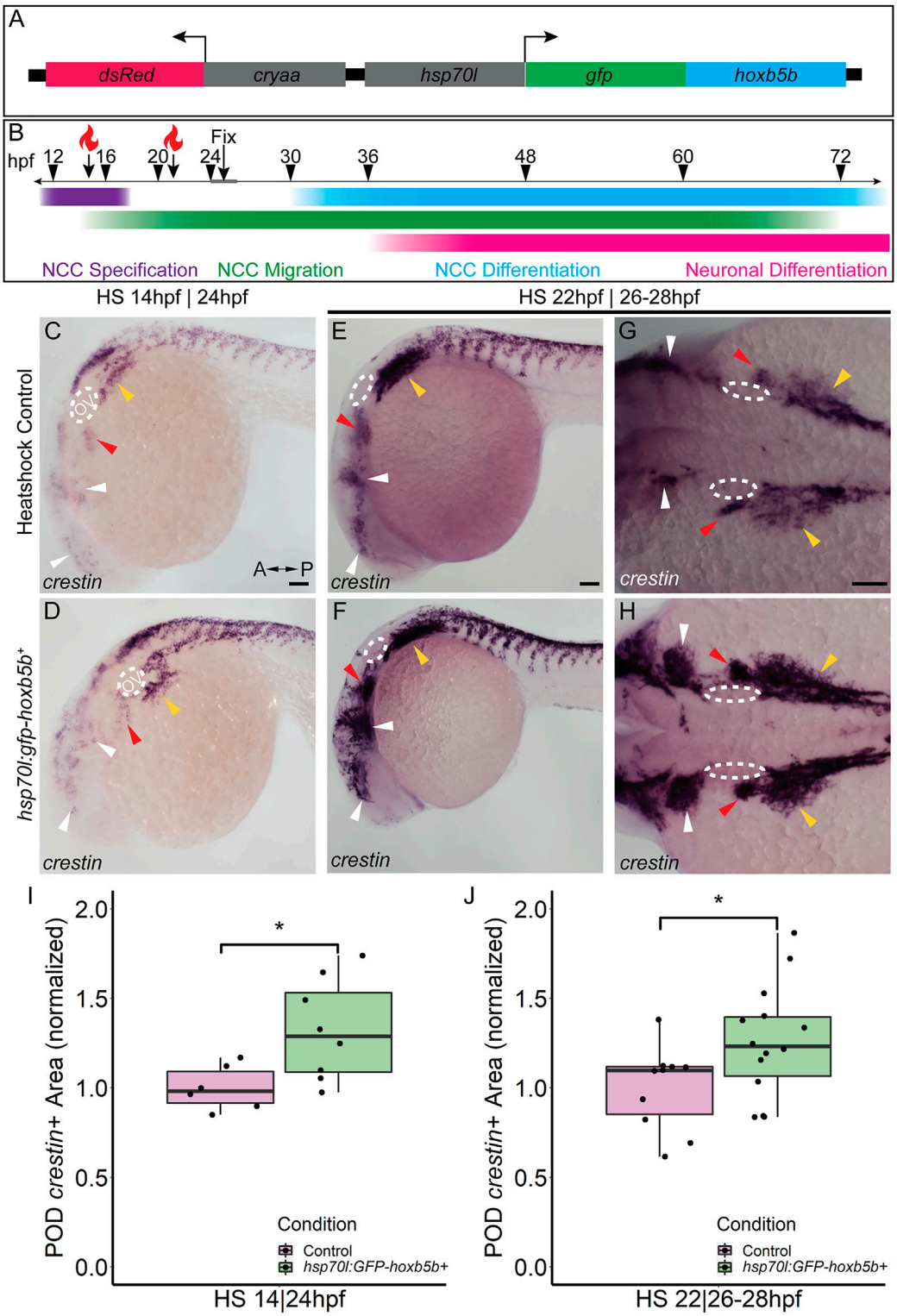


FIGURE 4 | Temporally controlled overexpression of Hoxb5b during the first day in development is sufficient to expand the neural crest pool. **(A)** Schematized model of the *hsp70l:gfp-hoxb5b;cryaa:dsRed* genetic construct. **(B)** Illustration depicting specific periods of heat shock for embryo groups relative to classical hallmarks of zebrafish NCC development. **(C,D)** *in situ* hybridization using a probe for *crestin* in *hsp70l:GFP-hoxb5b*$^{+}$ embryos heat shocked at 14 hpf and fixed at 24 hpf, compared to GFP^{-} sibling controls treated in parallel. Dramatic expansion of the POD can be seen (yellow arrowheads) in the Hoxb5b overexpressing embryos, while more subtle expansion is noted in the cranial NCC (white arrowheads) and pre-otic crest (red arrowheads). **(E–H)** Similar to **(C,D)** *in situ* hybridization with a *crestin*

(*Continued*)

FIGURE 4 | probe in *hsp70l:gfp-hoxb5b* and GFP- sibling controls heat shocked at 22 hpf and fixed at 26–28 hpf, examining both *crestin*$^+$ domains both laterally **(E,F)** and dorsally **(G,H)**. **(I,J)** Graphs depicting areas occupied by *crestin* staining in both GFP- control and Hoxb5b overexpressing embryos. Hoxb5b overexpression at 14 hpf was sufficient to expand NCC localization and qualitatively NCC number [**(I)**, control $n = 6$; *hsp70:GFP-hoxb5b*$^+$ $n = 8$; $p = 0.0234$]. A later heat shock at 22 hpf—fixation at 26–28 hpf, also expanded vagal NCC localization [**(J)**, control $n = 10$; *hsp70:GFP-hoxb5b*$^+$ $n = 14$; $p = 0.0398$], and increased *crestin* staining, indicative of an increase in NCC number. Scale bars **(C,E,G)**: 100 μM

downstream NCC derivatives were also affected. Therefore, we first investigated the effect of *vp16-hoxb5b* expression on vagal and trunk NCC-derived cell types; including, pigment cells (melanophores and iridophores) and neuronal derivatives. Embryos expressing *vp16-hoxb5b* were able to produce pigment cells without appreciable differences (Supplementary Figures S4A–D). Both dorsal root ganglia (DRG) and Superior Cervical Ganglion (SCG) are derived from the vagal/trunk NCC pool (Durbec et al., 1996). Neurons comprising the DRG and SCG were largely normal following *vp16-hoxb5b* injection (**Supplementary Figures S4E–H**).

We next assayed the NCC-derived cell population along the gut length (**Figure 5A**), enteric neural progenitors, which typically migrate to the midgut level by 2 dpf and will have colonized the hindgut by 3 dpf, giving rise to neurons of the enteric nervous system (Elworthy et al., 2005; Ganz, 2018). We found that by 50 hpf, Phox2b$^+$ enteric neural progenitors significantly increased after *vp16-hoxb5b* expression (**Figures 5C,D**), compared to the controls (**Figure 5B**). Counting Phox2b$^+$ cells in the POD and along the gut tract (**Figures 5B–D**; white dashes), revealed the number of cells trended with increasing amount of *vp16-hoxb5b* mRNA injected (**Figure 5E**), consistent with the phenotype at 32 hpf (**Figures 3H–J**). The supernumerary enteric neural progenitors were observed together with their accumulation along the foregut (**Figures 5B–D**; yellow arrowheads) and in the POD (**Figure 5F**), though an increase in the number of cells was also observed at the end of the enteric migration chain along the midgut (**Supplementary Figure S4I**; **Figures 5C,D**; white arrowheads). The increase in cells was uniform across the gut tract, with no change in the fraction of Phox2b$^+$ cells found in the POD or gut mesenchyme after Hoxb5b perturbation (**Supplementary Figure S4J**). These findings indicate that elevated Hoxb5b elicits a global increase in enteric neural progenitor number through the first 2 days of development.

To determine if the supernumerary enteric cells were capable of differentiating into neurons later in development, we utilized *-8.3phox2bb:kaede* transgenic embryos which label enteric progenitors during their early neuronal differentiation (Harrison et al., 2014). Surprisingly, despite the increase in enteric progenitors at 50 hpf, enteric neurons by the 3 dpf were dramatically decreased in *vp16-hoxb5b* expressing embryos compared to controls. Kaede$^+$ cells successfully colonized the gut length by 3 dpf in control embryos (**Figure 5G**; white arrowhead), many cells of which (42%) also co-expressed the pan neuronal marker Elavl3/4 (**Figures 5G,K,L**), signaling the onset of neuron differentiation. In contrast, *vp16-hoxb5b* expressing embryos at both doses displayed a drastic loss of Kaede$^+$ and Elavl3/4$^+$ enteric cells (**Figures 5J,K**), with the remaining Kaede$^+$ cells failing to localize past the level of the midgut (**Figures 5H,I**; white arrowheads). The fraction of Elavl3/4$^+$/Kaede$^+$ cells in both *vp16-hoxb5b* expressing conditions were reduced, at 0.31 and 0.28 respectively, when compared to control at 0.42 (**Figure 5L**). While not reaching significance, when these data are taken together with the significant reduction in total enteric cell numbers along the gut (**Figures 5J,K**), they likely indicate that general enteric progenitor pool depletion along the gut affects subsequent proper numbers of enteric neurons, following elevated Hoxb5b expression. Overall, these results suggest that while supernumerary enteric neural progenitors are present at and before 2 dpf after elevated Hoxb5b, they largely depleted by the 3 dpf.

Hoxb5b influences Enteric Colonization During Early Developmental Stages

In order to ascertain the timing during which excess Hoxb5b activity affects enteric nervous system development, we again leveraged the *hsp70l:GFP-hoxb5b* fish line. Embryos were heat shocked during NCC specification (14 hpf), migration (21 hpf), or differentiation (48 hpf), and all were fixed at 3 dpf, as schematized in **Figure 6A**. Embryos were assessed for enteric neuron abundance and localization via wholemount IHC, where Elavl3/4$^+$ cells were counted along the gut tract (same region as in **Figure 5A**, box). When heat shocked at 14 or 21 hpf, GFP-Hoxb5b$^+$ embryos formed significantly fewer Elavl3/4$^+$ cells, when compared to their GFP-Hoxb5b$^-$ sibling heat shock controls (**Figures 6B–E,H**). After heat shock at 48 hpf, GFP-Hoxb5b$^+$ embryos did not significantly vary in number of Elavl3/4$^+$ cells (**Figures 6F–H**). The distribution of Elavl3/4$^+$ cells was weighted more heavily toward the midgut, though cells could be detected along the entire length of the gut in GFP-Hoxb5b$^+$ embryos. This genetically encoded elevation of Hoxb5b activity during early NCC developmental phases corroborated the abrogation in Elavl3/4$^+$ cells resulting from *vp16-hoxb5b* mRNA injection. Overall, these data indicate that the ability of GFP-Hoxb5b to affect enteric neuron number is limited to early stages of NCC development, but not thereafter.

Excess Hoxb5b Leads to Stalled Enteric Nervous System Development

We had thus far discovered that Hoxb5b was sufficient to strongly increase NCCs at 30 and 50 hpf, but suppressed the number of enteric neural progenitor cells by 3 dpf. The loss in cells could easily be explained by an acute wave of cell death during enteric neural progenitor migration. We tested this hypothesis with whole mount IHC probing for activated Caspase-3, a marker for apoptotic cells (Sorrells et al., 2013), as well as Phox2b to label

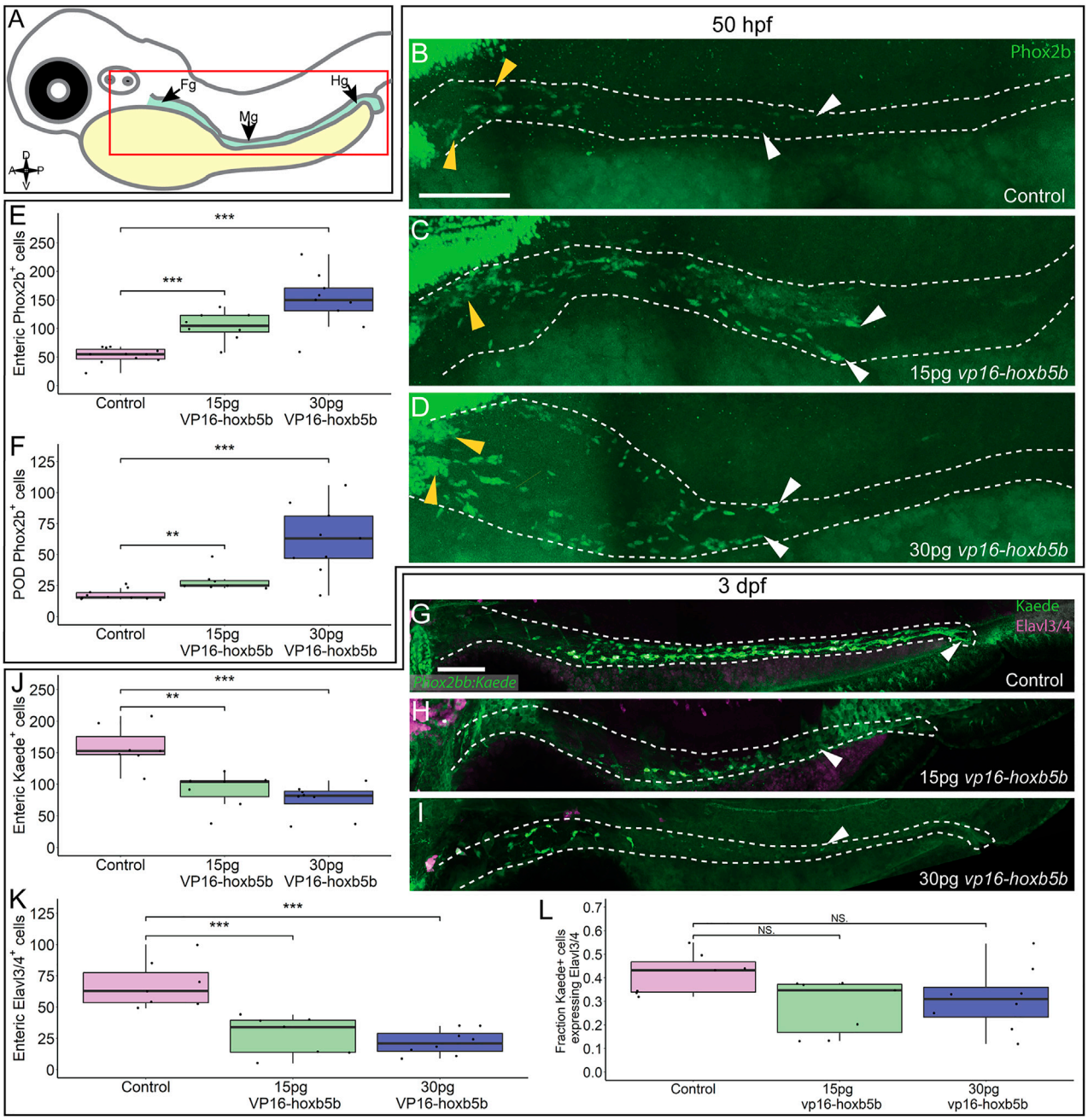


FIGURE 5 | Hoxb5b is sufficient to expand early enteric neural progenitors. **(A)** Schematized model of a zebrafish embryo highlighting the region of the gut tube, which is imaged in the following panels. **(B–D)** Whole mount IHC for Phox2b in 50 hpf control embryos [**(B)**, $n = 10$] compared to embryos injected either with 15 pg [**(C)**, $n = 7$] or 30 pg [**(D)**, $n = 9$] of *vp16-hoxb5b* mRNA. Yellow arrow heads indicated POD localized Phox2b$^+$ cells, white arrows designate terminal end of enteric NCC chain, which falls within the gut tract outlined with white dashes. **(E,F)** Quantified cell numbers from the same animals reveal a coordinate increase in Phox2b$^+$ along the gut axis at 50 hpf trending with increasing *vp16-hoxb5b* mRNA amounts [**(E)**, 15 pg: $p = 2.834 \times 10^{-5}$; 30 pg: $p = 0.00318$]. Additionally, the number of Phox2b$^+$ cells restricted to the POD also increased in response to elevated Hoxb5b activity [**(F)**, 15 pg: $p = 0.00328$; 30 pg: $p = 0.00068$]. **(G–I)** Whole mount IHC on 3 dpf -8.3*phox2bb:kaede* embryos with antibodies against Elavl3/4 and Kaede, marking the enteric NCC lineage cells in *vp16-hoxb5b* overexpressing animals (15 pg $n = 8$; 30 pg $n = 7$) compared to uninjected sibling controls ($n = 7$). **(J–L)** Quantification of the number of enteric neural progenitors [**(J)**, 15 pg: $p = 0.00011$; 30 pg: $p = 4.09 \times 10^{-5}$) and differentiating enteric neurons [**(K)**, 15 pg: $p = 0.001352$; 30 pg: $p = 0.0001042$] at 3 dpf show decreasing numbers of both cell populations. However, the total fraction of differentiating (Hu$^+$) NCC-derived Kaede$^+$ cells unchanged following elevated Hoxb5b activity [**(L)**, 15 pg: $p = 0.0.1282$; 30 pg: $p = 0.102$]. Scale Bar **(B,G)**: 100 μM.

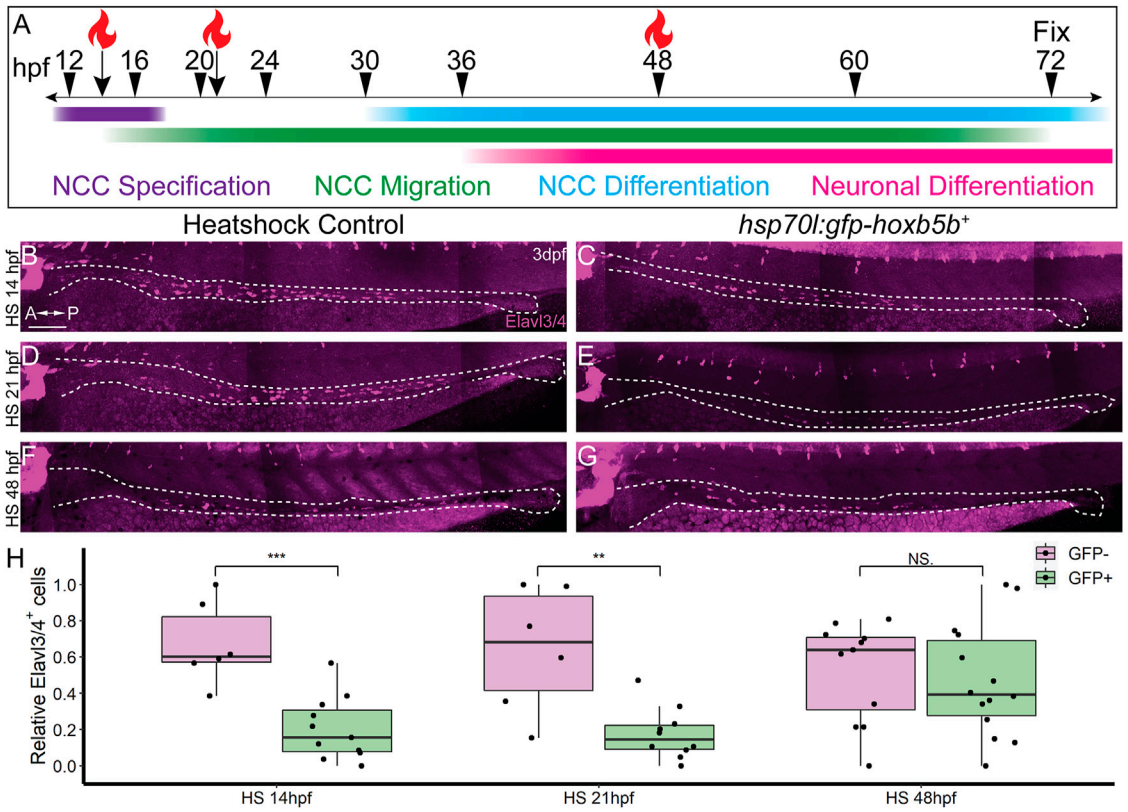


FIGURE 6 | NCC sensitivity to increased Hoxb5b activity is restricted to earlier stages of development. **(A)** Schematized model of when heat shocks occurred relative to standard stages of NCC development in zebrafish. **(B–H)** Whole mount immunolabeled *hsp70l:GFP-hoxb5b*$^{+}$ embryos and their GFP- sibling controls (HS14: n = 6, HS21: n = 6, HS48: n = 12) using an antibody against Elavl3/4. *hsp70l:GFP-hoxb5b*$^{+}$ embryos heat shocked either at 14 hpf (n = 11) or 21 hpf (n = 10) both exhibited fewer differentiating enteric neurons than controls, but not embryos heat shocked at 48 hpf (n = 14). Numbers of enteric neurons are quantified in (H, HS14: p = 0.000232; HS21: p = 0.001684; HS48: p = 0.5813). Scale Bar **(B)**: 100 μm.

enteric neural progenitors. In addition, we also conducted these experiments in the Tg (*-4.9sox10:eGFP*) embryos (hereafter referred to as *sox10:GFP*) (Carney et al., 2006), which marks migratory NCCs with cytoplasmic GFP, and relying on residual GFP signal post fixation to label the recently *sox10*$^{+}$ enteric neural progenitor cells, as we have previously (Howard et al., 2021). Notably, Caspase-3^{+} cells were rare in all controls tissues examined from 33 to 66 hpf (**Supplementary Figures S3A,C,E,G**). While small patches of apoptotic cells can be found proximal to the POD at 33 hpf and 55 hpf (**Supplementary Figures S3B,D**), there was not a detectable onset of cell death between 55 and 66 hpf along the entire vagal and gut region (**Supplementary Figures S3F,H**) to support the loss of enteric neural progenitors through apoptosis following Hoxb5b overexpression.

We next examined the progenitor state of enteric cells at 63 hpf, a window of development prior to the 3 dpf cut off, but after the 50 hpf NCC expansion noted previously. To this end, we asked if *sox10:GFP*$^{+}$ and/or Phox2b^{+} cell numbers were reduced along the gut at 63 hpf. Intriguingly, *vp16-hoxb5b* did not lead to a significant change in the total number of GFP^{+} cells along the gut tube at 63 hpf (**Figures 7A–C**). Similarly, there was no change in the number of Phox2b^{+} cells, compared with control embryos (**Figure 7C**). This was in contrast to our earlier observation that enteric Phox2B^{+} cells were increased at 50 hpf (**Figure 5**). Furthermore, the fraction of *sox10:GFP*$^{+}$ cells expressing Phox2b was unchanged (**Figure 7D**). Therefore, we found that by 63 hpf the enteric progenitors exhibited enteric differentiation capacity, despite their decreased abundance in the presence of excess Hoxb5b.

That we observed increased Phox2b^{+} cells at 32 and 50 hpf, yet saw a dampening of their numbers by 63 hpf, this suggested that the kinetics of enteric progenitor expansion may have been adversely affected following elevated Hoxb5b. Plotting the total number of Phox2b^{+} cells counted from each developmental stage assayed throughout this study (**Figure 7E**) revealed a steady increase in control Phox2b^{+} cells with developmental age, whereas Hoxb5b overexpressing animals presented a stalled curve at 63 hpf. As we previously have shown that the enteric cells are not cleared by apoptosis during the 63–80 hpf transition, these results indicated that Hoxb5b activity modulates enterically-fated NCC capacity to expand as a population along the gut. The Hoxb5b-dependent precocious expansion of NCC leads to a fixed number of available progenitors which are then unable to expand in sufficient numbers to lead to proper ENS formation. These findings position Hoxb5b as a fine-scale regulator of enteric NCC number.

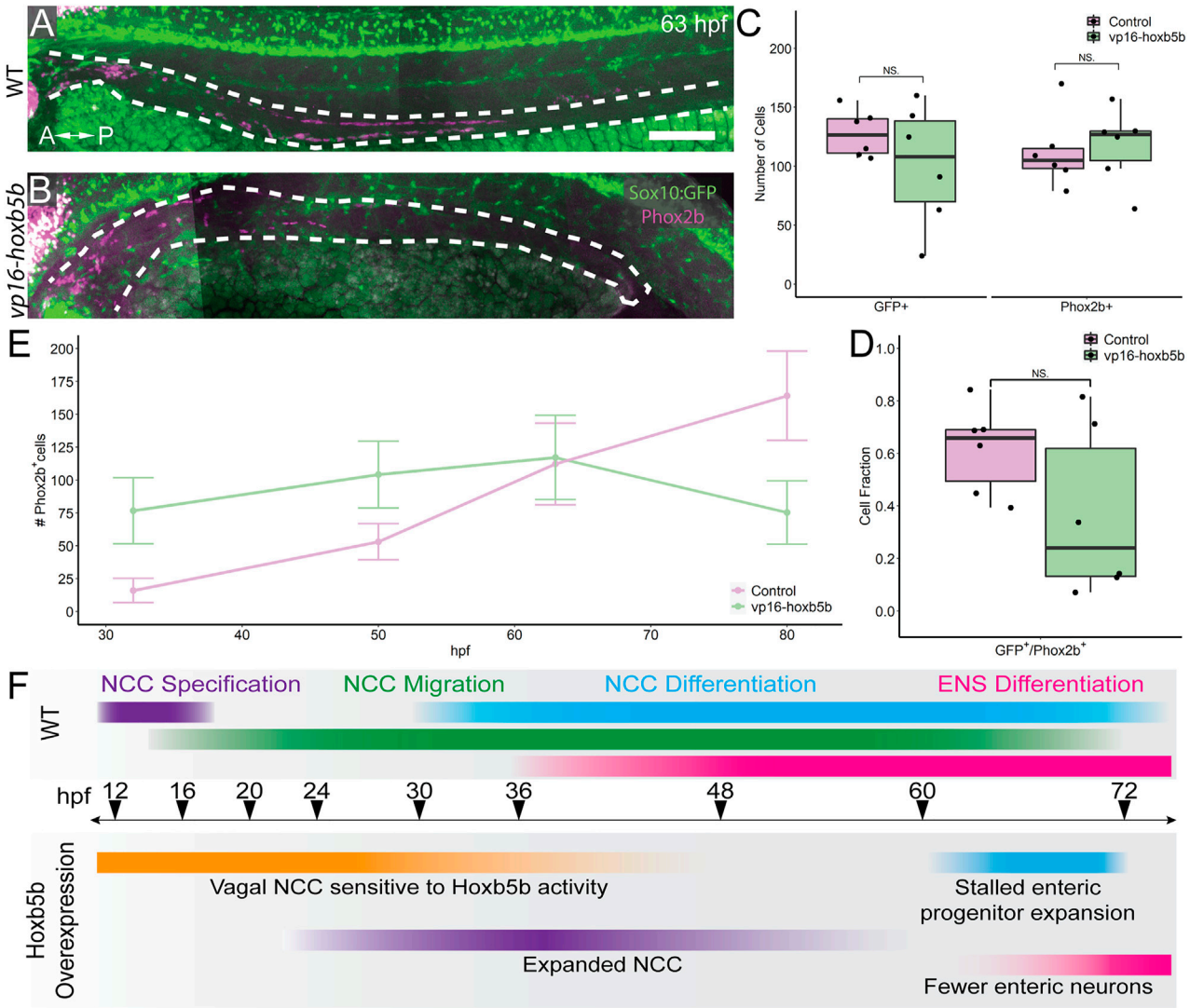


FIGURE 7 | Elevated Hoxb5b abrogates the expansion capacity of enteric NCC-derived neuronal lineage. **(A,B)** By 63 hpf, *sox10:GFP*$^{+}$ embryos immunolabeled for Phox2b show cells which have partially migrated along the gut in both control and *vp16-hoxb5b* injected animals. Gut tract outlined with white dashes. **(C)** No changes were found in the number of GFP^{+} NCC lineage or Phox2b^{+} cells along the gut between control embryos ($n = 6$) and those overexpressing *vp16-hoxb5b* ($n = 6$) (Enteric GFP^{+}: $p = 0.2619$; Enteric Phox2b^{+}: $p = 0.7896$). **(D)** Assessment of fraction of GFP^{+} cells which are also positive for Phox2b^{+} restricted to the gut axis from the same animals in **(C)**, as a measure of NCC which have initiated their differentiation programs shows the fraction of Sox10^{+} lineage cells which have turned on Phox2b expression is unchanged following elevated Hoxb5b ($p = 0.1253$). **(E)** Summary of the number of Phox2b^{+} cells for either WT control embryos or embryos injected with 15 pg *vp16-hoxb5b* mRNA as a function of age from animals used throughout this study. The trend shows while the number of Phox2b^{+} cells initially is greater than WT, eventually the WT Phox2b^{+} numbers continue to increase with the Phox2b^{+} population in Hoxb5b elevated embryos remains flat. Error bars reflect ± one standard deviation. **(F)** Graphical Model of the role of Hoxb5b in NCC development, such that early Hoxb5b expression grossly expands NCC numbers, while later in development Hoxb5b suppresses NCC enteric expansion. Scale bar **(A)**: 100 μm.

DISCUSSION

We discovered that throughout the course of NCC specification and migratory phases, *hoxb5b* is expressed within the vagal NCC domain along the post-otic/posterior zebrafish embryo. Enhancement of Hoxb5b activity was sufficient to dramatically expand NCC localization patterns, as well as their number, along the embryo. The expansion of NCCs was also accompanied by domain expansions in vagal NCC marker genes *phox2bb* and *foxd3*. Temporally-restricted pulses of ectopic Hoxb5b during early vagal NCC developmental phases was sufficient to swiftly expand vagal NCC populations, which also persisted well into the second day in development. While many vagal NCC derivatives were unaltered by 3 dpf, NCC derivatives along the enteric neural trajectory were dramatically impacted. While we observed an increase in enteric neural progenitors along the developing gut tube at 50 hpf, they failed to expand and colonize the gut efficiently, resulting in a marked decrease in the number of enteric neurons. The decrease in enteric neural progenitors following Hoxb5b induction appears to be due to a Hoxb5b-

dependent modulation of the colonization capacity of enteric neural progenitors. Cumulatively these data position Hoxb5b as a potent regulator of NCC patterning and number during early embryonic development (**Figure 7F**).

The potential involvement of Hoxb5b in zebrafish NCC development has been suggested by previous expression analyses, with discrete expression along the dorsolateral neural tube and in the post-otic domain, posterior to rhombomere 8 (Kudoh et al., 2001; Jarinova et al., 2008; Waxman et al., 2008; Barsh et al., 2017). Additional expression domains are found within the lateral plate mesoderm and foregut by 24 hpf (Dalgin and Prince, 2021). As identified in a single cell atlas of posterior zebrafish NCC lineages at 48–50 and 68–70 hpf, *hoxb5b* was among the most pervasively expressed transcripts encoding for a Hox transcription factor in both the NCC and in neural fated lineages (Howard et al., 2021). Extending this prior work, our HCRs show for the first time at high resolution the persistent expression of *hoxb5b* in zebrafish posterior (post-otic) NCC through the early course of their development.

In our zebrafish model, despite the regional potential exhibited by NCCs (Rocha et al., 2020), elevated Hoxb5b uniformly expanded cranial, vagal, and trunk NCC progenitors, though more robustly among the posterior NCC populations. Whether the Hoxb5b-dependent increase in NCCs is driven by increased NCC specification from the neural tube or through upregulation of NCC proliferation remains unresolved and was beyond the current scope of this study. Regardless of the underlying mechanism, these findings clearly indicate that Hoxb5b participates in NCC development from an early position in the NCC gene regulatory network (Simoes-Costa and Bronner, 2016; Martik and Bronner, 2017). Further, the rapid sensitivity of vagal NCC to temporally restricted pulses of elevated Hoxb5b suggests competency of these cells to abruptly respond to Hox activity early in their development. Indeed, early NCCs appear sensitive not only to Hox mediated activity but also the amount of Hoxb5b present. In our injection experiments, POD NCCs increased coordinately with amount of *vp16-hoxb5b* mRNA delivered. As such, the number of cells fated in select NCC lineages, such as the enteric NCCs, appear to be influenced not only by the activation of Hoxb5b-dependent activity, but also in part through the levels of Hoxb5b expression. From data derived across multiple animal models, Hoxb5b and its orthologues are known to be under control of several classical morphogenic signals, including WNTs (Lengerke et al., 2008), NOTCH (Hortopan and Baraban, 2011), and Retinoic Acid (Waxman et al., 2008), which enables fine spatiotemporal tuning of the levels of Hox expression. Summarily, our data thus illuminate a model in which Hoxb5b serves a potent regulator of posterior NCC identity and cell number, dependent on additional unspecified cofactors as well as its expression level.

Our findings regarding Hoxb5b in zebrafish NCC is complementary to and extends the developmental understanding of mammalian Hoxb5. For example, dominant negative suppression of Hoxb5 activity in NCC lineages led to a depletion of several NCC-derived cell populations including DRGs, pigment cells, as well as enteric neural progenitors (Lui et al., 2008; Kam et al., 2014b). Complementing these prior loss of function studies, our gain of function data indicates that elevated Hoxb5b activity is sufficient to induce expansion of vagal NCC progenitors—that paradoxically also leads to severe ENS hypoganglionosis. While we did not observe corresponding changes in DRG and pigment populations in our experiments; early enteric progenitors were dramatically increased in response to Hoxb5b activity along the gut, prior to the onset of neurogenesis, yet failed to properly execute enteric neuronal differentiation. Somewhat counterintuitively, the increase in POD NCCs following increased Hoxb5b activity did not correspondingly manifest a pan increase in vagal-derived NCC lineages. While many vagal-derived lineages exhibited no discernable phenotypic change, we observed a dramatic decrease in the number of enteric neurons in animals with elevated Hoxb5b activity. The shift in abundance of enterically fated cells in embryos overexpressing Hoxb5b was not the result of NCC-specific apoptosis or abrogated differentiation potential. Indeed, there was no change in fractions of enteric NCCs which had initiated differentiation programs following elevated Hoxb5b. Rather this tissue specific cell decrease appears to be caused by a late-onset suppression of enteric neuroblasts expansion. Assays for apoptosis-mediated cell death revealed limited tissue death, which was not restricted to NCC populations. While we did not test cell loss by other modes of death, such as necrosis or pyroptosis, the numbers of enteric NCC-derived cells were more consistent with the arrest of cell division as opposed to changes in cell survival. The insufficiency of ectopic Hoxb5b to expand the DRG and pigment lineages suggests a separate regulatory mechanism for Hoxb5b gain of function activity in enteric NCC.

Several possible frameworks may explain the differential role of Hoxb5b to expand early NCCs but abrogate their later expansion in the ENS. First, the differential role may be correlated with the temporally dynamic expression of additional co-factors, such as other Hox or TALE-family transcription factors within the ENS linage. TALE-family factors in particular, such as Meis3 (Uribe and Bronner, 2015), or other factors such as Pbx3 (Di Giacomo et al., 2006) are both expressed within the early enteric neural lineage and may cooperate with Hoxb5b to facilitate functional specificity. Additional reflection on the emerging numerous descriptions of combinatorial "Hox Codes" which define NCC identity (Parker et al., 2019; Soldatov et al., 2019; Howard et al., 2021), reported functional similarity in certain tissues (Jarinova et al., 2008), as well as complex regulatory relationship between various members of the Hox gene family (Zhu et al., 2017), the prospect of a shared sensitivity in enteric NCCs to Hoxb5b and other posterior Hox transcription factors is possible, though beyond the context of this study. Another alternative hypothesis to explain Hoxb5b's differential role may be related to the precociousness of the expansion of neural crest itself. As more NCCs accumulate earlier in development, temporally restricted signals may be insufficiently timed to signal for NCCs to continue to expand in number. While the underlying mechanism remains to be fully elucidated, when the data are considered together with mammalian suppression of activity studies referenced above,

our gain of function results suggest that the vertebrate embryo is exquisitely sensitive to perturbations in Hoxb5 activity, where either elevations or reductions in Hoxb5 lead to severe ENS defects. Collectively, we have discovered evidence in support of a model in which Hoxb5b plays an important role in NCC development, demonstrating the capacity to both expand vagal NCC localization and numbers. While additional questions still remain, these findings greatly inform our understanding of the role of posterior Hox genes in NCC development.

MATERIALS AND METHODS

Zebrafish Husbandry and Transgenic Lines

Synchronously staged embryos for each experiment were collected via controlled breeding of adult zebrafish. After collection, embryos were maintained in standard E3 media at 28°C until 24 h post fertilization (hpf), then transferred to 0.003% 1-phenyl 2-thiourea (PTU)/E3 solution (Karlsson et al., 2001), with the exception of larvae used to assay pigmentation, which were cultured in E3 media only. Transgenic embryos for the Tg (*-4.9sox10:EGFP*)^{ba2Tg} (Carney et al., 2006) and Tg (*-7.2sox10:mRFP*)^{vu234} (Kucenas et al., 2008) were generally sorted between 17–28 hpf for fluorescence while Tg (*hsp70l:EGFP-hoxb5b;acry:dsRed*)^{ci1014} and Tg (*-8.3phox2bb:kaede*) (Harrison et al., 2014) embryos were sorted for transgenic expression between 60–78 hpf. Tissue was collected from embryos out of their chorions at the stage noted in each experiment as described in (Ibarra-García-Padilla et al., 2021). All work was performed under protocols approved by, and in accordance with, the Rice University Institutional Animal Care and Use Committee (IACUC).

Generation of the *Tg* (*hsp70l:EGFP-Hoxb5b*) Transgenic Line

To generate EGFP-Hoxb5b, *egfp* was fused to the 5′-end of zebrafish *hoxb5b* with PCR. A sequence encoding a 7 amino acid linker was incorporated and the *hoxb5b* ATG was deleted to prevent alternative transcriptional initiation of *hoxb5b* downstream. To generate the *hsp70l:EGFP-Hoxb5b* transgene, standard Gateway methods were used (Kwan et al., 2007). The transgene includes a *EGFP-Hoxb5b* middle-entry vector and the reported *p5E-hsp70l* 5′-Entry and *p3E-polyA* 3′-entry vectors (Kwan et al., 2007), which were incorporated into the *pDestTol2-acry:dsRed* vector (Mandal et al., 2013). Sanger sequencing was used to confirm the proper orientation of the constructs within the destination vector and the sequence of *EGFP-hoxb5b*. Transgenic embryos were created by co-injecting wild-type embryos at the one-cell stage with 25 pg *hsp70l:EGFP-hoxb5b* vector and 25 pg of *Tol2* mRNA (Kawakami, 2004; Kawakami et al., 2004). Embryos were raised to adulthood and screened for the present of dsRed in the lens at ~3 days and the ability to induce robust EGFP expression following heat-shock (**Supplementary Figure S2**). Multiple founders for the Tg (*hsp70l:EGFP-hoxb5b;acry:dsRed*)^{ci1014} line were identified so the line that induced the most robust expression following heat-shock was retained. While some ectopic notochord expression was observed in non-heat shocked embryos they did not exhibit overt phenotypes and developed normally. For heat shock experiments, through routine outcrossing of transgenic animals to the wild type embryos of the AB/TL backgrounds, GFP-hoxb5b$^{-/-}$ siblings are produced with each subsequent breeding, which were heat shocked and processed in parallel.

Preparation and Injections of *hoxb5b* mRNA

Capped *vp16-hoxb5b* mRNA was prepared off a Not1 linearized pCS2+ plasmid containing the *vp16-hoxb5b* coding sequence using the Sp6 mMessage Kit (Ambion), as first reported in (Waxman et al., 2008). The *vp16-hoxb5b* construct encodes for a hyperactive form of Hoxb5b and allowed for lower doses of mRNA to be delivered (Waxman and Yelon, 2009). Embryos were injected prior to the four-cell stage with either 15 pg or 30 pg of mRNA and were empirically determined to produce similar phenotypes. Dosage of mRNA was determined per experimental condition. Uninjected wild type sibling embryos were cultured in parallel with injected animals and used for controls. Dead and grossly malformed embryos were removed from analysis.

