



**HAL**  
open science

# Décryptage des interactions moléculaires entre les protéines HOX et leurs partenaires

Amélie Dard

► **To cite this version:**

Amélie Dard. Décryptage des interactions moléculaires entre les protéines HOX et leurs partenaires. Biologie cellulaire. Université de Lyon, 2016. Français. NNT : 2016LYSE1177 . tel-01458348

**HAL Id: tel-01458348**

**<https://theses.hal.science/tel-01458348>**

Submitted on 6 Feb 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



**THESE de DOCTORAT DE L'UNIVERSITE DE LYON**

opérée au sein de

**l'Université Claude Bernard Lyon 1**

**Ecole Doctorale N°340**

**Biologie Moléculaire et Intégrative de la Cellule**

**Discipline : Sciences de la vie**

Soutenue publiquement le 13/10/2016, par :

**Amélie Dard**

---

**Décryptage des interactions moléculaires entre les  
protéines HOX et leurs partenaires.**

---

Devant le jury composé de :

M Yacine GRABA	Université Aix/Marseille	Rapporteur
M René REZSOHAZY	Université Catholique de Louvain	Rapporteur
Mme Françoise BLEICHER	Université de Claude Bernard Lyon I	Examinatrice
Mme Jacqueline DESCHAMPS	Hubrecht University	Examinatrice
M Samir MERABET	ENS Lyon	Directeur de thèse



RESUME .....	5
SUMMARY.....	6
REMERCIEMENTS .....	7
ABREVIATIONS .....	10
INTRODUCTION .....	11
1. DECOUVERTE DES PROTEINES HOX .....	11
2. ROLES DES PROTEINES HOX.....	12
2.1. LES PROTEINES HOX SONT DES FACTEURS DE TRANSCRIPTION.....	12
2.1.1. <i>LA REGULATION DE LA TRANSCRIPTION CHEZ LES EUCARYOTES</i> .....	12
2.1.1.1. <i>LES FACTEURS DE TRANSCRIPTION (FTs)</i> .....	12
2.1.1.2. <i>LES SEQUENCES CIS REGULATRICES</i> .....	13
2.1.1.3. <i>LES SITES DE RECONNAISSANCE A L'ADN : NOUVELLES DONNEES</i> .....	13
2.1.2. <i>ROLE DES PROTEINES HOX AU COURS DU DEVELOPPEMENT EMBRYONNAIRE</i> .....	13
2.1.3. <i>ROLE DES PROTEINES HOX POUR LE MAINTIEN DE L'HOMEOSTASIE CELLULAIRE CHEZ L'ADULTE</i> .....	14
2.2. LES PROTEINES HOX ONT AUSSI D'AUTRES FONCTIONS .....	17
2.2.1. <i>TRADUCTION</i> .....	17
2.2.1.1. <i>ACTIVATION DE LA TRADUCTION</i> .....	17
2.2.1.2. <i>ADENYLATION DES QUEUES POLY-A</i> .....	18
2.2.2. <i>REPLICATION DE L'ADN</i> .....	18
2.2.2.1. <i>INITIATION DE LA REPLICATION</i> .....	19
2.2.2.2. <i>REGULATION DE LA TRANSCRIPTION PAR UBIQUITINATION</i> .....	19
2.2.3. <i>DEGRADATION PROTEIQUE</i> .....	20
2.2.3.1. <i>DEGRADATION DE L'UBIQUITINE LIGASE RCHY1</i> .....	20
2.2.3.2. <i>DEGRADATION DE LA CYLCINE D1</i> .....	21
2.2.4. <i>REPARATION DE L'ADN</i> .....	21
2.2.5. <i>LES PROTEINES HOX POURRAIENT AVOIR ENCORE D'AUTRES FONCTIONS MOLECULAIRES</i> .....	21

3. STRUCTURE DES PROTEINES HOX .....	22
3.1. STRUCTURE GLOBALE DES PROTEINES HOX .....	22
3.1.1. HOMEODOMAINE (HD).....	22
3.1.2. LE MOTIF HEXAPEPTIDE (HX) .....	24
3.2. AUTRES MOTIFS DES PROTEINES HOX .....	25
3.2.1. QU'EST CE QU'UN « SHORT LINEAR MOTIF » ? .....	25
3.2.2. LE MOTIF SSYF .....	26
3.2.3. LE MOTIF UBDA .....	27
3.2.4. LE MOTIF TDWM.....	27
3.2.5. UTILISATION COMPLEXE DES SLIMS AU SEIN D'UNE PROTEINE HOX.....	28
4. LES PARTENAIRES DES PROTEINES HOX .....	28
4.1. LES TECHNIQUES DE VISUALISATION DES IPPS .....	28
4.1.1. QUELQUES TECHNIQUES CLASSIQUES .....	28
4.1.2. LA COMPLEMENTATION DE FLUORESCENCE BIMOLECULAIRE (BiFC).....	29
4.1.3. UTILISATIONS DE LA BiFC POUR LA RECHERCHE DE NOUVEAUX PARTENAIRES.....	31
4.2. LES COFACTEURS A HD DE LA FAMILLE TALE.....	31
4.2.1. LA FAMILLE TALE.....	32
4.2.2. LA FORMATION DES COMPLEXES HOX/PBX/MEIS .....	33
4.3. AUTRES PARTENAIRES DES PROTEINES HOX .....	34
4.3.1. DES FACTEURS DE TRANSCRIPTION .....	34
4.3.2. DES REGULATEURS DE LA CHROMATINE .....	34
4.3.3. AUTRES .....	34
<b>RESULTATS .....</b>	<b>37</b>
1. AUX ORIGINES DES COMPLEXES HOX-TALE .....	39
2. DES PETITS MOTIFS POUR REGULER LE POTENTIEL D'INTERACTION DES PROTEINES HOX .....	65
3. IDENTIFIER DE NOUVELLES INTERFACES POUR LES COFACTEURS TALE AU SEIN DES PROTEINES HOX HUMAINES .....	95
4. TROUVER DE NOUVEAUX PARTENAIRES DES PROTEINES HOX SPECIFIQUES D'UN CONTEXTE CANCEREUX .....	129
4.1. LE CANCER : UN CONTEXTE PROPICE A L'ETUDE DES PROTEINES HOX .....	139
4.1.1. SEIN.....	139
4.1.2. PROSTATE .....	139

4.1.3. OVAIRE .....	140
4.2. LE MODELE .....	139
4.2.1. LA BANQUE CCORF .....	142
4.2.2. EXPRESSION DE LA PROTEINE HOX DANS LES DIFFERENTES LIGNEES ETABLIES....	144
4.2.3. IDENTIFICATION DES CANDIDATS PAR SEQUENÇAGE .....	147
4.2.4. CONFIRMATION DES CANDIDATS EN BiFC.....	149
4.2.5. L'ECHEC DES PRECEDENTS CRIBLES .....	151
<b>DISCUSSION .....</b>	<b>153</b>
1. VERS DE NOUVELLES INTERFACES D'INTERACTION .....	153
1.1. POURQUOI LE MOTIF HX EST LONGTEMPS RESTE L'UNIQUE MOTIF D'INTERACTION DES COFACTEURS TALE AU SEIN DES PROTEINES HOX ?.....	153
1.2. LES PROTEINES HOX UTILISENT DES INTERFACES PARALOGUES SPECIFIQUES ADDITIONNELLES POUR INTERAGIR AVEC LES COFACTEURS TALE .....	153
1.3. VERS UNE VALIDATION FONCTIONNELLE DE LA DISSECTION MOLECULAIRE ...	155
1.4. INTERET DES DIFFERENTS MODES D'INTERACTION HOX-TALE .....	155
2. RECHERCHE DE NOUVEAUX PARTENAIRES DES PROTEINES HOX HUMAINES EN CONTEXTE ONCOGENIQUE .....	157
2.1. AMELIORER LA REPRESENTATIVITE DE LA BANQUE.....	157
2.2. ETENDRE LE CRIBLE A DIFFERENTES LIGNEES.....	157
2.3. REALISER LE CRIBLE AVEC DIFFERENTES PROTEINES HOX SAUVAGES OU MUTEES.....	158
2.4. APPORTER LE CANDIDAT PAR LES VLPs .....	159
2.5. LA BiFC AVEC LA SF GFP : RAPIDITE DE MATURATION.....	159
2.6. ETABLIR DES PARAMETRES DE SEQUENÇAGE OPTIMAUX .....	159
2.7. VERS UNE BANQUE ORDONNEE DE CANDIDATS .....	160
2.8. LE CHOIX DU PROMOTEUR .....	160
<b>MATERIELS ET METHODES.....</b>	<b>163</b>
<b>BIBLIOGRAPHIE.....</b>	<b>167</b>
<b>ANNEXES .....</b>	<b>185</b>





**THESE de DOCTORAT DE L'UNIVERSITE DE LYON**

opérée au sein de

**l'Université Claude Bernard Lyon 1**

**Ecole Doctorale N°340**

**Biologie Moléculaire et Intégrative de la Cellule**

**Discipline : Sciences de la vie**

Soutenue publiquement le 13/10/2016, par :

**Amélie Dard**

---

**Décryptage des interactions moléculaires entre les  
protéines HOX et leurs partenaires.**

---

Devant le jury composé de :

M Yacine GRABA	Université Aix/Marseille	Rapporteur
M René REZSOHAZY	Université Catholique de Louvain	Rapporteur
Mme Françoise BLEICHER	Université de Claude Bernard Lyon I	Examinatrice
Mme Jacqueline DESCHAMPS	Hubrecht University	Examinatrice
M Samir MERABET	ENS Lyon	Directeur de thèse





*Pour celle qui n'a qu'une seule aile.*

*La créativité est contagieuse, faites-la circuler.*

Albert Einstein



## RESUME

---

Les gènes Hox sont présents dans la grande majorité des espèces du règne animal, des cnidaires à l'Homme. Ils sont nécessaires à la différenciation coordonnée des cellules le long de différents axes longitudinaux au cours du développement embryonnaire. Ils sont aussi impliqués dans l'organogénèse et le maintien de l'homéostasie de nombreux tissus à l'âge adulte. Des mutations affectant leur expression et/ou leur fonction sont ainsi retrouvées dans de nombreux cancers chez l'Homme.

Les gènes Hox codent pour des facteurs de transcription reconnaissant des séquences nucléotidiques courtes et très similaires grâce à un domaine extrêmement conservé : l'homéodomaine. L'interaction avec deux autres classes de facteur de transcription, les protéines Pbx et Meis, permet aux protéines Hox de reconnaître des sites plus spécifiques. L'interaction directe établie entre les protéines Hox et le cofacteur Pbx a d'abord été décrite pour dépendre d'un petit motif très conservé au sein des protéines Hox, l'hexapeptide (HX). Plusieurs analyses fonctionnelles et moléculaires ont montré par la suite que ce mode d'interaction pouvait être bien plus complexe. Ainsi, le motif HX peut-être dispensable pour le recrutement du cofacteur Pbx en présence du partenaire Meis *in vitro*. Ces observations ont été confirmées *in vivo* par la méthode de complémentation de fluorescence bimoléculaire. Ces derniers résultats suggèrent que les protéines Hox pourraient utiliser d'autres motifs pour interagir avec les cofacteurs Pbx/Meis. De plus, l'extrême conservation du motif HX suggère que celui-ci pourrait aussi être important pour établir et/ou réguler d'autres interactions.

Mon travail de thèse s'inscrit dans la problématique du rôle du motif HX ainsi que d'autres motifs dans les interactions Hox-cofacteur. Plus particulièrement, j'ai développé deux projets principaux qui ont utilisé des protéines humaines et des lignées cellulaires dérivées de différents tissus comme système d'étude. Un premier projet a consisté à réaliser une analyse systématique du mode d'interaction de l'ensemble des représentants Hox des différents groupes de paralogie avec leurs cofacteurs Pbx/Meis. Cette analyse a permis de révéler de nouveaux modes d'interaction pour la grande majorité des protéines Hox, remettant en question la plupart des approches développées actuellement pour bloquer l'activité oncogénique des complexes Hox/Pbx/Meis.

Mon deuxième projet a consisté à mettre en place un nouveau système de crible moléculaire pour identifier des partenaires de la protéine humaine HoxA9 sauvage ou mutée dans son motif HX dans différentes lignées cellulaires.

Les résultats obtenus montrent l'extrême complexité du mode opératoire des protéines Hox humaines avec leurs cofacteurs Pbx/Meis. Cette complexité est certainement à la base des fonctions spécifiques de chaque complexe Hox/Pbx/Meis *in vivo*. En parallèle, j'ai pu observer des interactions préférentielles entre la protéine HoxA9 et d'autres partenaires transcriptionnels dans les lignées cellulaires utilisées, certaines dérivant de cancers différents. Cette plasticité moléculaire étonnante est probablement une propriété commune à l'ensemble des protéines Hox pour leur permettre d'acquies des fonctions tissu-spécifiques.