In Situ Hybridization

In situ hybridizations were performed similarly to the protocol of Jowett and Lettice, 1994, which should be referenced for specific details. Briefly, antisense digoxigenin-labeled riboprobes were generated from previously characterized plasmids containing sequences for *crestin* (Luo et al., 2001), *foxd3* (Odenthal and Nüsslein-Volhard, 1998; Hochgreb-Hagele and Bronner, 2013), *phox2bb* (Uribe and Bronner, 2015), and *hoxb5b* (Waxman et al., 2008). As per the protocol, PFA-fixed whole mount embryos stored in methanol were rehydrated in PBST, permeabilized with Proteinase K digestion (10 μg/ml), and post-fixed in 4% PFA. Embryos were incubated in probes overnight (~16 h) at 65°C and washed sequentially in graded SSCT buffers. Riboprobes solutions were recovered and stored at −20°C for reuse, with multiple uses leading to minimal loss of signal. Probed embryos were blocked for 1–2 h at ambient temperature in 5% Goat sera in PBST before detection overnight (~16 h) at 4°C using an anti-Digoxigenin-Fab fragments conjugated to Alkaline Phosphatase enzymes (1:1,000 dilution, Roche) in 5% Goat Sera in PBST. Finally, riboprobes were visualized with NBT/BCIP solution (3.5 μL each of NBT, BCIP stock solutions, Roche). Probes were validated prior to use on wildtype embryos to calibrate staining duration, with patterns compared to those curated on ZFIN (Howe et al., 2013; Ruzicka et al., 2019).

Whole Mount Immunohistochemistry, HCR, and WICHCR

Immunohistochemistry (IHC), Hybridization Chain Reaction (HCR), and Whole mount Immuno-Coupled Hybridization Chain Reaction (WICHCR) protocols all were conducted according to the methods published in Ibarra-García-Padilla et al., 2021. All IHC assays conducted in blocking 5% Goat Sera in 1X PBST. Primary antibodies against the following

TABLE 1 | Summary of Statistical Tests.

Figure	Panel	Condition	Age	Comparison to Control	Test	*p*-value	Control Animal N	Experimental Animal N
2	E	norm area POD Crestin + stain	32 hpf	vp16-hoxb5b	T-test	7.75E-07	27	21
2	E	norm area POD Crestin + stain	50 hpf	vp16-hoxb5b	T-test	1.14E-03	7	8
3	E	norm area POD foxd3 + stain	32 hpf	vp16-hoxb5b	Welch's T-test	5.71E-03	7	10
3	F	norm area hindbrain Phox2bb + stain	32 hpf	vp16-hoxb5b	Welch's T-test	0.01	7	8
3	G	norm area POD phox2bb + stain	32 hpf	vp16-hoxb5b	Welch's T-test	5.63E-03	7	8
3	O	POD phox2b + cells	32 hpf	15 pg vp16-hoxb5b	T-test	3.03E-05	7	9
3	O	POD phox2b + cells	32 hpf	30 pg vp16-hoxb5b	T-test	2.61E-05	7	5
4	I	norm area POD Crestin + stain	HS 14|24 hpf	hsp70:GFP-hoxb5b+	T-test	0.02	6	8
4	J	norm area POD Crestin + stain	HS 22|26-28 hpf	hsp70:GFP-hoxb5b+	T-test	0.04	10	14
5	E	# total Enteric Phox2b + cells	50 hpf	15 pg vp16-hoxb5b	T-test	1.83E-04	10	7
5	E	# total Enteric Phox2b + cells	50 hpf	30 pg vp16-hoxb5b	Welch's T-test	3.01E-04	10	9
5	F	# POD phox2b + cells	50 hpf	15 pg vp16-hoxb5b	Wilcoxon Test	3.28E-03	10	7
5	F	# POD phox2b + cells	50 hpf	30 pg vp16-hoxb5b	Kruskal-Wallis Test	6.80E-04	10	9
5	J	# Total Kaede + cells	3 dpf	15 pg vp16-hoxb5b	T-test	1.10E-04	7	8
5	J	# Total Kaede + cells	3 dpf	30 pg vp16-hoxb5b	T-test	4.09E-05	7	7
5	K	# Total Elavl3/4 + cells	3 dpf	15 pg vp16-hoxb5b	T-test	1.35E-03	7	8
5	K	# Total Elavl3/4 + cells	3 dpf	30 pg vp16-hoxb5b	T-test	1.04E-04	7	7
5	L	Copositive Elavl3/4+/Kaede + cells of Total Kaede + cells	3 dpf	15 pg vp16-hoxb5b	Wilcoxon Test	0.13	7	8
5	L	Copositive Elavl3/4 +/Kaede+ cells of Total Kaede + cells	3 dpf	30 pg vp16-hoxb5b	T-test	0.10	7	7
5S	I	# Gut Restricted Phox2b + Cells	50 hpf	15 pg vp16-hoxb5b	T-test	5.68E-04	10	7
5S	I	# Gut Restricted Phox2b + Cells	50 hpf	30 pg vp16-hoxb5b	Welch's T-test	3.98E-04	10	9
5S	J	Cell Fraction, POD Localized	50 hpf	15 pg vp16-hoxb5b	T-test	0.25	10	7
5S	J	Cell Fraction, POD Localized	50 hpf	30 pg vp16-hoxb5b	T-test	0.37	10	9
5S	J	Cell Fraction, Gut Localized	50 hpf	15 pg vp16-hoxb5b	T-test	0.25	10	7
5S	J	Cell Fraction, Gut Localized	50 hpf	30 pg vp16-hoxb5b	T-test	0.35	10	9
6	H	elavl3/4+ cells	HS 14|3 dpf	hsp70:GFP-hoxb5b+	T-test	2.32E-04	6	11
6	H	elavl3/4 + cells	HS 21|3 dpf	hsp70:GFP-hoxb5b+	T-test	1.68E-03	6	10
6	H	elavl3/4 + cells	HS 48|3 dpf	hsp70:GFP-hoxb5b+	T-test	0.58	12	14
7	C	# GFP + cells	63 hpf	Enteric GFP+	T-test	0.26	6	6
7	C	# Phox2b + cells	63 hpf	Enteric Phox2b+	T-test	0.79	6	6
7	D	Fraction Phox2b + copositive sox10:GFP cells	63 hpf	15 pg vp16-hoxb5b	T-test	0.13	6	6

proteins were used as follows: Phox2b (1:200, Santa Cruz, B-11), Kaede (1:500, MBL International, PM102M), Elavl3/4 (Same as HuC/D, 1:500, Invitrogen Molecular Probes, A21271), Activated Caspase-3 (1:200, BD Biosciences, 559565). Incubation in primary antibody solutions were conducted overnight at 4°C, except for assays with Phox2b or Caspase-3 antibodies which were allowed to incubate for 2 days at 4°C which provided optimal labeling. Corresponding secondary antibodies conjugated to spectrally distinct fluorophores were all used at 1:500 dilution, selected from the following depending on the experimental condition: Alexa Fluor 488 goat anti-rabbit IgG (ThermoFischer, A11008), Alexa Fluor 568 goat anti-rabbit IgG (ThermoFischer, A11011), Alexa Fluor 488 goat anti-mouse IgG1 (ThermoFischer, A21121), Alexa Fluor 594 goat anti-mouse IgG1 (ThermoFischer, A21125), Alexa Fluor 647 goat anti-mouse IgG1 (ThermoFischer, A21240), and Alexa Fluor 647 goat anti-mouse IgG2b (ThermoFischer, A21242). In the HCR and WICHCR assays, commercially designed probes were secured from Molecular Instruments as follows: *crestin* (B3, AF195881.1), *hoxb5b* (B2, BC078285.1), *phox2bb* (B1, NM_001014818.1), and used as prior (Ibarra-García-Padilla et al., 2021). Corresponding amplifiers were purchased from Molecular instruments and were used in experiments to include spectrally distinct fluorophores suitable for multiplexed imaging.

Heat Shock Induction of the *hsp70l:GFP-hoxb5b* Transgene

Adult zebrafish maintained as an outcross and positive for dsRed expression as larvae in the lens (see above method on description of the line) were bred to produce synchronously staged embryos. At the stage designated to begin the heat shock, embryos were rapidly transferred to 37°C E3 and maintained at that temperature. After a 1-h incubation, embryos were rapidly returned to 28°C. In the 1–3 h after heat shock GFP-Hoxb5b^{+} embryos were sorted from GFP^{-} siblings and cultured in parallel until the designated stage for tissue collection.

Imaging, Quantification, and Image Visualization

All embryos prior to imaging were cleared through graded washes of PBST/Glycerol to reach a final Glycerol content of 75%. Fluorescent Z-stacked images of IHC processed embryos were captured using an Olympus FV3000 point scanning confocal microscope supported by Fluoview Acquisition Software (version FV31S-SW). Images were stitched in FIJI (ImageJ version 1.53e) using the Grid/Collection applet as part of the Stitching plugin (Schindelin et al., 2012; Schneider et al., 2012; Rueden et al., 2017). Digital image files were converted with ImarisFileConverter Software (Bitplane) to three dimensional rendered images compatible with IMARIS (V9.4, V9.7, Bitplane). All images of fluorescent animals represented in this publication are derived from maximum intensity projections of the z-stacked image. Cells counts were conducted on volume images following an arithmetic background subtraction in IMARIS to ensure accurate counts, particularly when determining coincidence of labels. ISH processed embryos were imaged on a Nikon Eclipse Ni microscope equipped with a motorized stage. Z-stack images were acquired and extended depth of focus images were generated in the Nikon NIS-Elements BR software (v5.02.00). Areas of expression were measured in FIJI to include dark pixels in the post-otic or hindbrain domains. Quantifications were curated and analyzed in the Rstudio programming environment (v1.1.463). Images of pigmented embryos were captured similarly on the Nikon microscope with lateral illumination to distinguish iridophores, similar to (Petratou et al., 2021).

In vivo Confocal Microscopy

Embryos were sorted for RFP expression, anesthetized with 0.4% tricaine, and embedded in 1% low melting agarose in a 28.5°C chamber with a coverslip glass bottom. Care was taken to ensure embryo angled appropriately and proximal to the glass. Z-stack images were acquired approximately every half hour concurrently on the same Olympus FV3000 point scanning confocal microscope as above. Maximum intensity projections images were generated and exported from FIJI.

Statistical Analysis

An α of 0.05 was used as a cut off for all statistical tests. Normalcy of datasets was assessed by visual inspection of a density plot, a qqplot against a linear theoretical distribution, and Shapiro-Wilk test for Normalcy. Further, variance between each dataset was examined either with a Bartlett test for data which adhered to normalcy or with a Levene's test for non-normal data. Based on the normalcy and scedasticity conditions of the data, the appropriate statistical test was selected, as summarized in **Table 1**. All statistical analyses were carried out in the Rstudio (v1.1.463) programming environment, with key dependencies on the lawstat (v3.4) and stats (v3.6.3) packages. Plots were generated in Rstudio supported by the ggplot2 (v3.3.2) and ggsignif (v0.6.0) packages.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

The animal study was reviewed and approved by the Institutional Animal Care and Use Committee (IACUC) of Rice University and Cincinnati Children's Hospital Medical Center.

AUTHOR CONTRIBUTIONS

AH and RU designed the experiments and wrote the manuscript. RU and AA conceived the idea. AH, AN, JT, PR, ES, CL, GK, and RU performed the experiments. PR and JW created the zebrafish transgenic line. AH and RU performed the data analyses. All the

authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by Rice University, a Burroughs Wellcome Fund PDEP Award, Cancer Prevention and Research Institute of Texas (CPRIT) Recruitment of First-Time Tenure Track Faculty Members (No. CPRIT-RR170062) and the NSF CAREER Award (No. 1942019) awarded to RU as well as through the NIH NHLBI (Nos. HL137766, HL141186) awarded to JW.

ACKNOWLEDGMENTS

We offer our sincerest gratitude to Dr. Dan Wagner and to the entire Uribe Lab at Rice University for their insights and support throughout this project. We thank Dr. Budi Utama and the Rice University Shared Equipment Authority on IMARIS image analysis suite, which was indispensable toward this project. We thank Dr. Eric Bridenbaugh and Dr. Mariane Martinez at Olympus for their expert advice on confocal microscopy. We also thank Jonny Diaz and A. Augello Cook for technical assistance.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2021.803370/full#supplementary-material

Supplementary Figure S1 | Phox2b antibody labeling is synonymous with mRNA and transgenic labeling methods. **(A)** 3 dpf *-8.3phox2bb:kaede* larva co-labeled via WICHCR protocol (Ibarra-García-Padilla et al., 2021) showing Kaede labeled cells (green), *phox2bb* mRNA (magenta), and Phox2b protein (cyan). **(B)** Inset images in the hindbrain exhibit coincidence of all three probes along the Hindbrain (Hb), Pharyngeal (Pa), and PODs (arrowheads). Kaede$^+$ cells which are negative for both mRNA and protein are detected at the dorsal most aspect of the motor neuron complex (star) as well as in anterior regions of the presumptive CNS. **(C)** Images highlighting the midgut axis reveals a perfect coincidence of all three enteric labels with very little background, demonstrating the efficacy of the Phox2b antibody as an enteric neural crest label. While nearly every cell is coincident, selected cells are annotated to aid in comparison (arrowheads). Scale bar A: 100 μm.

Supplementary Figure S2 | Characterization of novel *hsp70l:GFP-hoxb5b* transgenic zebrafish line. **(A)** GFP-Hoxb5b$^+$ overexpressing embryos at 24 hpf are easily distinguishable from GFP$^-$ siblings 1 hour after heat shock. GFP$^-$ and GFP$^+$ embryos are cultured and assayed under identical conditions to control for variations induced by growth at an elevated temperature. **(B–C)** GFP-Hoxb5b construct is localized to cell nuclei following heat shock and persists at least a day after initial induction **(C)**. Scale bar B: 100 μM.

Supplementary Figure S3 | Depletion of NCC-derived cells in the gut is not the result of cell death. **(A–H)** Wholemount IHC embryos stained with an antibody against activated Caspase-3, a marker for apoptotic cells. NCCs are visualized either in the *sox10:GFP* or WT backgrounds in either uninjected or 15 pg *vp16-hoxb5b* injected embryos. Caspase-3$^+$ cells are very rare along the length of the vagal domain in all control embryos ranging from 33 hpf to 66 hpf (A,C,E,G). While discrete pockets of Caspase-3$^+$ cells are detected proximal to the POD NCC niche at 33 hpf **(B)** and 55 hpf **(D)**, very few co-positive cells are present at these time frames. Later in development at 63–66 hpf **(F,H)** almost no Caspase-3$^+$ cells remain along the posterior aspect of the embryo, with only ectopic labeling of the somatic myotome detectable.

Supplementary Figure S4 | Vagal NCC derivatives following Hoxb5b overexpression are largely unaffected. **(A–D)** Investigation of distinct NCC derived pigment lineages, melanophores (red arrowheads) and iridophores (yellow arrowheads), in 5 dpf embryos overexpressing Hoxb5b compared to controls. **(E,F)** Both control embryos and Hoxb5b overexpressing embryos by 50 hpf produced phenotypically wild type anterior dorsal root ganglia (DRG), as shown in whole mount immunolabeling with an antibody against Elavl3/4. **(G,H)** Position and development of the Superior Cervical Ganglion (SCG) occurs by the third day, as shown by immuno labeling for Elavl3/4 (red arrowhead). **(I)** Quantifications from animals in **Figure 5**, showing the number of Phox2b$^+$ cells localized along the gut length, excluding the POD increases in response to *vp16-hoxb5b* mRNA induction (I, 15 pg: $p = 9.92 \times 10^{-5}$; 30 pg: $p = 0.000404$). **(J)** The distribution of Phox2b$^+$ cells between the POD and the length of the gut was unchanged following elevated Hoxb5b activity [**(J)**, 15 pg in POD: $p = 0.2515$; 30 pg in POD: $p = 0.3673$; 15 pg in Gut: $p = 0.2306$; 30 pg in Gut: $p = 0.3496$]. Scale Bars **(E,F,G,H)**: 50 μm.

REFERENCES

Amores, A., Force, A., Yan, Y. L., Joly, L., Amemiya, C., Fritz, A., et al. (1998). Zebrafish Hox Clusters and Vertebrate Genome Evolution. *Science* 282, 1711–1714. doi:10.1126/science.282.5394.1711

Barsh, G. R., Isabella, A. J., and Moens, C. B. (2017). Vagus Motor Neuron Topographic Map Determined by Parallel Mechanisms of Hox5 Expression and Time of Axon Initiation. *Curr. Biol.* 27, 3812–3825. doi:10.1016/j.cub.2017.11.022

Breau, M. A., Wilkinson, D. G., and Xu, Q. (2013). A Hox Gene Controls Lateral Line Cell Migration by Regulating Chemokine Receptor Expression Downstream of Wnt Signaling. *Proc. Natl. Acad. Sci.* 110, 16892–16897. doi:10.1073/pnas.1306282110

Carney, T. J., Dutton, K. A., Greenhill, E., Delfino-Machín, M., Dufourcq, P., Blader, P., et al. (2006). A Direct Role for Sox10 in Specification of Neural Crest-Derived Sensory Neurons. *Development* 133, 4619–4630. doi:10.1242/dev.02668

Choi, H. M. T., Schwarzkopf, M., Fornace, M. E., Acharya, A., Artavanis, G., Stegmaier, J., et al. (2018). Third-generation *In Situ* Hybridization Chain Reaction: Multiplexed, Quantitative, Sensitive, Versatile, Robust. *Development* 145, dev165753. doi:10.1242/dev.165753

Choi, H. M. T., Calvert, C. R., Husain, N., Huss, D., Barsi, J. C., Deverman, B. E., et al. (2016). Mapping a Multiplexed Zoo of mRNA Expression. *Development* 143, 3632–3637. doi:10.1242/dev.140137

Choi, H. M. T., Chang, J. Y., Trinh, L. A., Padilla, J. E., Fraser, S. E., and Pierce, N. A. (2010). Programmable *In Situ* Amplification for Multiplexed Imaging of mRNA Expression. *Nat. Biotechnol.* 28, 1208–1212. doi:10.1038/nbt.1692

Condie, B. G., and Capecchi, M. R. (1993). Mice Homozygous for a Targeted Disruption of Hoxd-3 (Hox-4.1) Exhibit Anterior Transformations of the First and Second Cervical Vertebrae, the Atlas and the axis. *Development* 119, 579–595. doi:10.1242/dev.119.3.579

Condie, B. G., and Capecchi, M. R. (1994). Mice with Targeted Disruptions in the Paralogous Genes Hoxa-3 and Hoxd-3 Reveal Synergistic Interactions. *Nature* 370, 304–307. doi:10.1038/370304a0

Creuzet, S., Couly, G., Vincent, C., and Le Douarin, N. M. (2002). Negative Effect of Hox Gene Expression on the Development of the Neural Crest-Derived Facial Skeleton. *Development* 129, 4301–4313. doi:10.1242/dev.129.18.4301

Dalgin, G., and Prince, V. E. (2021). Midline Morphogenesis of Zebrafish Foregut Endoderm Is Dependent on Hoxb5b. *Developmental Biol.* 471, 1–9. doi:10.1016/j.ydbio.2020.12.001

Di Bonito, M., Glover, J. C., and Studer, M. (2013). Hoxgenes and Region-specific Sensorimotor Circuit Formation in the Hindbrain and Spinal Cord. *Dev. Dyn.* 242, 1348–1368. doi:10.1002/dvdy.24055

Di Giacomo, G., Koss, M., Capellini, T. D., Brendolan, A., Pöpperl, H., and Selleri, L. (2006). Spatio-temporal Expression of Pbx3 during Mouse Organogenesis. *Gene Expr. Patterns* 6, 747–757. doi:10.1016/j.modgep.2005.12.002

Durbec, P. L., Larsson-blomberg, L. B., Schuchardt, A., Costantini, F., and Pachnis, V. (1996). Common Origin and Developmental Dependence on C-Ret of

Subsets of Enteric and Sympathetic Neuroblasts. *Development* 122, 349–358. doi:10.1242/dev.122.1.349

Elworthy, S., Pinto, J. P., Pettifer, A., Cancela, M. L., and Kelsh, R. N. (2005). Phox2b Function in the Enteric Nervous System Is Conserved in Zebrafish and Is Sox10-dependent. *Mech. Development* 122, 659–669. doi:10.1016/j.mod.2004.12.008

Ganz, J. (2018). Gut Feelings: Studying Enteric Nervous System Development, Function, and Disease in the Zebrafish Model System. *Dev. Dyn.* 247, 268–278. doi:10.1002/dvdy.24597

Gouti, M., Briscoe, J., and Gavalas, A. (2011). Anterior Hox Genes Interact with Components of the Neural Crest Specification Network to Induce Neural Crest Fates. *Stem Cells* 29, 858–870. doi:10.1002/stem.630

Harrison, C., Wabbersen, T., and Shepherd, I. T. (2014). *In Vivo* visualization of the Development of the Enteric Nervous System Using aTg(−8.3bphox2b:Kaede) transgenic Zebrafish. *Genesis*. 52, 985–990. doi:10.1002/dvg.22826

Hochgreb-Hägele, T., and Bronner, M. E. (2013). A Novel FoxD3 Gene Trap Line Reveals Neural Crest Precursor Movement and a Role for FoxD3 in Their Specification. *Developmental Biol.* 374, 1–11. doi:10.1016/j.ydbio.2012.11.035

Hortopan, G. A., and Baraban, S. C. (2011). Aberrant Expression of Genes Necessary for Neuronal Development and Notch Signaling in an Epileptic Mind Bomb Zebrafish. *Dev. Dyn.* 240, 1964–1976. doi:10.1002/dvdy.22680

Howard, A. G. A., Baker, P. A., Ibarra-García-Padilla, R., Moore, J. A., Rivas, L. J., Tallman, J. J., et al. (2021). An Atlas of Neural Crest Lineages along the Posterior Developing Zebrafish at Single-Cell Resolution. *Elife* 10, e60005. doi:10.7554/elife.60005

Howe, D. G., Bradford, Y. M., Conlin, T., Eagle, A. E., Fashena, D., Frazer, K., et al. (2013). ZFIN, the Zebrafish Model Organism Database: Increased Support for Mutants and Transgenics. *Nucleic Acids Res.* 41, D854–D860. doi:10.1093/nar/gks938

Hutchins, E. J., Kunttas, E., Piacentino, M. L., Howard, A. G. A., Bronner, M. E., and Uribe, R. A. (2018). Migration and Diversification of the Vagal Neural Crest. *Dev. Biol.* 444 Suppl 1, S98–S30269. doi:10.1016/j.ydbio.2018.07.004

Ibarra-García-Padilla, R., Howard, A. G. A., Singleton, E. W., and Uribe, R. A. (2021). A Protocol for Whole-Mount Immuno-Coupled Hybridization Chain Reaction (WICHCR) in Zebrafish Embryos and Larvae. *STAR Protoc.* 2, 100709. doi:10.1016/j.xpro.2021.100709

Jarinova, O., Hatch, G., Poitras, L., Prudhomme, C., Grzyb, M., Aubin, J., et al. (2008). Functional Resolution of Duplicatedhoxb5genes in Teleosts. *Development* 135, 3543–3553. doi:10.1242/dev.025817

Jowett, T., and Lettice, L. (1994). Whole-mount *In Situ* Hybridizationon Zebrafish Embryos Using a Mixture of Digoxigenin- and Fluorescein-Labelled Probes. *Trends Genet.* 10, 73–74. doi:10.1016/0168-9525(94)90220-8

Kam, M. K. M., Cheung, M. C. H., Zhu, J. J., Cheng, W. W. C., Sat, E. W. Y., Tam, P. K. H., et al. (2014b). Perturbation of Hoxb5 Signaling in Vagal and Trunk Neural Crest Cells Causes Apoptosis and Neurocristopathies in Mice. *Cell Death Differ* 21, 278–289. doi:10.1038/cdd.2013.142

Kam, M. K. M., Cheung, M., Zhu, J. J., Cheng, W. W. C., Sat, E. W. Y., Tam, P. K. H., et al. (2014a). Homeobox B5 (Hoxb5) Regulates the Expression of Forkhead Box D3 Gene (Foxd3) in Neural Crest. *Int. J. Biochem. Cel Biol.* 55, 144–152. doi:10.1016/j.biocel.2014.09.002

Kam, M. K. M., and Lui, V. C. H. (2015). Roles of Hoxb5 in the Development of Vagal and Trunk Neural Crest Cells. *Develop. Growth Differ.* 57, 158–168. doi:10.1111/dgd.12199

Karlsson, J., Von Hofsten, J., and Olsson, P. E. (2001). Generating Transparent Zebrafish: A Refined Method to Improve Detection of Gene Expression during Embryonic Development. *Mar. Biotechnol. (Ny)* 3, 522–527. doi:10.1007/s1012601-0053-4

Kawakami, K., Takeda, H., Kawakami, N., Kobayashi, M., Matsuda, N., and Mishina, M. (2004). A Transposon-Mediated Gene Trap Approach Identifies Developmentally Regulated Genes in Zebrafish. *Developmental Cel.* 7, 133–144. doi:10.1016/j.devcel.2004.06.005

Kawakami, K. (2004). "Transgenesis and Gene Trap Methods in Zebrafish by Using the Tol2 Transposable Element," in *Methods in Cell Biology* (Academic Press), 201–222. doi:10.1016/s0091-679x(04)77011-9

Kucenas, S., Takada, N., Park, H.-C., Woodruff, E., Broadie, K., and Appel, B. (2008). CNS-derived Glia Ensheath Peripheral Nerves and Mediate Motor Root Development. *Nat. Neurosci.* 11, 143–151. doi:10.1038/nn2025

Kudoh, T., Tsang, M., Hukriede, N. A., Chen, X., Dedekian, M., Clarke, C. J., et al. (2001). A Gene Expression Screen in Zebrafish Embryogenesis. *Genome Res.* 11, 1979–1987. doi:10.1101/gr.209601

Kwan, K. M., Fujimoto, E., Grabher, C., Mangum, B. D., Hardy, M. E., Campbell, D. S., et al. (2007). The Tol2kit: A Multisite Gateway-Based Construction Kit forTol2 Transposon Transgenesis Constructs. *Dev. Dyn.* 236, 3088–3099. doi:10.1002/dvdy.21343

Le Douarin, N., and Kalcheim, C. (1999). *The Neural Crest*. Cambridge University Press.

Le Douarin, N. M. (1982). *The Neural Crest*. Cambridge: Cambridge University Press.

Lengerke, C., Schmitt, S., Bowman, T. V., Jang, I. H., Maouche-Chretien, L., McKinney-Freeman, S., et al. (2008). BMP and Wnt Specify Hematopoietic Fate by Activation of the Cdx-Hox Pathway. *Cell Stem Cell* 2, 72–82. doi:10.1016/j.stem.2007.10.022

Lister, J. A., Cooper, C., Nguyen, K., Modrell, M., Grant, K., and Raible, D. W. (2006). Zebrafish Foxd3 Is Required for Development of a Subset of Neural Crest Derivatives. *Developmental Biol.* 290, 92–104. doi:10.1016/j.ydbio.2005.11.014

Lui, V. C. H., Cheng, W. W. C., Leon, T. Y. Y., Lau, D. K. C., Garcia–Bareclo, M. M., Miao, X. P., et al. (2008). Perturbation of Hoxb5 Signaling in Vagal Neural Crests Down-Regulates Ret Leading to Intestinal Hypoganglionosis in Mice. *Gastroenterology* 134, 1104–1115. doi:10.1053/j.gastro.2008.01.028

Luo, R., An, M., Arduini, B. L., and Henion, P. D. (2001). Specific Pan-Neural Crest Expression of zebrafishCrestin throughout Embryonic Development. *Dev. Dyn.* 220, 169–174. doi:10.1002/1097-0177(2000)9999:9999<:aid-dvdy1097>3.0.co;2-1

Mandal, A., Rydeen, A., Anderson, J., Sorrell, M. R. J., Zygmunt, T., Torres-Vázquez, J., et al. (2013). Transgenic Retinoic Acid Sensor Lines in Zebrafish Indicate Regions of Available Embryonic Retinoic Acid. *Dev. Dyn.* 242, 989–1000. doi:10.1002/dvdy.23987

Martik, M. L., and Bronner, M. E. (2017). Regulatory Logic Underlying Diversification of the Neural Crest. *Trends Genet.* 33, 715–727. doi:10.1016/j.tig.2017.07.015

Minguillon, C., Nishimoto, S., Wood, S., Vendrell, E., Gibson-Brown, J. J., and Logan, M. P. O. (2012). Hox Genes Regulate the Onset of Tbx5 Expression in the Forelimb. *Development* 139, 3180–3188. doi:10.1242/dev.084814

Nagy, N., and Goldstein, A. M. (2017). Enteric Nervous System Development: A Crest Cell's Journey from Neural Tube to colon. *Semin. Cel Developmental Biol.* 66, 94–106. doi:10.1016/j.semcdb.2017.01.006

Odenthal, J., and Nüsslein-Volhard, C. (1998). fork Head Domain Genes in Zebrafish. *Development Genes Evol.* 208, 245–258. doi:10.1007/s004270050179

Parker, H. J., De Kumar, B., Green, S. A., Prummel, K. D., Hess, C., Kaufman, C. K., et al. (2019). A Hox-TALE Regulatory Circuit for Neural Crest Patterning Is Conserved across Vertebrates. *Nat. Commun.* 10, 1189. doi:10.1038/s41467-019-09197-8

Parker, H. J., Pushel, I., and Krumlauf, R. (2018). Coupling the Roles of Hox Genes to Regulatory Networks Patterning Cranial Neural Crest. *Developmental Biol.* 444, S67–S78. doi:10.1016/j.ydbio.2018.03.016

Pattyn, A., Morin, X., Cremer, H., Goridis, C., and Brunet, J.-F. (1999). The Homeobox Gene Phox2b Is Essential for the Development of Autonomic Neural Crest Derivatives. *Nature* 399, 366–370. doi:10.1038/20700

Pearson, J. C., Lemons, D., and McGinnis, W. (2005). Modulating Hox Gene Functions during Animal Body Patterning. *Nat. Rev. Genet.* 6, 893–904. doi:10.1038/nrg1726

Petratou, K., Spencer, S. A., Kelsh, R. N., and Lister, J. A. (2021). The MITF Paralog Tfec Is Required in Neural Crest Development for Fate Specification of the Iridophore Lineage from a Multipotent Pigment Cell Progenitor. *PLoS One* 16, e0244794. doi:10.1371/journal.pone.0244794

Rancourt, D. E., Tsuzuki, T., and Capecchi, M. R. (1995). Genetic Interaction between Hoxb-5 and Hoxb-6 Is Revealed by Nonallelic Noncomplementation. *Genes Dev.* 9, 108–122. doi:10.1101/gad.9.1.108

Rocha, M., Singh, N., Ahsan, K., Beiriger, A., and Prince, V. E. (2020). Neural Crest Development: Insights from the Zebrafish. *Developmental Dyn.* 249, 88–111. doi:10.1002/dvdy.122

Rueden, C. T., Schindelin, J., Hiner, M. C., Dezonia, B. E., Walter, A. E., Arena, E. T., et al. (2017). ImageJ2: ImageJ for the Next Generation of Scientific Image Data. *BMC Bioinformatics* 18, 529. doi:10.1186/s12859-017-1934-z

Ruzicka, L., Howe, D. G., Ramachandran, S., Toro, S., Van Slyke, C. E., Bradford, Y. M., et al. (2019). The Zebrafish Information Network: New Support for Non-coding Genes, Richer Gene Ontology Annotations and the Alliance of Genome Resources. *Nucleic Acids Res.* 47, D867–D873. doi:10.1093/nar/gky1090

Santini, S., Boore, J. L., and Meyer, A. (2003). Evolutionary Conservation of Regulatory Elements in Vertebrate Hox Gene Clusters. *Genome Res.* 13, 1111–1122. doi:10.1101/gr.700503

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., et al. (2012). Fiji: an Open-Source Platform for Biological-Image Analysis. *Nat. Methods* 9, 676–682. doi:10.1038/nmeth.2019

Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 Years of Image Analysis. *Nat. Methods* 9, 671–675. doi:10.1038/nmeth.2089

Simoes-Costa, M., and Bronner, M. E. (2016). Reprogramming of Avian Neural Crest Axial Identity and Cell Fate. *Science* 352, 1570–1573. doi:10.1126/science.aaf2729

Soldatov, R., Kaucka, M., Kastriti, M. E., Petersen, J., Chontorotzea, T., Englmaier, L., et al. (2019). Spatiotemporal Structure of Cell Fate Decisions in Murine Neural Crest. *Science* 364, eaas9536. doi:10.1126/science.aas9536

Sorrells, S., Toruno, C., Stewart, R. A., and Jette, C. (2013). Analysis of Apoptosis in Zebrafish Embryos by Whole-Mount Immunofluorescence to Detect Activated Caspase 3. *JoVE* 82, 51060. doi:10.3791/51060

Tang, W., Li, Y., Li, A., and Bronner, M. E. (2021). Clonal Analysis and Dynamic Imaging Identify Multipotency of Individual Gallus gallus Caudal Hindbrain Neural Crest Cells toward Cardiac and Enteric Fates. *Nat. Commun.* 12, 1894. doi:10.1038/s41467-021-22146-8

Tennyson, V. M., Gershon, M. D., Sherman, D. L., Behringer, R. R., Raz, R., Crotty, D. A., et al. (1993). Structural Abnormalities Associated with Congenital Megacolon in Transgenic Mice that Overexpress the Hoxa-4 Gene. *Dev. Dyn.* 198, 28–53. doi:10.1002/aja.1001980105

Uribe, R. A., and Bronner, M. E. (2015). Meis3 Is Required for Neural Crest Invasion of the Gut during Zebrafish Enteric Nervous System Development. *MBoC* 26, 3728–3740. doi:10.1091/mbc.e15-02-0112

van der Velden, Y. U., Wang, L., van Lohuizen, M., and Haramis, A.-P. G. (2012). The Polycomb Group Protein Ring1b Is Essential for Pectoral Fin Development. *Development* 139, 2210–2220. doi:10.1242/dev.077156

Waxman, J. S., Keegan, B. R., Roberts, R. W., Poss, K. D., and Yelon, D. (2008). Hoxb5b Acts Downstream of Retinoic Acid Signaling in the Forelimb Field to Restrict Heart Field Potential in Zebrafish. *Developmental Cel.* 15, 923–934. doi:10.1016/j.devcel.2008.09.009

Waxman, J. S., and Yelon, D. (2009). Increased Hox Activity Mimics the Teratogenic Effects of Excess Retinoic Acid Signaling. *Dev. Dyn.* 238, 1207–1213. doi:10.1002/dvdy.21951

Wolgemuth, D. J., Behringer, R. R., Mostoller, M. P., Brinster, R. L., and Palmiter, R. D. (1989). Transgenic Mice Overexpressing the Mouse Homoeobox-Containing Gene Hox-1.4 Exhibit Abnormal Gut Development. *Nature* 337, 464–467. doi:10.1038/337464a0

Zhu, K., Spaink, H. P., and Durston, A. J. (2017). Collinear Hox-Hox Interactions Are Involved in Patterning the Vertebrate Anteroposterior (A-P) axis. *PLoS One* 12, e0175287–15. doi:10.1371/journal.pone.0175287

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Cell and Developmental Biology

ORIGINAL RESEARCH
published: 26 April 2022
doi: 10.3389/fcell.2022.806545

Hoxa5 Activity Across the Lateral Somitic Frontier Regulates Development of the Mouse Sternum

Kira Mitchel[1†], *Jenna M. Bergmann*[1†], *Ava E. Brent*[1†*], *Tova M. Finkelstein*[1], *Kyra A. Schindler*[1], *Miriam A. Holzman*[1], *Lucie Jeannotte*[2,3] *and Jennifer H. Mansfield*[1*]

[1]*Department of Biology, Barnard College, Columbia University, New York, NY, United States,* [2]*Department of Molecular Biology, Medical Biochemistry and Pathology, Faculty of Medicine, Université Laval, Québec, QC, Canada,* [3]*Centre de Recherche sur le Cancer de l'Université Laval, CRCHU de Québec-Université, Laval (Oncology Axis), Québec, QC, Canada*

The skeletal system derives from multiple embryonic sources whose derivatives must develop in coordination to produce an integrated whole. In particular, interactions across the lateral somitic frontier, where derivatives of the somites and lateral plate mesoderm come into contact, are important for proper development. Many questions remain about genetic control of this coordination, and embryological information is incomplete for some structures that incorporate the frontier, including the sternum. Hox genes act in both tissues as regulators of skeletal pattern. Here, we used conditional deletion to characterize the tissue-specific contributions of *Hoxa5* to skeletal patterning. We found that most aspects of the *Hoxa5* skeletal phenotype are attributable to its activity in one or the other tissue, indicating largely additive roles. However, multiple roles are identified at the junction of the T1 ribs and the anterior portion of the sternum, or presternum. The embryology of the presternum has not been well described in mouse. We present a model for presternum development, and show that it arises from multiple, paired LPM-derived primordia. We show evidence that HOXA5 expression marks the embryonic precursor of a recently identified lateral presternum structure that is variably present in therians.

Keywords: *Hoxa5*, sternum, somites, lateral plate mesoderm, lateral somitic frontier, skeletal patterning

OPEN ACCESS

Edited by:
Marie Kmita,
Montreal Clinical Research Institute (IRCM), Canada

Reviewed by:
Moises Mallo,
Gulbenkian Institute of Science (IGC), Portugal
Raquel P. Andrade,
Universidade do Algarve, Portugal

****Correspondence:***
Ava E. Brent
abrent@barnard.edu
Jennifer H. Mansfield
jmansfie@barnard.edu

[†]*These authors contributed equally to this work*

Specialty section:
This article was submitted to Morphogenesis and Patterning, a section of the journal Frontiers in Cell and Developmental Biology

Received: *31 October 2021*
Accepted: *03 March 2022*
Published: *26 April 2022*

Citation:
Mitchel K, Bergmann JM, Brent AE, Finkelstein TM, Schindler KA, Holzman MA, Jeannotte L and Mansfield JH (2022) Hoxa5 Activity Across the Lateral Somitic Frontier Regulates Development of the Mouse Sternum. Front. Cell Dev. Biol. 10:806545. doi: 10.3389/fcell.2022.806545

INTRODUCTION

The post-cranial musculoskeletal system derives primarily from two embryonic sources: the somites and the lateral plate mesoderm (LPM). These mesodermal populations have distinct developmental programs, genetic control, and evolutionary histories, yet they must develop in coordination to permit the functional integration of their derivatives. Somites and LPM each give rise to skeletal tissue, including cartilage and bone, and to connective tissue such as tendons, ligaments, and muscle connective tissue (reviewed in (Christ et al., 2007; Prummel et al., 2020)). In contrast, all skeletal muscle is somite-derived, with the exception of some cranial muscles (reviewed in (Yahya et al., 2020)).

Musculoskeletal structures can be categorized as primaxial or abaxial depending on the source of their connective tissues (Burke & Nowicki, 2003). In primaxial structures, which include the vertebral column and proximal ribs, all musculoskeletal tissues are entirely somite-derived. Abaxial structures include the limbs, most of the limb girdles (shoulder and pelvis), the sternum, and in mice, the distal T1 rib (Durland et al., 2008). In these structures, all connective and most skeletal tissue is LPM-derived, but muscles arise from progenitors that migrate ventrally from the somites into LPM

territory. In some abaxial structures, such as the distal T1 rib, the cartilage also arises from ventrally migrating somite-derived cells. The border between the primaxial and abaxial domains was first mapped in chick and has been termed the lateral somitic frontier (LSF (Nowicki & Burke, 2000)). The LSF has also been mapped in mouse embryos using the *Prx1-Cre* transgene (Durland et al., 2008). A few structures, described as incorporating the frontier, have both primaxial and abaxial portions. In mice, this includes the scapula (whose dorso-medial border is somite-derived), the intercostal muscles, and the junction of the ribs and sternum (described further below) (Durland et al., 2008).