L'ensemble de mon travail de thèse ouvre ainsi de nouvelles perspectives sur notre compréhension du mode moléculaire d'action des protéines Hox et de leurs cofacteurs au sens large du terme, que cela soit en contexte développemental normal ou pathologique.

## SUMMARY

---

Hox genes are present in the vast majority of the animal kingdom, from cnidarians to human. These genes are required for the differentiation of several longitudinal axes during embryogenesis. There are also involved in the formation of different organs and for the homeostasis of several tissues in the adult organism. Accordingly, mutations affecting their expression and/or function are found in numerous human cancers.

Hox genes encode for transcription factors that recognize short and highly similar DNA-binding sites by using an extremely conserved DNA-binding domain called the homeodomain. The interaction with two different classes of transcription factors, the Pbx and Meis proteins, allows Hox proteins to recognize longer therefore more specific DNA-binding sites. The direct interaction between Hox proteins and the Pbx cofactor was first described to rely on a common short Hox protein motif called hexapeptide (HX). However, subsequent functional and molecular analyses showed that this interaction mode could be much more complicated. In particular, the HX motif can be dispensable for the interaction with Pbx in the presence of the third Meis partner *in vitro*. These observations were raised from the study with cnidarian, fly and mouse Hox proteins and further confirmed *in vivo* by the bimolecular fluorescence complementation approach. These last results strongly suggest that Hox proteins could use different motifs to interact with the Pbx/Meis cofactors. Still, only two such motifs have been described in only two different *Drosophila* Hox proteins to date. Moreover, the strong conservation of the HX motif in animal Hox proteins suggests that it could be involved for the interaction with other partners than Pbx.

My thesis work is dedicated to the issue of the role of the HX motif and other short motifs in Hox-cofactor interactions. More particularly, I developed two main projects using human Hox proteins and cell lines derived from different tissues as a model system. My first project consisted in the systematic analysis of the interaction property of all Hox paralogs with the Pbx/Meis cofactors. This work revealed new Hox-Pbx-Meis interaction properties for the vast majority of Hox proteins, which challenges the current molecular strategies deployed to block the activity of Hox/Pbx/Meis complexes.

My second project consisted in establishing a new molecular screen to identify transcriptional partners of the wild type or HX-mutated human HoxA9 protein in different cell lines.

Results obtained show the extreme complexity of the molecular mode of action of Hox proteins with their Pbx/Meis cofactors. This complexity is likely at the basis of the specific function of each Hox/Pbx/Meis complex *in vivo*. In addition, I identified different preferential interactions between HoxA9 and transcriptional partners, depending on the cell line considered. Two of these cell lines derived from different cancers. This striking molecular plasticity is probably a common property of Hox proteins, allowing them to acquire tissue-specific functions.

Overall, my thesis work opens new perspectives into our understanding of the molecular mode of action of Hox proteins and their cofactors, in a normal or pathological developmental context.

## REMERCIEMENTS

---

Je tenais d'abord à remercier **Samir** pour son encadrement, sa présence et son endurance tout au long de ma thèse. Quelques coups de pieds au cul, des doutes et du stress mais finalement pas mal de résultats et beaucoup, beaucoup, beaucoup de clonages. C'était un projet casse-gueule, du cancer humain avec un chef drosophiliste, seule pendant deux ans et demi, mais on y est arrivés et je suis fière de ce que nous avons accomplis ensemble. Ce projet de brevet qui se concrétise sonne la récompense de ces quatre années de labeur. Merci d'avoir cru en moi dès le M2, de m'avoir soutenue dans mes projets de thèse, et de m'avoir emmenée dans tes valises pour cette aventure lyonnaise. Je t'ai vu devenir un chef d'équipe au fil de ces 6 ans (eh oui, 6 ans...) avec des hauts, des bas, et je ne peux que te souhaiter une belle suite pour ta carrière, que ton ange gardien reste au-dessus de ton épaule comme il l'a toujours si bien fait.

Je voulais ensuite remercier mes rapporteurs **Yacine Graba** et **René Rezsöhazi**. Yacine pour m'avoir accueillie dans son laboratoire quand je n'étais qu'une petite M2 à Marseille. C'était pour moi une conclusion plus que logique à ma thèse que de t'inviter à ma soutenance. René pour les discussions au cours du congrès COST, avec toi et toute ton équipe. Ton expertise dans le domaine m'a aussi poussée à te demander de participer à ce grand moment. Je remercie aussi **Jacqueline Deschamps** d'avoir accepté de venir assister à ma thèse pour une rencontre lyonnaise entre chercheurs Hox.

Je remercie **Vincent Laudet** sans qui je n'aurais peut-être jamais pu faire de thèse à l'IGFL. Je sais combien il a poussé pour que je passe le concours BMIC et que j'obtienne mon financement. Petit souvenir mémorable d'une bière avec le film « Nos jours heureux » en salle de réunion.

Merci à toute l'équipe administrative de l'IGFL, **Martine, Joanne, Fabienne et Sonia** qui sont d'une efficacité redoutable.

Je remercie les membres de mon comité de suivi de thèse. **Els Verhoeyen** pour ses conseils sur les cellules hématopoïétiques et son écoute ; **Catherine Lavau** pour toute son aide sur les projets souris, sans elle il manquerait une grosse partie de mes résultats ; **François Morlé** pour ses remarques et son aide sur le projet ; et **Philippe Mangeot** pour toute sa science, ses idées extraordinaires, son infatigable créativité en matière d'outils biologiques.

Merci à mon équipe, **Marilyne, Leïore, Françoise**, pour l'aide aux manips, l'écoute, les questions et les critiques. Françoise en particulier dont l'arrivée dans l'équipe m'a soulagée et beaucoup aidée. Et puis aussi pour sa maison dans le Pilat et les bons légumes du jardin.

Je n'oublie pas mon armée de furbys (*oui, ces stagiaires qui, lorsqu'on les place face à face, discutent sans discontinuer*) : **Lorette, Adrien, Claudine, Claire-Emmanuelle, Hélène, Claire et Thomas, Jonathan, Solène**. Parce que certains, même des mois après la fin de leur stage, me font encore des gâteaux (*Miss Caillot, aka La Différence*). **Supernathan**, merci d'avoir été le furby de conclusion de tout ça, j'ai eu de la chance de tomber sur toi (*Ton nom est Puanteur !*).

J'ai forcément une pensée pour **Guillaume**, collègue de galère, de la manip qui foire, du shift qui gondole, de la radioac merdique et des clonages daubés. Guillaume le thésard toujours actif avec son calepin à la main, son bac polystyrène, et ses éternels TOCs. Moi aussi je n'oublie pas la réponse universelle, le **52**, mais je n'oublie pas non plus les roulades krav maga, le petit chou, la quiche lorraine, Her Holzer, la poubelle jaune, le lancer de gants, les salons du vigneron, Marie-Antoinette et super platy moche.