The primaxial/abaxial distinction has facilitated understanding of developmental phenotypes, which often differ in or are confined to derivatives of one domain or another, suggesting that these are at least partially independent developmental fields (Burke & Nowicki, 2003). It is also consistent with an instructive role for connective tissue in musculoskeletal patterning, growth and homeostasis (reviewed in (Sefton & Kardon, 2019)). Finally, the primaxial/abaxial distinction is important for interpreting patterns and potential constraint of morphological evolution due to required interaction between the two tissues (reviewed in (Shearman & Burke, 2009)).

Indeed, somites and LPM have distinct evolutionary histories (Liem et al., 2001). For example, somites arose in basal (invertebrate) chordates and likely gave rise to the axial system of muscles and the connective tissues that attached them to the body wall (laterally) and notochord (medially), permitting locomotion (reviewed in Willey, 1894). Somites also gave rise to the earliest-evolved post-cranial skeleton: the vertebral column. Paired appendages evolved later, with skeletal tissue developing instead from LPM, but with somite-derived muscle migrating to populate them (reviewed in (Tanaka & Onimaru, 2012)). In tetrapods, the elaboration and diversification of limbs and limb girdles expanded the contribution of LPM to musculoskeletal development, and would have necessitated novel interactions across the lateral somitic frontier to preserve an integrated musculoskeletal system.

One region where such integration must occur is at the junction of the sternum and the ribs. The sternum evolved in tetrapods as an extension of the pectoral girdle. It functions as an attachment site for pectoral muscles that facilitate the transfer of body weight to the forelimbs. Rib-sternum articulation evolved secondarily in some tetrapods, serving to strengthen the ribcage, and has also been adapted for respiratory function. Indeed, sternal anatomy varies across tetrapod groups and reflects diverse modes of locomotion and respiration (Liem et al., 2001; Scaal, 2021). In most living mammals, the sternum contains three segments: the anterior sternum, or presternum (sometimes called the manubrium), the mesosternum and the xiphoid process. The presternum articulates anteriorly with the clavicles (in species where clavicles are present) and posteriorly with the second pair of ribs. The first pair of ribs joins the presternum laterally and this attachment is morphologically and functionally unique among rib attachments. The mesosternum extends from the second rib to the last attached (true) rib and is segmented in most mammals, made up of ossified sternebrae alternating with cartilaginous joints at the points of rib attachment; these provide flexibility for diaphragm-based respiration. The xiphoid process is a thickened plate at the posterior end the sternum that serves as an attachment point for diaphragm muscles.

The mesosternum develops from paired structures called sternal bars, and their embryology has been well-described in classical studies of mammals and birds (for example, (Gladstone & Wakeley, 1932; Fell, 1939; Chen, 1952a, 1952b, 1953; Chevallier, 1975). The sternal bars arise in the axillary region and derive from the *Tbx5*-expressing forelimb field (Gladstone & Wakeley, 1932; Chen, 1952a; Bickley & Logan, 2014). They migrate to the midline of the ventral body wall, where they fuse to form the sternum and displace existing mesenchyme, some of which undergoes cell death (Chen, 1952a). Rib anlagen, which grow ventrally from the somites into the body wall, fuse with the sternal bars during their migration. Sternal bars subsequently provide force that "pulls" the rib primordia toward the midline (Chen, 1952b), and conditional deletion of *Tbx5* in mouse LPM results in a complete loss of the sternal bars, and secondary failure of the ribcage to close ventrally (Bickley & Logan, 2014). The segmented structure of the mammalian sternum arises secondarily because the rib primordia inhibit sternum ossification at their attachment points via an unknown signal (Chen, 1953).

The embryology of the presternum has been less studied. However, there is evidence that in mammals it arises from the fusion of multiple embryonic cartilage condensations (Gladstone & Wakeley, 1932; Rodríguez-Vázquez et al., 2013; Buchholtz et al., 2020). These include the anterior sternal bars and an additional midline condensation proposed homologous to the interclavicle, which is an unpaired midline bone present in most synapsids but that is not present as a separate skeletal structure in therian mammals. The presence of an additional paired lateral element at the position of rib 1 articulation was proposed by a comparative study of presternum anatomy from extant and fossil mammals and from medical CT scans showing variably present lateral skeletal structures in human presternae (Buchholtz et al., 2020). Lineage tracing confirmed the LPM origin of the R1 attachment site (Durland et al., 2008) and a neural crest contribution has also been reported in the anterior sternum (Matsuoka et al., 2005). However, a thorough description of presternum development in a mammalian model system has been lacking. Thus, the number and origin of its primordia, as well as their embryological tissue of origin and relationship to the lateral somitic frontier, is unknown, as are potential signals acting across the frontier to coordinate rib-sternum interactions. Hox genes are a good candidate for playing a role in this latter activity.

Hox genes globally pattern anterior-posterior fates, including in both somite and LPM derivatives, and loss-of-function studies have demonstrated patterning roles in both primaxial and abaxial structures (reviewed in (Mallo et al., 2010)). Vertebrate Hox genes are expressed in a nested, colinear pattern in somites, but their expression boundaries are less regular with respect to cluster organization in the LPM, and often differ from those in somites (Burke et al., 1995). Heterotopic transplantation of presomitic mesoderm reveals that both segmental identity and Hox expression become determined prior to somite segmentation;

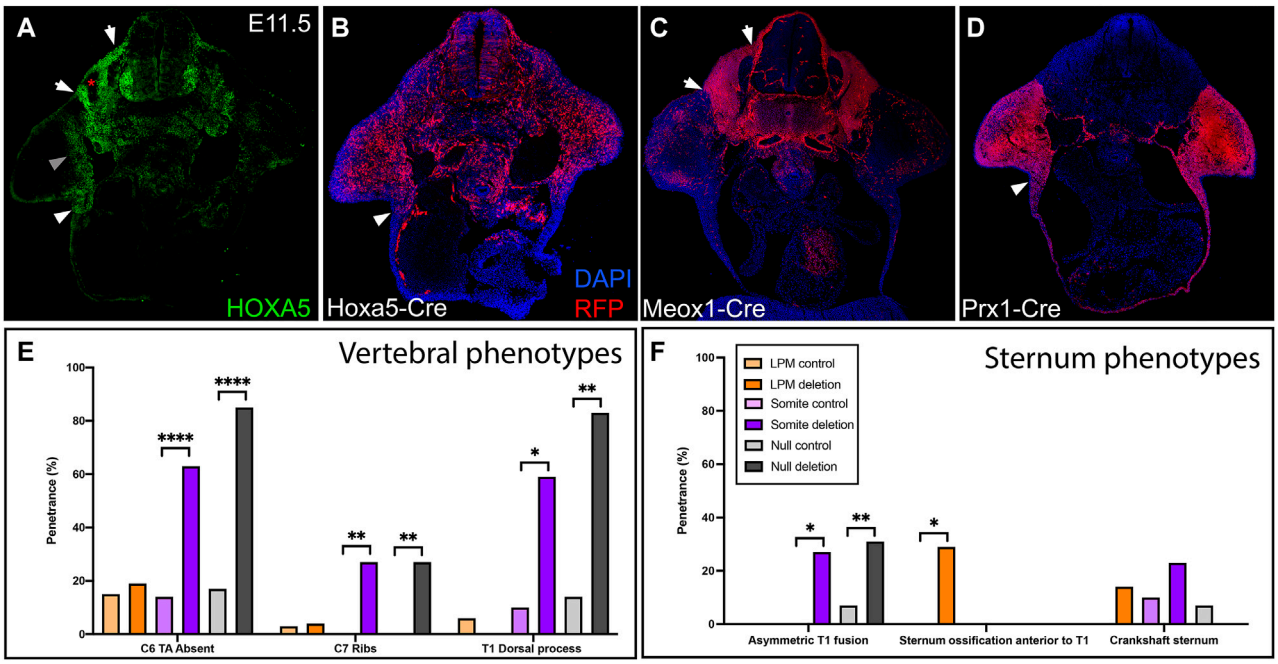


FIGURE 1 | *Hoxa5* expression in somites and LPM. **(A)** HOXA5 protein expression and **(B)** domain of cells with an expression history of *Hoxa5* at E11.5, prior to migration of axial skeletal progenitors across the lateral somatic frontier. Compare to the location of cells derived from **(C)** somites, labelled with Meox1-Cre, and **(D)** LPM, labelled with Prx1-Cre. White arrows mark the lateral (left) and medial (right) borders of the somites. White arrowheads indicate the axillary region known to contain sternal bar progenitors. Grey arrowhead indicates additional HOXA5 expression in the limb bud, and asterisk marks the myotome (see text). **(E,F)** Summary of skeletal phenotypes following conditional *Hoxa5* deletion. Frequency of *Hoxa5* associated phenotypes in the **(E)** C6-T1 vertebrae and **(F)** sternum at E18.5. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{****}p < 0.0001$, Fisher's exact test. Complete genotypes of control and experimental groups are given in **Supplementary Table S1**.

primaxial structures develop according to their original location and maintain their own Hox code after transplantation. In contrast, the segmental identity of abaxial structures is governed by Hox expression in the LPM: muscle and skeletal progenitors that migrate across the lateral somitic frontier following heterotopic transplantation adopt the morphology and Hox code associated with the surrounding LPM (Kieny et al., 1972; Murakami & Nakamura, 1991; Nowicki & Burke, 2000).

In order to better understand how developmental programs are coordinated across the lateral somitic frontier, and specifically at the point of forelimb attachment, we took two approaches. In the first, we examined the tissue-specific requirements for *Hoxa5* in somites vs. LPM for mouse skeletal development. *Hoxa5* is a good model because of its non-redundant skeletal phenotypes affecting forelimb attachments and vertebral segments spanning the cervical-thoracic transition (reviewed in (Jeannotte et al., 2016)). Our results showed that *Hoxa5* null-associated skeletal phenotypes can be reproduced by tissue-specific *Hoxa5* deletion in somites or LPM, identifying the tissues in which it is required. Interestingly, LPM-specific deletion produced a novel phenotype, suggesting that coordinated HOXA5 expression across the frontier may be necessary for some aspects of its role in presternum development. In a second approach, and to better contextualize these genetic results, we characterized mouse presternum development at a series of stages with reference to the cartilage condensations, tissues of origin, and the relative role of *Hoxa5*.

Together, our results confirm and extend previous findings about the origin of the presternum as a composite structure. We show that the mouse presternum, including what is commonly referred to as the manubrium, is composed of at least five independent, paired mesenchymal condensations, all of which are primarily LPM-derived and lack contribution from somites. Molecular subdivision of one presternal element is provided by the differential expression of HOXA5 specifically at the points of rib 1 attachment, and this same region is disrupted in *Hoxa5* mutants. Thus, we propose that HOXA5 expression molecularly marks a previously-identified lateral element of the therian presternum. Together, our results shed light on the development of a structure arising at the lateral somitic frontier, and provide a genetic dissection of *Hox* activity spanning this junction.

RESULTS

Distinct Phenotypes Result From Tissue-Specific Deletion of *Hoxa5* in Somites or LPM

HOXA5 protein is expressed in both somites and LPM, in adjacent domains that flank the lateral somitic frontier. This can be observed in a brachial somite at E11.5, just prior to the migration of rib and axial muscle progenitors, by comparison of HOXA5 expression and Cre-based lineage label (**Figures 1A,B**)

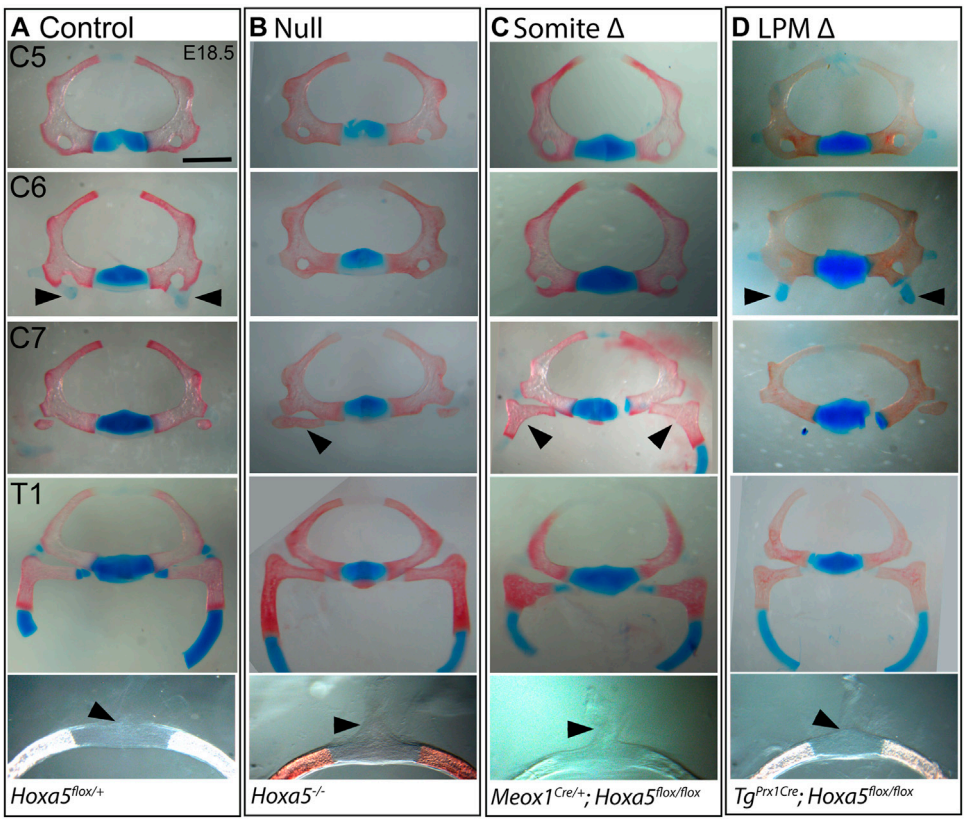


FIGURE 2 | Vertebral phenotypes associated with conditional *Hoxa5* deletion at E18.5. Alcian blue and Alizarin red staining of vertebrae from control **(A)**, null **(B)**, somite deleted **(C)**, and LPM-deleted **(D)** embryos. Arrowheads indicate tuberculum anterior (TA) on C6, ribs on C7 and the position of the dorsal process on T1 (absence of a dorsal process indicated by a tilted arrowhead). Note that the T1 ribs have been cut within the cartilaginous, distal portion for photography. Scale bar 1 mm (top four rows) or 0.47 mm (bottom row).

to position-matched sections from embryos lineage-labeled for somites and LPM (**Figures 1C,D**) using Meox1-Cre (Jukkola et al., 2005) and Prx1-Cre (Logan et al., 2002), respectively. A comparison of these latter two images illustrates the position of the lateral somitic frontier. As previously described (Holzman et al., 2018), HOXA5 is expressed broadly in somites (between the two arrows, **Figures 1A,C**), with the exception of prospective skeletal muscle (asterisk and (Holzman et al., 2018)). In LPM, expression is observed in the limb field, including mesenchyme of the proximal forelimb (grey arrowhead) and the axillary region (white arrowheads in **Figures 1A,B,D**) where the sternal progenitors are found (Bickley & Logan, 2014). In contrast, few HOXA5-expressing cells are found in the LPM-derived ventral body wall mesenchyme. Similar to HOXA5 protein, Hoxa5-Cre activity visualized with an RFP reporter reveals a similar but somewhat broader spatial domain of cells with a *Hoxa5* expression history (especially in the forelimb) reflecting the dynamic nature of *Hoxa5* expression (**Figure 1B**; Bérubé-Simard & Jeannotte, 2014; Holzman et al., 2018). Together, these expression data raise the question of which domain(s) of *Hoxa5* activity mediate skeletal patterning, and whether somite and LPM activity is functionally independent for skeletal structures that incorporate the frontier.

To test these questions, we deleted *Hoxa5* in somites or in LPM with the $Hoxa5^{flox}$ conditional allele (Tabariès et al., 2007) and tissue-specific Cre lines. Skeletal phenotypes were examined at E18.5, with a focus on the sixth cervical to first thoracic (C6-T1) segments and on the sternum: regions that span the cervical-thoracic transition, that contain a mixture of primaxial, abaxial, and transitional strucutres, and that also include the most penetrant *Hoxa5* phenotypes (Jeannotte et al., 1993). Results are summarized in **Figures 1E,F** and detailed in **Supplementary Table S1**. To account for possible effects of genetic background introduced from the Cre lines, we included data for littermate controls from each cross.

Homeotic Transformations Involve Somitic *Hoxa5* Activity

The vertebrae are somite-derived and, except for the distal T1 rib, are entirely primaxial (Durland et al., 2008). Consistent with this, all vertebral phenotypes previously associated with *Hoxa5* loss-of-function (Jeannotte et al., 1993) were recapitulated by conditional deletion of *Hoxa5* in somites with Meox1-Cre (**Figures 1E**, **2**). These included loss of the tuberculum anterior on C6, which is considered an anterior transformation of C6 to C5. Ectopic C7 ribs were also

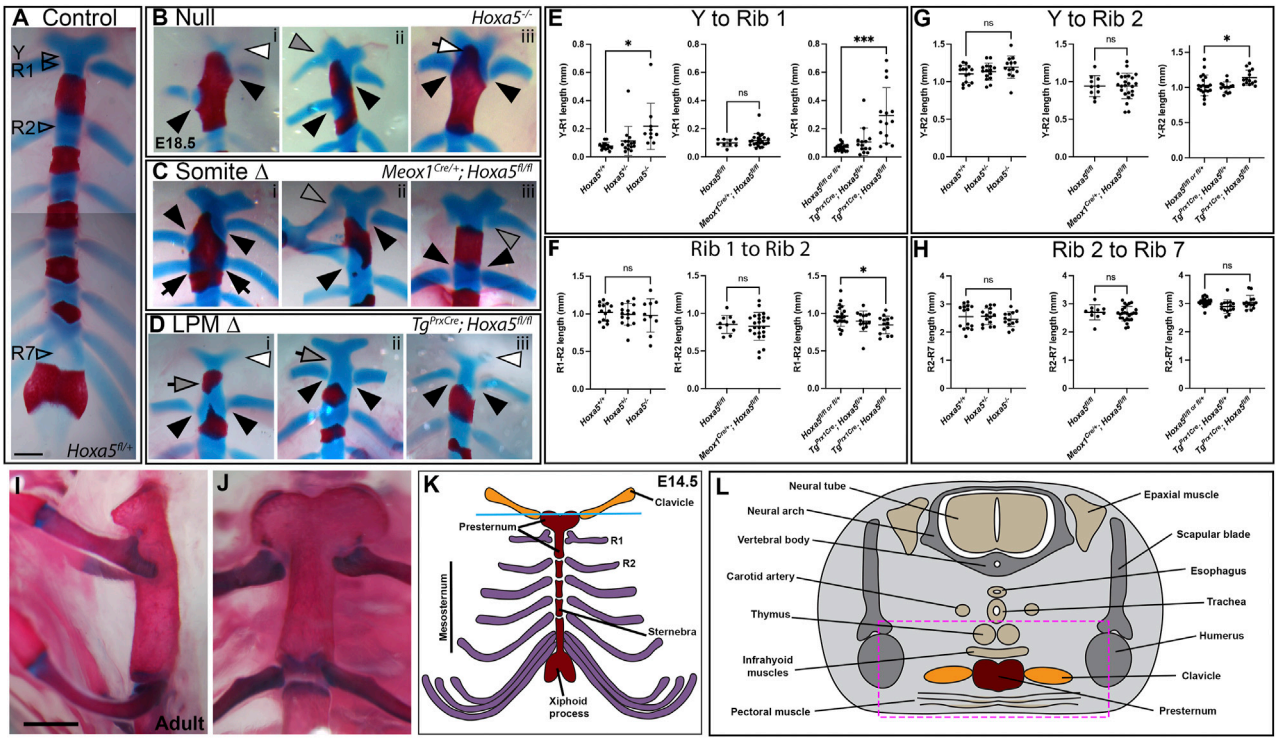


FIGURE 3 | Sternum phenotypes associated with conditional *Hoxa5* deletion at E18.5. **(A)** Control sternum at E18.5. Arrowheads indicate the positions at which measurements were taken for data shown in panels **(E–H)**. **(B–D)** Examples of anterior sternum phenotypes. In all panels, filled black arrowheads indicate T1 ribs, black arrows indicate T2 ribs, and grey arrowheads indicate C7 ribs. White arrowheads indicate Y-shaped cartilage reduction, and white arrow in **(Biii)** indicates ectopic ossification of the Y-shaped cartilage. Grey arrows in **(Di,ii)** indicate an elongated sternum region between the Y-shaped cartilage and T1 attachment position, which was sometimes associated with an ectopic ossification. Scale bar: 500 μm. **(E–H)** Comparative lengths of sternal regions between arroweads indicated in **(A)**. In each panel, null crosses are graphed on the left, somite-specific deletion in the middle and LPM-specific deletion on the right. Bars and brackets indicate the mean and standard deviation, respectively. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, n.s. not significant, Welch's *t*-test. **(I,J)** Adult sternum with T1 and T2 attachments shown in side **(I)** or front **(J)** views. **(K)** Schematic of E14.5 sternum and ribs attachments. **(L)** Schematic of a cross-sections through an E14.5 embryo at the axial level indicated by the blue line in **(K)**.

observed, interpreted as a posterior transformation of C7 to T1. As was previously reported for the null allele, C7 ribs could be unilateral or bilateral and varied in extent and in whether they were free or fused to the T1 rib or sternum (shown below). Cartilage nodules were frequent in all genotypes and were not considered ribs (see control embryo in **Figure 2A** and not shown). Finally, somite-specific *Hoxa5* deletion led to ectopic formation of dorsal processes on T1, considered a posterior transformation of T1 to T2 (**Figures 1E**, **2C**). Note that at E18.5 the dorsal process was not yet positive for Alcian blue, but mesenchymal condensations could be observed. In contrast, none of these vertebral phenotypes were observed following LPM-specific *Hoxa5* deletion with Prx1-Cre (**Figures 1F**, **2D**). These results indicate that, as expected, somite-specific *Hoxa5* activity is responsible for patterning of primaxial skeletal structures.

Rib-Sternum Attachment Phenotypes Involve Somitic *Hoxa5* Activity

In addition to homeotic transformations, *Hoxa5* null mutants present rib fusions, bifurcations and asymmetric sternal attachment involving the first (T1) ribs and in some cases ectopic C7 ribs (Jeannotte et al., 1993). These phenotypes are consistent with a defect in rib guidance during segmental outgrowth across the lateral somitic frontier, and/or with altered recognition between T1 and a defined location on the presternum. Interestingly, these rib defects were all reproduced by conditional deletion of *Hoxa5* in somites but not by deletion of *Hoxa5* in LPM.

In wild-type embryos, the T1 ribs fuse symmetrically and at a consistent position on the presternum, at the base of the Y-shaped cartilage (**Figure 3A**). Rib attachment defects are shown in **Figure 3** and **Supplementary Figure S1**. Asymmetric attachment of T1 ribs to the sternum was frequent in both null embryos and following somite-specific deletion (**Figures 3B,C, i,ii** black arrowheads indicate T1 ribs in all panels). A T1 bifurcation following conditional deletion of *Hoxa5* in somites is shown in **Supplementary Figure S1Civ**. When they formed, C7 ribs were sometimes free (**Supplementary Figure S1**, compare **A** to **Bi,Ci**), but others fused with the T1 rib (**Supplementary Figure S1Cii**), or were complete ribs, fused to the sternum in the position normally occupied by T1 (**Supplementary Figure S1Ciii**, grey arrowheads indicate C7 ribs in all panels). All of these phenotypes were previously described for *Hoxa5* null embryos (Jeannotte et al., 1993). In contrast, no rib asymmetries, fusions or bifurcations were observed following *Hoxa5* conditional deletion in LPM.

We hypothesized that asymmetric T1 fusion could be caused by physical displacement of one rib due to presence an extra cervical rib on the same side. This was not the case, however, because there was no correlation between aberrant T1 rib fusion and an ipsilateral C7 rib. Of embryos with asymmetric T1 rib fusions, 2/4 null embryos and 2/7 following somite-specific *Hoxa5* deletion lacked an ectopic C7 rib on the same side (although we cannot rule out presence of ectopic intercostal muscle or connective tissue). Of those that did have ipsilateral C7 ribs, some were free, others fused to T1, and others fused to the sternum. Thus, the symmetric outgrowth and fusion of T1 ribs is apparently an independent phenotype from the presence of C7 ribs, and consistent with hypothesized loss of positional information either during rib migration and/or during presternum fusion. In one embryo with somite-specific deletion, in which a C7 rib was fused to the sternum, both T1 ribs shifted their attachment symmetrically to the normal T2 position (**Figure 3Ciii**; **Supplementary Figure S1Ciii**). This phenotype has been previously described in null embryos (Jeannotte et al., 1993) but was not observed in our null samples.

The posterior fusion of T1 ribs, particularly in the latter example, could be interpreted as a homeotic transformation of T1 to T2. This did not appear to be the case, however, based on sternum morphology. In wild-type mice (and other mammals), the mesosternum segments at the positions rib attachment (T2–T7), but not at the T1 attachment (**Figures 3I,J**). At E18.5, the future mesosternal joints are visible as thin discontinuities in Alcian blue staining, which are never observed at T1. Following somite-specific *Hoxa5* deletion, ectopic mesosternum-like segmentation was not observed at the position of T1 fusion, even when T1 ribs fused at the T2 position (not shown). This indicates that the joint maintained at least this aspect of T1 identity and was not transformed to a mesosternal-type joint. The T1 distal ribs are also morphologically distinct from all other ribs in adults because they make a double-contact with the sternum (**Figures 3I,J**, and see below). However, this morphology has not developed at E18.5 and thus could not be scored with Alcian blue staining.

Finally, it was previously reported that the *Hoxa5* skeletal phenotype involves asymmetric fusion of all or most ribs, resulting in bifurcated or fused sternebrae, also called a crankshaft sternum (Jeannotte et al., 1993). This phenotype has been hypothesized to result from asymmetric fusion of the ribs to the sternal bars, or from asymmetric fusion of the sternal bars. This in turn subsequently disrupts the regular alternation of ossified sternebrae and cartilaginous joints because the ribs are known to inhibit sternum ossification at points of contact (Chen, 1953), and likely does not reflect a patterning change in the sternum itself. We observed the crankshaft sternum phenotype at low frequency in both conditional knockout crosses, but occurrence was statistically indistinguishable from controls (**Figure 1F**; **Supplementary Figure S1Cv**).

A Novel Presternum Phenotype Involves *Hoxa5* Activity in LPM

A novel phenotype not described in *Hoxa5* null embryos was produced by LPM-specific deletion of *Hoxa5* with Prx1-Cre: an additional ossification formed within the presternum, anterior to the T1 attachment (**Figures 1F**, **3Di**; **Supplementary Figure S1D**, grey arrows). Further examination revealed that some embryos showed presternum elongation even when no extra ossification was present (**Figure 3Dii**, grey arrow). We therefore measured absolute and relative lengths of different sternal regions in all samples (measurement positions marked in **Figure 3A**). This confirmed that the distance from the base of the Y-shaped cartilage to the center of the T1 attachment was significantly increased following LPM-specific deletion (**Figure 3E**). Conversely, the distance between the T1 and T2 attachments, which is typically greater than that between mesosternum rib pairs, was significantly reduced (**Figure 3F**). However, the overall distance between Y and T2 was significantly increased (**Figure 3G**) confirming an overall elongation of the presternum. In contrast, the length of the mesosternum, measured between the T2 and T7 attachments, was not affected (**Figure 3H**). Similar to the case above, the extra sternebra above the T1 attachment might indicate a posterior shift in identity. However, as in the examples above, even when an extra ossification was present anterior to the T1 attachment, the sternum retained a T1-like morphology because it did not segment to form a joint (not shown).

Sternum lengths were also measured following somite-specific or complete *Hoxa5* deletion (here, specimens with asymmetric T1 fusion were excluded from measurements involving T1 position but included in Y-T2 and T2–T7 measurements). No significant differences in the lengths of sternal regions were observed following somite-specific deletion of *Hoxa5* (**Figures 3E–H**). However, the Y to T1 length was significantly increased in the null embryos, even excluding those with T1 asymmetric fusion. However, the elongation was less severe than following LPM-specific deletion and never involved an ectopic ossification anterior to T1 (**Figure 3E**).

Together, these results show that *Hoxa5* activity specific to the LPM is responsible for the presternum phenotypes. The observation that the elongation of the presternum is more severe following LPM-specific deletion of *Hoxa5* compared to complete loss-of-function implies that *Hoxa5* activities in somites and LPM are not completely independent. Rather, this could indicate that they act coordinately to pattern the presternum; if so, a mismatch in positional information following conditional deletion could be expected to produce such a novel phenotype relative to complete deletion. This possibility is discussed further below.

Finally, the Y-shaped cartilage was often reduced following either LPM-specific deletion or in null embryos (**Figures 3Bi,Di,iii**, white arrowheads) but not following somite-specific deletion.

Sternum Ossification Phenotypes are Most Prevalent in *Hoxa5* Null Embryos

In wildtype embryos, rib primordia inhibit ossification of the sternum at contact points (Chen 1953). In E18.5 embryos, the sternum anterior to T1 is completely cartilaginous (**Figure 3A**) although in adults this region does ossify (**Figures 3I,J**). In *Hoxa5*

null embryos, early sternal ossification occurred frequently at points of T1 rib contact (12/26 T1 ribs; for example, **Figures 3Bi,iii**). Ectopic ossification often extended anteriorly into the Y-shaped cartilage (7/13 null embryos; **Figure 3Biii** and **Supplementary Figure S1Bii**). This phenotype was observed in one embryo following somite-specific deletion (**Figure 3Ci**) and involved one T1 and both T2 ribs, but never following LPM-specific deletion. Together, this implicates *Hoxa5* in negatively regulating ossification of the presternum, but does not clearly identify the location of its activity in this role.

Embryonic Development of the Mouse Presternum and the Role of *Hoxa5*

The results above showed that *Hoxa5* plays distinct and largely independent roles in the somites and LPM. However, a novel phenotype following LPM conditional deletion indicates that tissue-specific roles may not be purely additive in the presternum. Overall, the function of *Hoxa5* in presternum development could be better understood if described in reference to presternum embryology, which is incompletely characterized.

Presternum Development is Disrupted by *Hoxa5* Loss of Function

In an adult skeletal specimen, the unique nature of the anterior sternum and T1 articulation is apparent, including the absence of sternal segmentation at the T1 joint and the morphology of the distal T1 ribs, which form a double attachmentto the sternum, unlike the mesosternal ribs. (**Figures 3I,J**, and diagrammed in **Figure 3K**).

A recent description in human embryos suggests that multiple cartilage condensations contribute to the presternum (Rodríguez-Vázquez et al., 2013). This study showed that interclavicular (IC) mesenchyme arises between and continuous with the clavicle condensations and was proposed to have neural crest origin, similar to the endochondral portion of the clavicle (Matsuoka et al., 2005). IC mesenchyme later expands caudally and bilaterally toward the first ribs, ultimately forming a cartilaginous continuity with the first ribs. This bilateral domain extending from the posterior end of the IC mesenchyme to the site of T1 attachment was referred to as the intercostoclavicular mesenchyme (ICC). We adopt the IC and ICC terminology below to indicate the anterior-posterior organization of the presternum. The complex composed of the first appears to join posteriorly with the sternal bars, making up the region of the presternum between the T1 and T2 ribs. Throughout development, the mesenchymal condensations that contribute to the presternum were found to have distinct associations with developing muscles, further supporting their different identities. Distinct muscles from the infrahyoid group were found to be adjacent to the IC or ICC mesenchyme from the earliest stages observed, while the pectoralis was associated with the sternal bars.

To determine whether the mouse presternum develops in a similar manner, we first examined embryos at E14.5, when sternal bar closure is mostly complete. Using Sox9 expression to mark cartilage condensations, we analyzed serial cross-sections, starting at the point of contact between the clavicles and sternum (blue bar in **Figure 3K** and diagrammed in **Figure 3L**), and continuing through attachment of the T3 rib. Results are shown in **Figure 4**. We were able to identify the presternum regions corresponding to the IC and ICC mesenchyme previously described in human embryos (Rodríguez-Vázquez et al., 2013). At the axial level of the IC (**Figures 4A,B**), Sox9 marks clavicle as well as sternal structures. At this anterior-most point, the presternum appears to contain at least three components: paired dorsal elements that contact the clavicles (blue arrow marks the right side of the pair); separate paired elements that lie ventral to the clavicle (pink arrow marks the right side), and a ventral-midline element where pectoral muscle attachment occurs (white arrow). All three elements continue posterior to the clavicles, in the presternum ICC mesenchyme (**Figures 4C–F**, arrows as described above). At the axial level of the ICC, the ventral paired elements are larger and bar-shaped, and the level of SOX9 expression within them is increased compared the level of SOX9 observed in the dorsal IC and ICC mesenchyme (compare dorsal and ventral expression in **Figures 4A–F**). Moving posteriorly through the ICC mesenchyme, these ventral bar-shaped elements change shape, with the lateral ends curving dorsally and the medial portion forming a bulge (**Figure 4F**) that is more similar in shape to the mesosternum. In contrast, the paired dorsal ICC elements remain rounded and composed of a less dense mesenchyme with lower SOX9 expression compared to the ventral element. In **Figure 4G**, rib 1 is seen making first contact with the ventral ICC element (red arrow in **Figure 4G′**). As observed in the adult skeleton (**Figures 3I,J**), rib 1 makes a double contact. The first is with this ventral element, while the second, more posterior contact occurs in the position where the sternum shifts to a more rounded morphology and no longer contains laterally-extended bars (**Figures 4I,J**; blue arrow in 4J indicates where the second attachment begins to form. See also **Figure 3K**). The rounded sternum morphology continues posterior to T1 (**Figures 4J–M**), including the first intercostal region where the sternum is rounded dorsally and is elongated in the dorso-ventral plane relative to more anterior regions (**Figure 4K**). This elongated region evidently corresponds to paired, fusing sternal bars at points of rib 2 (**Figure 4L**) and rib 3 (**Figure 4M**) attachment. Continuity between ventral ICC and the posterior mesosternum suggests that ventral ICC might represent the anterior ends of the sternal bars.

Transplantation and Prx1-Cre lineage-labeling has shown that the sternal bars are LPM-derived (Bickley and Logan, 2014; Chevallier, 1975; Durland et al., 2008; Fell, 1939) but the origins the presternum components described above are not known. Using the same Prx1-Cre transgene, we found that all of these presternum elements are primarily, if not entirely, of LPM origin (**Supplementary Figure S2**). We observed Prx1-Cre-RFP in the clavicles (**Supplementary Figures S2A",B"**), in the dorsal and ventral elements of the presternum (**Supplementary Figures S2C"–J"**), and in the mesosternum (**Supplementary Figures S2K",L"**; see SOX9 staining in **Supplementary Figure**

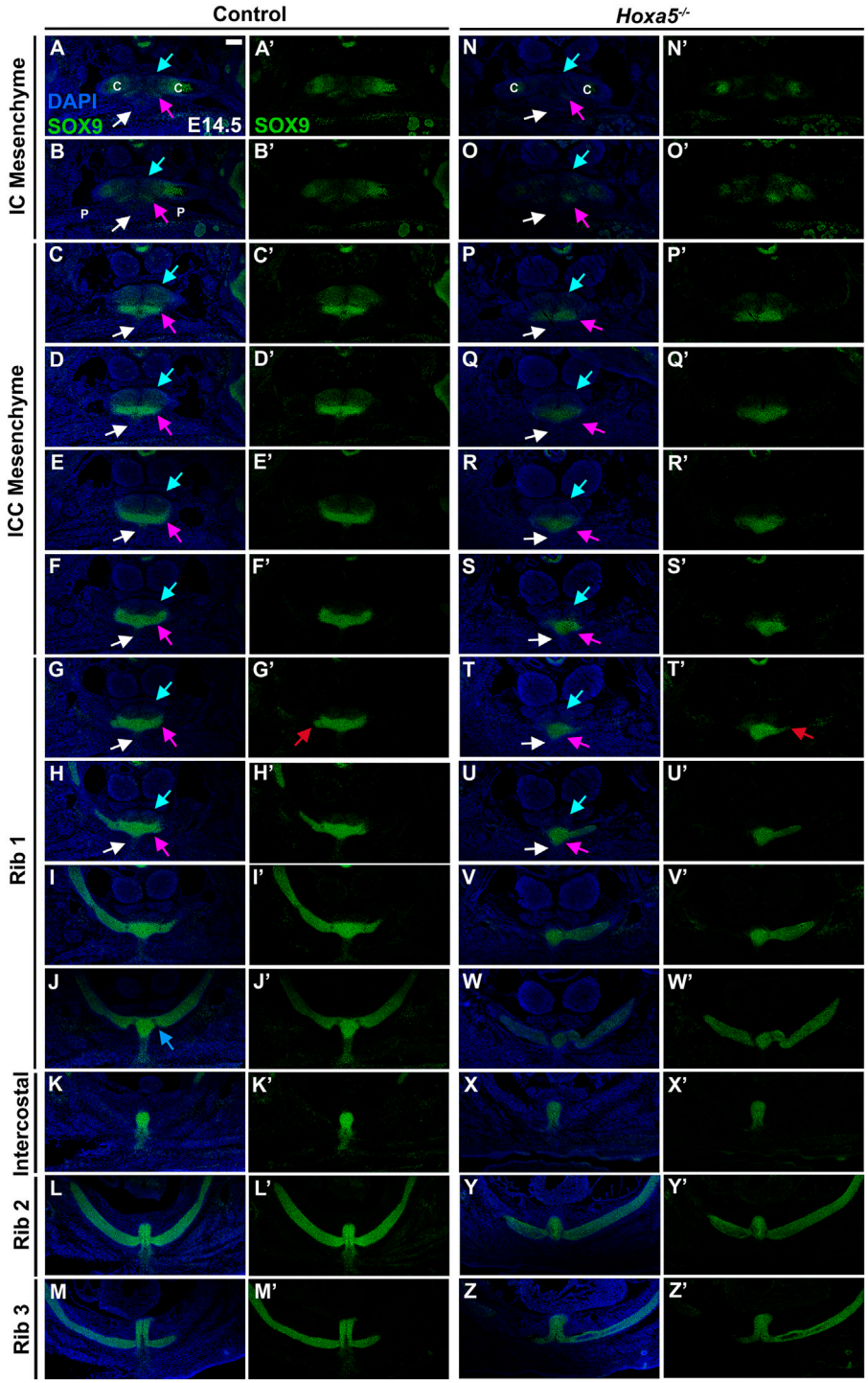


FIGURE 4 | Development of the presternum is disrupted in *Hoxa5* null mutants. **(A–M′)** SOX9 expression, shown with **(A–M)** and without **(A′–M′)** DAPI labelling, in serial sections from the anterior presternum through T1–T3 attachments of an E14.5 WT embryo. **(N,N′–Z,Z′)** SOX9 expression in serial sections through the presternum and T1–T3 attachment of a Hoxa5$^{-/-}$ null littermate. **(A–H)** and **(N–U)** Blue arrows mark dorsal IC/ICC; pink arrows indicate ventral ICC; white arrows point to the site of pectoral muscle attachment. Red arrows in **G′** and **T′** indicate first contact of T1 with the sternum. Blue arrow in **J** indicates the beginning of the second contact of T1 with the sternum. Scale bar: 200 μM. C, clavicle; IC, interclavicular; ICC, intercostoclavicular; P, pectoral muscle.

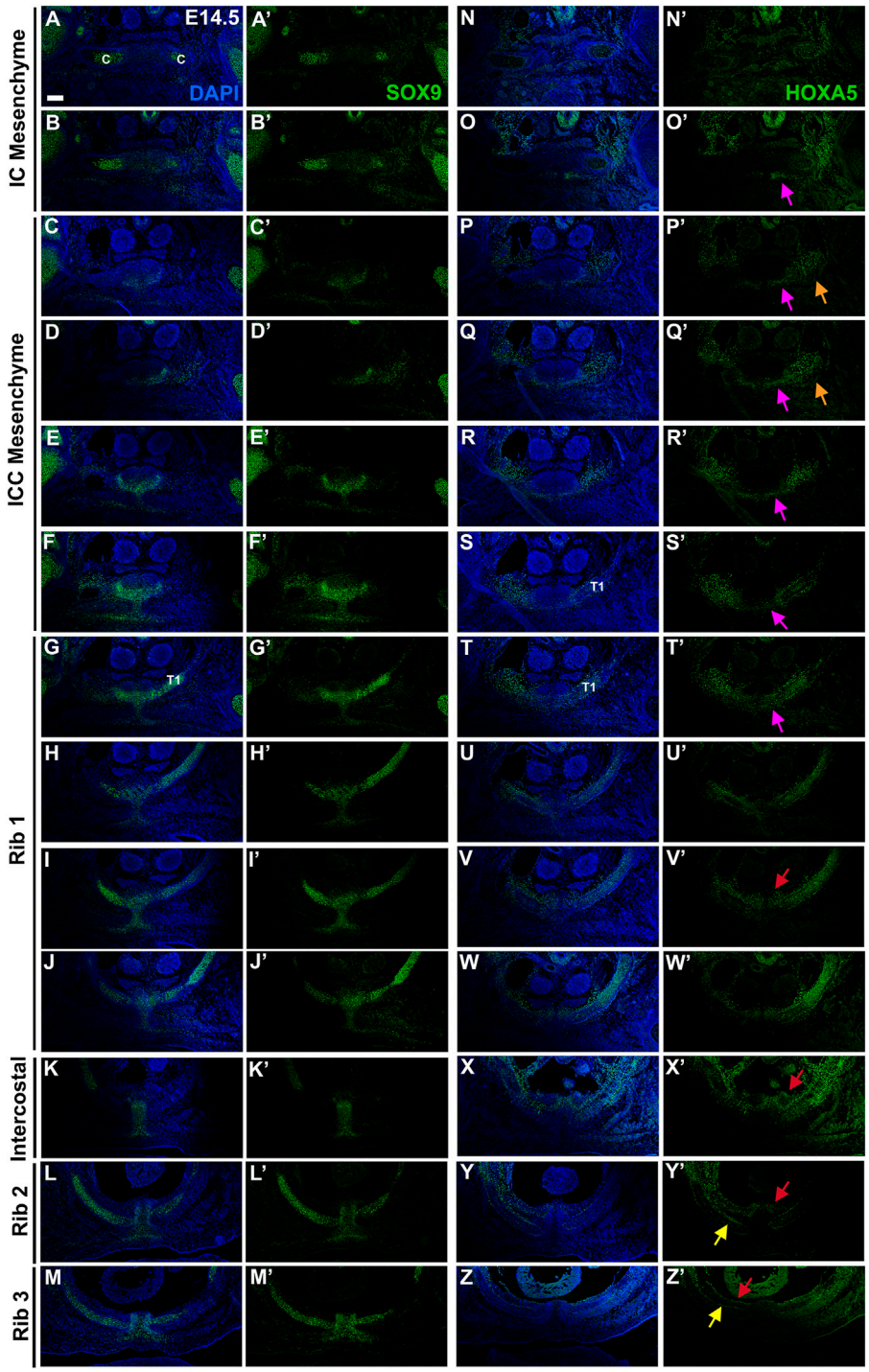


FIGURE 5 | HOXA5 is specifically expressed in ventral elements of the presternum. SOX9 **(A–M′)** and HOXA5 **(N–Z′)** expression in alternate sections of a WT E14.5, from the anterior presternum through T1–T3 attachments, shown with **(A–M, N–Z)** and without **(A′–M′, N′–Z′)** DAPI. C, clavicle. Pink arrow in **(O′–T′)** identifies expression of HOXA5 in the paired ventral elements of the ICC mesenchyme. Orange arrows in **P′** and **Q′** mark HOXA5 expression lateral to, but continuous with, the presternum. Red arrow in **V′** indicates HOXA5 expression within the sternum as it shifts to a more rounded morphology. Red arrows in **X′** and **Y′** point to a shift in Hoxa5 expression more dorsally, likely marking the connective tissue of an infrahyoid muscles. HOXA5 can also be seen in the perichondrium of ribs 2 and 3 (yellow arrows in **Y′** and **Z′**), and in the body wall mesothelium (red arrow in **Z′**). Scale bar: 200 μM. C, clavicle; IC, interclavicular; ICC, intercostoclavicular; T1, thoracic rib 1.

S2 for identification of each cartilage condensation). These observations do not exclude contribution to the IC/ICC mesenchyme and clavicles from other tissue sources, and neural crest was previously reported to contribute to the clavicle and manubrium (Matsuoka et al., 2005). However, we were able to rule out any somitic contribution to the presternum by performing a similar analysis with Meox1-Cre (**Figures 6**,**7** and data not shown).

Having identified presternum and T1 skeletal phenotypes in *Hoxa5* mutants, we next examined *Hoxa5* null embryos at E14.5 (**Figures 4N–Z**, $n = 4$), with references to the condensations described above. While development of the clavicles and IC mesenchyme was normal (**Figures 4N,O**, blue arrows), the ventral paired elements of the ICC mesenchyme were noticeably smaller than those of littermate controls (compare **Figures 4E,F** to **Figures 4R,S**, pink arrows). In contrast, the medial portion of the ventral ICC appears intact. The T1 rib contact is thus perturbed: instead of attaching to a bar of SOX9-positive ventral ICC mesenchyme, which is absent in *Hoxa5* nulls (**Figure 4G**), rib 1 appears to make contact with the rounded, medial ICC. In this particular *Hoxa5* null specimen, the T1 ribs attach symmetrically and at approximately their normal position on the presternum, comparable to the skeletal examples shown in **Figures 3Biii**. Moving further posteriorly, the medial element where rib 1 attaches in control embryos becomes more mesosternum-like in its morphology (**Figures 4I,J**). Here, the *Hoxa5* null sternum remains misshapen (**Figures 4V,W**). By contrast, the morphology of the sternum at intercostal (**Figure 4X**), rib 2 (**Figure 4Y**), and rib 3 (**Figure 4Z**) levels is similar to those of the control (**Figures 4K–M**). To determine whether changes in either cell death or cell proliferation were responsible for the reduction of the lateral portions of the ventral ICC in *Hoxa5* mutants, we examined expression of cleaved Caspase 3 and PCNA in E14.5 embryos, but observed no noticeable differences in either (**Supplementary Figure S5**; $n = 4$).

Together, analysis at E14.5 reveals that *Hoxa5* null embryos have altered sternum morphology from an early stage of development, even in specimens with a relatively normal T1 attachment position. Specifically, the lateral bars of the ventral ICC element are reduced, and their contact with rib 1 is perturbed. In contrast, the IC, dorsal ICC, and all elements posterior to the T1 rib appear normal.

HOXA5 is Expressed in Ventral ICC of the Presternum at and Anterior to the Site of Rib 1 Attachment

Having determined how *Hoxa5* loss of function affects presternum development, we next inquired whether these phenotypes correspond to a domain of HOXA5 expression in LPM-derived sternal elements. HOXA5 expression was examined in E14.5 embryos and compared to SOX9 in alternate sections from the same embryo (**Figure 5**) or by double labelling (**Supplementary Figure S3**). Sternal HOXA5 expression is first detected faintly in the paired ventral condensations at the axial level of the IC mesenchyme (pink arrow, **Figure 5O′** and **Supplementary Figures S3B,H**), and then intensifies within the ICC mesenchyme of the ventral sternum (**Figures 5P′–S′**; **Supplementary Figures S3C,D,I,J**). We confirmed that this HOXA5 expression domain was derived from LPM by comparing expression of HOXA5 to Prx1-Cre-RFP (**Supplementary Figure S2**). Interestingly, while SOX9 marks the entire ventral ICC, HOXA5 labels the lateral edges but is excluded from the midline-most domain (compare **Figures 5E′,R′**, red arrows in **Supplementary Figures S3C′,D′**). This pattern of midline exclusion of HOXA5 continues, even as the ventral sternum shifts to a more rounded, mesosternum-like shape at the anterior point of T1 attachment (**Figure 5T′**). The expression domain of HOXA5 encompasses the same lateral portions of the ventral ICC that appear reduced in *Hoxa5* mutant embryos (**Figures 4R,S**). At all anterior-posterior levels of the presternum, HOXA5 is excluded from the SOX9-expressing dorsal IC and ICC elements (blue arrows in **Figure 4** and **Supplementary Figure S3**).

In addition to the ventral ICC, HOXA5 is expressed in a domain lateral to but continuous with the developing sternum (**Figures 5P′,Q′** and **Supplementary Figures S3C,D** orange arrows) throughout the ICC region between the clavicle and T1 rib. These HOXA5-expressing cells are also LPM-derived (**Supplementary Figure S2F‴**, red arrow), and may either reflect cells that are migrating toward incorporation into the presternum (discussed below), or cells marking LPM-derived connective tissue anterior to rib 1. Interestingly, it is this component of the presternum that is elongated in *Hoxa5* mutants, and forms an ectopic ossification following LPM-specific deletion (**Figures 3Di,ii,E**). Initiation of rib 1 attachment is seen in **Figure 5T**. Importantly, both rib 1 and the ventral presternum express HOXA5 (**Figure 5T′**), but while expression looks continuous, the sternal HOXA5 domain is LPM-derived, and the rib 1 domain arises from the somites (compare **Supplementary Figures S2H″,H‴**).

Expression of HOXA5 in the ventral-lateral ICC extends throughout the ICC and T1 rib contact. Posterior to this, a dorsal LPM expression domain is also observed surrounding infrahyoid muscles at the T1, intercostal and T2 levels (**Figures 5V′–Z′**, red arrows and **Supplementary Figure S2L′**, yellow arrow). These muscles were previously described to associate with ICC mesenchyme (Rodríguez-Vázquez et al., 2013). HOXA5 can also be detected in the perichondrium of ribs 2 and 3 (**Figures 5Y′,Z′**, yellow arrows), and in a thin layer of expression along the body wall mesothelium at the level of rib 3 (**Figure 5Z′**, red arrow). Importantly, HOXA5 is not expressed in any part of the sternum posterior to the T1 site at E14.5.

We next characterized earlier stages of presternum development, including the elements identified above, and tracked both HOXA5 and SOX9 expression in this region. At E13.5 (**Figure 6** and **Supplementary Figure S4**) all of the presternal condensations observed at E14.5 can be distinguished. In the ventral ICC element, SOX9 expression reveals paired bar-like condensations similar to E14.5 except that fusion is apparently not yet complete (red arrows, **Figure 6D′** and **Supplementary Figure S4E**), and thus the medial bulge of the ventral ICC is not present. However,

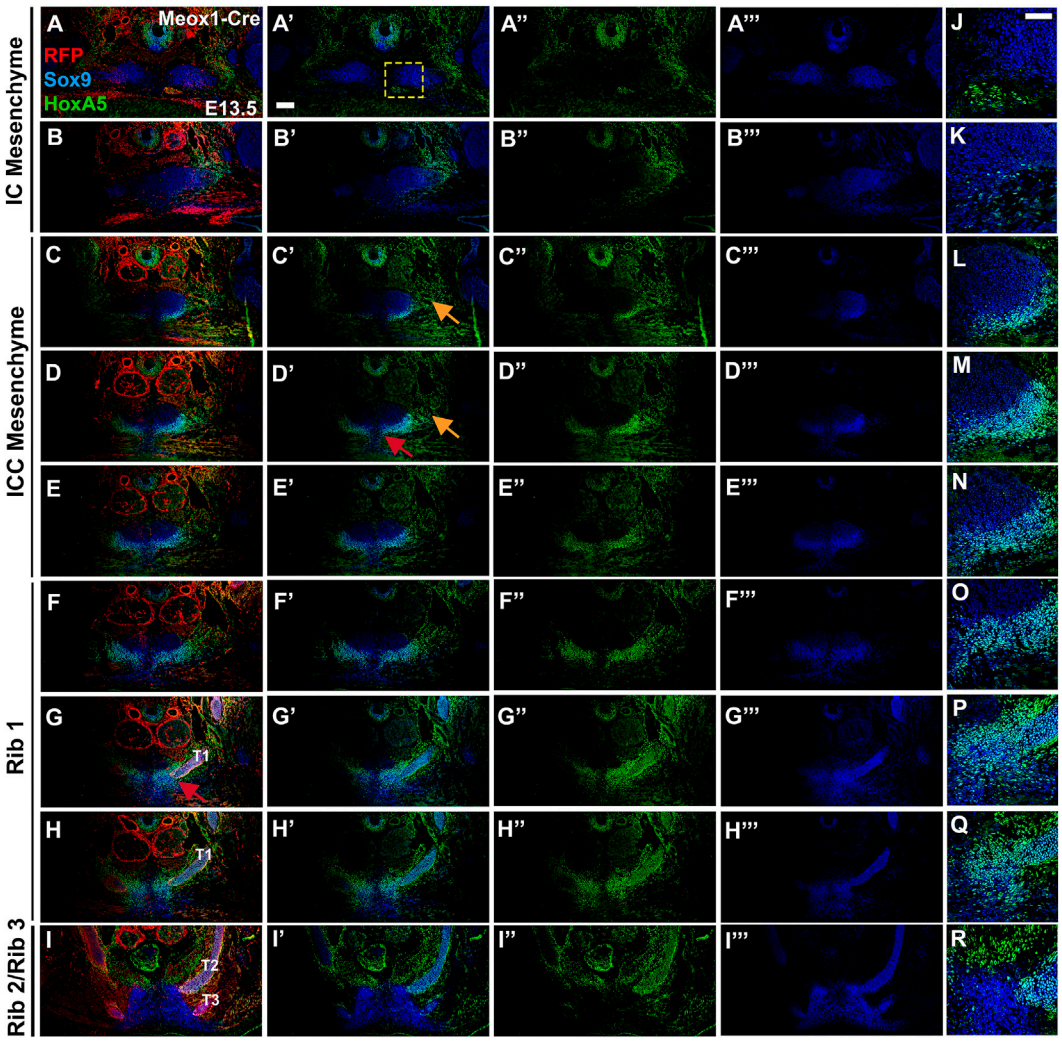


FIGURE 6 | At E13.5, HOXA5 and SOX9 co-expression identifies a component of the presternum that is continuous with the sternal bars, and is the site of rib 1 attachment. **(A–I‴)** E13.5 Meox1-Cre-RFP embryo co-labelled for HOXA5 (green **(A′,A″–I′)**, SOX9 (blue, **(A′,A‴–I′,I‴)**), and RFP (red **(A–I)**). As observed at E14.5, HOXA5, and SOX9 are co-expressed (cyan cells, **(A,A′–I,I′)** in the ventral ICC element of the presternum—with HOXA5 expression extending more broadly ventral to SOX9 **(C′–E′**; **L–N)**. **(J–R)** At higher magnification (corresponding to yellow square in **(A′)**), co-expression is more clearly visualized. SOX9, but not HOXA5, is detected at lower levels in the dorsal IC and ICC **(A′–E′)**, while both are co-expressed in ribs 1 and 2 **(F′–I′)** and spanning the site of rib 1-to-sternum fusion **(F′–H′**; **O–Q)**. HOXA5 is additionally found in the perichondrium of ribs 1 and 2 **(G′–I′)**, as well as in the connective tissue of the infrahyoid muscles, dorsal to the sternum **(I″)**. Red arrow in **(D′)** indicates exclusion of HOXA5 from the most medial region of the ventral ICC mesenchyme. Orange arrows in **(C′,D′)** point to portion of HOXA5 lateral domain that does not co-express with SOX9. Red arrow in **(G)** indicates site of rib 1 attachment. Scale bars: 200 μM T1, thoracic rib 1; T2, thoracic rib 2; T3, thoracic rib 3.

contact between rib 1 and with the ventral ICC bars has already occurred (**Figure 6G‴**; **Supplementary Figure S4F**). Interestingly, sternal bar fusion is also incomplete at this stage, thus the transition during rib 1 attachment from the ventral element of the presternum to a more mesosternum-like, sternal bar structure is quite distinct: in **Figure 6H‴** and **Supplementary Figure S4H**, the sternum appears as unfused sternal bars, with a loose collection of SOX9-expressing cells extending between them. And more posteriorly, ribs 2 and 3 attach to unfused sternal bars also connected by a stream of SOX9-positive cells (**Figure 6I‴**; **Supplementary Figures S4I,J**). As at E14.5, all presternum elements were composed largely if not entirely from Prx1-Cre labelled, and thus LPM derived cells (**Supplementary Figures S4K–T**), and none of these presternum elements showed contribution from Meox1-Cre labelled somites (**Figures 6A–I**).

A comparison of SOX9 and HOXA5 coexpression (**Figures 6A′–I′,J–R**) or in alternate sections at E13.5 (**Supplementary Figures S4K–T**) also revealed a pattern similar to that observed a day later at E14.5, with HOXA5 expression largely absent from IC mesenchyme but observed in the ventral ICC mesenchyme of the presternum, where it is localized to a lateral domain that does not express SOX9 (orange arrows, **Figures 6C′,D′** and **Supplementary Figures S4M‴,N‴**) and is LPM-derived (orange arrow, **Supplementary Figures S4N′, S6B**). To determine whether this lateral domain might

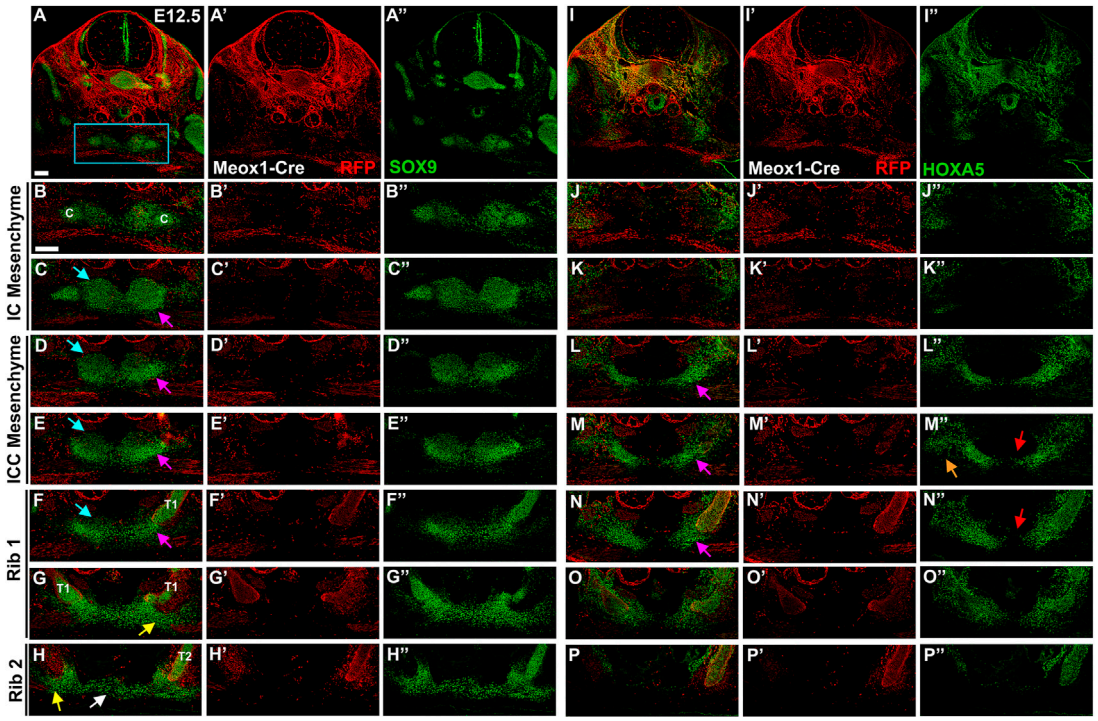


FIGURE 7 | At E12.5, HOXA5 occupies a lateral domain as well as marking the site of rib 1 attachment to the sternum. E12.5 Meox1-Cre-RFP embryo in which alternate sections were labelled for SOX9 **(A–H″)** or HOXA5 **(I–P″)**. Whole cross-sections are shown in **(A)** and **(I)**, with remaining images from the region indicated by blue rectangle in **(A)**. Images are shown with RFP and either SOX9 **(A–H)** or HOXA5 **(I–P)**; RFP alone (**A′–H′**; **I′–P′**), or SOX9 **(A″–H″)** or HOXA5 **(I″–P″)** alone. Pink arrows in **(C–F)** and **(L–N)** indicate ventral ICC. Blue arrows in **(C–F)** point to dorsal IC/ICC. Orange arrow in **(M″)** indicates lateral HOXA5 expression domain. Red arrows in **(M″,N″)** point to exclusion of HOXA5 from the medial-most region of the ventral ICC mesenchyme. Yellow arrow in **(G,H)** marks triangle of SOX9 expression contacting the rib. White arrow in **(H)** marks stream of cells crossing the midline. Scale bar: 100 μM. C, clavicle; T1, thoracic rib 1; T2, thoracic rib 2.

include cells that contribute to the connective tissue components of the developing sternum, we co-labeled E14.5 embryo sections for HOXA5 and either Tenascin, a marker of the extracellular matrix associated with tendons and ligaments (**Supplementary Figure S6A′**) or EBF3, a transcription factor expressed in connective tissue of the developing sternum (Kuriki et al., 2020) (**Supplementary Figure S6B′**). HOXA5 and Tenascin were not expressed in the HOXA5-expressing cells lateral to the IC/ICC mesenchyme of the presternum (**Supplementary Figure S6A**, orange arrow, and **Supplementary Figure S6A′**), suggesting these cells do not correspond to tendon or ligament progenitors. By contrast, HOXA5 and EBF3 are coexpressed in these lateral cells (**Supplementary Figure S6B**, orange arrow). While this may indicate future contribution of these cells to the sternum based on EBF3 loss-of-function phenotypes (Kuriki et al., 2020), we also observe that EBF3 is expressed broadly in muscle connective tissue (within LPM in **Supplementary Figure S6B**, and data not shown). HOXA5 has also been previously shown to be expressed in muscle connective tissue (Holzman et al., 2018), thus raising the possibility that this lateral domain acts in establishing the muscle connective tissue of intercostal muscles.

The anterior of the two Rib 1 contacts with the presternum occurs within this region of the HOXA5-expressing ventral-lateral ICC (red arrows, **Figure 6G** and **Supplementary Figure S4P‴**). Further, like at E14.5, Rib 1 contact with the presternum continues posteriorly into the region where sternal morphology shifts to unfused sternal bars (**Figure 6H‴**; **Supplementary Figure S4Q‴**). At this axial level, HOXA5 expression is also present dorsally and thus surrounds the point of rib 1 contact (**Figure 6H”**; **Supplementary Figure S4R‴**), similar to E14.5. However, it does not mark the remainder of the sternal bar. Continuity of SOX9 expression again suggests that the ventral ICC likely represents the anterior-most portion of the sternal bars.

In summary, E13.5 analysis showed the same elements and HOXA5 expression pattern as at E14.5 and additionally showed that the T1 ribs contact the presternum at a transitional point, encompassing both ventral ICC and sternal bars and the junction between them. This region is HOXA5-positive and LPM-derived. Further, at this stage it is clear that pre-sternal elements arise as paired progenitors that flank the midline.

HOXA5-Expressing Ventral Presternum may Arise From a Lateral Population of Cells

Observation that the presternum appears to be composed of multiple paired elements at E13.5 and E14.5, some of which express HOXA5 and some of which do not, led us to further examine sternum development at earlier stages, including E12.5

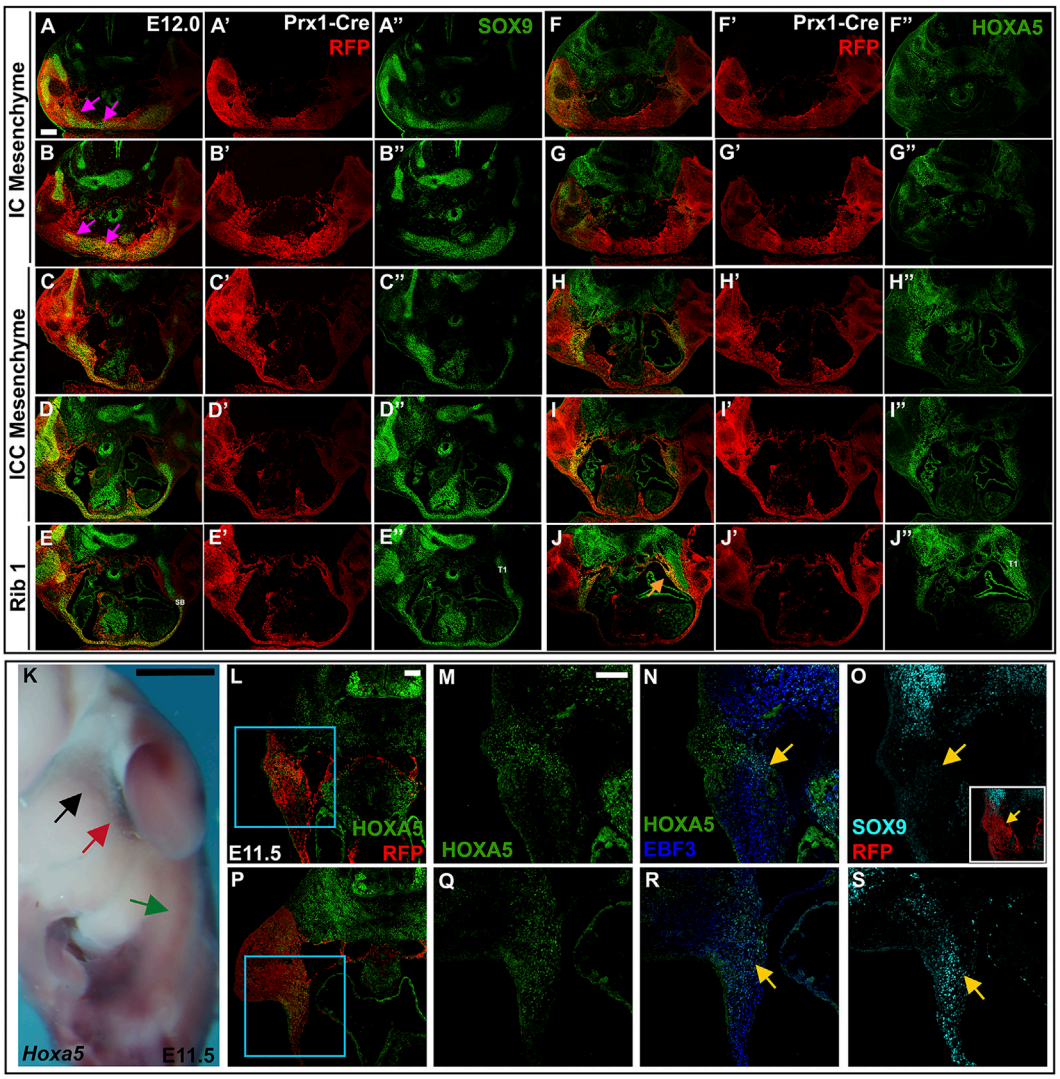


FIGURE 8 | LPM-derived HOXA5 expression can be seen laterally as early as E11.5 and E12. **(A–J″)** E12.5 Prx1-Cre-RFP embryo in which alternate sections were labelled for SOX9 **(A–E″)** or HOXA5 **(F–J″)**. Images are shown with RFP and either SOX9 **(A–E)** or HOXA5 **(F–J)**; RFP alone **(A′–E′**; **F′–J′)**, or SOX9 **(A″–E″)** or HOXA5 **(F″–J″)** alone. Pink arrows in **(A,B)** indicate SOX9 expression associated with developing clavicles and IC elements. **(K)** Whole mount *in situ* hybridization for *Hoxa5* in an E11.5 embryo. Black arrow indicates LPM expression anterior to the forelimb; red arrow points to LPM expression adjacent to the forelimb; green arrow marks additional HOXA5 expression that remains untranslated. **(L,P)** RFP and HOXA5 expression in sections through an E11.5 Prx1-Cre RFP embryo at the level indicated in **(K)** by black arrow **(L)**, or red arrow **(P)**. **(M,N)** expression of Hoxa5 **(M)** or EBF3 and HOXA5 **(N)** in the area indicated by blue square in **(L)**. **(O)** SOX9 expression in an alternate section of same embryo shown in **(L)**. **(Q,R)** expression of HOXA5 **(Q)** or EBF3 and HOXA5 **(R)** in area indicated by blue square in **(P)**. **(S)** SOX9 expression in an alternate section of same embryo shown in **(P)**. Yellow arrow in **(N,R)** indicates region of HOXA5 and EBF3 co-expression. Yellow arrow in **(O,S)** shows corresponding region of SOX9 expression. Inset in **(O)** shows SOX9 expression in the context of Prx1-Cre-RFP. Scale bar: **(A–J)**, 200 μM; **(K)**, 25 mm; **(L–S)**, 100 μM. SB, sternal bar; T1, thoracic rib 1.

(**Figure 7**), E12 (**Figure 8**), and E11.5 (**Figure 8**), with the goal of identifying the axial source of these individual primordia. At E12.5, SOX9 expression in the presternum is very similar to that of E13.5 (**Figures 7A"–H"**): expression is seen in separate dorsal and ventral presternal elements (**Figures 7B"–E"**, blue and pink arrows, respectively). Rib 1 contacts the paired ventral ICC elements that have not yet fused or fully condensed into the bar shape observed later (**Figure 7F**, pink arrow). More posteriorly, the site of rib 1 attachment transitions to the plow-shaped, dorso-ventrally elongated morphology characteristic of the sternal bars (**Figure 7G**, yellow arrow). Rib 2 similarly contacts the unfused sternal bars, and the bars are connected by a stream of Sox9-expressing cells (**Figure 7H**, white arrow). Meox1-Cre dependent expression of RFP marks all cells derived from the somites, demonstrating that only the ribs are somite derived, while all components of the sternum are RFP-negative (**Figures 7A–H**).

HOXA5 immunofluorescence in alternate sections of the same embryo revealed expression in a domain laterally adjacent to but not overlapping the SOX9 positive IC mesenchyme (**Figure 7J**). Further, this domain of HOXA5 expressing cells is LPM-derived (not shown) and is not labeled with Meox-Cre (**Figure 7J′**). In

contrast, posterior to the clavicle in the ICC mesenchyme, HOXA5 is strongly expressed in the ventral elements of the ICC (**Figure 7L"**) but remains excluded from the dorsal ICC elements. A comparison with SOX9 reveals that as for later stages, HOXA5 expression is excluded from the most medial cells of the ventral ICC, but overlaps with SOX9 in the lateral part of these condensations (**Figures 7M",N"**, pink arrows). In addition, HOXA5 is expressed in a LPM-derived lateral domain that does not stain for SOX9 (orange arrow, **Figure 7M"**). The SOX9-positive and negative domains of HOXA5 expression here appear continuous, suggesting that the lateral population of HOXA5 positive cells may migrate medially to populate the ventral presternum, and activate chondrogenesis. This broad lateral domain, inside and lateral to the ventral ICC cartilage condensation is observed anterior to rib 1, as well as throughout the region of rib 1 attachment (**Figure 7N**). As described above, HOXA5 expression in the sternum is restricted to the region around rib 1 attachment, and shifts dorsally when rib 2 appears (**Figure 7P**). As observed at later stages, HOXA5 expression suggests an important role in the double attachment of rib 1 to the sternum.

Half a day earlier, at 12 days of development, SOX9 is expressed in a broad LPM-derived domain around the midline, in the area that will form the clavicles and the IC mesenchyme, although these structures are not yet morphologically distinct from each other (**Figures 8A,B** pink arrows). In the IC and ICC regions, the SOX9-positive mesenchyme cannot be resolved as separate dorsal and ventral domains, as it can at E12.5. Instead, a population of HOXA5-positive cells are located lateral to SOX9 domain at the IC and ICC levels (**Figures 8F,G**). More posteriorly, HOXA5 expression remains lateral, but includes the lateral-most edges of the SOX9 domain, which now extends in a stream of cells across the midline (**Figures 8C,D,H,I**). At the level of rib1, HOXA5 marks a population of LPM-derived cells lining the medial edge of the incoming rib (**Figure 8J**, orange arrow). At this stage, the sternal bars appear as a Sox9-positive triangle of cells in contact with the rib (**Figure 8E**), with a stream of cells reaching across the midline, as previously described (**Figure 8E**) (Chen, 1952a). These observations suggest that the HOXA5-expressing ventral ICC presternum may arise from this population of LPM-derived cells that develop lateral to the dorsal sternum elements. However, they do not rule out the possibility that HOXA5 expression initiates in the ventral ICC presternum only after those elements form at the midline.

At 11.5 days of development, prior to ventral body wall closure and sternal bar migration, we detected *Hoxa5* transcripts in the LPM, just anterior to and adjacent to the forelimb bud (**Figure 8K**, black and red arrows, respectively). More posteriorly, additional *Hoxa5* mRNA was detected (green arrow, **Figure 8K**), but it was previously shown that this corresponds to alternative *Hoxa5* transcripts that are not translated, and no HOXA5 protein is produced posteriorly (Coulombe et al., 2010). In section, LPM-derived HOXA5 expressing cells were observed anterior to the forelimb (**Figures 8L,M**, black arrow in **Figure 8K**). Counterstaining was performed for EBF3, a transcription factor previously shown to be expressed in LPM-derived connective tissue precursors as early as E10.5, and required within the LPM for sternum ossification (Kuriki et al., 2020). This revealed co-expression of HOXA5 and EBF3 in a subdomain of HOXA5-positive cells (yellow arrow, **Figure 8N**). This raises the possibility that HOXA5 marks the progenitors of connective tissue associated with the rib-sternum attachment point in addition to the prospective cartilage itself. In an alternate section, a low level of SOX9 was detected in the same region (yellow arrow, **Figure 8O**). At the level of the forelimb bud (red arrow in **Figure 8K**), a similar pattern was found. LPM-specific expression of HOXA5 was seen ventral to the forelimb (**Figures 8P,Q**), a region previously shown to contain sternal bar progenitors (Bickley & Logan, 2014). Co-expression with EBF3 was again observed in a subset of the HOXA5 domain (yellow arrow, **Figure 8R**), and in an alternate section, SOX9 was expressed in this region as well (yellow arrow, **Figure 8S**). While we cannot yet determine if either or both of these HOXA5 domains will ultimately contribute to development of the *Hoxa5*-dependent ventral presternum domain that arises later in development, it is clear that LPM-specific expression partially overlaps both SOX9 and EBF3 as early as E11.5.

DISCUSSION

In this work, we characterized and compared the roles of *Hoxa5* in somites versus LPM, in order to address the broader question of how skeletal development is coordinated between these tissues. Our results led to an analysis of presternum development, which had not been well characterized in mice. The presternum is both morphologically and developmentally distinct from the mesosternum. Anteriorly, it encompasses the sternoclavicular joint--a structure that provides the single point of skeletal articulation between the shoulder girdle and axial column via contact with the clavicles--as well as the attachment site for rib 1. In humans, this clavicle-presternum-T1 unit is further strengthened by the presence of the costoclavicular ligament, which attaches the clavicles to T1. Here, we first present a model for mouse presternum development. In following sections, we discuss evidence that HOXA5 expression marks the embryonic precursor of a previously identified lateral presternum structure that is variably present in therians (Buchholtz et al., 2020) and discuss *Hoxa5* phenotypes with reference the model for presternum development, as well as the combinatorial action of other *Hox* genes. Finally, implications for understanding sternum evolution in the mammalian lineage are discussed.

The Mouse Presternum Develops From Multiple LPM-Derived Elements

Analysis across a time-series allowed a detailed description of presternum development in mouse. We propose that there are at least five elements contributing to the presternum, each of which arises as paired primordia and is derived primarily from LPM with no contribution from the somites. Further, we propose that

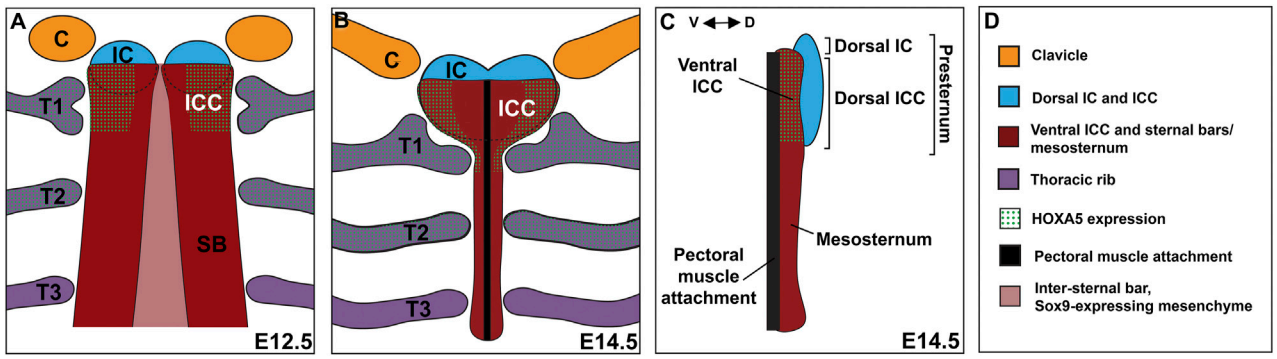


FIGURE 9 | The mammalian presternum is a composite structure with dorsal-ventral organization and molecularly distinct domains. **(A)** Schematic of sternum development in an E12.5 embryo. The clavicles (orange) and dorsal IC/ICC mesenchyme form as paired condensations that express low levels of SOX9. The ventral ICC of the presternum appears as a region of high SOX9 expression that is continuous with the posterior sternal bars (SB). HOXA5 is expressed in both the ventral ICC and rib 1, marking the site of rib 1 attachment to the sternum. Rib 2 also expresses high levels of HOXA5. SOX9 expression can be seen in a stream of cells that cross the midline between the sternal bars. **(B)** Schematic of sternum development in an E14.5 embryo. By E14.5 sternal bar fusion is mostly complete and the presternum can be seen to have a morphology distinct from the rest of the sternum. The clavicles articulate with the dorsal IC, while the first rib contacts the HOXA5-expressing ventral ICC in two places. The pectoral muscles attach along the ventral midline of the sternum. **(C)** Side view of an E14.5 sternum showing dorsal ventral regionalization. The dorsal IC and ICC (blue) form a distinct element, while the ventral ICC appears continuous with the sternal bars. HOXA5 specifically marks the ventral ICC. **(D)** Key for coloring in **(A–C)**.

there are at least two progenitor domains for these five elements (and more are possible). This model is summarized in **Figure 9**. The existence of an additional, lateral and LPM-derived progenitor domain contributing to the presternum at the point of T1 attachment is also supported by our results (see below).