**Melanie, Coraline**, mes amies de craquage fin de thèse. Merci pour votre folie, tellement normale finalement, qui fait du bien et qui rassure. Vous êtes deux licornes, l'une qui fait de la neige, l'autre avec une fleur dans la crinière. *Dans mon monde à moi...* Et je n'oublie pas non plus notre ex-







## ABREVIATIONS

---

**BiFC** : Complémentation de Fluorescence Bimoléculaire

**CC** : Cerulean C-terminal

**CN** : Cerulean N-terminal

**FT** : Facteur de transcription

**HD** : Homéodomaine

**HX** : Hexapeptide

**IPP** : Interaction Protéine-Protéine

**MEIS** : Myeloid Ecotropic Integration Site

**MOI** : Multiplicity of Infection

**ORF** : Open Reading Frame

**PBC** : Pré-B cell leukemia transcription factor class

**PBX** : Pré-B cell leukemia homéobox

**SLiM** : Short Linear Motif

**TALE** : Three amino acid loop extension

**VC** : Venus C-terminal

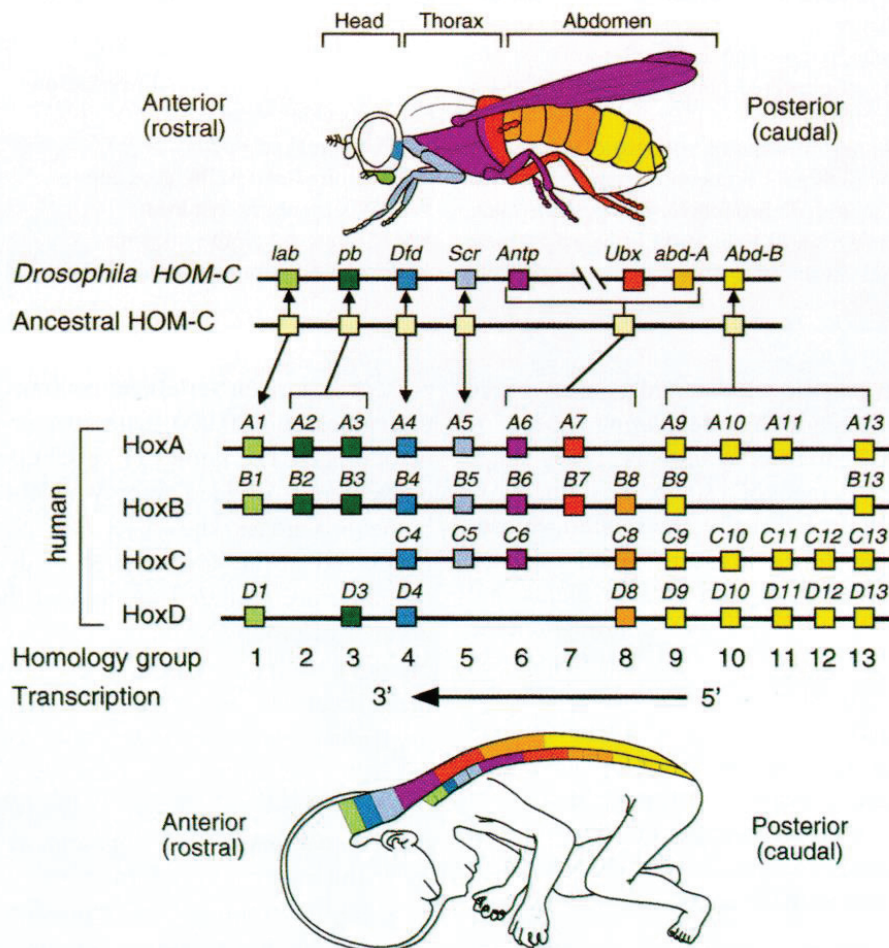
**VLP** : Virus Like Particule

**VN** : Venus N-terminal

# INTRODUCTION

## 1. DECOUVERTE DES PROTEINES HOX

La découverte d'un petit répertoire de molécules très conservées et capables de contrôler le développement embryonnaire des métazoaires fut l'une des plus grandes avancées de la biologie. Parmi celles-ci, les protéines HOX jouent un rôle primordial dans la spécification des différentes lignées cellulaires le long de l'axe antéro-postérieur de tous les animaux ayant une symétrie bilatérale. Elles fascinent ainsi les biologistes par leur extrême conservation évolutive et aussi par la caractérisation de mutants homéotiques impressionnants (Lewis, 1978 ; Bender *et al.*, 1983 ; Krumlauf, 1994 ; Maconochie *et al.*, 1996). De plus, elles possèdent une propriété que l'on appelle **colinéarité spatiale** (Figure 1) : en effet, les protéines HOX ont un profil d'expression régionalisé le long de l'axe antéro-postérieur qui peut suivre l'organisation des gènes sur le chromosome. (Harding *et al.*, 1985 ; McGinnis et Krumlauf, 1992).



**Figure 1 :** Organisation génomique et profil d'expression des gènes *HOX* De la *Drosophile* et de l'Homme. (Mark *et al.*, 1997).

Ce profil d'expression régionalisé est important pour le rôle des protéines HOX. Ce sont des facteurs de transcription (FTs) dont les fonctions sont spécifiques et différentes le long de

l'axe antéro-postérieur. Cette spécificité fonctionnelle est démontrée par une analyse transcriptomique de six protéines HOX différentes au cours de l'embryogénèse de la drosophile : 68,9% des gènes identifiés ne sont régulés que par une seule protéine Hox et seulement 1,3% sont des cibles communes à toutes les protéines HOX (**Hueber, 2007**). Cette découverte est d'autant plus déroutante que le domaine de reconnaissance à l'ADN des protéines HOX est lui-même très conservé et capable de reconnaître des séquences nucléotidiques très similaires (**Noyes et al., 2008**).

Leurs rôles, aussi variés que spécifiques, fascinent et interrogent les biologistes. Comment des protéines aussi conservées sont capables de réguler des processus cellulaires aussi variés avec une finesse extraordinaire ? Comment, alors qu'elles ne sont capables que de reconnaître un site quasi identique, deux protéines différentes peuvent réguler des gènes drastiquement différents ? Quels partenaires protéiques peuvent expliquer cette richesse du vivant qui fait de ces architectes ancestraux des acteurs majeurs de l'immense variété du monde animal ?

## 2. ROLES DES PROTEINES HOX

### 2.1. LES PROTEINES HOX SONT DES FACTEURS DE TRANSCRIPTION

Le principal rôle décrit des protéines HOX est celui de FT. Nous allons dans cette partie développer quelques exemples de ce rôle au cours de l'embryogénèse mais aussi au cours de la vie adulte. En effet, si les protéines HOX sont souvent présentées comme les architectes des animaux au cours des stades précoces de l'embryogénèse pour la mise en place des axes de symétrie, ainsi que pour le développement des différents organes, elles sont aussi essentielles au maintien de l'homéostasie cellulaire chez l'adulte. Des dérèglements de ces FTs cruciaux entraînent de nombreuses pathologies, au premier rang desquelles, des cancers.

#### 2.1.1. *La régulation de la transcription chez les eucaryotes*

##### 2.1.1.1. *Les facteurs de transcription*

Les FTs sont des protéines caractérisées par un domaine de liaison à l'ADN. Celles-ci se lient avec une affinité forte ( $KD < 10^{-8}M$ ) sur des séquences nucléotidiques spécifiques par deux mécanismes distincts : par la reconnaissance de structures formées par les bases du **sillon mineur** de l'ADN, ou par la formation de liaisons hydrogènes entre les résidus du domaine de liaison et des nucléotides du **sillon majeur** de l'ADN (**Rohs, 2009**).

Il existe de nombreuses familles de FTs chacune caractérisée par l'identité de son domaine de liaison à l'ADN : les protéines leucine Zipper, les protéines à doigts de zinc (ZnFinger), *etc.* Une des familles les plus importantes en termes de nombre de représentants chez les animaux est la famille des protéines à homéodomaine (HD). L'HD fait en général 60 acides aminés de long et contient trois hélices alpha ainsi qu'un bras N terminal. La capacité d'interagir avec l'ADN est portée par la troisième hélice ainsi que le bras N-terminal de la structure. Les protéines HOX font partie de cette famille (**Bürklin et Affolter, 2015**). On les retrouve dans la classe ANTP qui est l'une des 16 classes de FTs à HD (ANTP, PRD-LIKE,

PRD, ZF, LIM, POU, CUT, HNF, CERS, PROS, SIX/SO, MKX, IRO, TGIF, MEIS et PBC). Nous décrirons ce domaine plus en détail dans le chapitre III.

A l’instar de l’hexapeptide chez les protéines HOX que nous décrirons aussi dans le chapitre III, les FTs possèdent d’autres motifs en plus de leur domaine de liaison à l’ADN. Ces motifs permettent des interactions protéine-protéines (IPPs) nécessaires à la régulation fine et précise des gènes cibles. En effet un FT peut se fixer sur son site de reconnaissance à l’ADN, puis recruter des partenaires afin de réguler l’expression du gène cible ou recruter en premier des partenaires capables de participer au remodelage de la chromatine afin de rendre des sites accessibles ou inaccessibles à la polymérase et d’autres FTs, ou simplement recruter ou relâcher le complexe de polymérisation au niveau des séquences promotrices (**Bobola et Merabet, en préparation**).

#### 2.1.1.2. Les séquences cis-régulatrices

Il existe trois types de séquences cis-régulatrices. Les Enhancers dont la position est variable par rapport au site d’initiation de la transcription vont délivrer une information d’augmentation de la transcription. Les Silencers, à l’instar des Enhancers vont avoir une position variable mais vont donner une information contraire : elles vont faire diminuer la transcription. Et enfin les Insulators vont moduler les deux précédentes en étant placées entre celles-ci et le promoteur. Le rôle de ces séquences va être modulé par le contexte protéique.

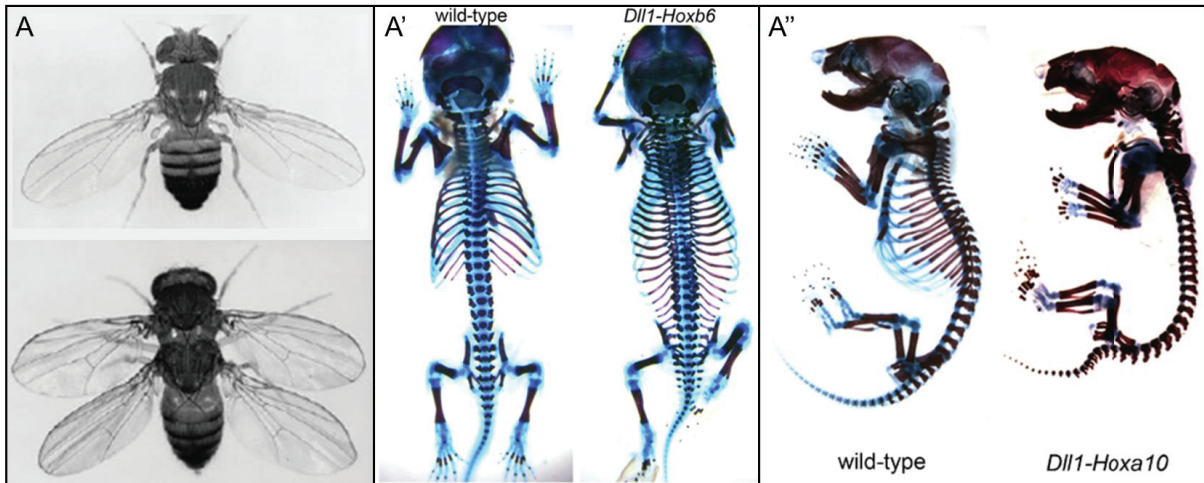
#### 2.1.1.3. Les sites de reconnaissance à l’ADN : nouvelles données.

Les FTs reconnaissent préférentiellement différents types de séquences sur l’ADN en fonction de leur domaine de liaison. Par exemple les protéines à HD se fixent préférentiellement *in vitro* sur de courtes séquences avec un cœur TAAT (**Noyes et al., 2008**) alors que les protéines à leucine zipper vont préférentiellement reconnaître le motif AGCT (**Lüscher et Larsson, 1999**).

Plus récemment il a été décrit que les FTs pouvaient aussi utiliser des sites de moindre affinité *in vivo* (**Crocker et al., 2015**). Les FTs allant et venant pour réguler de façon transitoire certains gènes, le fait qu’ils soient capables de fixer des séquences différentes avec une moindre affinité peut leur permettre d’assurer de manière plus efficace l’aspect transitoire de ces événements. Les sites de moindre affinité peuvent aussi permettre à des protéines d’une même famille de reconnaître des sites différents pour réguler des gènes différents au sein d’une même cellule.