By E14.5, five SOX9-positive condensations of the presternum were observed (**Figures 9B,C**): 1) the IC mesenchyme between the clavicles continuous with 2) dorsal ICC mesenchyme, which is a heart-shaped, loose mesenchyme anterior to and flanking the T1 rib attachment; 3) ventral ICC mesenchyme, which is a bar shaped dense mesenchyme anterior to and flanking the T1 rib; 4) ventral midline mesenchyme likely forming the connective tissue of the pectoral muscle attachment. and 5) the sternal bars, which have fused at the midline by this stage and are clearly visible from the posterior edge of the T1 attachment and extending posteriorly through the rest of the sternum. Because the IC and dorsal ICC appear morphologically continuous they are represented as a single dorsal component (colored blue in **Figure 9**). Similarly, the ventral ICC is morphologically continuous with the sternal bars and they are represented as a second component (colored red in **Figure 9**).

Several points from past and the present work support these two regions as having distinct developmental origins. Indeed, other descriptions of presternum development in mammals agreed that the presternum incorporates anterior sternal bars (Gladstone & Wakeley, 1932; Rodríguez-Vázquez et al., 2013) and references therein). Further, the anterior presternum reportedly consists of dermal bone while the more posterior part of the presternum is endochondral, like the mesosternum (Gladstone & Wakeley, 1932). The former region has been proposed to be homologous to the interclavicle, which is a midline component of the presternum in most tetrapods, but absent in therian mammals (for example, (Gladstone & Wakeley, 1932; Rodríguez-Vázquez et al., 2013; Buchholtz et al., 2020), and references therein).

Our examination at stages prior to E14.5 also suggest at least two progenitor domains for the presternum and extends previous characterizations of the ICC and T1 attachment point. At E11.5, progenitors of sternal bars are known to be clustered in a triangle of cells ventral to the forelimb bud, many of which, as previously reported, express EBF3 and/or SOX9 (Bickley & Logan, 2014; Kuriki et al., 2020). In addition, beginning as early as E11.5 and until fusion of the sternal bars, we observed a continuous line of SOX9-postitive cells reaching across the midline from one sternal bar to the other. These may be the "stream of flattened cells" previously described as migrating with the outgrowing intercostal muscles (Chen, 1952a). No other sternal progenitors are known at that stage. However, at E12, we first observed a paired, SOX9 positive condensation between and largely continuous with the clavicle condensations, and extending posteriorly into the ICC region, similar to what was recently reported in human embryos (Rodríguez-Vázquez et al., 2013). At this stage, the sternal bars are in a separate and more dorso-lateral location, have already fused with the outgrowing rib anlage, and migrated ventrally from their point of origin in the axillary region. One half day later, at E12.5, all five elements described above are distinct (**Figure 9A**). A more closely spaced time series between E12-E12.5, coupled with fate mapping, could further resolve the origin of the five regions above.

Observations across this time series indicate that the ICC region at and just anterior to the T1 attachment represents a unique AP domain in which separate dorsal and ventral ICC elements overlap with one another. Further, the lateral-most portion of the ventral ICC is molecularly distinct based on its differential expression with HOXA5 (discussed below).

Additionally, our time series shows that all of these five presternal regions originate from paired progenitor domains, and thus likely migrate toward and/or fuse across the midline. Finally, lineage-labeling with Prx1-Cre shows that all five

presternal elements are primarily if not entirely LPM-derived, consistent with previous observations for the sternal bars (Fell, 1939; Chevallier, 1975; Durland et al., 2008; Bickley & Logan, 2014) and the sternum at the point of T1 attachment (Durland et al., 2008). Further, Meox1-Cre labeling shows that none of these elements contain contribution from the somites, which definitively marks the position of the lateral somitic frontier at the points of rib attachment. It was previously reported that post-otic neural crest contributes to the clavicle and anterior presternum (Matsuoka et al., 2005), which is almost certainly restricted to the IC mesenchyme.

HOXA5 Expression Identifies a Molecularly Distinct Region of the Presternum and T1 Rib Attachment

HOXA5 is expressed in the developing T1-T3 ribs (**Figure 5**; Holzman et al., 2018). Here we show it is also expressed in the presternum, but specifically at both points of T1 rib contact as well as anterior to rib 1 (green hatching, **Figure 9**). This domain of expression corresponds to the ventral-lateral ICC. HOXA5 is also expressed in presternum at the second, more posterior point of T1 contact, where there is sternal bar-like morphology. The posterior border of HOXA5 expression in the sternum marks the posterior edge of the T1 attachment. Finally, in addition to expression in these lateral presternum cartilage condensations (as marked by SOX9 expression), at E12.5–E13.5 we also observed HOXA5 in laterally adjacent, SOX9-negative cells. It is possible that these lateral cells contain progenitors that will contribute to the presternum condensation, or alternatively they may become connective or other tissues of the surrounding sternoclavicular joint. In either case, the position of HOXA5-expressing cells in LPM specifically surrounding the point of rib 1 contact suggests that HOXA5 marks the embryonic anlage of a recently-defined lateral presternum element. This was identified as a separate lateral presternum ossification present at low penetrance in human medical scans, and also variably present in some groups of therians (Buchholtz et al., 2020). These results suggest that the embryonic primordium of this region is present in mice and thus may be common to all therians. HOXA5 expression marks this lateral embryonic element and HOXA5 expression was absent from other regions of the developing presternum (**Figure 9**).

It is notable that the presternum HOXA5 expression domains spatially coincide so precisely with the HOXA5-expressing T1 ribs. Indeed, expression in the LPM is notably specific to regions immediately adjacent to somite-derived and HOXA5-positive tissue from the earliest (E11.5, **Figure 1**) to the latest (E14.5, **Figure 5**) stages we examined. As deletion of *HoxA5* in either LPM or somitic mesoderm affects attachment of the T1 rib, this bridge of HOXA5 expression across the lateral somitic frontier may allow for proper communication between these two distinct but cooperating mesoderm sources. In this context, it would be interesting to know whether expression of *Hoxa5* is co-regulated in somites and LPM. In fact, the cis-regulatory elements that control both forelimb bud and somitic transcription of the single *Hoxa5* coding transcript have been localized to a 2.1 kb fragment termed the mesodermal enhancer (MES) (Larochelle et al., 1999). However, reporter assays suggest that CREs necessary for somitic vs. limb bud expression can be at least partially separated but reside within the same ~900 bp fragment (Tabariès et al., 2007). Additionally, CDX1 binding sites within the same region define the sharp posterior boundary of somitic expression. It would be interesting to further characterize this CRE from mouse and other mammals, to better understand how this intricate pattern of expression is achieved across the lateral somitic frontier.

Tissue-Specific Roles for *Hoxa5* in Presternum and Rib Development

Results from conditional deletion of *Hoxa5* in somites vs. LPM show that *Hoxa5* plays largely independent roles in these two tissues: the conditional deletion phenotypes for the most part additively explain the *Hoxa5* null skeletal phenotype.

Somite-specific *Hoxa5* deletion reproduced all of the vertebral (primaxial) skeletal phenotypes associated with the null allele, including homeotic transformations of C7-T1. This included both vertebral changes and a gain of ectopic C7 ribs, which could be unilateral or bilateral, and could be free, fused to the T1 rib, or fused to the sternum. While the proximal ribs are primaxial, the distal T1 rib is abaxial (Durland et al., 2008). However, LPM-specific *Hoxa5* activity was dispensable for T1 (and C7) rib phenotypes including symmetrical guidance and fusion to the presternum.

Somite-specific *Hoxa5* activity is also necessary for proper positioning of the T1 fusion to the presternum, and Meox1-Cre conditional deletion reproduced the asymmetric T1 fusion and T1 bifurcation phenotypes associated with the null allele. As mentioned above, these phenotypes could be explained in at least two ways, which are not mutually exclusive. First, somitic *Hoxa5* activity may be required for rib progenitors (and those of associated intercostal muscle and connective tissue) to migrate segmentally across the lateral somitic frontier. This segmental migration is known to rely on signaling among somite derivatives and to involve guidance molecules such as Ephs/Ephrins (Compagni et al., 2003). Second, it is also possible that symmetrical T1 fusion is achieved because the presternum attachment point is molecularly unique. HOXA5 expression in somitic tissues may be necessary for recognition and attachment to this unique region, and attachment becomes more random in its absence resulting in altered an asymmetric rib fusions. HOXA5 itself could contribute to molecularly defining this point of attachment, although its sternal expression is clearly dispensable for symmetry of T1 fusion. However, it is required for morphogenesis of the presternum; in null embryos, we observed that the lateral, HOXA5-expressing portions of the presternum are misshapen and thus the morphology of this attachment is altered even in embryos with symmetrical rib fusion. Unfortunately, the perinatal lethality of the *Hoxa5* mutation (due to respiratory defects) prevents analysis at stages when this region is further developed.

LPM-specific activity of *Hoxa5* is shown here to include negative regulation of presternum extension anterior to T1; in both null embryos and in those with LPM-specific *Hoxa5* deletion this region was rostro-caudally elongated. The phenotype was more extreme in the latter case. We suspect that it is the ventral

ICC that is perturbed, as these are the only LPM-derived cells in the affected region that express HOXA5. HOXA5 could negatively regulate recruitment of mesenchyme to the lateral ICC condensation. Indeed, it is expressed both within the SOX9 condensation and in mesenchyme lateral to it. Alternatively or in addition, HOXA5 could regulate the progress of chondrogenic differentiation which would also alter the morphology of the resulting structure. Interestingly, both roles were previously described for *Hoxa5* in the acromion (Aubin et al., 2002), which is the portion of the scapula that articulates with the lateral end of the clavicle.

This presternum elongation was the only phenotype that may be non-additive. The elongation of the presternum was far more pronounced in LPM-deleted embryos than in *Hoxa5* null embryos; further, in the former, several embryos were observed with a complete additional ossification between T1 and the Y-shaped cartilage. This could not be considered a homeotic transformation by the criterion available at E18.5: in these embryos, the sternum does not segment at the point of T1 attachment, in contrast to mesosternal joints where segmentation is apparent. Rather, this phenotype could arise simply by failure to limit recruitment or proliferation of cells into the ventro-lateral ICC, or to regulate differentiation, as described above. It is possible that the more severe phenotype in conditional compared to complete knockouts results from a requirement for communication between somites and LPM, and that a mismatch in their Hox code exacerbates the effect. However, we cannot rule out contribution of mixed genetic background introduced from the Prx-1-Cre line altering the expressivity of the phenotype. Countering this, the *Hoxa5* skeletal phenotype has been characterized in different backgrounds and in no case were ectopic presternum ossifications ever observed (Aubin et al., 2002). Finally, we cannot totally reconcile two different phenotypes but note they are not mutually exclusive: in whole-mount skeletons there is evident AP elongation of the presternum corresponding to the ventral ICC. In sections, we also observe a truncation in the medio-lateral direction; this is more spatially restricted to the T1 attachment point.

In contrast to the presternum, there was no effect of *Hoxa5* deletion (either conditional or complete) on the length of the mesosternum, consistent with its lack of expression posterior to T3.

Finally, while HOXA5 expression is largely excluded from the IC and dorsal ICC, we noted that the Y-shaped cartilage that develops from this region is often smaller in *Hoxa5* null or LPM-specific mutants. Further, in null embryos both the Y-shaped cartilage and the point of T1 attachment often ossify prematurely. This latter phenotype was not produced by LPM-specific deletion, and only once following somite-specific deletion. This could indicate that either domain of *Hoxa5* is sufficient to regulate ossification of the presternum, or alternatively, that *Hoxa5* acts in a different tissue (such as neural crest) to mediate this phenotype.

A Presternum Hox Code

It is well established that Hox genes act combinatorially to pattern somites and neural tube derivatives, with nested, colinear expression important for their function (reviewed in (Krumlauf, 2018)). With the exception of the limbs, colinear Hox expression is not well-reproduced in the body wall LPM. However, genetic evidence indicates a combinatorial Hox code is also operating in the presternum.

For example, while *Hoxa5* mutants show an elongated presternum anterior to the T1 rib, in contrast, the entire presternum anterior to the T2 attachment is absent in *Hoxa5; Hoxb5; Hoxc5* triple mutants (McIntyre et al., 2007). This phenotype is consistent with a loss of both the dorsal and ventral ICC elements described here. It is also consistent with other *Hoxa5* phenotypes including homeotic vertebral transformations that show opposite effects in *Hoxa5* single mutants compared to *Hoxb5* or *Hoxa5; Hoxb5; Hoxc5* triple mutants (Jeannotte et al., 1993; McIntyre et al., 2007). Together, this highlights the often antagonistic activity of *Hoxa5* and *Hoxb5*.

The *Hox4* and *Hox6* paralog groups also pattern in the presternum. Following triple knockout of all *Hox6* paralogs, the distance between T1-T2 rib attachments is greatly reduced and fails to ossify, but the Y-shaped cartilage anterior to T1 forms(McIntyre et al., 2007). This could be interpreted as disruption specifically of the ventral ICC posterior to T1, a region that is HOXA5-negative and has a sternal-bar like morphology. In contrast, *Hoxb4* mutants show disruption of the anterior presternum: the Y-shaped cartilage is absent, but 2 strips of unfused cartilage articulate with the T1-T2 ribs (Ramirez-Solis et al., 1993). This could possibly result from disruption of the IC/dorsal ICC, but retention of the ventral ICC (which fails to fuse). All of these *Hox* phenotypes are consistent with this model of presternum development, and combinatorial effects on its different components.

Symmetrical rib attachment to the sternum is a genetically separable phenotype from the growth of the various presternum components. This is shown for *Hoxa5* above, where *Hoxa5* in somites is necessary symmetrical T1 fusion. Although HOXA5 is also expressed in rib 2, *Hoxa5* single knockouts do not have T2 phenotypes. In contrast, in *Hoxb5* single mutants and *Hoxb5; Hoxb6* double mutants, the second rib often bifurcates and one arm fuses at the T1 rib position (in the same animals the T1 rib is either missing or shortened (Rancourt et al., 1995). *Hox9* compound mutants also have aberrant fusion of T2 at the T1 position (McIntyre et al., 2007). Together, this implicates multiple *Hox* genes in positional information located either on the sternum and/or in the migrating distal ribs or associated tissues.

Implications for Sternum Evolution

The sternum arose in early tetrapod evolution as an adaptation to life on land, and sternum morphology varies with mode of locomotion and respiration, making it a good system in which to investigate the developmental mechanisms underlying evolutionary change. In addition, the sternum is of interest because it is the site of interaction between mesoderm derivatives with different embryonic origins. In this way, the site of contact between the two, the lateral somitic frontier, requires the evolution of signals that function across it.

The mammalian sternum has been reduced in size and skeletal structure number over the course of evolution. On the basis of fossil evidence, it has been proposed that this reduction is either due to the loss of the primitive interclavicle, the ancestral structure that constitutes the entire sternum in basal synapsids, in the transition to basal therians, or to the incorporation of the interclavicle into a single presternum structure that articulates with clavicles anteriorly, rib 1 laterally, and rib 2 posteriorly. Loss or incorporation of the interclavicle has thus been thought of as key step in sternum evolution. Using fossil and anatomical imaging, evidence has been found suggesting that the mammalian presternum is in fact a composite structure containing the ancestral interclavicle, as well as the anterior ends of the sternal bars. In addition, Buchholtz et al (2020) provide evidence for the existence of lateral elements that are the site of rib 1 attachment. Our observations of presternum development in mice support this hypothesis that the mammalian presternum is a composite structure. We show that the presternum has dorsal-ventral organization, and that it contains LPM-derived IC and ICC elements (**Figure 8**). Our analysis of SOX9 expression indicates that the dorsal IC and ICC may represent the more primitive interclavicle element. The ventral ICC likely arises from the anterior ends of the sternal bars, and represents what is referred to as the manubrium in the fossil record (e.g., Luo et al., 2007), in agreement with previous studies (Rodriguez Vázquez et al., 2013; Buchholtz et al., 2020). Moreover, while morphologically continuous with the sternal bars, analysis of *Hoxa5* function and expression reveals that the ventral ICC is molecularly distinct, and may indeed contain HOXA5-expressing lateral elements that provide the site of rib 1 attachment. Finally, our finding that HOXA5 expression spans the lateral somitic frontier, marking both somite-derived ribs and LPM-derived sternum, suggests that *Hoxa5* may play an important role in allowing for communication between both mesoderm types at the lateral somitic frontier, thereby permitting establishment of the unique rib 1-sternum attachment site.

METHODS

Mouse Strains and Genotyping

The following mouse strains were used: *Hoxa5-Cre: Tg(Hoxa5-cre)447BLjea* (Bérubé-Simard & Jeannotte, 2014); *Hoxa5^{flox}*: *Hoxa5$^{tm1.1Ljea}$* (Tabariès et al., 2007); *Prx1-Cre*: *B6.Cg-Tg(Prrx1-cre)1Cjt/J* (Logan et al., 2002); *Rosa6^{tdtomato}*; *B6.Cg-Gt(ROSA)26Sortm9(CAG-tdTomato)Hze/J* (Madisen et al., 2010); *Hoxa5^{-}*: *Hoxa5tm1Rob* and Meox1Cre: *Meox1tm1(cre) Jpa* (Jukkola et al., 2005). Conditional knockout or RFP-labeling was carried out by crossing a male harboring a Cre allele to a female with a loxP allele. Timed-pregnant females were sacrificed to collect embryos at the stages indicated. Genotyping of tail snip DNA was performed with primers for Cre, RFP, *Hoxa5^{+}* or *Hoxa5^{-}* alleles. All procedures were performed in accordance with the NIH Guide for Care and Use of Laboratory Animals and approved by the Columbia University IACUC. Genotyping primers (5′ to 3′) were as follows. Cre: forward GCGGTCTGGCAGTAAAAACTATC, reverse GTGAAACAG CATTGCTGTCACTT; tdTOMATO/RFP: forward CTGTTC CTGTACGGCATGG, reverse GGCATTAAAGCAGCGTAT CC; *Hoxa5^{+}* allele: forward ACTGGGAGGGCAGTGCCCCCA CTTAGGACA, reverse CTGCCGCGGCCATACTCATGCTTT TCAGCT; *Hoxa5^{-}* allele forward ACTGGGAGGGCAGTGCCC CCACTTAGGACA, reverse GGCTACCTGCCCATTCGACCA CCAAGCGAA.

Skeletal Staining

Alcian Blue and Alizarin Red staining was performed on E18.5 embryonic skeletons as previously described (Lufkin et al., 1992).

Immunofluorescence

Immunostaining was performed as described (McGlinn et al., 2019). Briefly, embryos were fixed 2 h to overnight in 4% paraformaldehyde at 4°C and embedded in OCT. 8–10 μm sections were cut, tissue was permeabilized and blocked in 5% Normal Donkey Serum/0.3% Triton-X/PBS and incubated in primary antibodies overnight at 4°C, washed in PBS, and incubated for 3 h RT with Alexa-488 or 594-conjugated secondary antibodies (Jackson Immunoresearch, diluted 1:400). Slides were washed in PBS counterstained with DAPI, and mounted in Prolong Diamond. Primary antibodies: HoxA5 (Phillipidou et al., 2012; 1:5000); Sox 9 (Millipore-Sigma AB5535, 1:500 or RnD Systems AF3075-SP 1: 500); RFP (Chromotek 5F8, 1:1000); Ebf3 (RnD Systems AF5166, 1: 1000); Tenascin (Sigma-Aldrich T3413, 1: 100); PCNA (Santa Cruz sc-56 1:200), Cleaved Caspase 3 (CST 9661 1:200).

In situ Hybridization

Whole-mount *in situ* hybridization for *Sox9* was performed as previously described (Brent et al., 2005).

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

ETHICS STATEMENT

The animal study was reviewed and approved by Columbia University IACUC.

AUTHOR CONTRIBUTIONS

JM, LJ, and AB designed the study. KM, JB, AB, MH, TF, KS, and JM conducted experiments. JM and AB wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was funded by the NIH (NICHD 1R15HD092179-01 to JM) and NSF (IOS-2019537 to JM and LJ). Acquisition of a confocal microscope at Barnard was funded by an NSF MRI (DBI-1828264). Undergraduate summer fellowships were funded by the Amgen Scholars program (to KM and TF), and by Barnard College via the Columbia/Barnard Summer Undergraduate Research Fellowship program and the Barnard Summer Research Institute (to KM, JB, TF, KS).

ACKNOWLEDGMENTS

We thank Juha Partanen for the *Meox1Cre* mouse line and Jeremy Dasen for the HOXA5 antibody, as well as Abigail Ryckman for excellent technical assistance and lab management. We thank Emily Buchholtz for advice about functional anatomy and sternum evolution, and we are grateful to Emily Buchholtz and the reviewers for insightful comments and suggestions that improved the manuscript. This work was funded by the NIH (NICHD 1R15HD092179-01 to JM) and NSF (IOS-2019537 to JM and LJ). Acquisition of a confocal microscope at Barnard was funded by an NSF MRI (DBI-1828264). Undergraduate summer fellowships were funded by the Amgen Scholars program (to KM and TF), and by Barnard College via the Columbia/Barnard Summer Undergraduate Research Fellowship program and the Barnard Summer Research Institute (to KM, JB, TF, and KS).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2022.806545/full#supplementary-material

Supplementary Figure 1 | Rib and sternum defects following conditional deletion of *Hoxa5*. **(A)** control sternum showing proper T1 and T2 fusion. **(B–D)** Examples of rib and sternum phenotypes, described in the text. In all panels, black arrowheads indicate T1 ribs, black arrows indicate T2 ribs, and grey arrowheads indicate (ectopic) C7 ribs. Grey arrow indicates (ectopic) ossified sternebra anterior to T1. White arrowhead in Bii indicates ossification of the Y-shaped cartilage, White arrow in Cv indicates fused sternebrae, also called crankshaft sternum. Most of the whole-mount views shown here correspond to specimens shown in **Figure 3** after sternum dissection. Correspondence of images is as follows: **Supplementary Figures S1A to Figure 3A**, control embryo; **Supplementary Figures S1Bi to Figure 3Bii** *Hoxa5* null embryo; **Supplementary Figures S1Ci-iii to Figure 3Ci-iii**; three somite-conditionally deleted embryos; **Supplementary Figures S1D to Figure 3Di**, LPM conditionally deleted embryo. Specimens shown in **Supplementary Figures S1Bii,Civ,v** are not shown in **Figure 3**.

Supplementary Figure 2 | Medial clavicles and sternum are derived from LPM. Analysis of E14.5 Prx1-Cre-RFP embryo in which alternate sections were labelled for SOX9 or HOXA5. Images are shown with RFP, SOX9 or HOXA5, and DAPI **(A–L)**, without DAPI **(A′–L′)**, RFP alone **(A″–L″)**, and with SOX9 or HOXA5 alone **(A‴–L‴)**. Red arrow in **(F‴)** marks lateral HOXA5 expression. Yellow arrow in **(L′)** points to LPM-derived connective tissue in an infrahyoid group muscle. Scale bar: 200 μm. C, clavicle; IC, interclavicular; ICC, intercostoclavicular.

Supplementary Figure 3 | At E14.5, HOXA5 and SOX9 are co-expressed in the cells of the ventral ICC mesenchyme. **(A–F)** E14.5 embryo co-labelled for HOXA5 (green, **A′–F′**) and SOX9 (red, **A″–F″**). **(G–L)**. Higher magnification of the region corresponding to the yellow square in **(A)**, reveals co-expression (yellow cells). Co-expression is seen in the ventral ICC **(I,J)**, with the HOXA5 domain extending more broadly ventral to SOX9. HOXA5 and SOX9 additionally co-express in ribs 1 and 2 **(K,L)** and spanning the site of rib 1-to-sternum fusion **(K)**, while HOXA5 alone is detected in the perichondrium of ribs 1 and 2 **(E,F)**, and in the connective tissue of the infrahyoid muscles, dorsal to the sternum **(E,F)**. Pink arrows in **(B,H)** mark first HOXA5 expression in the ventral mesenchyme. Red arrows in **(C′,D′)** indicate exclusion of HOXA5 from the most medial region of the ventral ICC mesenchyme. Orange arrows in **(C,D)** point to HOXA5 lateral domain that does not co-express with SOX9. Scale bars: 200 μM.

Supplementary Figure 4 | Hoxa5 marks the ventral presternum at E13.5. **(A–J′)** SOX9 expression in serial sections of a WT embryo after E13.5 days of development, from the presternum to attachment of rib 3, shown with **(A–J)** or without **(A′–J′)** DAPI staining. All of the presternal condensations observed at E14.5 can be distinguished: the paired dorsal elements are present in IC mesenchyme at the site of clavicle attachment **(A,A′)** and extend posteriorly through the ICC region. Just posterior to the clavicles, the paired ventral elements that express HOXA5 at E14.5 are also visible (pink arrow, **B**) and extends through the ICC region **(C–E)**. The ventral-most SOX9-positive mesenchyme, which articulates with pectoral muscles, is also apparent (white arrow, **C′**). Red arrow in **(E)** points to the space between the paired ventral elements that have not yet fused at the midline. To confirm that each presternum element was derived from the LPM, we again compared SOX9 expression to that of RFP under control of Prx1-Cre **(K–T)‴**. E13.5 Prx1-Cre-RFP embryo in which alternate sections were labelled for SOX9 or HOXA5. Images shown with RFP, SOX9 or HOXA5, and DAPI **(K–T)**, without DAPI **(K′–T′)**, RFP alone **(K″–T″)**, and with SOX9 or HOXA5 alone **(K‴–T‴)**. While the ribs were clearly RFP-negative as expected (R), all SOX9-expressing sternal elements were RFP-positive **(K–T)**. As at E14.5, none of these presternum elements showed contribution form somites (not shown). A comparison of SOX9 and HOXA5 in alternate sections at E13.5 revealed a pattern similar to that observed a day later at E14.5. HOXA5 expression is largely absent from the presternum in the IC mesenchyme **(L‴)**, but it is observed in the ventral ICC mesenchyme in a lateral domain (orange arrows, **N′**, **N‴**) that also expresses SOX9 (orange arrow, **M‴**) and is LPM-derived (orange arrow, **N′**). Additionally, HOXA5 is expressed in the ventral presternum of the ICC region (pink arrows, **N‴**, **P‴**) anterior to T1. Note: the plane of section in this example is uneven, hence the right side of the embryo is slightly more posterior than the left side. Anteriorly, Rib1 makes contact with the HOXA5-expressing ventral-lateral presternum (red arrow, **P‴**). Like at E14.5, Rib 1 contact with the presternum also continues posteriorly into the region where sternal morphology shifts to unfused sternal bars (**Q‴**). At this axial level, HOXA5 expression is also present dorsally and thus surrounds the point of rib 1 contact (**R‴**). However, it does not mark the remainder of the sternal bar (green arrow, **R‴**) On the embryo's left side, which is slightly anterior to the right, HOXA5 is expressed in its lateral, LPM-derived domain (orange arrow, **R′**) like at E14.5. **T‴** shows the sites of rib 1 and rib 2 attachment: sternal HOXA5 marks the point of rib 1 contact on the embryo's left side (red arrow, **T‴**), but shifts to its more dorsal domain at the point of rib 2 contact (green arrow, **T‴**). Pink arrows in **B**, **N‴**, and **P‴** mark the ventral ICC that expresses Hoxa5. Orange arrows in **M‴** and **N‴** indicate lateral SOX9 or HOXA5 expression. Red arrows in **T‴** and **P‴** point to LPM-HOXA5 expression at the point of rib 1 fusion. Green arrow in **T‴** marks the point of rib 2 attachment, where HOXA5 has now shifted dorsally. Scale bar: 200 μM. C, clavicle; IC, interclavicular; ICC, intercostoclavicular; P, pectoral muscle; SB, sternal bar; T1, thoracic rib 1; T2, thoracic rib 2; T3, thoracic rib 3.

Supplementary Figure 5 | Reduction in size of lateral ventral ICC mesenchyme in *Hoxa5* mutants is not accompanied by change in either cell death or cell proliferation. Cell death and cell proliferation at E14.5 were examined at E14.5 in *Hoxa5* mutants ($n = 4$) and WT littermate controls ($n = 4$). To identify changes in cell death, cleaved Caspase 3 (green), SOX9 (red), and muscle Actin (blue) were analyzed in WT **(A–D)**, in both the ventral ICC **(A,B)** and at rib 1 attachment **(C,D)**. Cell death appears minimal in the sternum at both axial levels. Inset in **(A)** indicates positive cleaved Caspase 3 in the dorsal root ganglion. In *Hoxa5* mutants, **(E–H)**, cleaved Caspase 3 in the sternum is comparable to that of WT, while an increase in cell death is visible in the pectoral muscles ventral to the sternum **(G,H)**. To identify changes in cell proliferation, expression of PCNA (green), SOX9 (red), and muscle Actin (blue) was examined in WT **(I–L)**, in the ventral ICC **(I,J)** and at rib 1 attachment **(K,L)**. At this stage, cell proliferation appears minimal in the sternum at both axial levels, while PCNA levels are higher in developing muscles (pink arrow in **(J)**). In *Hoxa5* mutants, **(M–P)**, PCNA expression in the sternum is comparable to that of WT. Scale bar: 200 μM T1, thoracic rib 1.

Supplementary Figure 6 | Lateral HOXA5 domain co-expresses EBF3, but not Tenascin. **(A–A‴)** E14.5 embryo co-labelled for muscle Actin (blue, **A**), HOXA5 (red, **A′**,**A‴**), and Tenascin (green, **A′**,**A″**). HOXA5 and Tenascin are not co-expressed in lateral domain **(A′)**. **(B–B‴)** E14.5 Prx1-Cre-expressing embryo co-labelled for RFP (blue, **B**), HOXA5 (red, **B′**,**B‴**), and EBF3 (green, **B′**,**B′′**). HOXA5 and EBF3 are co-expressed in lateral domain (**B′**, yellow cells). Orange arrows in **(A,B)** point to HOXA5 lateral domain that does not co-express with SOX9, and corresponds to region indicated in (**Supplementary Figure S3C**). IC, interclavicular mesenchyme; ICC, intercostoclavicular mesenchyme; PM, pectoral muscle. Scale bar: 200 μM.

REFERENCES

Aubin, J., Lemieux, M., Moreau, J., Lapointe, J., and Jeannotte, L. (2002). Cooperation of Hoxa5 and Pax1 Genes during Formation of the Pectoral Girdle. *Developmental Biol.* 244 (1), 96–113. doi:10.1006/dbio.2002.0596

Bérubé-Simard, F.-A., and Jeannotte, L. (2014). Hoxa5/Cre Transgenic Mice: Novel Tools for Regional Deletion along the Anterior-Posterior axis. *Genesis* 52 (2), 149–156. doi:10.1002/dvg.22733

Bickley, S. R. B., and Logan, M. P. O. (2014). Regulatory Modulation of the T-Box Gene Tbx5 Links Development, Evolution, and Adaptation of the Sternum. *Proc. Natl. Acad. Sci. U.S.A.* 111 (50), 17917–17922. doi:10.1073/pnas.1409913111

Brent, A. E., Braun, T., and Tabin, C. J. (2005). Genetic Analysis of Interactions between the Somitic Muscle, Cartilage and Tendon Cell Lineages during Mouse Development. *Development* 132 (3), 515–528. doi:10.1242/dev.01605

Buchholtz, E. A., Yozgyur, Z. M., Feldman, A., Weaver, A. A., and Gaudin, T. J. (2020). The Therian Sternum at the Lateral Somitic Frontier: Evolution of a Composite Structure. *J. Zool* 315, 19–28. doi:10.1111/jzo.12809

Burke, A. C., Nelson, C. E., Morgan, B. A., and Tabin, C. (1995). Hox Genes and the Evolution of Vertebrate Axial Morphology. *Development* 121 (2), 333–346. doi:10.1242/dev.121.2.333

Burke, A. C., and Nowicki, J. L. (2003). A New View of Patterning Domains in the Vertebrate Mesoderm. *Developmental Cel* 4 (2), 159–165. doi:10.1016/S1534-5807(03)00033-9

Chen, J. M. (1952b). Studies on the Morphogenesis of the Mouse Sternum. II. Experiments on the Origin of the Sternum and its Capacity for Self-Differentiation *In Vitro*. *J. Anat.* 86 (4), 387–401. Availableat: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1273691/.

Chen, J. M. (1953). Studies on the Morphogenesis of the Mouse Sternum. III. Experiments on the Closure and Segmentation of the Sternal Bands. *J. Anat.* 87 (2), 130–149. Availableat: http://www.ncbi.nlm.nih.gov/pubmed/13044725%0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1244579.

Chen, J. M. (1952a). Studies on the Morphogenesis of the Mouse Sternum. I. Normal Embryonic Development. *J. Anat.* 86 (4), 373–386.

Chevallier, A. (1975). Role du me'soderme somitique dans le de'veloppement de la cage thoracique de l'embryon d'oiseau. I. Origine du segment sternal et me'canismes de la diffe'rentiation des côtes. *J. Embryol. Exp. Morphol.* 33 (2), 291–311. doi:10.1242/dev.33.2.291

Christ, B., Huang, R., and Scaal, M. (2007). Amniote Somite Derivatives. *Dev. Dyn.* 236 (9), 2382–2396. doi:10.1002/dvdy.21189

Compagni, A., Logan, M., Klein, R., and Adams, R. H. (2003). Control of Skeletal Patterning by EphrinB1-EphB Interactions. *Developmental Cel* 5 (2), 217–230. doi:10.1016/S1534-5807(03)00198-9

Coulombe, Y., Lemieux, M., Moreau, J., Aubin, J., Joksimovic, M., Bérubé-Simard, F.-A., et al. (2010). Multiple Promoters and Alternative Splicing: Hoxa5 Transcriptional Complexity in the Mouse Embryo. *PLoS ONE* 5 (5), e10600. doi:10.1371/journal.pone.0010600

Durland, J. L., Sferlazzo, M., Logan, M., and Burke, A. C. (2008). Visualizing the Lateral Somitic Frontier in the Prx1Cre Transgenic Mouse. *J. Anat.* 212 (5), 590–602. doi:10.1111/j.1469-7580.2008.00879.x

Fell, H. B. (1939). The Origin and Developmental Mechanics of the Avian Sternum. *Phil. Trans. R. Soc. Lond. B* 229 (563), 407–463. doi:10.1098/RSTB.1939.0002

Gladstone, R. J., and Wakeley, C. P. (1932). The Morphology of the Sternum and its Relation to the Ribs. *J. Anat.* 66 (Pt 4), 508–564. Availableat: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1248911/.

Holzman, M. A., Bergmann, J. M., Feldman, M., Landry-Truchon, K., Jeannotte, L., and Mansfield, J. H. (2018). HOXA5 Protein Expression and Genetic Fate Mapping Show Lineage Restriction in the Developing Musculoskeletal System. *Int. J. Dev. Biol.* 62 (11–12), 785–796. doi:10.1387/ijdb.180214jm

Jeannotte, L., Gotti, F., and Landry-Truchon, K. (2016). Hoxa5: A Key Player in Development and Disease. *J. Dev. Biol.* 4 (2). doi:10.3390/jdb4020013

Jeannotte, L., Lemieux, M., Charron, J., Poirier, F., and Robertson, E. J. (1993). Specification of Axial Identity in the Mouse: Role of the Hoxa-5 (Hox1.3) Gene. *Genes Dev.* 7 (11), 2085–2096. doi:10.1101/gad.7.11.2085

Jukkola, T., Trokovic, R., Maj, P., Lamberg, A., Mankoo, B., Pachnis, V., et al. (2005). Meox1Cre: A Mouse Line Expressing Cre Recombinase in Somitic Mesoderm. *Genesis* 43 (3), 148–153. doi:10.1002/gene.20163

Kieny, M., Mauger, A., and Sengel, P. (1972). Early Regionalization of the Somitic Mesoderm as Studied by the Development of the Axial Skeleton of the Chick Embryo. *Developmental Biol.* 28 (1), 142–161. doi:10.1016/0012-1606(72)90133-9

Krumlauf, R. (2018). Hox Genes, Clusters and Collinearity. *Int. J. Dev. Biol.* 62 (11–12), 659–663. doi:10.1387/ijdb.180330rr

Kuriki, M., Sato, F., Arai, H. N., Sogabe, M., Kaneko, M., Kiyonari, H., et al. (2020). Transient and Lineage-Restricted Requirement of Ebf3 for Sternum Ossification. *Development (Cambridge)* 147 (9). doi:10.1242/dev.186239

Liem, K. F., Bemis, W. E., Walker, W. F., and Grande, L. (2001). *Functional Anatomy of the Vertebrates: An Evolutionary Perspective*. 3rd ed. Philadelphia, PA: Harcourt Brace College Publishers.

Logan, M., Martin, J. F., Nagy, A., Lobe, C., Olson, E. N., and Tabin, C. J. (2002). Expression of Cre Recombinase in the Developing Mouse Limb Bud Driven by a Prxl Enhancer. *Genesis* 33, 77–80. doi:10.1002/gene.10092

Lufkin, T., Mark, M., Hart, C. P., Dollé, P., LeMeur, M., and Chambon, P. (1992). Homeotic Transformation of the Occipital Bones of the Skull by Ectopic Expression of a Homeobox Gene. *Nature* 359 (6398), 835–841. doi:10.1038/359835a0

Luo, Z.-X., Ji, Q., and Yuan, C.-X. (2007). Convergent Dental Adaptations in Pseudo-tribosphenic and Tribosphenic Mammals. *Nature* 450, 93–97. doi:10.1038/nature06221

Madisen, L., Zwingman, T. A., Sunkin, S. M., Oh, S. W., Zariwala, H. A., Gu, H., et al. (2010). A Robust and High-Throughput Cre Reporting and Characterization System for the Whole Mouse Brain. *Nat. Neurosci.* 13 (1), 133–140. doi:10.1038/nn.2467

Mallo, M., Wellik, D. M., and Deschamps, J. (2010). Hox Genes and Regional Patterning of the Vertebrate Body Plan. *Developmental Biol.* 344 (1), 7–15. doi:10.1016/j.ydbio.2010.04.024

Matsuoka, T., Ahlberg, P. E., Kessaris, N., Iannarelli, P., Dennehy, U., Richardson, W. D., et al. (2005). Neural Crest Origins of the Neck and Shoulder. *Nature* 436 (7049), 347–355. doi:10.1038/nature03837

McGlinn, E., Holzman, M. A., and Mansfield, J. H. (2019). Detection of Gene and Protein Expression in Mouse Embryos and Tissue Sections. *Methods Mol. Biol.* 1920. doi:10.1007/978-1-4939-9009-2_12

McIntyre, D. C., Rakshit, S., Yallowitz, A. R., Loken, L., Jeannotte, L., Capecchi, M. R., et al. (2007). Hox Patterning of the Vertebrate Rib Cage. *Development* 134 (16), 2981–2989. doi:10.1242/dev.007567

Murakami, G., and Nakamura, H. (1991). Somites and the Pattern Formation of Trunk Muscles: A Study in Quail-Chick Chimera. *Arch. Histology Cytol.* 54 (3), 249–258. doi:10.1679/aohc.54.249

Nowicki, J. L., and Burke, A. C. (2000). Hox Genes and Morphological Identity: Axial versus Lateral Patterning in the Vertebrate Mesoderm. *Development* 127 (19), 4265–4275. doi:10.1242/dev.127.19.4265

Philippidou, P., Walsh, C. M., Aubin, J., Jeannotte, L., and Dasen, J. S. (2012). Sustained Hox5 Gene Activity Is Required for Respiratory Motor Neuron Development. *Nat. Neurosci.* 15 (12), 1636–1644. doi:10.1038/nn.3242

Prummel, K. D., Nieuwenhuize, S., and Mosimann, C. (2020). The Lateral Plate Mesoderm, *Development* 147, 12. doi:10.1242/dev.175059

Ramírez-Solis, R., Zheng, H., Whiting, J., Krumlauf, R., and Bradley, A. (1993). Hoxb-4 (Hox-2.6) Mutant Mice Show Homeotic Transformation of a Cervical Vertebra and Defects in the Closure of the Sternal Rudiments. *Cell* 73 (2), 279–294. doi:10.1016/0092-8674(93)90229-J

Rancourt, D. E., Tsuzuki, T., and Capecchi, M. R. (1995). Genetic Interaction between Hoxb-5 and Hoxb-6 Is Revealed by Nonallelic Noncomplementation. *Genes Dev.* 9 (1), 108–122. doi:10.1101/gad.9.1.108

Rodríguez-Vázquez, J. F., Verdugo-López, S., Garrido, J. M., Murakami, G., and Kim, J. H. (20132007). Morphogenesis of the Manubrium of Sternum in Human Embryos: a New Concept. *Anat. Rec.* 296 (2), 279–289. doi:10.1002/ar.22623

Scaal, M. (2021). Development of the Amniote Ventrolateral Body wall. *Developmental Dyn.* 250 (1), 39–59. doi:10.1002/dvdy.193

Sefton, E. M., and Kardon, G. (2019). Connecting Muscle Development, Birth Defects, and Evolution: An Essential Role for Muscle Connective Tissue. *Curr. Top. Developmental Biol.* 132, 137–176. doi:10.1016/bs.ctdb.2018.12.004

Shearman, R. M., and Burke, A. C. (2009). The Lateral Somitic Frontier in Ontogeny and Phylogeny. *J. Exp. Zool.* 312B (6), 603–612. doi:10.1002/jez.b.21246

Tabariès, S., Lemieux, M., Aubin, J., and Jeannotte, L. (2007). Comparative Analysis ofHoxa5 Allelic Series. *Genesis* 45 (4), 218–228. doi:10.1002/dvg.20292

Tanaka, M., and Onimaru, K. (2012). Acquisition of the Paired Fins: a View from the Sequential Evolution of the Lateral Plate Mesoderm. *Evol. Development* 14 (5), 412–420. doi:10.1111/j.1525-142X.2012.00561.x

Willey, A. (1894). *Amphioxus and the Ancestry of the Vertebrates*. New York: Macmillan. doi:10.5962/bhl.title.55924

Yahya, I., Morosan-Puopolo, G., and Brand-Saberi, B. (2020). The CXCR4/SDF-1 Axis in the Development of Facial Expression and Non-somitic Neck Muscles. *Front. Cel Dev. Biol.* 8, 1674. doi:10.3389/FCELL.2020.615264

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Cell and Developmental Biology

REVIEW
published: 13 May 2022
doi: 10.3389/fcell.2022.798812

Diversification and Functional Evolution of HOX Proteins

Narendra Pratap Singh[1†] *and Robb Krumlauf*[1,2*]

[1]*Stowers Institute for Medical Research, Kansas City, MO, United States,* [2]*Department of Anatomy and Cell Biology, Kansas University Medical Center, Kansas City, KS, United States*

Gene duplication and divergence is a major contributor to the generation of morphological diversity and the emergence of novel features in vertebrates during evolution. The availability of sequenced genomes has facilitated our understanding of the evolution of genes and regulatory elements. However, progress in understanding conservation and divergence in the function of proteins has been slow and mainly assessed by comparing protein sequences in combination with *in vitro* analyses. These approaches help to classify proteins into different families and sub-families, such as distinct types of transcription factors, but how protein function varies within a gene family is less well understood. Some studies have explored the functional evolution of closely related proteins and important insights have begun to emerge. In this review, we will provide a general overview of gene duplication and functional divergence and then focus on the functional evolution of HOX proteins to illustrate evolutionary changes underlying diversification and their role in animal evolution.