#### 2.1.2. Rôle des protéines HOX au cours du développement embryonnaire

Chez tous les animaux, la mise en place des différents axes au cours de l’embryogenèse est régulée en partie par les protéines HOX. Elles sont des régulateurs clés de la diversité morphologique le long de l’axe antéro-postérieur (**McGinnis et Krumlauf, 1992 ; Pearson et al., 2005**). Cette fonction est parfaitement illustrée par les phénotypes homéotiques issus d’une expression anormale des protéines HOX. Ce terme a été pour la première fois utilisé en 1894 (**Bateson, 1894**) pour désigner la transformation d’un membre ou partie du corps en une autre partie du corps.



**Figure 2.1.2A : Les mutants homéotiques.** (A) La perte d'expression de la protéine Ubx dans l'haltère chez la *Drosophila* entraîne la formation d'une aile. (A') La surexpression ectopique de la protéine Hoxb6 entraîne la transformation des segments lombaires en segments thoraciques avec apparitions de côtes chez la souris. (A'') A l'inverse la surexpression ectopique de la protéine Hoxa10 transforme les segments thoraciques en segments lombaires avec perte des côtes chez la souris (Vinagre *et al.*, 2010).

On trouve ainsi l'exemple de la protéine Ultrabithorax (Ubx) chez la drosophile dont la mutation entraîne l'apparition d'une paire d'ailes supplémentaire à la place des haltères, organes balanciers (Figure 2.1.2A ; Lewis, 1978). Les mutations des gènes *HOX* ont souvent des phénotypes moins marqués chez les vertébrés du fait de la duplication des groupes de paralogie (Figure 1). En effet, les profils d'expression des différentes protéines d'un même groupe de paralogie vont se superposer et ces protéines vont présenter une redondance fonctionnelle qui minimisera les effets d'un simple mutant (Kmita *et al.*, 2005). On peut cependant avoir des phénotypes drastiques par une sur-expression de certains gènes *HOX* comme observé chez la souris (Figure 2.1.2A' ; Vinagre *et al.*, 2010).

Des cribles génétiques à grande échelle ont permis de révéler un certain nombre de gènes cibles des protéines HOX en particulier chez la drosophile. Ces cribles ne permettent cependant pas de déterminer si ces gènes cibles sont directs ou non. Les gènes cibles directs les mieux caractérisés proviennent d'études individuelles de régulation, comme les gènes *Distalless* (*Dll*), *rhomboid* (*rho*), *reaper* (*rp*), ou *Hoxb1*. Il faut noter qu'un certain nombre de ces gènes cibles sont les gènes *HOX* eux-mêmes (Pearson *et al.*, 2005 ; Hueber et Lohmann, 2008). Les approches de type transcriptomique ou de précipitation de la chromatine (ChIP) suivies de séquençages massifs haut débit (-seq) ont permis d'affiner notre compréhension de l'identité des gènes cibles des protéines HOX (Figure 2.1.2B issue de Hueber et Lohmann, 2008).

La nature des gènes cibles est ainsi révélée par exemple lors d'une expérience de ChIP-seq contre la protéine Hox de drosophile Ubx dans un tissu spécifique : l'haltère (Slattery *et al.*, 2011 ; Choo *et al.*, 2011). La nature des gènes cibles révèle la grande diversité d'action des protéines HOX à différentes échelles : contrôle de l'expression de FTs, de gènes effecteurs impliqués dans la migration cellulaire ou encore la prolifération cellulaire, etc.

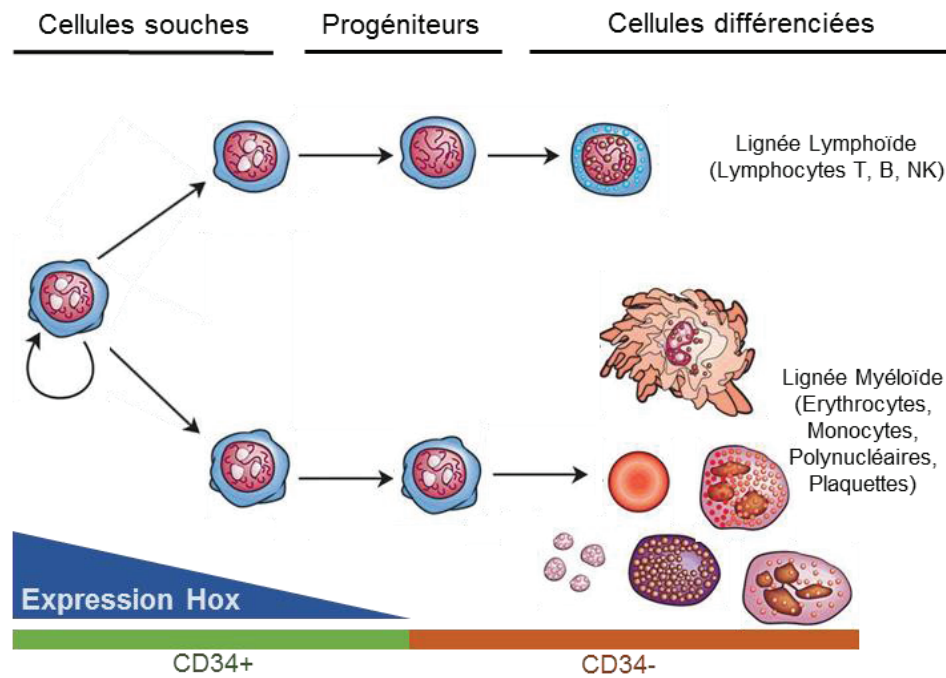
References	Hox genes	Organism	Tissue	Stage	#Targets
Leemans et al., 2001 <sup>(117)</sup>	Lab	<i>Drosophila</i>	whole embryo	embryonic stage 10-17	96
Mohit et al., 2006 <sup>(92)</sup>	Ubx	<i>Drosophila</i>	haltere and wing disc	3rd instar larvae	541
Hersh et al., 2007 <sup>(93)</sup>	Ubx	<i>Drosophila</i>	haltere and wing disc	3rd instar larvae	447
Hueber et al., 2007 <sup>(27)</sup>	Dfd, Scr, Antp, Ubx, Abd-A, Abd-B	<i>Drosophila</i>	whole embryo	embryonic stage 11+ 12	1508
Shen et al., 2000 <sup>(97)</sup>	HoxA1	mouse	cell culture – teratocarcinoma		28
Zhao and Potter, 2001 <sup>(120)</sup>	HoxA13	mouse	uterus and cervix tissue	4.5 weeks old	unclear
Valerius et al., 2002 <sup>(119)</sup>	HoxA11	mouse	kidney tissue	embryonic stage 18.5	10
Hedlund et al., 2004 <sup>(34)</sup>	HoxD10	mouse	spinal cord tissue	embryonic stage 12.5	69
Martinez-Ceballos et al., 2005 <sup>(96)</sup>	HoxA1	mouse	cell culture – embryonic blastocysts		145
Lei et al., 2005 <sup>(95)</sup>	HoxC8	mouse	cell culture – embryonic fibroblasts		34
Cobb et al., 2005 <sup>(33)</sup>	HoxD cluster genes	mouse	mouse tissue of limbs and genitalia	embryonic stage 12.5	16
Williams et al., 2005 <sup>(35)</sup>	HoxA13	mouse	cell culture – embryonic fibroblasts		68
Schwab et al., 2006 <sup>(99)</sup>	HoxA11 + HoxD11	mouse	whole embryonic kidneys and urogenital tissue	embryonic stage 11.5, 12.5 13.5, 16.5 + adult	1518
Rohrschneider et al., 2007 <sup>(98)</sup>	HoxB1a	zebrafish	whole embryo	19-20 hours post fertilization	471
Ferrell et al., 2005 <sup>(94)</sup>	HoxA10	human	cell culture – umbilical cord cells		115
<b>Large-scale identification of Hox response elements</b>					
References	Hox genes	Organism	Approach		
Ebner et al., 2005 <sup>(41)</sup>	Lab	<i>Drosophila</i>	<i>in silico</i> prediction		
Hueber et al., 2007 <sup>(27)</sup>	Dfd	<i>Drosophila</i>	<i>in silico</i> prediction		
McCabe et al., 2005 <sup>(101)</sup>	HoxA13 + HoxD13	mouse	ChIP		

**Figure 2.1.2B :** Identification des gènes des cibles de protéines HOX avec des cribles à large échelle. (A partir de Hueber et Lohmann, 2008).

### 2.1.3. Rôle des protéines HOX pour le maintien de l'homéostasie cellulaire chez l'adulte

Au-delà de ce rôle primordial lors de l'embryogénèse, on trouve aussi les protéines HOX dans les tissus adultes. En effet, ces FTs sont responsables du maintien de l'homéostasie de certains tissus, veillant à la correcte différenciation des lignées cellulaires ainsi qu'à la régulation de leur division.

L'un des exemples les plus caractéristiques est celui de la lignée hématopoïétique. Au cours de l'hématopoïèse chez l'adulte l'expression de différents gènes *HOX* va être activée ou inhibée en fonction de l'état de différenciation des cellules (**Figure 2.1.3A**). Les cellules souches hématopoïétiques porteuses d'un marqueur CD34 (cellules CD34+) expriment les protéines HOX antérieures (Hox 1 à 6) plus précocement, puis cette expression diminue à la faveur des protéines HOX plus postérieures lors de leur différenciation en progéniteurs (**Sauvageau et al., 1994**). Enfin, l'expression des protéines HOX est éteinte chez les cellules différenciées de la moelle osseuse que l'on identifie aussi par la perte du marqueur CD34 (cellules CD34- ; **Pineault et al., 2002**). Les gènes cibles régulés par les protéines HOX ne sont pas encore tous bien décrits mais une partie d'entre eux a pu d'ores et déjà être découverte. La majeure partie des travaux dans cette optique a été menée sur la protéine Hoxa9 (**Faber et al., 2009**) démontrant qu'elle est capable d'activer l'expression d'autres protéines HOX (HoxA7, HoxA10) mais aussi de ses partenaires Meis1 (**Hu et al., 2009**) et Pbx3 ainsi que d'autres gènes tels que Foxp1 et Sox4 entre autres (**Huang et al., 2012**). Une longue liste de gènes régulés par Hoxa9 et induisant la prolifération cellulaire a ainsi été dressée (**Alharbi et al., 2013**). Il en va de même pour la protéine Hoxb4 (**Oshima et al., 2011 ; Fan et al., 2012**) ou la protéine Hoxa10 (**Shah et al., 2011**). Ces FTs sont donc nécessaires au maintien du pouvoir prolifératif des progéniteurs du sang.



**Figure 2.1.3A :** Expression des gènes *HOX* au cours de l'hématopoïèse normale. Hématopoïèse simplifiée à partir de documents libres de droits.

Les pertes artificielles de fonction des protéines HOX au sein de la lignée hématopoïétique ne se traduisent pas par des phénotypes drastiques. En effet, la redondance fonctionnelle des différents FTs au sein de cette famille permet de sauver cette perte qui pourra se traduire par des organes hématopoïétiques réduits mais une différenciation cellulaire non affectée (**Bjornsson *et al.*, 2003**). Cependant, un gain de fonction par le maintien de l'expression ou la surexpression d'une protéine Hox dans des progéniteurs va entraîner une hyper prolifération cellulaire et/ou un blocage de la différenciation des cellules CD34+. Ce phénomène a été décrit pour les protéines Hoxb6 (**Fischbach *et al.*, 2005**), Hoxb3 (**Sauvageau *et al.*, 1997**), Hoxc4 (**Daga *et al.*, 2000**), HoxA5 (**Crooks *et al.*, 1999**), Hoxa10 (**Buske *et al.*, 2001**), Hoxb4 (**Sauvageau *et al.*, 1995**) et Hoxa9 (**Kroon *et al.*, 1998**).

Lors de phénomènes de translocation chromosomique au cours de la division cellulaire, le gène Hoxa9 a été retrouvé fusionné avec une nucléoporeine : Nup98 (**Nakamura *et al.*, 1996**). Il en résulte une surexpression de la protéine Hoxa9 localisée dans le noyau. Celle-ci aura alors un rôle oncogénique en maintenant l'expression de gènes impliqués dans la prolifération cellulaire. Cette dérégulation est à l'origine de leucémies myéloïdes aiguës qui sont de très mauvais pronostic.

Cet exemple d'implication des protéines HOX dans le maintien de l'homéostasie d'un tissu montre l'importance de ces FTs aussi bien dans le tissu sain que dans les cas pathologiques. Ils sont primordiaux pour la régulation du renouvellement de cellules souches et de leur différenciation. Dans le sein par exemple, la protéine HoxA5 va jouer un rôle suppresseur de tumeur par l'activation de l'expression du gène *TP53* (**Raman *et al.*, 2000**) alors que les protéines HoxB5 et HoxB7 jouent un rôle oncogénique si elles sont surexprimées en activant le gène *FGF2* (**Care *et al.*, 1998**) ainsi qu'en déclenchant une transition épithélio-mésenchymale (**Wu *et al.*, 2006**).