Keywords: gene duplication and divergence, protein evolution, HOX proteins, *Drosophila*, mouse

OPEN ACCESS

Edited by:
Edwina McGlinn,
Monash University, Australia

Reviewed by:
René Rezsohazy,
Catholic University of Louvain, Belgium
Jennifer Mansfield,
Columbia University, United States

****Correspondence:***
Robb Krumlauf
rek@stowers.org

†Present address:
Narendra Pratap Singh,
Laboratory of Systems Genetics, Systems Biology Center, National Heart Lung and Blood Institute, Bethesda, MD, United States

Specialty section:
This article was submitted to Morphogenesis and Patterning, a section of the journal Frontiers in Cell and Developmental Biology

Received: *20 October 2021*
Accepted: *08 April 2022*
Published: *13 May 2022*

Citation:
Singh NP and Krumlauf R (2022) Diversification and Functional Evolution of HOX Proteins. Front. Cell Dev. Biol. 10:798812. doi: 10.3389/fcell.2022.798812

INTRODUCTION

Evolution has brought an incredible range of morphological and physiological novelties to diverse animals. Centuries of classical research has served to catalog diverse novelties in 1.2 million species and sub-divide them into ~36 phyla and 107 classes, 500 orders, 5500 families and 110000 genera (Mora et al., 2011). These efforts have uncovered the emergence of novelties during the progressive evolution of animals, but we know relatively little about the genetic and genomic changes and mechanisms that underlie this diversity. Technological advances which enabled the systematic sequencing of animal genomes has reenergized this field of research and provided an opportunity for comparative genomics of the diverse animals to probe the underlying genetic causes of their morphological and physiological differences (Rogers and Gibbs, 2014). These genome-wide analyses have highlighted common origins and similar physiological functions but have found it challenging to uncover the genetic changes and mechanisms that underlie animal diversity. Comparative genomic analyses reveal a very similar number of genes in diverse animals (**Figure 1**), indicating that the total gene number does not reflect diversity (Hahn and Wray, 2002; Copley, 2008). Furthermore, many of the same genes and gene families are present in a broad range of animal species, suggesting there is a shared or common "gene toolkit."

The discovery of a very similar number of genes and a common "gene toolkit" in animal genomes appeared to refute the hypothesis that gene duplication and diversification is a major contributor to animal diversity. This led to a shift in the focus of research from analyses of coding regions to identifying and characterizing diversification of *cis*-regulatory (non-coding) regions and gene regulatory networks embedded in the genome (Biemont and Vieira, 2006). A broad array of

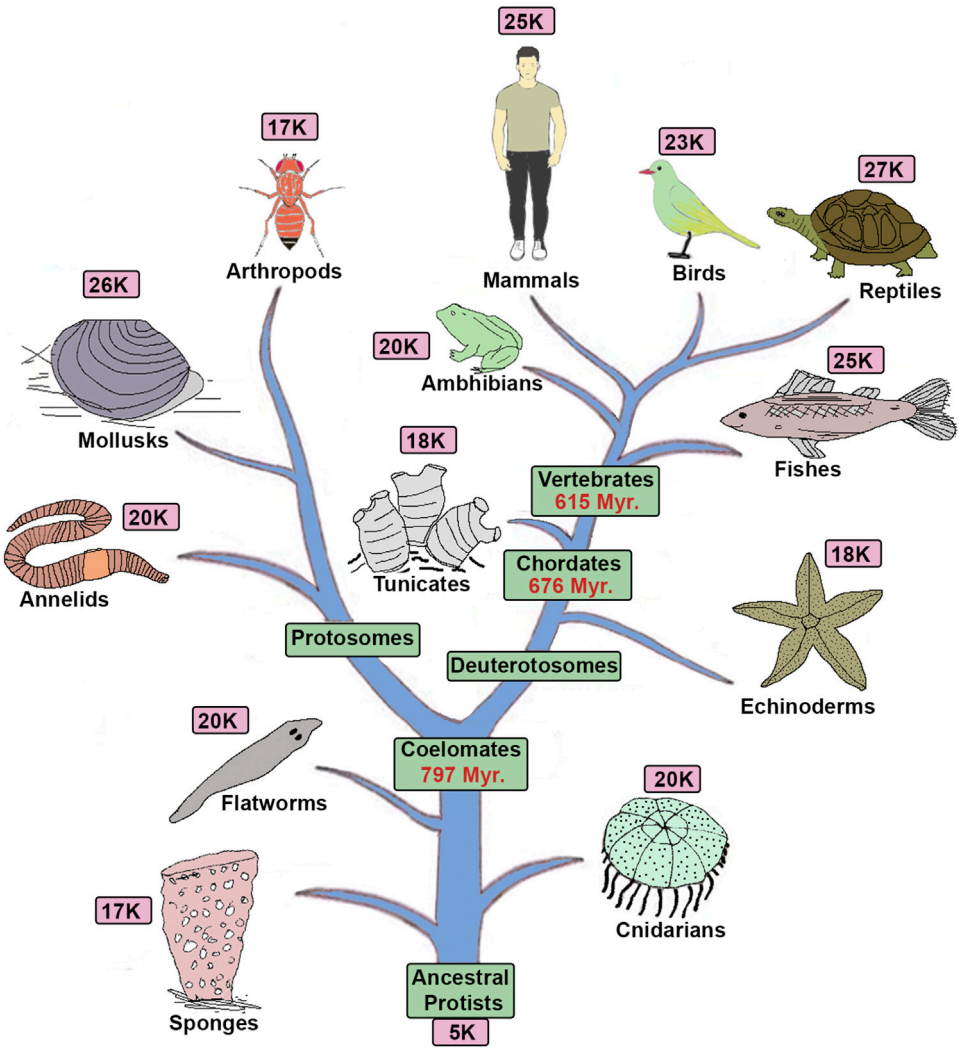


FIGURE 1 | Tree of life and gene complexity: The diagram shows a phylogenetic tree of animals with the approximate total number of genes present in each genome based on sequencing of their representative genomes (Hedges et al., 2015).

technological advances have also revolutionized this field and enhanced our ability to identify and functionally validate the *cis*-regulatory code embedded in the genome by the integration of comparative genomics, transgenic analyses, CRISPR/Cas9 genome modifications and genome-wide approaches (i.e., ChIP-seq, ATAC-seq, massive parallel reporter assays and single cell transcriptomics) (Zhen and Andolfatto, 2012; Paul et al., 2014; He et al., 2015; Gasperskaja and Kucinskas, 2017; Avsec et al., 2021). Application of these approaches, has revealed that diversity in non-coding *cis*-regulatory regions of the genome has played a major role in the emergence of animal diversity (Carroll, 2008; Rubinstein and de Souza, 2013; Reilly and Noonan, 2016; Franchini and Pollard, 2017; Xie et al., 2019; Roberts Kingman et al., 2021a; Roberts Kingman et al., 2021b).

Evaluating how changes in protein sequence impact *in vivo* function following gene duplications in animal evolution has been challenging to investigate in the absence of technologies for precise manipulation of endogenous genes and quantitative and qualitative functional assays to evaluate activity in an *in vivo* context. The development of CRISPER/Cas9 gene editing technology for precise manipulation of endogenous genes in the genome of diverse animals has opened the door for more direct cross-species comparisons of homologous protein functions (Doudna and Charpentier, 2014). This approach has the advantage of expressing the proteins being compared at the same physiological levels in their normal spatial, temporal and tissue-specific contexts under control of regulatory components of the endogenous loci. In the past, functional studies have primarily relied on comparing the degree of conservation in amino acid sequences, *in silico* structure predictions, *in vitro* assays for activity and ectopic over-expression assays *in vivo*. The *in vitro* assays of protein activity, such as ligand binding, enzymatic activity, transcription factor binding properties, can be limited by *ex-vivo* conditions, which often lack key co-factors

or components important *in vivo*. Hence, they may provide a limited perspective on a subset of functional activities relevant to their *in vivo* roles. To overcome some of the limitations of *in vitro* assays, transgenic approaches have been used to ectopically express genes in the animals and compare the *in vivo* properties of candidate proteins (Mcginnis et al., 1990; Quiring et al., 1994; Hanks et al., 1998). However, this approach often involved broad over-expression of proteins at high levels and ectopic sites, making it difficult to compare activities in normal physiological and developmental contexts. CRISPR/Cas9 technologies now offer possibilities to manipulate the genome to precisely compare function of homologous genes *in vivo*. As a result, cross-species analyses of gene function are beginning to uncover unexpected changes and mechanisms that contribute to conservation and divergence of protein functions contributing to animal diversity (Enard et al., 2009; Truong and Boeke, 2017; Laurent et al., 2020; Singh et al., 2020). In addition, recent advances in cryo-EM and computational approaches for predicting protein structures are rapidly changing our ability to analyze and compare the properties of proteins (Assaiya et al., 2021; Jumper et al., 2021).

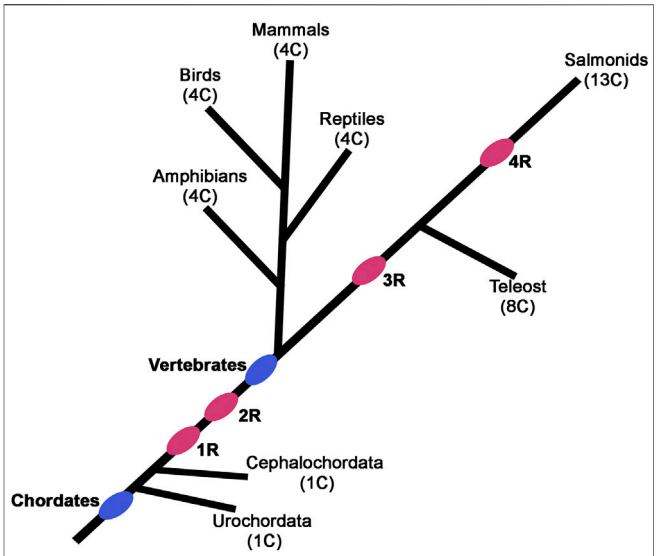


FIGURE 2 | Phylogenetic tree depicting the position of postulated whole-genome duplication (R) events during vertebrate evolution. The pink ovals indicate the progressive rounds of whole-genome duplication. The number of Hox clusters (C) in the different species is indicated in parentheses.

GENE DUPLICATION AND DIVERGENCE

Sequence analyses have revealed a high level of conservation of many domains in proteins with very diverse functions across the animal phylum (Laity et al., 2001; Ponting and Russell, 2002; Noyes et al., 2008). This implies that during evolution, the appearance of novel functions is not associated with widespread *de novo* evolution of new genes and that novel functional activities most likely arose by diversification of existing genes (Holland et al., 1994; Hughes, 1994; Friedman and Hughes, 2001; Blomme et al., 2006). Altering the function of an essential gene could be detrimental to the survival or fitness of a species, but gene duplication events provide a mechanism to circumvent this limitation. Generating multiple copies of a gene provides a range of opportunities to maintain essential functions, releasing selective pressure on a single essential gene, while also producing new substrates that can diversify and evolve novel functions. Analyses of gene sets across the animal kingdom revealed that vertebrate genomes have multiple copies of many invertebrate genes, including those that regulate development, differentiation and physiological processes, such as transcription factors, cell signaling pathways, odorant receptor genes etc. (Paps and Holland, 2018; Richter et al., 2018; Fernandez and Gabaldon, 2020). Large genome duplication events followed by gene losses are considered as a critical step in the emergence and evolution of vertebrates. Susumu Ohno suggested that two rounds of whole-genome duplications (2R-WGDs) could be a major source of gene amplification and functional diversification in vertebrate lineage (Ohno, 1970; Panopoulou et al., 2003; Vandepoele et al., 2004; Dehal and Boore, 2005). While this hypothesis was used to explain a major cause of gene duplications, comparing all homologous gene families between *Drosophila* and humans showed that less than 5% of these families display a predicted 1:4 gene ratio (Friedman and Hughes, 2001). This was not consistent with Ohno's hypothesis and lead to an alternative idea, suggesting that selective regional duplications (segmental duplications) created a mixed repertoire of duplicated genes in the genome.

Despite these conflicting models, which remain challenging to resolve, there is evidence for whole genome duplications in many plants, yeast, and vertebrate species, such as teleost fish (3R), salmonid fish (4R) and *Xenopus laevis* (Wendel, 2000; De Bodt et al., 2005; Scannell et al., 2006; Pascual-Anaya et al., 2013; Session et al., 2016) (**Figure 2**). Advocates of the 2R hypothesis have further refined the model to suggest that during vertebrate evolution frequent gene loss after duplication, as a consequence of redundancy, contributed to the observed digression from the expected 1:4 gene ratio. Analyses of many known genome duplications indicate that gene loss is the most common fate of duplicated genes (Nadeau and Sankoff, 1997; Albalat and Canestro, 2016). Loss of many duplicated genes and nonessential genes could also explain how a similar number of total genes are present in higher animal genomes despite differences in the whole genome or segmental duplication events. Despite the controversy on the underlying mechanisms for widespread amplification of gene families (Braasch et al., 2018; Sandve et al., 2018), the importance of gene duplication in creating a major substrate for functional divergence and emergence of novel functions is widely accepted in the field (Lynch and Conery, 2000).

FATE OF THE DUPLICATED GENES

The contribution of gene duplication and divergence to the emergence of novel protein functions in animal evolution has gained more traction as gene functions have been characterized in different animals across the phylogenetic spectrum (Blomme

et al., 2006). Comparing classes of genes based on their functions revealed that genes involved in animal development and cell signaling pathways are highly amplified during the evolution, while many other classes of genes have been lost. Based on these analyses several hypotheses came to explain the forces that shaped the future of a duplicated genes. Loss of function is the most common fate of the duplicated genes. Studies have shown that only half of duplicated genes are retained while others lose functional activities by processes that include deletion, rearrangement, point mutations, and pseudogene formation (Nadeau and Sankoff, 1997; Blomme et al., 2006; Scannell et al., 2006; Albalat and Canestro, 2016; Guijarro-Clarke et al., 2020a; b). Even in the absence of duplicated genes, a large number of genes (90% in bacteria, 80% in yeast, 65% in *C. elegans* and 85% in *Drosophila*) are dispensable for animal survival, which provides a large set of substrates for evolutionary change. Gene duplication events further expand the repertoire of substrates for change and generate opportunities for functional redundancy that allows for non-deleterious functional diversification of genes.

Essential genes and their functional roles need to be retained to ensure animal survival and fitness. Hence, at least one member of a duplicated gene family must retain the key ancestral functions during evolution. Other members are free to accumulate mutations that potentiates diversification and the emergence of novel activities, which is called neofunctionalization (Ohno, 1970; Clark, 1994; Holland et al., 1994; Lundin, 1999; Friedman and Hughes, 2001; Mazet and Shimeld, 2002; Sandve et al., 2018). A variation of this idea is that the ancestral functions of essential genes maybe collectively retained by partitioning sub-sets of the functional roles between different duplicated family members, which is termed subfunctionalization (Ohno, 1970; Force et al., 1999; Lynch and Force, 2000; Sandve et al., 2018). These are not mutually exclusive processes. A study by Le and Zhang demonstrated that neofunctionalization or subfunctionalization alone do not adequately explain the diversification of protein function. They proposed that many duplicated genes may go through a combination of subfunctionalization and neofunctionalization to produce duplicated genes that possess new and retain some ancestral roles (He and Zhang, 2005; Marcussen et al., 2010). Various theoretical and functional studies have explored these ideas and confirmed that the retention of duplicated genes appears to be mediated by a varying combination of these processes (Force et al., 1999; Lynch and Force, 2000; Vandenbussche et al., 2003; Walsh, 2003; Burki and Kaessmann, 2004; Escriva et al., 2006; Perry et al., 2007; Kleinjan et al., 2008; Innan, 2009; Truong and Boeke, 2017; Zimmer et al., 2018; Singh et al., 2020).

HOX GENES AS A PARADIGM FOR DUPLICATION AND DIVERGENCE OF FUNCTION

Hox genes, encode a broadly conserved family of transcription factors in animals, and represent an interesting paradigm for examining the duplication and divergence of gene functions. The HOX proteins are involved in patterning and specification of the anterior-posterior (AP) axis of all bilaterian animals (Mcginnis and Krumlauf, 1992; Krumlauf, 1994; Carroll, 1995; Pearson et al., 2005). The temporal and spatial order of *Hox* gene expression and function across the embryo is "colinear" and correlated with their organization along the chromosome (Lewis, 1978; Duboule and Dolle, 1989; Graham et al., 1989; Duboule, 1998; Kmita and Duboule, 2003). These genes are typically found to be tightly clustered in the genome except for some animals where evolution has led to the disintegration of the ancestral complex (Kaufman et al., 1980; Akam et al., 1994; Seo et al., 2004; Sekigami et al., 2017). Each gene in a cluster specifies distinct cellular identities along AP axis during very early embryonic development, which ultimately patterns tissues and structures in adult animals. Evidence of *Hox* genes in animal genomes is traced back to Cnidarians, however, their role in patterning the AP axis is observed only in bilaterians, as they have roles in patterning radial segmentation in cnidarians (Pascual-Anaya et al., 2013; Arendt, 2018; He et al., 2018). Mutations that affect the expression and function of *Hox* genes in bilaterians lead to homeotic transformation of one part of the body into another (Lewis, 1994). Furthermore, diversification of *Hox* gene number and function correlates with increased diversity in the evolution of animals (Wagner et al., 2003; Lemons and Mcginnis, 2006). There are fewer *Hox* genes in lower invertebrates as compared to higher invertebrates, chordates, and vertebrates, as illustrated by the 5 *Hox* genes in nematodes (*C. elegans*), 8 in arthropods (*Drosophila*), 14 in chordates, and 39 in mammals (Human) (Ikuta, 2011; Pascual-Anaya et al., 2013; Irie et al., 2018) (**Figure 3A**).

Despite their diverse and distinct functional activities, sequence analysis indicates that *Hox* gene cluster was generated by tandem duplication of a single gene. Analysis of *Hox* gene clusters in invertebrate genomes suggests that they expanded from a common ancestral gene through tandem regional duplication events to form a maximum of fifteen alloparalogs in cephalochordates (ex. *Branchiostoma floridae)* (Garcia-Fernandez and Holland, 1994; Koonin, 2005). There are four Hox clusters in mammalian genomes with a maximum of 14 genes in each cluster suggesting two rounds of whole complex duplication (2R) from a common invertebrate ancestor with 14 genes (Maconochie et al., 1996; Hoegg and Meyer, 2005; Lemons and Mcginnis, 2006; Duboule, 2007; Kuraku and Meyer, 2009; Pascual-Anaya et al., 2013; Holland and Ocampo Daza, 2018; Smith et al., 2018). However, there is a maximum of 13 *Hox* genes in each cluster of tetrapod genomes, such as mouse and human, indicating a loss of 14th paralog during vertebrate diversification (**Figure 3A**). There is also evidence for two additional lineage specific whole-genome duplications events in vertebrates, one in teleosts (3R) (Pascual-Anaya et al., 2013) and an additional round (4R) in salmonid lineages to further amplify HOX genes (**Figure 2**) (Soshnikova et al., 2013; Vieux-

Rochas et al., 2013). These data suggest that genome duplication events have dramatically increased the number of *Hox* genes in vertebrate genomes and provide opportunities for evolution of novel functions. Further, functional compensation or redundancy after gene duplication events allowed for loss and diversification of these genes which have a critical role in patterning the AP axis and the properties of tissues in a manner that has been remarkably conserved across the bilateral animals (Wagner et al., 2003).

CONSERVATION AND DIVERSIFICATION OF *HOX* GENE FUNCTION

The correlation between the expansion of *Hox* genes, a master regulator of development, and greater complexity of vertebrates highlights the importance of studying the functional evolution of HOX proteins. There are reports to suggest both conservation and diversification of *Hox* gene function through evolution (Lawrence and Morata, 1994; Saurin et al., 2018). Early work

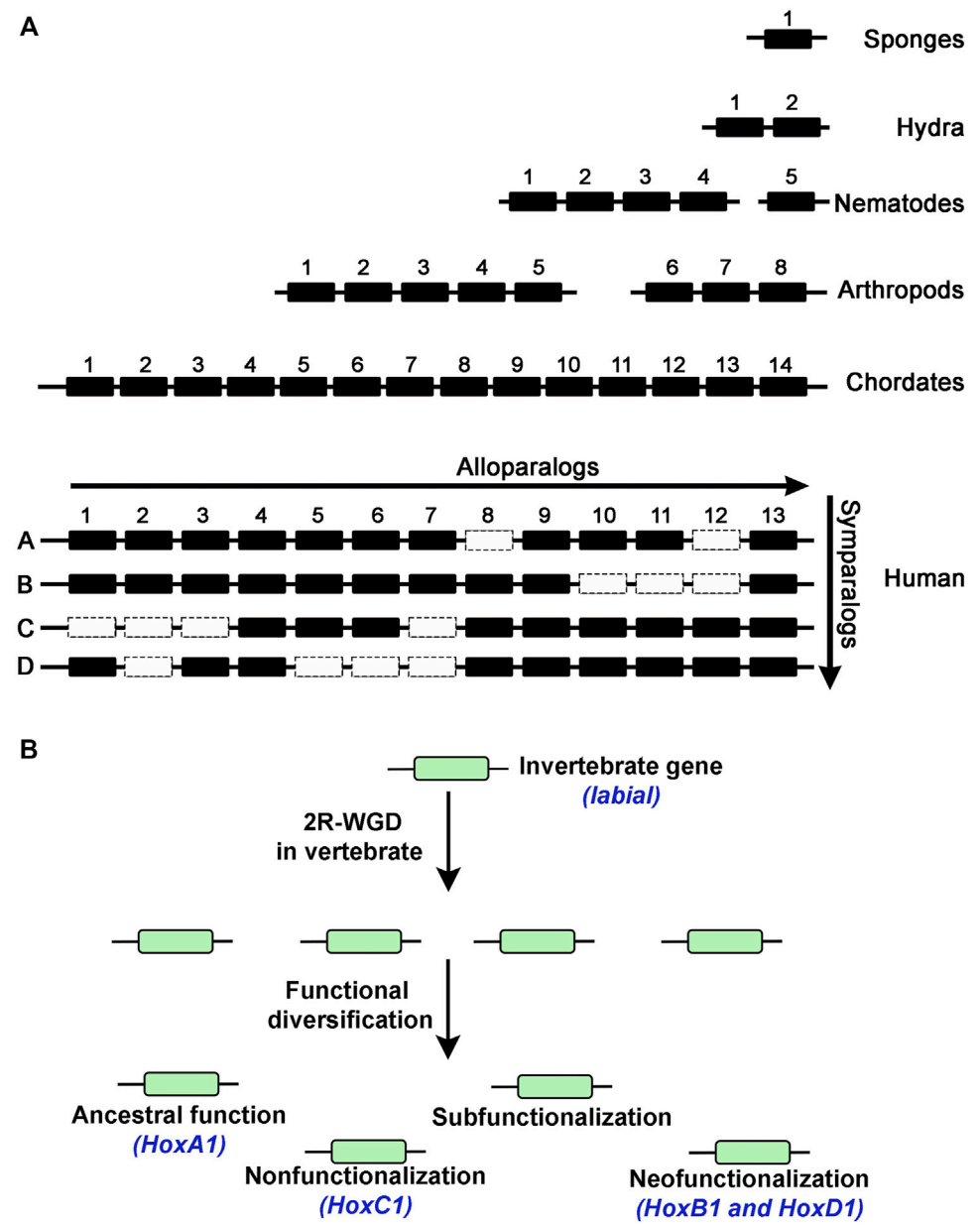


FIGURE 3 | (A) Diagram shows number of *Hox* genes in different animals indicating expansion of the Hox cluster into alloparalogs. Duplication events in vertebrates generated 4 Hox clusters to create symparalogs, Human genome has 39 *Hox* genes suggesting loss of some genes (dotted box) during evolution. **(B)** Possibilities of functional diversification of duplicated genes in the vertebrate genome after duplication. An example of the invertebrate *labial* gene from *Drosophila*, duplication in vertebrates and functional diversification of the homologous *Hox1* genes (blue text).

in the field has demonstrated that HOX proteins from different alloparalogs groups defines distinct morphological features along the AP body axis implying that each regulates a distinct set of gene regulatory networks and developmental pathways (Garcia-Bellido et al., 1973; Mcginnis and Krumlauf, 1992; Lawrence and Morata, 1994; Averof, 2002; Maeda and Karch, 2006). However, HOX proteins appear to share many similar biochemical properties. They display nearly identical DNA binding specificities *in vitro*, as a consequence of the presence of a highly conserved 60 amino acid long homeodomain (HD) (Scott et al., 1989; Gehring et al., 1994a; Piper et al., 1999). HOX proteins have the ability to interact with a PBC class (e.g., PBX, MEIS, PREP) of homeodomain transcription factors, which serve as cofactors in binding DNA and modulating gene expression (Desplan et al., 1988; Gehring et al., 1994a; Knoepfler and Kamps, 1995; Mann and Chan, 1996; Noyes et al., 2008). The interaction of HOX proteins with PBC factors is primarily mediated by a conserved six amino acid domain, referred to as the hexapeptide (HP). HOX-PBX cofactor interactions modify the affinity and specificity of HOX DNA binding on target sites in the genome (Slattery et al., 2011; Merabet and Mann, 2016). These generic DNA binding properties of HOX proteins and interactions with shared cofactors, such as PBX, make it difficult to explain the paralog-specific functions of HOX proteins. Hence, the distinct differences in the functional roles and genome-wide binding preferences of HOX proteins *in vivo* is likely to be a consequence of additional unidentified features and interactions of HOX proteins that impact and modulate their context-dependent activities.

Conservation

Early studies on *Hox* genes explored the evolutionary conservation of functions among homologous genes across the animal evolution by using transgenesis to express vertebrate *Hox* genes in *Drosophila* (Malicki et al., 1990; Mcginnis et al., 1990; Zhao et al., 1993; Lutz et al., 1996). These studies uncovered deep conservation of *Hox* gene function in specifying regional identity along the AP body axis over 600 million years of animal evolution. Protein sequences of alloparalogs, formed by tandem duplication of the ancestral *Hox* genes, show high diversification and each of them are known to drive distinct developmental and differentiation programs to regulate regional identities in specific tissues along the AP axis (Lamka et al., 1992; Mclain et al., 1992; Yokouchi et al., 1995; Carapuco et al., 2005). Despite these differences in functional roles, many alloparalogs show redundancy or overlaps of function in regulating developmental events in some specific tissues (Saurin et al., 2018). For example, it has been observed that ubiquitous expression of many *Hox* genes is important in *Drosophila* larval fat body cells, which appears distinctly different from their roles along the AP axis facilitated by their nested collinear expression patterns in embryos. Furthermore, functional analyses in autophagy inhibition uncovered no paralog specificity and suggested redundant functions of many HOX proteins (Banreti et al., 2014). Similarly, *Drosophila Hox* genes have been also shown to have redundant activity in specification of tritocerebrum identity, endocrine ring gland development, dorsal DA3 muscle lineage specification and head repression etc. (Hirth et al., 2001; Coiffier et al., 2008; Enriquez et al., 2010; Sanchez-Higueras et al., 2014). The redundant role of *Hox* genes is not limited to *Drosophila*, vertebrate HOX proteins have also been shown to have overlapping or redundant functions during development of several tissue types (Young et al., 2009; Lacombe et al., 2013; Denans et al., 2015). For example, vertebrate *Hox6* paralogs (*Hoxa6*, *Hoxc6*, and *Hoxb6*) are required to specify lateral motor column motoneurons and functional studies displayed that Hox paralogs 5, 7, and 8 can all substitute for this function (Lacombe et al., 2013). These observations suggest that despite evidence for sequence and functional diversification among HOX alloparalogs, which underlies their distinct roles in axial patterning, some aspects of their functional activities have been conserved during evolution and play roles in specific tissue contexts during development.

As expected, the functional redundancy among the *Hox* genes is more common among the symparalogs formed more recently after whole cluster duplications in vertebrate lineage (**Figure 3A**). Deletion of a single gene or even a whole cluster does not show dramatic consequences on embryonic development, consistent with the idea of extensive functional redundancy between *Hox* genes (Medina-Martinez et al., 2000; Suemori and Noguchi, 2000; Spitz et al., 2001; van Den Akker et al., 2001; Hunter and Prince, 2002; Soshnikova et al., 2013). Gene swap experiments in mouse models have also demonstrated that symparalogs, formed after duplication of ancestral invertebrate cluster, are functionally equivalent (Horan et al., 1995; Manley and Capecchi, 1997; Greer et al., 2000; Tvrdik and Capecchi, 2006; Iacovino et al., 2009). An interesting example is a gene swap of mouse *Hoxa3* and *Hoxd3* that resulted in adult mice with no detectable developmental defects (Greer et al., 2000). These observations suggests that HOX paralogs have retained similarity in their activity through millions of years of evolution, which has been attributed to conservation of DNA binding properties of the homeodomain and a shared hexapeptide domain that mediates interaction with PBC factors.

Diversification

Over expression studies have provided evidence that Hox alloparalogs regulate development of specific organs across the body axis, indicating that they can drive distinct gene regulatory networks (Schneuwly et al., 1987; Lamka et al., 1992; Mclain et al., 1992; Yokouchi et al., 1995; Carapuco et al., 2005). Comparative analyses of gene expression profiles upon ubiquitous expression of *Drosophila Hox* genes show a very small number (1.3%) of common changes, and most of changes are unique to each gene, suggesting they individually regulate distinct set of targets (Hueber et al., 2007). This data implies that despite the high conservation of homeodomain region, HOX proteins regulate distinct set of genes to specify unique cellular identities across the AP body axis during animal development (Mcginnis and Krumlauf, 1992; Averof and Akam, 1993; Lawrence and Morata, 1994; Patterson et al., 2001; Averof, 2002). The functional diversity among the HOX proteins may be

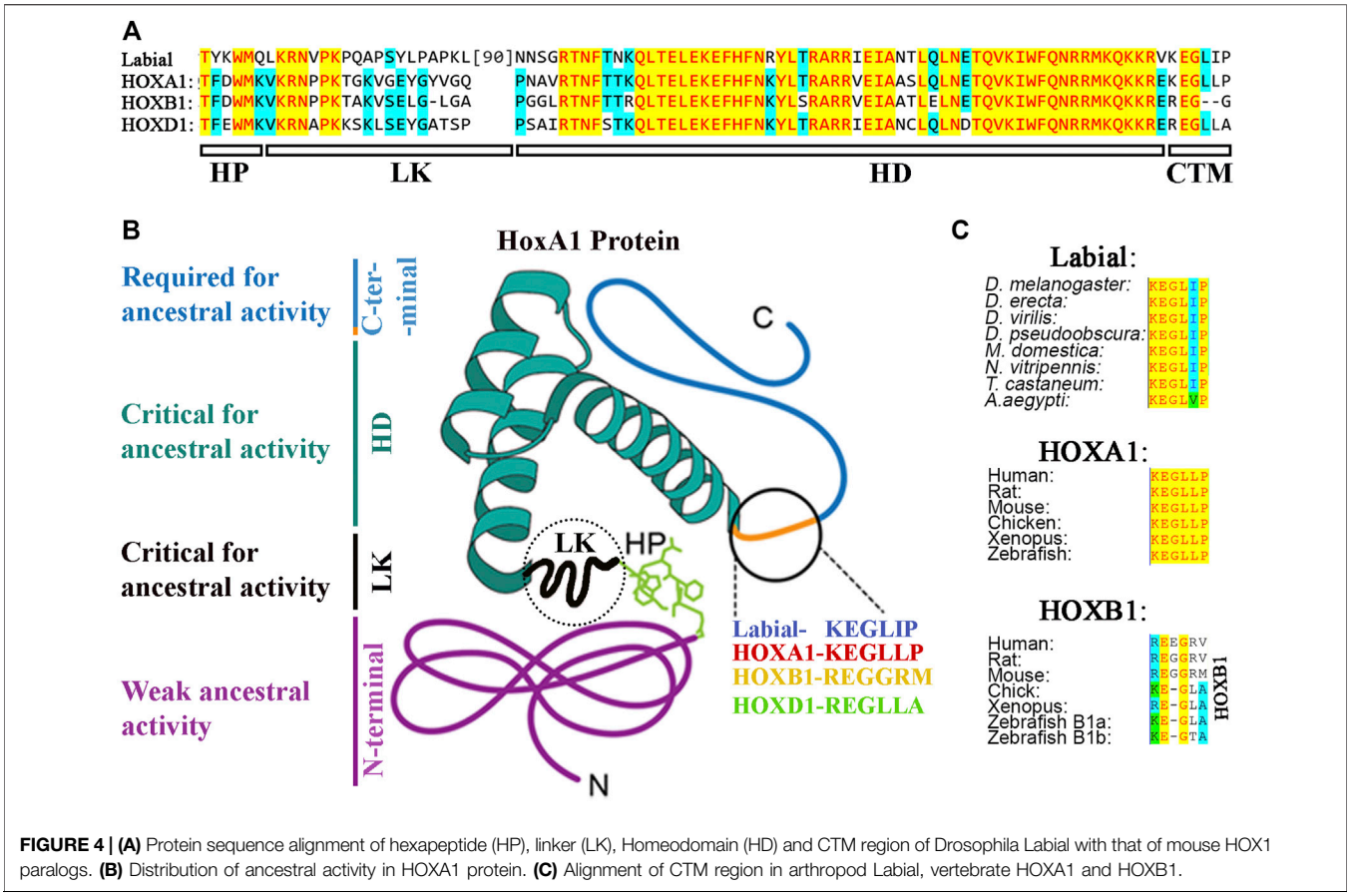


FIGURE 4 | (A) Protein sequence alignment of hexapeptide (HP), linker (LK), Homeodomain (HD) and CTM region of Drosophila Labial with that of mouse HOX1 paralogs. **(B)** Distribution of ancestral activity in HOXA1 protein. **(C)** Alignment of CTM region in arthropod Labial, vertebrate HOXA1 and HOXB1.

generated in multiple ways. Domain swaps of the HD, the most conserved region of HOX proteins, can result in functionally distinct activity in some developmental contexts. This suggests that even a small number of changes can lead to diversification of the DNA binding properties and transcriptional activity of HOX proteins (Zhao and Potter, 2001; 2002). In addition, studies have shown that amino acid differences among the Hox alloparalogs may not alter DNA binding preference but change their ability to recruit different coactivators or corepressors (Li and Mcginnis, 1999; Gebelein et al., 2004; Joshi et al., 2010). Another mechanism for generating diversity among HOX proteins is through its interaction with PBC group of cofactors, which alters DNA binding specificity. In fact, a high throughput study revealed that DNA binding specificities of HOX-Exd complex are only revealed upon heterodimerization (Slattery et al., 2011). These observations suggest that HOX proteins are subject to a variety of ways that can diversify or modulate their functional properties.

As discussed earlier, there are many examples of functional redundancy among the HOX symparalogs (Horan et al., 1995; Chen and Capecchi, 1997; Greer et al., 2000; Wahba et al., 2001; Tvrdik and Capecchi, 2006; Iacovino et al., 2009). However, several studies have also found functional diversification among symparalogs (Fromental-Ramain et al., 1996; Miguez et al., 2012; Singh et al., 2020; Singh et al., 2021). Analyses of *Hoxa9* and *Hoxd9* mutants in mouse revealed that these two symparalogous genes have both specific and redundant functions in lumbosacral axial skeleton patterning and in limb morphogenesis (Fromental-Ramain et al., 1996). Similarly, *Hoxa2* and *Hoxb2* symparalogs show synergistic interactions in regulation of gene expression in the hindbrain, however, during oligodendrogenesis in the mouse hindbrain *Hoxb2* antagonizes *Hoxa2* function (Davenne et al., 1999; Miguez et al., 2012). These analyses indicate that symparalogs have retained a lot of overlapping functions during evolution, but they have also diversified their functional roles thought changes in the patterns of expression and protein structure. An important understudied question in the field is how changes in amino acid sequences of symparalogous after duplication from the ancestor homolog relate to altered functions. Analyses of HOX1 proteins in Xenopus suggests they have redundant roles of HOXA1, B1 and D1 in hindbrain development (Mcnulty et al., 2005). Gene swaps of HoxA1 and HoxB1 in mouse also suggest they have largely overlapping or redundant roles (Tvrdik and Capecchi, 2006). However, in a recent study, we utilized CRISPR/Cas9 technology to replace the *Drosophila Hox* gene *labial* with its mouse homologs *HoxA1*, *HoxB1* and *HoxD1* to investigate conservation of ancestral functions and assess diversification during evolution (Singh et al., 2020). Despite similar degrees of protein sequence diversification of mouse HOX1 proteins from *Drosophila* Labial, our results revealed that among the HOX1 symparalogs (HOXA1, HOXB1 and HOXD1) only *HOXA1* is able to rescue *labial* function, as HOXB1 and HOXD1 failed to do

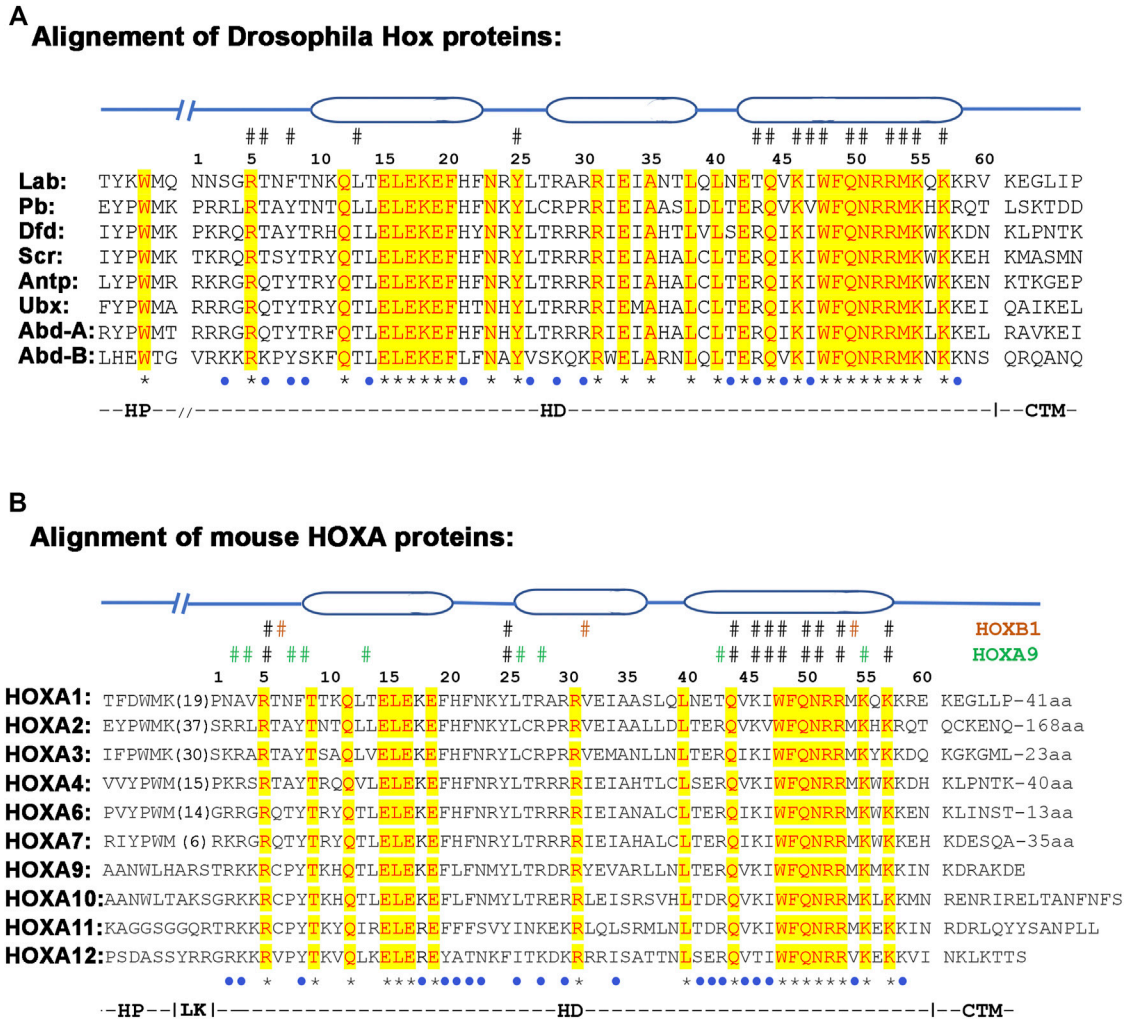


FIGURE 5 | Alignment of hexapeptide (HP), linker (LK), homeodomain (HD) and CTM regions from Drosophila **(A)** and mouse **(B)** HOX proteins. The three ovel regions on the top of the alignment shows the position of alpha-helices. The # sign above the alignment depict amino acids that contact DNA, while star (*) and blue dot under the alignment show conservative and semi conservative amino acids respectively.

so (**Figures 3B**, **4A**). This demonstrates a remarkable conservation of ancestral activity by HOXA1 and indicates that HOXB1 and HOXD1 have diversified through 600 million years of evolution (**Figure 3B**). Furthermore, consistent with their share ancestral activities, comparative genome-wide DNA binding properties revealed that HOXA1 and Labial have similar patterns of binding in mouse genome, while HOXB1 binds to a distinctly different set of targets. This adds support for neofunctionalization of HOXB1 by regulating a district gene regulatory program (**Figure 3B**) (Singh et al., 2020; Singh et al., 2021). Studies on mouse HOXD1 have shown that it has lost ancestral activities and appears to have undergone neofunctionalization through expression in new tissue types and altered activities in regulating novel gene regulatory programs (Guo et al., 2011). These distinct functional properties of mouse HOX1 symparalogs illustrate the diversification of HOXB1 and HOXD1 function and loss of ancestral activity in mammalian lineage. Mapping of the protein sequences that underlie functional diversification of HOXA1 and HOXB1 proteins revealed that a small number of changes across the protein cause functional diversification. It will be important to examine similar changes in other paralogy groups of HOX proteins to determine if this is a common means for modulating functional activities.

CRITICAL REGIONS UNDERLY HOX PROTEIN FUNCTION AND FUNCTIONAL DIVERSIFICATION

The deep conservation of HOX protein function across bilaterians and their redundant role in some tissue types suggest that the essential roles of HOX proteins in AP patterning has restricted their diversification. Preservation of

many of their ancestral functions is likely mediated through conservation of the homeodomain and hexapeptide regions. The homeodomain of *Drosophila* HOX proteins is known to bind directly to DNA (Galant et al., 2002; Merabet et al., 2003). However, on many *in vivo* target sites, binding affinity and specificity is enhanced by interaction with PBC factors (Loker et al., 2021). The regions outside these domains are highly diversified among the HOX proteins that can impact activation or repression of transcription of potential target genes (Li et al., 1999; Li and Mcginnis, 1999). In this section of the review, we will discuss evidences on what is known about various domains of HOX proteins and how they may have diversified to adopt novel functions during evolution.

Homeodomain

The HD stands out as the most conserved region of all HOX proteins (**Figure 5**). Three-dimensional structural studies using X-ray crystallography and NMR spectroscopy have revealed the presence of three alpha-helix regions in the HDs (Gehring et al., 1994b; Passner et al., 1999; Piper et al., 1999). The third helix, also known as the recognition helix, directly contacts the DNA through the major groove of DNA, while the region between the first and second helices establishes specific contact with the phosphate backbone. In addition, sequences adjacent to the N-terminus of the HD, referred to as the N-terminal extension region, also makes specific contact with DNA in the minor groove. These regions that insert into the minor groove have been shown to confer specificity to the HDs of *Drosophila* Ultrabithorax and Sex combs reduced Hox proteins imparting distinct DNA recognition properties (Joshi et al., 2007). The interactions of the N-terminal extension along with contacts mediated by the third helix are the primary determinates of DNA binding specificity and serve as a key constraint, maintaining the high level of sequence conservation.

Genetic and biochemical data suggest that the HD of HOX proteins is not freely interchangeable with each other. Functional assays of chimeric HOX proteins with swapped homeodomains of paralogs have been found to alter their functions, suggesting that homeodomains are not equivalent to each other (Zhao and Potter, 2001; Zhao and Potter, 2002). This indicates that the small differences between HD sequences contribute to differences in their DNA binding properties and functional activities (Furukubo-Tokunaga et al., 1993; Phelan et al., 1994; Zappavigna et al., 1994; Noyes et al., 2008; Breitinger et al., 2012). For example, 28 out of 60 amino acids are identical between all *Drosophila* HOX proteins and 18/60 are identical between all mouse HOXA paralogs (**Figures 5A,B**). In addition, many differences among the paralogs are conservative amino acid replacements that would be expected to preserve biochemical properties. This means that there are a series of non-conservative amino acid replacements or changes (19/60 in *Drosophila* and 22/60 in mouse HOXA paralogs), that have the potential to alter the DNA binding properties of HDs. For example, the crystal structures of HOXB1 and HOXA9 display many paralog specific amino acid contacts with DNA (**Figure 5B**) (Piper et al., 1999; LaRonde-LeBlanc and Wolberger, 2003). On the basis of sequence alignments and specific amino-acid residues within the homeodomain along with their relative positions in *Hox* clusters, *Hox* genes have been assigned to 14 different alloparalogous groups (PG1-14). These fall into three general classes: the anterior class contains PG1-5, the central class PG6-8 and the posterior class genes (PG9-14) (Domsch et al., 2015; Frobius and Funch, 2017). Hence, despite high conservation between homeodomain sequences there are paralog-specific differences that may alter DNA binding properties and preferences of HDs to modulate HOX function. The paralog specific changes in homeodomains may also alter the interaction of conserved amino acids with the DNA (Gruschus et al., 1997).

Hexapeptide

The PBC group of proteins, *Drosophila* Extradenticle (Exd) and Homothorax (Hth) proteins and their vertebrate orthologs Pre-B-cell leukemia transcription factor (PBX1, PBX 2, PBX3 and PBX4) and myeloid ectopic leukemia virus integration site (MEIS1, MEIS2 and MEIS3) respectively, are well characterized cofactors of HOX proteins (Chan et al., 1994; Chang et al., 1995; Lu et al., 1995; Mann, 1995; Ryoo et al., 1999). Interactions with PBC proteins change DNA binding affinity and specificity of HOX monomer proteins (Slattery et al., 2011). *In vitro* binding analyses have demonstrated that the interaction between most HOX proteins and PBX are highly dependent on the HP motif, which contains the four core residues YPWM. This interaction alters the DNA binding properties of each protein, leading to the recognition of a bipartite HOX-PBC consensus site, and it enhances affinity and specificity of HOX target selection (van Dijk and Murre, 1994; Chang et al., 1995; Phelan et al., 1995; Popperl et al., 1995; Chan and Mann, 1996; Hudry et al., 2012).

The presence of the HP is important for HOX function, as loss of this motif has been implicated in the eliminating homeotic functions in several genes (*fushi-tarazu*, *zerknüllt*, *bicoid*) formed after duplication of *Hox* genes (Alonso et al., 2001; Panfilio and Akam, 2007). There is also evidence for sequence diversity among HP regions of different HOX proteins, which may impact the DNA binding properties and function of HOX proteins (Chang et al., 1995; Neuteboom et al., 1995; Medina-Martinez and Ramirez-Solis, 2003; Remacle et al., 2004). While the four core amino acids in the HP motif have been shown to be critical for its activity, only the tryptophan (W) residue in the 4^{th} position is highly conserved. The rest of the amino acids are highly variable among the HOX protein paralogs, which may alter their interactions with PBX (**Figures 5A,B**) (Neuteboom et al., 1995; Neuteboom and Murre, 1997; Mann et al., 2009). The divergence of the HP region away from YPWM motif in HOX proteins from paralogous groups 8–13 suggests a progressive change in their function through evolution of this motif (**Figure 5B**).

Consistent with the divergence of the HP regions, *in vitro* DNA binding experiments show that HOX proteins from paralog groups 1–10 mainly interact with PBX, while paralogy groups 11-13 preferentially interact with MEIS (Chan et al., 1994; Shen W. F. et al., 1997; Shen W.-F. et al., 1997). HOX proteins also form trimeric complex with PBX and MEIS proteins upon binding to target genes (Berthelsen et al., 1998; Shanmugam et al., 1999;

Ferretti et al., 2000; Penkov et al., 2013; Amin et al., 2015). The presence of MEIS proteins have also been shown to induce remodeling of HOX-PBX interactions that leads to changes in the requirements for motifs that drive trimeric complex formation (Dard et al., 2019). High throughput SELEX-seq technology, which measures the relative affinities of transcription factor complexes with all possible DNA sequences, has shown that interaction of HOX proteins with Exd-Hth dimers alters the DNA binding properties of all eight *Drosophila* HOX proteins. Based on this data, the DNA binding specificities of HOX proteins can be subdivided into three main classes with similar DNA binding preferences: 1) anterior, containing Labial and Proboscipedia; 2) middle with Deformed and Sex comb reduced; and posterior, with Antennapedia, Ultrabithorax, Abdominal-A and Abdominal-B. These observations are consistent with the idea that the nature of interactions with PBX and MEIS varies among different HOX proteins and this directly impacts their DNA binding properties and functions.

Recent research has opened new insights into the complex nature of interactions between HOX and PBC proteins. Studies have revealed that novel HOX-PBX cofactor interactions arise through the loss or gain of other interacting domains beyond the HP and these have the potential to further modify HOX DNA binding properties and functions (Shen W. F. et al., 1997; Liu et al., 2008; Noro et al., 2011; Slattery et al., 2011; Hudry et al., 2012; Merabet and Mann, 2016; Dard et al., 2019; Singh et al., 2020). Use of sensitive *in vivo* assays to quantify HOX-PBX interactions have revealed that, in the presence of MEIS, the HP motif is dispensable in all HOX proteins except those from anterior paralog groups 1 and 2 (Dard et al., 2018). Furthermore, detailed analyses uncovered alternative PBC interaction motifs in human HOXB3, HOXA7 and HOXC8 proteins that are critical for HOX-PBC interaction in specific cell contexts and DNA-binding site topologies. While in the case of HOXA9, the HOXA9-PBX-MEIS interaction is dependent on the activity of the HP motif and two paralog-specific residues of the homeodomain region. These observations suggest that HOX-PBC interactions are not rigid and may behave in a dynamic manner that vary based on specific cellular and genomic contexts. The array of HOX- PBC interactions might have evolved independently in novel ways as a common regulatory node or mechanism to diversify DNA binding properties of HOX proteins. The highly conserved W residue in the HP region has been shown to be required for binding on HOX–PBX consensus motifs, while HOX protein binding on non-consensus motifs and low-affinity binding sites may be altered through HP diversity in combination with other novel interaction domains (Foos et al., 2015; Dard et al., 2018; Singh et al., 2020).

Linker Region

HP region of HOX proteins is connected to the N-terminal of the HD through a linker (LK) region (**Figures 5A,B**). The sequence and size of the linker region is highly variable among the HOX paralogs, ranging from 3 to 50 amino acids in the vertebrate HOX proteins. Since the HP region of HOX proteins interacts with a highly conserved three amino acid loop extension domain in the HD region of PBX cofactors, the size and sequence of the linker region may constrain these interactions and further modulate DNA binding properties. Structural studies have revealed that the LK region of the *Drosophila* Sex Combs Reduced (Scr) HOX protein is critical for binding at some target sites, but on other binding targets it is disordered and makes minimal contribution to binding (Joshi et al., 2007). Converting the LK region of *Drosophila* Antp to that of Scr changes the DNA binding preference of the protein such that it binds very similar targets to those of Scr (Abe et al., 2015). Furthermore, comparative functional studies of mouse HOXA1 with HOXB1 proteins, Drosophila Ubx with Abd-A and Dfd with Scr show that the linker region is required for some aspects of the paralog specific functions of these HOX proteins (Merabet et al., 2003; Joshi et al., 2010; Singh et al., 2020). This indicates that the LK sequence has a role in modulating DNA binding preferences of HOX proteins and diversity in this region may be a determinant that underlies aspects of the paralog-specific functions of HOX proteins.

C-Terminal Region

The C-terminal region flanking the homeodomain of HOX proteins is highly variable in size and sequence. The importance of this part of HOX proteins has been generally ignored because of the high degree of variability among the HOX paralogs. This region has also been left out of the studies that analyzed the three-dimensional structure of HOX proteins with PBX and DNA. Sequence analyses show that this region varies from 7 amino acids in HOXA13 to 168 amino acids in HOXA2 (**Figure 5B**). This is also highly variable between the HOXA alloparalogs (HOXA1 to HOXA13) and symparalogs among Hox A, B, C and D clusters. Our recent cross-species functional analyses revealed that a highly conserved CTM motif (KEGLLP) is a key determinant involved in maintaining the homologous ancestral functions of *Drosophila* Labial in the mouse HOXA1 protein (**Figures 4A,C**) (Singh et al., 2020). Diversification of this motif in mammalian homologs, HOXB1 and HOXD1led to a loss of the ancestral activity. Furthermore, structure prediction analyses suggested that the CTM region may establish another interaction domain with PBX1 on DNA. Consistent with this idea, *in vitro* DNA binding analyses revealed that the CTM region is not sufficient for HOX1 interaction with PBX1, but it can modulate the ability of HOXA1 to interact with PBX1 when bound to a target site (Singh et al., 2020). Similarly, small, conserved regions in the C-terminal domains of HOX proteins have also been observed in *Drosophila* HOX proteins Ubx and Abd-A (Lelli et al., 2011). *In vitro* DNA binding analyses and three-dimensional structures show that this region in Ubx (called Ubd-A region) is required for direct physical interaction with Exd, a homolog of PBX1, and affects DNA binding properties (Foos et al., 2015). Hence, it may play an analogous role to the CTM, identified in HOXA1. The UBD-A region is highly conserved among insect orthologs of Ubx, but it is absent from other arthropods and onychophorans (Galant and Carroll, 2002; Ronshaugen et al., 2002). Functional analyses of this domain with a transgenic reporter line displayed a repressive role on a Distal-less (*Dll*) *cis*-regulatory element that is involved in promoting limb development. These results suggests

that evolution of the UBD-A domain suppressed limb formation in abdominal segments and provided an evolutionary transition to hexapod limb pattern. Furthermore, *in vitro* DNA binding and *in vivo* reporter assays show cooperation between linker and UBD-A region of the Ubx protein, which suggests subtle changes in HOX–PBC complexes have played a major role in the diversification of HOX protein function in evolution (Saadaoui et al., 2011). These data illustrate that flexible extensions outside of the HD helix have the potential to mediate additional contacts between HOX proteins and their cofactors in concert with those mediated by the HP motif on the opposing side of the DNA. Together these findings demonstrate that small differences in sequences outside of HDs, which do not contact DNA themselves, may be a common mechanism for modulating protein-protein interactions that impact DNA binding specificity of HOX proteins.

SUMMARY

A small number of changes in key amino acids may affect DNA binding properties and protein-protein interactions of transcription factors that can influence their DNA binding targets and potential for transcriptional activation or repression (Lai et al., 2001; Johnson et al., 2003; Chi, 2005; Sakazume et al., 2007; Shoubridge et al., 2012; Webb et al., 2012; Jubb et al., 2017; Singh et al., 2020; Chi, 2005). This indicates that sequence conservation alone may be a poor determinant in predicting the functions of HOX proteins. Genome-wide binding and gene expression analyses have revealed both overlapping and paralog-specific targets of HOX proteins (Hueber et al., 2007; Sorge et al., 2012; Beh et al., 2016; Bulajic et al., 2020; Singh et al., 2020). This suggests that there could be many common downstream targets in the genome, but the unique targets might have evolved by diversification of HOX proteins, resulting in selective alterations in their downstream target genes and inputs into novel gene regulatory programs. Paralog specific binding at unique target sites could arise through small differences in DNA binding domain and associated regions that alter interaction with cofactors such as PBX. Studies have shown that HOX-PBX interactions have diversified by altering interactions through the HP region and evolving novel contact points beyond it (Merabet et al., 2011; Saadaoui et al., 2011; Rivas et al., 2013). Altered HOX-PBX interactions may affect both DNA binding specificity of the HOX proteins and the transcriptional state of the target site. The diversification in the function of HOX proteins can be also introduced by changes outside the homeodomain and hexapeptide region (Chauvet et al., 2000; Gebelein et al., 2002; Singh et al., 2020). Several conserved short linear motifs (SLiMs) have been identified in HOX proteins that can often restrain the interaction potential of HOX proteins (Baeza et al., 2015). Deletion of SLiM motifs leads to loss, gain or interestingly enhanced interaction with cofactors that can alter regulatory potential of HOX proteins in a context specific manner. These dynamic changes in interaction with cofactors may alter Hox activity in tissue and cell type-specific manners which vary depending upon the cellular context (Capovilla et al., 1994; Joshi et al., 2010; Jung et al., 2014). These observations illustrate that a small number or subtle changes in multiple regions of HOX proteins can have a dramatic effect on their activity and may be an important feature that underlies the paralog specific functions by modifying DNA binding specificity and/or protein-protein interactions. Investigating the *in vivo* functional roles and evolution of other domains of HOX proteins beyond the HD should help to unravel how such similar proteins can exert diverse functions and be relevant in determining if this is a general mechanism used by other transcription factor families in the generation of diversity and evolution of novel functional activities of proteins.

AUTHOR CONTRIBUTIONS

NS and RK researched, wrote, edited and organized the manuscript.

FUNDING

This work was supported by funds from the Stowers Institute for Medical Research (Grant no. 1001) to RK.

REFERENCES

Abe, N., Dror, I., Yang, L., Slattery, M., Zhou, T., Bussemaker, H. J., et al. (2015). Deconvolving the Recognition of DNA Shape from Sequence. *Cell* 161, 307–318. doi:10.1016/j.cell.2015.02.008

Akam, M., Averof, M., Castelli-Gair, J., Dawes, R., Falciani, F., and Ferrier, D. (1994). The Evolving Role of Hox Genes in Arthropods. *Dev. Suppl.* 1994, 209–215. doi:10.1242/dev.1994.supplement.209

Albalat, R., and Cañestro, C. (2016). Evolution by Gene Loss. *Nat. Rev. Genet.* 17, 379–391. doi:10.1038/nrg.2016.39

Alonso, C. R., Maxton-Kuechenmeister, J., and Akam, M. (2001). Evolution of Ftz Protein Function in Insects. *Curr. Biol.* 11, 1473–1478. doi:10.1016/s0960-9822(01)00425-0

Amin, S., Donaldson, I. J., Zannino, D. A., Hensman, J., Rattray, M., Losa, M., et al. (2015). Hoxa2 Selectively Enhances Meis Binding to Change a Branchial Arch Ground State. *Develop. Cel* 32, 265–277. doi:10.1016/j.devcel.2014.12.024

Arendt, D. (2018). Hox Genes and Body Segmentation. *Science* 361, 1310–1311. doi:10.1126/science.aav0692

Assaiya, A., Burada, A. P., Dhingra, S., and Kumar, J. (2021). An Overview of the Recent Advances in Cryo-Electron Microscopy for Life Sciences. *Emerg. Top. Life Sci.* 5, 151–168. doi:10.1042/etls20200295

Averof, M., and Akam, M. (1993). *HOM/Hox* Genes of *Artemia*: Implications for the Origin of Insect and Crustacean Body Plans. *Curr. Biol.* 3, 73–78. doi:10.1016/0960-9822(93)90158-k

Averof, M. (2002). Arthropod Hox Genes: Insights on the Evolutionary Forces that Shape Gene Functions. *Curr. Opin. Genet. Develop.* 12, 386–392. doi:10.1016/s0959-437x(02)00314-3

Avsec, Ž., Weilert, M., Shrikumar, A., Krueger, S., Alexandari, A., Dalal, K., et al. (2021). Base-resolution Models of Transcription-Factor Binding Reveal Soft Motif Syntax. *Nat. Genet.* 53, 354–366. doi:10.1038/s41588-021-00782-6

Baëza, M., Viala, S., Heim, M., Dard, A., Hudry, B., Duffraisse, M., et al. (2015). Inhibitory Activities of Short Linear Motifs Underlie Hox Interactome Specificity *In Vivo*. *Elife* 4, 6034. doi:10.7554/eLife.06034

Banreti, A., Hudry, B., Sass, M., Saurin, A. J., and Graba, Y. (2014). Hox Proteins Mediate Developmental and Environmental Control of Autophagy. *Develop. Cel* 28, 56–69. doi:10.1016/j.devcel.2013.11.024

Beh, C. Y., El-Sharnouby, S., Chatzipli, A., Russell, S., Choo, S. W., and White, R. (2016). Roles of Cofactors and Chromatin Accessibility in Hox Protein Target Specificity. *Epigenetics. Chromatin* 9, 1. doi:10.1186/s13072-015-0049-x

Berthelsen, J., Zappavigna, V., Ferretti, E., Mavilio, F., and Blasi, F. (1998). The Novel Homeoprotein Prep1 Modulates Pbx-Hox Protein Cooperativity. *Embo J.* 17, 1434–1445. doi:10.1093/emboj/17.5.1434

Biémont, C., and Vieira, C. (2006). Junk DNA as an Evolutionary Force. *Nature* 443, 521–524. doi:10.1038/443521a

Blomme, T., Vandepoele, K., De Bodt, S., Simillion, C., Maere, S., and Van De Peer, Y. (2006). The Gain and Loss of Genes during 600 Million Years of Vertebrate Evolution. *Genome Biol.* 7, R43. doi:10.1186/gb-2006-7-5-r43

Braasch, I., Bobe, J., Guiguen, Y., and Postlethwait, J. H. (2018). Reply to: 'Subfunctionalization versus Neofunctionalization after Whole-Genome Duplication'. *Nat. Genet.* 50, 910–911. doi:10.1038/s41588-018-0163-3

Breitinger, C., Maethner, E., Garcia-Cuellar, M.-P., and Slany, R. K. (2012). The Homeodomain Region Controls the Phenotype of HOX-Induced Murine Leukemia. *Blood* 120, 4018–4027. doi:10.1182/blood-2011-10-384685

Bulajic, M., Srivastava, D., Dasen, J. S., Wichterle, H., Mahony, S., and Mazzoni, E. O. (2020). Differential Abilities to Engage Inaccessible Chromatin Diversify Vertebrate Hox Binding Patterns. *Development* 147, dev194761. doi:10.1242/dev.194761

Burki, F., and Kaessmann, H. (2004). Birth and Adaptive Evolution of a Hominoid Gene that Supports High Neurotransmitter Flux. *Nat. Genet.* 36, 1061–1063. doi:10.1038/ng1431

Capovilla, M., Brandt, M., and Botas, J. (1994). Direct Regulation of Decapentaplegic by Ultrabithorax and its Role in Drosophila Midgut Morphogenesis. *Cell* 76, 461–475. doi:10.1016/0092-8674(94)90111-2

Carapuço, M., Nóvoa, A., Bobola, N., and Mallo, M. (2005). Hox Genes Specify Vertebral Types in the Presomitic Mesoderm. *Genes Dev.* 19, 2116–2121. doi:10.1101/gad.338705

Carroll, S. B. (2008). Evo-devo and an Expanding Evolutionary Synthesis: a Genetic Theory of Morphological Evolution. *Cell* 134, 25–36. doi:10.1016/j.cell.2008.06.030

Carroll, S. B. (1995). Homeotic Genes and the Evolution of Arthropods and Chordates. *Nature* 376, 479–485. doi:10.1038/376479a0

Chan, S.-K., Jaffe, L., Capovilla, M., Botas, J., and Mann, R. S. (1994). The DNA Binding Specificity of Ultrabithorax Is Modulated by Cooperative Interactions with Extradenticle, Another Homeoprotein. *Cell* 78, 603–615. doi:10.1016/0092-8674(94)90525-8

Chan, S. K., and Mann, R. S. (1996). A Structural Model for a Homeotic Protein-Extradenticle-DNA Complex Accounts for the Choice of HOX Protein in the Heterodimer. *Proc. Natl. Acad. Sci. U.S.A.* 93, 5223–5228. doi:10.1073/pnas.93.11.5223

Chang, C. P., Shen, W. F., Rozenfeld, S., Lawrence, H. J., Largman, C., and Cleary, M. L. (1995). Pbx Proteins Display Hexapeptide-dependent Cooperative DNA Binding with a Subset of Hox Proteins. *Genes Dev.* 9, 663–674. doi:10.1101/gad.9.6.663

Chauvet, S., Merabet, S., Bilder, D., Scott, M. P., Pradel, J., and Graba, Y. (2000). Distinct Hox Protein Sequences Determine Specificity in Different Tissues. *Proc. Natl. Acad. Sci. U.S.A.* 97, 4064–4069. doi:10.1073/pnas.070046997

Chen, F., and Capecchi, M. R. (1997). Targeted Mutations inHoxa-9andHoxb-9Reveal Synergistic Interactions. *Develop. Biol.* 181, 186–196. doi:10.1006/dbio.1996.8440

Chi, Y.-I. (2005). Homeodomain Revisited: a Lesson from Disease-Causing Mutations. *Hum. Genet.* 116, 433–444. doi:10.1007/s00439-004-1252-1

Clark, A. G. (1994). Invasion and Maintenance of a Gene Duplication. *Proc. Natl. Acad. Sci. U.S.A.* 91, 2950–2954. doi:10.1073/pnas.91.8.2950

Coiffier, D., Charroux, B., and Kerridge, S. (2008). Common Functions of central and Posterior Hox Genes for the Repression of Head in the Trunk of Drosophila. *Development* 135, 291–300. doi:10.1242/dev.009662

Copley, R. R. (2008). The Animal in the Genome: Comparative Genomics and Evolution. *Phil. Trans. R. Soc. B* 363, 1453–1461. doi:10.1098/rstb.2007.2235

Dard, A., Jia, Y., Reboulet, J., Bleicher, F., Lavau, C., and Merabet, S. (2019). The Human HOXA9 Protein Uses Paralog-specific Residues of the Homeodomain to Interact with TALE-Class Cofactors. *Sci. Rep.* 9, 5664. doi:10.1038/s41598-019-42096-y

Dard, A., Reboulet, J., Jia, Y., Bleicher, F., Duffraisse, M., Vanaker, J.-M., et al. (2018). Human HOX Proteins Use Diverse and Context-dependent Motifs to Interact with TALE Class Cofactors. *Cel Rep.* 22, 3058–3071. doi:10.1016/j.celrep.2018.02.070

Davenne, M., Maconochie, M. K., Neun, R., Pattyn, A., Chambon, P., Krumlauf, R., et al. (1999). *Hoxa2* and *Hoxb2* Control Dorsoventral Patterns of Neuronal Development in the Rostral Hindbrain. *Neuron* 22, 677–691. doi:10.1016/s0896-6273(00)80728-x

Debodt, S., Maere, S., and Vandepeer, Y. (2005). Genome Duplication and the Origin of Angiosperms. *Trends Ecol. Evol.* 20, 591–597. doi:10.1016/j.tree.2005.07.008

Dehal, P., and Boore, J. L. (2005). Two Rounds of Whole Genome Duplication in the Ancestral Vertebrate. *Plos Biol.* 3, e314. doi:10.1371/journal.pbio.0030314

Denans, N., Iimura, T., and Pourquié, O. (2015). Hox Genes Control Vertebrate Body Elongation by Collinear Wnt Repression. *Elife* 4. doi:10.7554/eLife.04379

Desplan, C., Theis, J., and O'farrell, P. H. (1988). The Sequence Specificity of Homeodomain-DNA Interaction. *Cell* 54, 1081–1090. doi:10.1016/0092-8674(88)90123-7

Domsch, K., Papagiannouli, F., and Lohmann, I. (2015). The HOX-Apoptosis Regulatory Interplay in Development and Disease. *Curr. Top. Dev. Biol.* 114, 121–158. doi:10.1016/bs.ctdb.2015.07.014

Doudna, J. A., and Charpentier, E. (2014). The New Frontier of Genome Engineering with CRISPR-Cas9. *Science* 346, 1258096. doi:10.1126/science.1258096

Duboule, D., and Dollé, P. (1989). The Structural and Functional Organization of the Murine *HOX* Gene Family Resembles that of *Drosophila* Homeotic Genes. *EMBO J.* 8, 1497–1505. doi:10.1002/j.1460-2075.1989.tb03534.x

Duboule, D. (2007). The Rise and Fall of Hox Gene Clusters. *Development* 134, 2549–2560. doi:10.1242/dev.001065

Duboule, D. (1998). Vertebrate Hox Gene Regulation: Clustering And/or Colinearity? *Curr. Opin. Genet. Develop.* 8, 514–518. doi:10.1016/s0959-437x(98)80004-x

Enard, W., Gehre, S., Hammerschmidt, K., Hölter, S. M., Blass, T., Somel, M., et al. (2009). A Humanized Version of Foxp2 Affects Cortico-Basal Ganglia Circuits in Mice. *Cell* 137, 961–971. doi:10.1016/j.cell.2009.03.041

Enriquez, J., Boukhatmi, H., Dubois, L., Philippakis, A. A., Bulyk, M. L., Michelson, A. M., et al. (2010). Multi-step Control of Muscle Diversity by Hox Proteins in the Drosophila Embryo. *Development* 137, 457–466. doi:10.1242/dev.045286

Escriva, H., Bertrand, S., Germain, P., Robinson-Rechavi, M., Umbhauer, M., Cartry, J., et al. (2006). Neofunctionalization in Vertebrates: the Example of Retinoic Acid Receptors. *Plos Genet.* 2, e102. doi:10.1371/journal.pgen.0020102

Fernández, R., and Gabaldón, T. (2020). Gene Gain and Loss across the Metazoan Tree of Life. *Nat. Ecol. Evol.* 4, 524–533. doi:10.1038/s41559-019-1069-x

Ferretti, E., Marshall, H., Pöpperl, H., Maconochie, M., Krumlauf, R., and Blasi, F. (2000). Segmental Expression of Hoxb2 in R4 Requires Two Separate Sites that Integrate Cooperative Interactions between Prep1, Pbx and Hox Proteins. *Development* 127, 155–166. doi:10.1242/dev.127.1.155

Foos, N., Maurel-Zaffran, C., Maté, M. J., Vincentelli, R., Hainaut, M., Berenger, H., et al. (2015). A Flexible Extension of the Drosophila Ultrabithorax Homeodomain Defines a Novel Hox/PBC Interaction Mode. *Structure* 23, 270–279. doi:10.1016/j.str.2014.12.011

Force, A., Lynch, M., Pickett, F. B., Amores, A., Yan, Y.-l., and Postlethwait, J. (1999). Preservation of Duplicate Genes by Complementary, Degenerative Mutations. *Genetics* 151, 1531–1545. doi:10.1093/genetics/151.4.1531

Franchini, L. F., and Pollard, K. S. (2017). Human Evolution: the Non-coding Revolution. *BMC Biol.* 15, 89. doi:10.1186/s12915-017-0428-9

Friedman, R., and Hughes, A. L. (2001). Pattern and Timing of Gene Duplication in Animal Genomes. *Genome Res.* 11, 1842–1847. doi:10.1101/gr.200601

Fröbius, A. C., and Funch, P. (2017). Rotiferan Hox Genes Give New Insights into the Evolution of Metazoan Bodyplans. *Nat. Commun.* 8, 9. doi:10.1038/s41467-017-00020-w

Fromental-Ramain, C., Warot, X., Lakkaraju, S., Favier, B., Haack, H., Birling, C., et al. (1996). Specific and Redundant Functions of the Paralogous Hoxa-9 and Hoxd-9 Genes in Forelimb and Axial Skeleton Patterning. *Development* 122, 461–472. doi:10.1242/dev.122.2.461

Furukubo-Tokunaga, K., Flister, S., and Gehring, W. J. (1993). Functional Specificity of the Antennapedia Homeodomain. *Proc. Natl. Acad. Sci. U.S.A.* 90, 6360–6364. doi:10.1073/pnas.90.13.6360

Galant, R., and Carroll, S. B. (2002). Evolution of a Transcriptional Repression Domain in an Insect Hox Protein. *Nature* 415, 910–913. doi:10.1038/nature717

Galant, R., Walsh, C. M., and Carroll, S. B. (2002). Hox Repression of a Target Gene: Extradenticle-independent, Additive Action through Multiple Monomer Binding Sites. *Development* 129, 3115–3126. doi:10.1242/dev.129.13.3115

Garcia-Bellido, A., Ripoll, P., and Morata, G. (1973). Developmental Compartmentalisation of the wing Disk of Drosophila. *Nat. New Biol.* 245, 251–253. doi:10.1038/newbio245251a0

Garcia-Fernàndez, J., and Holland, P. W. H. (1994). Archetypal Organization of the Amphioxus Hox Gene Cluster. *Nature* 370, 563–566. doi:10.1038/370563a0

Gasperskaja, E., and Kučinskas, V. (2017). The Most Common Technologies and Tools for Functional Genome Analysis. *Aml* 24, 1–11. doi:10.6001/actamedica.v24i1.3457

Gebelein, B., Culi, J., Ryoo, H. D., Zhang, W., and Mann, R. S. (2002). Specificity of Distalless Repression and Limb Primordia Development by Abdominal Hox Proteins. *Develop. Cel* 3, 487–498. doi:10.1016/s1534-5807(02)00257-5

Gebelein, B., Mckay, D. J., and Mann, R. S. (2004). Direct Integration of Hox and Segmentation Gene Inputs during Drosophila Development. *Nature* 431, 653–659. doi:10.1038/nature02946

Gehring, W. J., Affolter, M., and Bürglin, T. (1994a). Homeodomain Proteins. *Annu. Rev. Biochem.* 63, 487–526. doi:10.1146/annurev.bi.63.070194.002415

Gehring, W. J., Qian, Y. Q., Billeter, M., Furukubo-Tokunaga, K., Schier, A. F., Resendez-Perez, D., et al. (1994b). Homeodomain-DNA Recognition. *Cell* 78, 211–223. doi:10.1016/0092-8674(94)90292-5

Graham, A., Papalopulu, N., and Krumlauf, R. (1989). The Murine and *Drosophila* Homeobox Gene Complexes Have Common Features of Organization and Expression. *Cell* 57, 367–378. doi:10.1016/0092-8674(89)90912-4

Greer, J. M., Puetz, J., Thomas, K. R., and Capecchi, M. R. (2000). Maintenance of Functional Equivalence during Paralogous Hox Gene Evolution. *Nature* 403, 661–665. doi:10.1038/35001077

Gruschus, J. M., Tsao, D. H. H., Wang, L.-H., Nirenberg, M., and Ferretti, J. A. (1997). Interactions of the vnd/NK-2 Homeodomain with DNA by Nuclear Magnetic Resonance Spectroscopy: Basis of Binding Specificity. *Biochemistry* 36, 5372–5380. doi:10.1021/bi9620060

Guijarro-Clarke, C., Holland, P. W. H., and Paps, J. (2020a). Publisher Correction: Widespread Patterns of Gene Loss in the Evolution of the Animal Kingdom. *Nat. Ecol. Evol.* 4, 661. doi:10.1038/s41559-020-1159-9

Guijarro-Clarke, C., Holland, P. W. H., and Paps, J. (2020b). Widespread Patterns of Gene Loss in the Evolution of the Animal Kingdom. *Nat. Ecol. Evol.* 4, 519–523. doi:10.1038/s41559-020-1129-2

Guo, T., Mandai, K., Condie, B. G., Wickramasinghe, S. R., Capecchi, M. R., and Ginty, D. D. (2011). An Evolving NGF-Hoxd1 Signaling Pathway Mediates Development of Divergent Neural Circuits in Vertebrates. *Nat. Neurosci.* 14, 31–36. doi:10.1038/nn.2710

Hahn, M. W., and Wray, G. A. (2002). The G-Value Paradox. *Evol. Dev.* 4, 73–75. doi:10.1046/j.1525-142x.2002.01069.x