La vaste implication des protéines HOX dans les cancers (**Tableau 2.1.3B** ; **Bhatlekar et al., 2014**) ainsi que le manque de connaissances quant à leurs gènes cibles impliqués dans ces mécanismes de transformation oncogénique en font des sujets d'étude privilégiés.

	HOXA													HOXB													HOXC													HOXD												
	1	2	3	4	5	6	7	9	10	11	13	1	2	3	4	5	6	7	8	9	13	4	5	6	8	9	10	11	12	13	1	3	4	8	9	10	11	12	13													
Colon				X				X				X	X			X	X	X	X	X	X					X	X				X	X	X	X		X			X													
Sein	X	X	X		X	X		X		X		X		X	X	X	X	X	X	X	X	X				X		X		X	X	X	X	X	X	X	X		X													
Prostate														X							X	X	X	X	X																											
Poumon					X					X			X	X								X		X	X	X				X				X		X																
Cerveau								X	X																										X																	
Thyroïde										X	X		X							X	X									X									X													
Ovaire				X			X	X	X	X								X			X																															
Urinaire							X		X											X	X	X	X	X					X									X														
Peau																													X																							

**Tableau 2.1.3B : Expressions aberrantes des gènes HOX dans différentes tumeurs solides.**

## 2.2. LES PROTEINES HOX ONT AUSSI D'AUTRES FONCTIONS

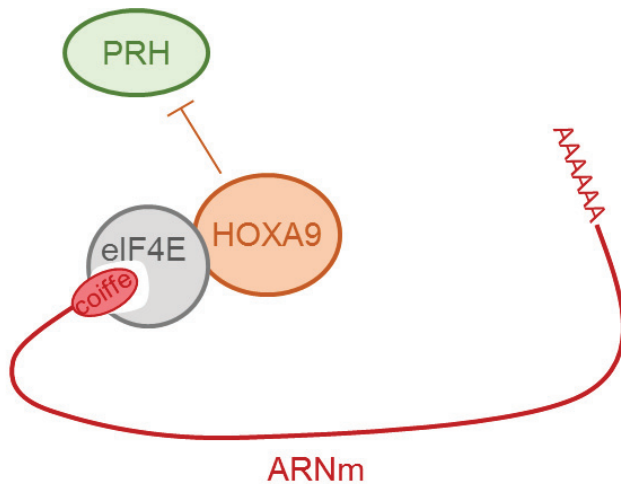
Les protéines HOX ne sont pas seulement impliquées dans la régulation de la transcription. En effet, il a été montré qu'elles jouaient un rôle dans d'autres processus biologiques au sein de la cellule parmi lesquels on trouve la réplication et la réparation de l'ADN, la dégradation des protéines, mais aussi la traduction des ARNm (**Rezsohazy, 2013**). Ces fonctions mises en avant par les nombreux nouveaux interacteurs découverts lors de vastes cribles donnent une nouvelle dimension au rôle des protéines HOX. Nous allons ici donner quelques exemples qui mettent l'accent sur les activités moléculaires des protéines HOX en les éloignant drastiquement de leur description première de FT.

### 2.2.1. Traduction

#### 2.2.1.1. Activation de la traduction

La protéine HoxA9 est capable de se lier à eIF4E afin de stimuler son activité grâce à un motif YxxxxL $\phi$  où x peut être n'importe quel résidu et  $\phi$  un résidu hydrophobe (**Sonenberg 1998, Dostie 2000**) situé en N-terminal de la protéine. En effet, HoxA9 va favoriser le transport nucléo-cytoplasmique des ARNm en agissant comme antagoniste de PRH (**Figure 2.2.1.1**) qui est quant à elle capable de bloquer le transport nucléo-cytoplasmique des ARNm dans un contexte hématopoïétique (**Topisirovic et al., 2003**). La protéine HoxA9 active aussi la traduction médiée par les complexes eIF (**Topisirovic et al., 2005**). Cependant les mécanismes qui articulent la compétition entre les effets inhibiteurs et activateurs de PRH et HoxA9 respectivement sur la protéine eIF4E restent encore à élucider à l'heure actuelle.

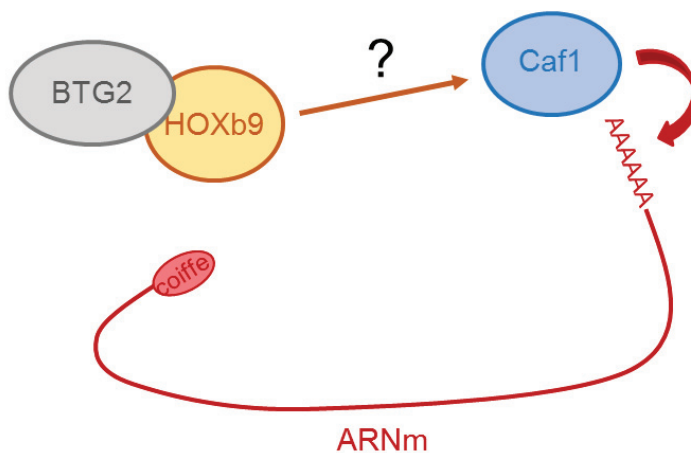




**Figure 2.2.1.1 :** Rôle de la protéine HOXA9 dans l'activation de la traduction. En interagissant avec la protéine eIF4E (qui reconnaît la coiffe des ARNm) la protéine HOXA9 va libérer la protéine PRH par compétition et ainsi lever l'inhibition de celle-ci sur l'export nucléaire. L'interaction entre la protéine HOXA9 et la protéine eIF4E permet aussi l'activation de la traduction par des mécanismes encore non élucidés.

### 2.2.1.2. Adenylation des queues Poly-A

La régulation de la traduction nécessite trois actions conjointes : le niveau d'export des ARNm, la régulation de la formation de la coiffe et celle de l'adénylation de la queue poly-A. Nous avons vu que certaines protéines HOX pouvaient être impliquées dans l'un des deux premiers processus. Il est spéculé que la protéine Hoxb9 pourrait réguler l'adénylation des queues poly-A via l'interaction avec une protéine BTG2 (**Figure 2.2.1.2**). Effectivement, Prevot *et al.* démontrent que les protéines BTG1 et BTG2 sont des partenaires de la protéine Hoxb9 (**Prevot *et al.*, 2000**), et il est mis en évidence que BTG2 est capable d'interagir avec des déadénylases et de les stimuler (**Doidge *et al.*, 2012**).



**Figure 2.2.1.2 :** La protéine HOXB9 interagit avec la protéine BTG2 susceptible de réguler l'adénylation des queues poly-