Hanks, M. C., Loomis, C. A., Harris, E., Tong, C. X., Anson-Cartwright, L., Auerbach, A., et al. (1998). Drosophila Engrailed Can Substitute for Mouse Engrailed1 Function in Mid-hindbrain, but Not Limb Development. *Development* 125, 4521–4530. doi:10.1242/dev.125.22.4521

He, Q., Johnston, J., and Zeitlinger, J. (2015). ChIP-nexus Enables Improved Detection of *In Vivo* Transcription Factor Binding Footprints. *Nat. Biotechnol.* 33, 395–401. doi:10.1038/nbt.3121

He, S., Del Viso, F., Chen, C.-Y., Ikmi, A., Kroesen, A. E., and Gibson, M. C. (2018). An Axial Hox Code Controls Tissue Segmentation and Body Patterning in *Nematostella vectensis*. *Science* 361, 1377–1380. doi:10.1126/science.aar8384

He, X., and Zhang, J. (2005). Rapid Subfunctionalization Accompanied by Prolonged and Substantial Neofunctionalization in Duplicate Gene Evolution. *Genetics* 169, 1157–1164. doi:10.1534/genetics.104.037051

Hedges, S. B., Marin, J., Suleski, M., Paymer, M., and Kumar, S. (2015). Tree of Life Reveals Clock-like Speciation and Diversification. *Mol. Biol. Evol.* 32, 835–845. doi:10.1093/molbev/msv037

Hirth, F., Loop, T., Egger, B., Miller, D. F. B., Kaufman, T. C., and Reichert, H. (2001). Functional Equivalence of Hox Gene Products in the Specification of the Tritocerebrum during Embryonic Brain Development of Drosophila. *Development* 128, 4781–4788. doi:10.1242/dev.128.23.4781

Hoegg, S., and Meyer, A. (2005). Hox Clusters as Models for Vertebrate Genome Evolution. *Trends Genet.* 21, 421–424. doi:10.1016/j.tig.2005.06.004

Holland, L. Z., and Ocampo Daza, D. (2018). A New Look at an Old Question: when Did the Second Whole Genome Duplication Occur in Vertebrate Evolution? *Genome Biol.* 19, 209. doi:10.1186/s13059-018-1592-0

Holland, P. W. H., Garcia-Fernàndez, J., Williams, N. A., and Sidow, A. (1994). Gene Duplications and the Origins of Vertebrate Development. *Dev. Suppl.* 1994, 125–133. doi:10.1242/dev.1994.supplement.125

Horan, G. S., Ramírez-Solis, R., Featherstone, M. S., Wolgemuth, D. J., Bradley, A., and Behringer, R. R. (1995). Compound Mutants for the Paralogous *Hoxa-4, Hoxb-4*, and *Hoxd-4* Genes Show More Complete Homeotic Transformations and a Dose-dependent Increase in the Number of Vertebrae Transformed. *Genes Dev.* 9, 1667–1677. doi:10.1101/gad.9.13.1667

Hudry, B., Remacle, S., Delfini, M.-C., Rezsohazy, R., Graba, Y., and Merabet, S. (2012). Hox Proteins Display a Common and Ancestral Ability to Diversify Their Interaction Mode with the PBC Class Cofactors. *Plos Biol.* 10, e1001351. doi:10.1371/journal.pbio.1001351

Hueber, S. D., Bezdan, D., Henz, S. R., Blank, M., Wu, H., and Lohmann, I. (2007). Comparative Analysis of Hox Downstream Genes inDrosophila. *Development* 134, 381–392. doi:10.1242/dev.02746

Hughes, A. L. (1994). The Evolution of Functionally Novel Proteins after Gene Duplication. *Proc. Biol. Sci.* 256, 119–124. doi:10.1098/rspb.1994.0058

Hunter, M. P., and Prince, V. E. (2002). Zebrafish Hox Paralogue Group 2 Genes Function Redundantly as Selector Genes to Pattern the Second Pharyngeal Arch. *Develop. Biol.* 247, 367–389. doi:10.1006/dbio.2002.0701

Iacovino, M., Hernandez, C., Xu, Z., Bajwa, G., Prather, M., and Kyba, M. (2009). A Conserved Role for Hox Paralog Group 4 in Regulation of Hematopoietic Progenitors. *Stem Cell Develop.* 18, 783–792. doi:10.1089/scd.2008.0227

Ikuta, T. (2011). Evolution of Invertebrate Deuterostomes and Hox/ParaHox Genes. *Genomics. Proteomics. Bioinformatics* 9, 77–96. doi:10.1016/s1672-0229(11)60011-9

Innan, H. (2009). Population Genetic Models of Duplicated Genes. *Genetica* 137, 19–37. doi:10.1007/s10709-009-9355-1

Irie, N., Satoh, N., and Kuratani, S. (2018). The Phylum Vertebrata: a Case for Zoological Recognition. *Zoolog. Lett* 4, 32. doi:10.1186/s40851-018-0114-y

Johnson, D., Kan, S.-h., Oldridge, M., Trembath, R. C., Roche, P., Esnouf, R. M., et al. (2003). Missense Mutations in the Homeodomain of HOXD13 Are Associated with Brachydactyly Types D and E. *Am. J. Hum. Genet.* 72, 984–997. doi:10.1086/374721

Joshi, R., Passner, J. M., Rohs, R., Jain, R., Sosinsky, A., Crickmore, M. A., et al. (2007). Functional Specificity of a Hox Protein Mediated by the Recognition of Minor Groove Structure. *Cell* 131, 530–543. doi:10.1016/j.cell.2007.09.024

Joshi, R., Sun, L., and Mann, R. (2010). Dissecting the Functional Specificities of Two Hox Proteins. *Genes Dev.* 24, 1533–1545. doi:10.1101/gad.1936910

Jubb, H. C., Pandurangan, A. P., Turner, M. A., Ochoa-Montaño, B., Blundell, T. L., and Ascher, D. B. (2017). Mutations at Protein-Protein Interfaces: Small Changes over Big Surfaces Have Large Impacts on Human Health. *Prog. Biophys. Mol. Biol.* 128, 3–13. doi:10.1016/j.pbiomolbio.2016.10.002

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., et al. (2021). Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* 596, 583–589. doi:10.1038/s41586-021-03819-2

Jung, H., Mazzoni, E. O., Soshnikova, N., Hanley, O., Venkatesh, B., Duboule, D., et al. (2014). Evolving Hox Activity Profiles Govern Diversity in Locomotor Systems. *Develop. Cel* 29, 171–187. doi:10.1016/j.devcel.2014.03.008

Kaufman, T. C., Lewis, R., and Wakimoto, B. (1980). Cytogenetic Analysis of Chromosome 3 in Drosophila Melanogaster: The Homoeotic Gene Complex in Polytene Chromosome Interval 84a-B. *Genetics* 94, 115–133. doi:10.1093/genetics/94.1.115

Kleinjan, D. A., Bancewicz, R. M., Gautier, P., Dahm, R., Schonthaler, H. B., Damante, G., et al. (2008). Subfunctionalization of Duplicated Zebrafish Pax6 Genes by Cis-Regulatory Divergence. *Plos Genet.* 4, e29. doi:10.1371/journal.pgen.0040029

Kmita, M., and Duboule, D. (2003). Organizing Axes in Time and Space; 25 Years of Colinear Tinkering. *Science* 301, 331–333. doi:10.1126/science.1085753

Knoepfler, P. S., and Kamps, M. P. (1995). The Pentapeptide Motif of Hox Proteins Is Required for Cooperative DNA Binding with Pbx1, Physically Contacts Pbx1,

and Enhances DNA Binding by Pbx1. *Mol. Cel Biol* 15, 5811–5819. doi:10.1128/mcb.15.10.5811

Koonin, E. V. (2005). Orthologs, Paralogs, and Evolutionary Genomics. *Annu. Rev. Genet.* 39, 309–338. doi:10.1146/annurev.genet.39.073003.114725

Krumlauf, R. (1994). *Hox* Genes in Vertebrate Development. *Cell* 78, 191–201. doi:10.1016/0092-8674(94)90290-9

Kuraku, S., and Meyer, A. (2009). The Evolution and Maintenance of Hox Gene Clusters in Vertebrates and the Teleost-specific Genome Duplication. *Int. J. Dev. Biol.* 53, 765–773. doi:10.1387/ijdb.072533km

Lacombe, J., Hanley, O., Jung, H., Philippidou, P., Surmeli, G., Grinstein, J., et al. (2013). Genetic and Functional Modularity of Hox Activities in the Specification of Limb-Innervating Motor Neurons. *Plos Genet.* 9, e1003184. doi:10.1371/journal.pgen.1003184

Lai, C. S. L., Fisher, S. E., Hurst, J. A., Vargha-Khadem, F., and Monaco, A. P. (2001). A Forkhead-Domain Gene Is Mutated in a Severe Speech and Language Disorder. *Nature* 413, 519–523. doi:10.1038/35097076

Laity, J. H., Lee, B. M., and Wright, P. E. (2001). Zinc finger Proteins: New Insights into Structural and Functional Diversity. *Curr. Opin. Struct. Biol.* 11, 39–46. doi:10.1016/s0959-440x(00)00167-6

Lamka, M. L., Boulet, A. M., and Sakonju, S. (1992). Ectopic Expression of UBX and ABD-B Proteins during Drosophila Embryogenesis: Competition, Not a Functional Hierarchy, Explains Phenotypic Suppression. *Development* 116, 841–854. doi:10.1242/dev.116.4.841

Laronde-Leblanc, N. A., and Wolberger, C. (2003). Structure of HoxA9 and Pbx1 Bound to DNA: Hox Hexapeptide and DNA Recognition Anterior to Posterior. *Genes Dev.* 17, 2060–2072. doi:10.1101/gad.1103303

Laurent, J. M., Garge, R. K., Teufel, A. I., Wilke, C. O., Kachroo, A. H., and Marcotte, E. M. (2020). Humanization of Yeast Genes with Multiple Human Orthologs Reveals Functional Divergence between Paralogs. *Plos Biol.* 18, e3000627. doi:10.1371/journal.pbio.3000627

Lawrence, P. A., and Morata, G. (1994). Homeobox Genes: Their Function in Drosophila Segmentation and Pattern Formation. *Cell* 78, 181–189. doi:10.1016/0092-8674(94)90289-5

Lelli, K. M., Noro, B., and Mann, R. S. (2011). Variable Motif Utilization in Homeotic Selector (Hox)-Cofactor Complex Formation Controls Specificity. *Proc. Natl. Acad. Sci. U.S.A.* 108, 21122–21127. doi:10.1073/pnas.1114118109

Lemons, D., and Mcginnis, W. (2006). Genomic Evolution of Hox Gene Clusters. *Science* 313, 1918–1922. doi:10.1126/science.1132040

Lewis, E. B. (1978). A Gene Complex Controlling Segmentation in Drosophila. *Nature* 276, 565–570. doi:10.1038/276565a0

Lewis, E. B. (1994). Homeosis: the First 100 Years. *Trends Genet.* 10, 341–343. doi:10.1016/0168-9525(94)90117-1

Li, X., and Mcginnis, W. (1999). Activity Regulation of Hox Proteins, a Mechanism for Altering Functional Specificity in Development and Evolution. *Proc. Natl. Acad. Sci. U.S.A.* 96, 6802–6807. doi:10.1073/pnas.96.12.6802

Li, X., Murre, C., and Mcginnis, W. (1999). Activity Regulation of a Hox Protein and a Role for the Homeodomain in Inhibiting Transcriptional Activation. *Embo J.* 18, 198–211. doi:10.1093/emboj/18.1.198

Liu, Y., Matthews, K. S., and Bondos, S. E. (2008). Multiple Intrinsically Disordered Sequences Alter DNA Binding by the Homeodomain of the Drosophila Hox Protein Ultrabithorax. *J. Biol. Chem.* 283, 20874–20887. doi:10.1074/jbc.m800375200

Loker, R., Sanner, J. E., and Mann, R. S. (2021). Cell-type-specific Hox Regulatory Strategies Orchestrate Tissue Identity. *Curr. Biol.* doi:10.1016/j.cub.2021.07.030

Lu, Q., Knoepfler, P., Scheele, J., Wright, D., and Kamps, M. (1995). Both Pbx1 and E2A-Pbx1 Bind the DNA Motif ACCAATCAA Cooperatively with the Products of Mutliple Murine HOX Genes, Some of Which Themselves Are Oncogenes. *Mol. Cel Biol* 15, 3786. doi:10.1128/mcb.15.7.3786

Lundin, L. G. (1999). Gene Duplications in Early Metazoan Evolution. *Semin. Cel Develop. Biol.* 10, 523–530. doi:10.1006/scdb.1999.0333

Lutz, B., Lu, H. C., Eichele, G., Miller, D., and Kaufman, T. C. (1996). Rescue of *Drosophila Labial* Null Mutant by the Chicken Ortholog *Hoxb-1* Demonstrates that the Function of *Hox* Genes Is Phylogenetically Conserved. *Genes Dev.* 10, 176–184. doi:10.1101/gad.10.2.176

Lynch, M., and Conery, J. S. (2000). The Evolutionary Fate and Consequences of Duplicate Genes. *Science* 290, 1151–1155. doi:10.1126/science.290.5494.1151

Lynch, M., and Force, A. (2000). The Probability of Duplicate Gene Preservation by Subfunctionalization. *Genetics* 154, 459–473. doi:10.1093/genetics/154.1.459

Maconochie, M., Nonchev, S., Morrison, A., and Krumlauf, R. (1996). Paralogous *Hox* Genes: Function and Regulation. *Annu. Rev. Genet.* 30, 529–556. doi:10.1146/annurev.genet.30.1.529

Maeda, R. K., and Karch, F. (2006). The ABC of the BX-C: the Bithorax Complex Explained. *Development* 133, 1413–1422. doi:10.1242/dev.02323

Malicki, J., Schughart, K., and Mcginnis, W. (1990). Mouse Hox-2.2 Specifies Thoracic Segmental Identity in Drosophila Embryos and Larvae. *Cell* 63, 961–967. doi:10.1016/0092-8674(90)90499-5

Manley, N. R., and Capecchi, M. R. (1997). *Hox* Group 3 Paralogous Genes Act Synergistically in the Formation of Somitic and Neural Crest-Derived Structures. *Develop. Biol.* 192, 274–288. doi:10.1006/dbio.1997.8765

Mann, R. S., and Chan, S.-K. (1996). Extra Specificity from Extradenticle: the Partnership between HOX and PBX/EXD Homeodomain Proteins. *Trends Genet.* 12, 258–262. doi:10.1016/0168-9525(96)10026-3

Mann, R. S., Lelli, K. M., and Joshi, R. (2009). Chapter 3 Hox Specificity. *Curr. Top. Dev. Biol.* 88, 63–101. doi:10.1016/s0070-2153(09)88003-4

Mann, R. S. (1995). The Specificity of Homeotic Gene Function. *Bioessays* 17, 855–863. doi:10.1002/bies.950171007

Marcussen, T., Oxelman, B., Skog, A., and Jakobsen, K. S. (2010). Evolution of Plant RNA Polymerase IV/V Genes: Evidence of Subneofunctionalization of Duplicated NRPD2/NRPE2-like Paralogs in Viola (Violaceae). *BMC Evol. Biol.* 10, 45. doi:10.1186/1471-2148-10-45

Mazet, F., and M. Shimeld, S. (2002). Gene Duplication and Divergence in the Early Evolution of Vertebrates. *Curr. Opin. Genet. Develop.* 12, 393–396. doi:10.1016/s0959-437x(02)00315-5

Mcginnis, N., Kuziora, M. A., and Mcginnis, W. (1990). Human Hox-4.2 and Drosophila Deformed Encode Similar Regulatory Specificities in Drosophila Embryos and Larvae. *Cell* 63, 969–976. doi:10.1016/0092-8674(90)90500-e

Mcginnis, W., and Krumlauf, R. (1992). Homeobox Genes and Axial Patterning. *Cell* 68, 283–302. doi:10.1016/0092-8674(92)90471-n

Mclain, K., Schreiner, C., Yager, K. L., Stock, J. L., and Steven Potter, S. (1992). Ectopic Expression of Hox-2.3 Induces Craniofacial and Skeletal Malformations in Transgenic Mice. *Mech. Develop.* 39, 3–16. doi:10.1016/0925-4773(92)90021-b

Mcnulty, C. L., Peres, J. N., Bardine, N., Van Den Akker, W. M. R., and Durston, A. J. (2005). Knockdown of the Complete Hox Paralogous Group 1 Leads to Dramatic Hindbrain and Neural Crest Defects. *Development* 132, 2861–2871. doi:10.1242/dev.01872

Medina-Martinez, O., Bradley, A., and Ramírez-Solis, R. (2000). A Large Targeted Deletion of Hoxb1-Hoxb9 Produces a Series Single-Segment Anterior Homeotic Transformations. *Develop. Biol.* 222, 71–83. doi:10.1006/dbio.2000.9683

Medina-Martínez, O., and Ramírez-Solis, R. (2003). *In Vivo* mutagenesis of the Hoxb8 Hexapeptide Domain Leads to Dominant Homeotic Transformations that Mimic the Loss-Of-Function Mutations in Genes of the Hoxb Cluster. *Develop. Biol.* 264, 77–90. doi:10.1016/j.ydbio.2003.07.020

Merabet, S., Kambris, Z., Capovilla, M., Bérenger, H., Pradel, J., and Graba, Y. (2003). The Hexapeptide and Linker Regions of the AbdA Hox Protein Regulate its Activating and Repressive Functions. *Develop. Cel* 4, 761–768. doi:10.1016/s1534-5807(03)00126-6

Merabet, S., Litim-Mecheri, I., Karlsson, D., Dixit, R., Saadaoui, M., Monier, B., et al. (2011). Insights into Hox Protein Function from a Large Scale Combinatorial Analysis of Protein Domains. *Plos Genet.* 7, e1002302. doi:10.1371/journal.pgen.1002302

Merabet, S., and Mann, R. S. (2016). To Be Specific or Not: The Critical Relationship between Hox and TALE Proteins. *Trends Genet.* 32, 334–347. doi:10.1016/j.tig.2016.03.004

Miguez, A., Ducret, S., Di Meglio, T., Parras, C., Hmidan, H., Haton, C., et al. (2012). Opposing Roles for Hoxa2 and Hoxb2 in Hindbrain Oligodendrocyte Patterning. *J. Neurosci.* 32, 17172–17185. doi:10.1523/jneurosci.0885-12.2012

Mora, C., Tittensor, D. P., Adl, S., Simpson, A. G. B., and Worm, B. (2011). How many Species Are There on Earth and in the Ocean? *Plos Biol.* 9, e1001127. doi:10.1371/journal.pbio.1001127

Nadeau, J. H., and Sankoff, D. (1997). Comparable Rates of Gene Loss and Functional Divergence after Genome Duplications Early in Vertebrate Evolution. *Genetics* 147, 1259–1266. doi:10.1093/genetics/147.3.1259

Neuteboom, S. T., and Murre, C. (1997). Pbx Raises the DNA Binding Specificity but Not the Selectivity of Antennapedia Hox Proteins. *Mol. Cel Biol* 17, 4696–4706. doi:10.1128/mcb.17.8.4696

Neuteboom, S. T., Peltenburg, L. T., Van Dijk, M. A., and Murre, C. (1995). The Hexapeptide LFPWMR in Hoxb-8 Is Required for Cooperative DNA Binding with Pbx1 and Pbx2 Proteins. *Proc. Natl. Acad. Sci. U.S.A.* 92, 9166–9170. doi:10.1073/pnas.92.20.9166

Noro, B., Lelli, K., Sun, L., and Mann, R. S. (2011). Competition for Cofactor-dependent DNA Binding Underlies Hox Phenotypic Suppression. *Genes Dev.* 25, 2327–2332. doi:10.1101/gad.175539.111

Noyes, M. B., Christensen, R. G., Wakabayashi, A., Stormo, G. D., Brodsky, M. H., and Wolfe, S. A. (2008). Analysis of Homeodomain Specificities Allows the Family-wide Prediction of Preferred Recognition Sites. *Cell* 133, 1277–1289. doi:10.1016/j.cell.2008.05.023

Ohno, S. (1970). *Evolution by Gene Duplication*. Springer-Verlag Berlin Heidelberg, 160.

Panfilio, K. A., and Akam, M. (2007). A Comparison of Hox3 and Zen Protein Coding Sequences in Taxa that Span the Hox3/zen Divergence. *Dev. Genes Evol.* 217, 323–329. doi:10.1007/s00427-007-0133-8

Panopoulou, G., Hennig, S., Groth, D., Krause, A., Poustka, A. J., Herwig, R., et al. (2003). New Evidence for Genome-wide Duplications at the Origin of Vertebrates Using an Amphioxus Gene Set and Completed Animal Genomes. *Genome Res.* 13, 1056–1066. doi:10.1101/gr.874803

Paps, J., and Holland, P. W. H. (2018). Reconstruction of the Ancestral Metazoan Genome Reveals an Increase in Genomic novelty. *Nat. Commun.* 9, 1730. doi:10.1038/s41467-018-04136-5

Pascual-Anaya, J., D'Aniello, S., Kuratani, S., and Garcia-Fernàndez, J. (2013). Evolution of Hox Gene Clusters in Deuterostomes. *BMC Develop. Biol.* 13, 26. doi:10.1186/1471-213x-13-26

Passner, J. M., Ryoo, H. D., Shen, L., Mann, R. S., and Aggarwal, A. K. (1999). Structure of a DNA-Bound Ultrabithorax-Extradenticle Homeodomain Complex. *Nature* 397, 714–719. doi:10.1038/17833

Patterson, L. T., Pembaur, M., and Potter, S. S. (2001). Hoxa11andHoxd11regulate Branching Morphogenesis of the Ureteric Bud in the Developing Kidney. *Development* 128, 2153–2161. doi:10.1242/dev.128.11.2153

Paul, D. S., Soranzo, N., and Beck, S. (2014). Functional Interpretation of Non-coding Sequence Variation: Concepts and Challenges. *Bioessays* 36, 191–199. doi:10.1002/bies.201300126

Pearson, J. C., Lemons, D., and Mcginnis, W. (2005). Modulating Hox Gene Functions during Animal Body Patterning. *Nat. Rev. Genet.* 6, 893–904. doi:10.1038/nrg1726

Penkov, D., San Martín, D. M., Fernandez-Díaz, L. C., Rosselló, C. A., Torroja, C., Sánchez-Cabo, F., et al. (2013). Analysis of the DNA-Binding Profile and Function of TALE Homeoproteins Reveals Their Specialization and Specific Interactions with Hox Genes/proteins. *Cel Rep.* 3, 1321–1333. doi:10.1016/j.celrep.2013.03.029

Perry, G. H., Dominy, N. J., Claw, K. G., Lee, A. S., Fiegler, H., Redon, R., et al. (2007). Diet and the Evolution of Human Amylase Gene Copy Number Variation. *Nat. Genet.* 39, 1256–1260. doi:10.1038/ng2123

Phelan, M. L., Rambaldi, I., and Featherstone, M. S. (1995). Cooperative Interactions between HOX and PBX Proteins Mediated by a Conserved Peptide Motif. *Mol. Cel Biol* 15, 3989–3997. doi:10.1128/mcb.15.8.3989

Phelan, M. L., Sadoul, R., and Featherstone, M. S. (1994). Functional Differences between HOX Proteins Conferred by Two Residues in the Homeodomain N-Terminal Arm. *Mol. Cel. Biol.* 14, 5066–5075. doi:10.1128/mcb.14.8.5066

Piper, D. E., Batchelor, A. H., Chang, C.-P., Cleary, M. L., and Wolberger, C. (1999). Structure of a HoxB1-Pbx1 Heterodimer Bound to DNA. *Cell* 96, 587–597. doi:10.1016/s0092-8674(00)80662-5

Ponting, C. P., and Russell, R. R. (2002). The Natural History of Protein Domains. *Annu. Rev. Biophys. Biomol. Struct.* 31, 45–71. doi:10.1146/annurev.biophys.31.082901.134314

Pöpperl, H., Bienz, M., Studer, M., Chan, S.-K., Aparicio, S., Brenner, S., et al. (1995). Segmental Expression of Hoxb-1 Is Controlled by a Highly Conserved Autoregulatory Loop Dependent upon Exd/pbx. *Cell* 81, 1031–1042. doi:10.1016/s0092-8674(05)80008-x

Quiring, R., Walldorf, U., Kloter, U., and Gehring, W. J. (1994). Homology of the Eyeless Gene of Drosophila to the Small Eye Gene in Mice and Aniridia in Humans. *Science* 265, 785–789. doi:10.1126/science.7914031

Reilly, S. K., and Noonan, J. P. (2016). Evolution of Gene Regulation in Humans. *Annu. Rev. Genom. Hum. Genet.* 17, 45–67. doi:10.1146/annurev-genom-090314-045935

Remacle, S., Abbas, L., De Backer, O., Pacico, N., Gavalas, A., Gofflot, F., et al. (2004). Loss of Function but No Gain of Function Caused by Amino Acid Substitutions in the Hexapeptide of Hoxa1 *In Vivo*. *Mol. Cel Biol* 24, 8567–8575. doi:10.1128/mcb.24.19.8567-8575.2004

Richter, D. J., Fozouni, P., Eisen, M. B., and King, N. (2018). Gene Family Innovation, Conservation and Loss on the Animal Stem Lineage. *Elife* 7, 34226. doi:10.7554/eLife.34226

Rivas, M. L., Espinosa-Vázquez, J. M., Sambrani, N., Greig, S., Merabet, S., Graba, Y., et al. (2013). Antagonism versus Cooperativity with TALE Cofactors at the Base of the Functional Diversification of Hox Protein Function. *Plos Genet.* 9, e1003252. doi:10.1371/journal.pgen.1003252

Roberts Kingman, G. A., Lee, D., Jones, F. C., Desmet, D., Bell, M. A., and Kingsley, D. M. (2021a). Longer or Shorter Spines: Reciprocal Trait Evolution in Stickleback via Triallelic Regulatory Changes in Stanniocalcin2a. *Proc. Natl. Acad. Sci. U S A.* 118, 694118. doi:10.1073/pnas.2100694118

Roberts Kingman, G. A., Vyas, D. N., Jones, F. C., Brady, S. D., Chen, H. I., Reid, K., et al. (2021b). Predicting Future from Past: The Genomic Basis of Recurrent and Rapid Stickleback Evolution. *Sci. Adv.* 7, 5285. doi:10.1126/sciadv.abg5285

Rogers, J., and Gibbs, R. A. (2014). Comparative Primate Genomics: Emerging Patterns of Genome Content and Dynamics. *Nat. Rev. Genet.* 15, 347–359. doi:10.1038/nrg3707

Ronshaugen, M., Mcginnis, N., and Mcginnis, W. (2002). Hox Protein Mutation and Macroevolution of the Insect Body Plan. *Nature* 415, 914. doi:10.1038/nature716

Rubinstein, M., and De Souza, F. S. J. (2013). Evolution of Transcriptional Enhancers and Animal Diversity. *Phil. Trans. R. Soc. B* 368, 20130017. doi:10.1098/rstb.2013.0017

Ryoo, H. D., Marty, T., Casares, F., Affolter, M., and Mann, R. S. (1999). Regulation of Hox Target Genes by a DNA Bound Homothorax/Hox/Extradenticle Complex. *Development* 126, 5137–5148. doi:10.1242/dev.126.22.5137

Saadaoui, M., Merabet, S., Litim-Mecheri, I., Arbeille, E., Sambrani, N., Damen, W., et al. (2011). Selection of Distinct Hox-Extradenticle Interaction Modes fine-tunes Hox Protein Activity. *Proc. Natl. Acad. Sci. U.S.A.* 108, 2276–2281. doi:10.1073/pnas.1006964108

Sakazume, S., Sorokina, E., Iwamoto, Y., and Semina, E. V. (2007). Functional Analysis of Human Mutations in Homeodomain Transcription Factor PITX3. *BMC Mol. Biol.* 8, 84. doi:10.1186/1471-2199-8-84

Sánchez-Higueras, C., Sotillos, S., and Castelli-Gair Hombría, J. (2014). Common Origin of Insect Trachea and Endocrine Organs from a Segmentally Repeated Precursor. *Curr. Biol.* 24, 76–81. doi:10.1016/j.cub.2013.11.010

Sandve, S. R., Rohlfs, R. V., and Hvidsten, T. R. (2018). Subfunctionalization versus Neofunctionalization after Whole-Genome Duplication. *Nat. Genet.* 50, 908–909. doi:10.1038/s41588-018-0162-4

Saurin, A. J., Delfini, M. C., Maurel-Zaffran, C., and Graba, Y. (2018). The Generic Facet of Hox Protein Function. *Trends Genet.* 34, 941–953. doi:10.1016/j.tig.2018.08.006

Scannell, D. R., Byrne, K. P., Gordon, J. L., Wong, S., and Wolfe, K. H. (2006). Multiple Rounds of Speciation Associated with Reciprocal Gene Loss in Polyploid Yeasts. *Nature* 440, 341–345. doi:10.1038/nature04562

Schneuwly, S., Klemenz, R., and Gehring, W. J. (1987). Redesigning the Body Plan of Drosophila by Ectopic Expression of the Homoeotic Gene Antennapedia. *Nature* 325, 816–818. doi:10.1038/325816a0

Scott, M. P., Tamkun, J. W., and Hartzell III, G. W. (1989). 3dThe Structure and Function of the Homeodomain. *Biochim. Biophys. Acta (Bba) - Rev. Cancer* 989, 25–48. doi:10.1016/0304-419x(89)90033-4

Sekigami, Y., Kobayashi, T., Omi, A., Nishitsuji, K., Ikuta, T., Fujiyama, A., et al. (2017). Hox Gene Cluster of the Ascidian, Halocynthia Roretzi, Reveals Multiple Ancient Steps of Cluster Disintegration during Ascidian Evolution. *Zoolog. Lett* 3, 17. doi:10.1186/s40851-017-0078-3

Seo, H.-C., Edvardsen, R. B., Maeland, A. D., Bjordal, M., Jensen, M. F., Hansen, A., et al. (2004). Hox Cluster Disintegration with Persistent Anteroposterior Order of Expression in *Oikopleura dioica*. *Nature* 431, 67–71. doi:10.1038/nature02709

Session, A. M., Uno, Y., Kwon, T., Chapman, J. A., Toyoda, A., Takahashi, S., et al. (2016). Genome Evolution in the Allotetraploid Frog *Xenopus laevis*. *Nature* 538, 336–343. doi:10.1038/nature19840

Shanmugam, K., Green, N. C., Rambaldi, I., Saragovi, H. U., and Featherstone, M. S. (1999). PBX and MEIS as Non-DNA-binding Partners in Trimeric Complexes with HOX Proteins. *Mol. Cel Biol* 19, 7577–7588. doi:10.1128/mcb.19.11.7577

Shen, W.-F., Rozenfeld, S., Lawrence, H. J., and Largman, C. (1997b). The Abd-B-like Hox Homeodomain Proteins Can Be Subdivided by the Ability to Form Complexes with Pbx1a on a Novel DNA Target. *J. Biol. Chem.* 272, 8198–8206. doi:10.1074/jbc.272.13.8198

Shen, W. F., Montgomery, J. C., Rozenfeld, S., Moskow, J. J., Lawrence, H. J., Buchberg, A. M., et al. (1997a). AbdB-like Hox Proteins Stabilize DNA Binding by the Meis1 Homeodomain Proteins. *Mol. Cel Biol* 17, 6448–6458. doi:10.1128/mcb.17.11.6448

Shoubridge, C., Tan, M. H., Seiboth, G., and Gecz, J. (2012). ARX Homeodomain Mutations Abolish DNA Binding and lead to a Loss of Transcriptional Repression. *Hum. Mol. Genet.* 21, 1639–1647. doi:10.1093/hmg/ddr601

Singh, N. P., De Kumar, B., Paulson, A., Parrish, M. E., Scott, C., Zhang, Y., et al. (2021). Genome-Wide Binding Analyses of HOXB1 Revealed a Novel DNA Binding Motif Associated with Gene Repression. *J. Dev. Biol.* 9, 6. doi:10.3390/jdb9010006

Singh, N. P., De Kumar, B., Paulson, A., Parrish, M. E., Zhang, Y., Florens, L., et al. (2020). A Six-Amino-Acid Motif Is a Major Determinant in Functional Evolution of HOX1 Proteins. *Genes Dev.* 34, 1680–1696. doi:10.1101/gad.342329.120

Slattery, M., Riley, T., Liu, P., Abe, N., Gomez-Alcala, P., Dror, I., et al. (2011). Cofactor Binding Evokes Latent Differences in DNA Binding Specificity between Hox Proteins. *Cell* 147, 1270–1282. doi:10.1016/j.cell.2011.10.053

Smith, J. J., Timoshevskaya, N., Ye, C., Holt, C., Keinath, M. C., Parker, H. J., et al. (2018). The Sea Lamprey Germline Genome Provides Insights into Programmed Genome Rearrangement and Vertebrate Evolution. *Nat. Genet.* 50, 270–277. doi:10.1038/s41588-017-0036-1

Sorge, S., Ha, N., Polychronidou, M., Friedrich, J., Bezdan, D., Kaspar, P., et al. (2012). Thecis-regulatory Code of Hox Function inDrosophila. *EMBO J.* 31, 3323–3333. doi:10.1038/emboj.2012.179

Soshnikova, N., Dewaele, R., Janvier, P., Krumlauf, R., and Duboule, D. (2013). Duplications of Hox Gene Clusters and the Emergence of Vertebrates. *Develop. Biol.* 378, 194–199. doi:10.1016/j.ydbio.2013.03.004

Spitz, F., Gonzalez, F., Peichel, C., Vogt, T. F., Duboule, D., and Zákány, J. (2001). Large Scale Transgenic and Cluster Deletion Analysis of the HoxD Complex Separate an Ancestral Regulatory Module from Evolutionary Innovations. *Genes Dev.* 15, 2209–2214. doi:10.1101/gad.205701

Suemori, H., and Noguchi, S. (2000). Hox C Cluster Genes Are Dispensable for Overall Body Plan of Mouse Embryonic Development. *Develop. Biol.* 220, 333–342. doi:10.1006/dbio.2000.9651

Truong, D. M., and Boeke, J. D. (2017). Resetting the Yeast Epigenome with Human Nucleosomes. *Cell* 171, 1508–1519. doi:10.1016/j.cell.2017.10.043

Tvrdik, P., and Capecchi, M. R. (2006). Reversal of Hox1 Gene Subfunctionalization in the Mouse. *Develop. Cel* 11, 239–250. doi:10.1016/j.devcel.2006.06.016

Van Den Akker, E., Fromental-Ramain, C., De Graaff, W., Le Mouellic, H., Brulet, P., Chambon, P., et al. (2001). Axial Skeletal Patterning in Mice Lacking All Paralogous Group 8 Hox Genes. *Development* 128, 1911–1921. doi:10.1242/dev.128.10.1911

Van Dijk, M. A., and Murre, C. (1994). *Extradenticle* Raises the DNA Binding Specificity of Homeotic Selector Gene Products. *Cell* 78, 617–624. doi:10.1016/0092-8674(94)90526-6

Vandenbussche, M., Theissen, G., Van De Peer, Y., and Gerats, T. (2003). Structural Diversification and Neo-Functionalization during floral MADS-Box Gene Evolution by C-Terminal Frameshift Mutations. *Nucleic Acids Res.* 31, 4401–4409. doi:10.1093/nar/gkg642

Vandepoele, K., De Vos, W., Taylor, J. S., Meyer, A., and Van De Peer, Y. (2004). Major Events in the Genome Evolution of Vertebrates: Paranome Age and Size Differ Considerably between ray-finned Fishes and Land Vertebrates. *Proc. Natl. Acad. Sci. U.S.A.* 101, 1638–1643. doi:10.1073/pnas.0307968100

Vieux-Rochas, M., Mascrez, B., Krumlauf, R., and Duboule, D. (2013). Combined Function of HoxA and HoxB Clusters in Neural Crest Cells. *Develop. Biol.* 382, 293–301. doi:10.1016/j.ydbio.2013.06.027

Wagner, G. P., Amemiya, C., and Ruddle, F. (2003). Hox Cluster Duplications and the Opportunity for Evolutionary Novelties. *Proc. Natl. Acad. Sci. U.S.A.* 100, 14603–14606. doi:10.1073/pnas.2536656100

Wahba, G. M., Hostikka, S. L., and Carpenter, E. M. (2001). The Paralogous Hox Genes Hoxa10 and Hoxd10 Interact to Pattern the Mouse Hindlimb Peripheral Nervous System and Skeleton. *Develop. Biol.* 231, 87–102. doi:10.1006/dbio.2000.0130

Walsh, B. (2003). Population-genetic Models of the Fates of Duplicate Genes. *Genetica* 118, 279–294. doi:10.1007/978-94-010-0229-5_16

Webb, B. D., Shaaban, S., Gaspar, H., Cunha, L. F., Schubert, C. R., Hao, K., et al. (2012). HOXB1 Founder Mutation in Humans Recapitulates the Phenotype of Hoxb1 Mice. *Am. J. Hum. Genet.* 91, 171–179. doi:10.1016/j.ajhg.2012.05.018

Wendel, J. F. (2000). Genome Evolution in Polyploids. *Plant Mol. Biol.* 42, 225–249. doi:10.1007/978-94-011-4221-2_12

Xie, K. T., Wang, G., Thompson, A. C., Wucherpfennig, J. I., Reimchen, T. E., Maccoll, A. D. C., et al. (2019). DNA Fragility in the Parallel Evolution of Pelvic Reduction in Stickleback Fish. *Science* 363, 81–84. doi:10.1126/science.aan1425

Yokouchi, Y., Nakazato, S., Yamamoto, M., Goto, Y., Kameda, T., Iba, H., et al. (1995). Misexpression of *Hoxa-13* Induces Cartilage Homeotic Transformation and Changes Cell Adhesiveness in Chick Limb Buds. *Genes Dev.* 9, 2509–2522. doi:10.1101/gad.9.20.2509

Young, T., Rowland, J. E., Van De Ven, C., Bialecka, M., Novoa, A., Carapuco, M., et al. (2009). Cdx and Hox Genes Differentially Regulate Posterior Axial Growth in Mammalian Embryos. *Develop. Cel* 17, 516–526. doi:10.1016/j.devcel.2009.08.010

Zappavigna, V., Sartori, D., and Mavilio, F. (1994). Specificity of HOX Protein Function Depends on DNA-Protein and Protein-Protein Interactions, Both Mediated by the Homeo Domain. *Genes Dev.* 8, 732–744. doi:10.1101/gad.8.6.732

Zhao, J. J., Lazzarini, R. A., and Pick, L. (1993). The Mouse Hox-1.3 Gene Is Functionally Equivalent to the Drosophila Sex combs Reduced Gene. *Genes Dev.* 7, 343–354. doi:10.1101/gad.7.3.343

Zhao, Y., and Potter, S. S. (2002). Functional Comparison of the Hoxa 4, Hoxa 10, and Hoxa 11 Homeoboxes. *Develop. Biol.* 244, 21–36. doi:10.1006/dbio.2002.0595

Zhao, Y., and Potter, S. S. (2001). Functional Specificity of theHoxa13homeobox. *Development* 128, 3197–3207. doi:10.1242/dev.128.16.3197

Zhen, Y., and Andolfatto, P. (2012). Methods to Detect Selection on Noncoding DNA. *Methods Mol. Biol.* 856, 141–159. doi:10.1007/978-1-61779-585-5_6

Zimmer, C. T., Garrood, W. T., Singh, K. S., Randall, E., Lueke, B., Gutbrod, O., et al. (2018). Neofunctionalization of Duplicated P450 Genes Drives the Evolution of Insecticide Resistance in the Brown Planthopper. *Curr. Biol.* 28, 268–274. doi:10.1016/j.cub.2017.11.060

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership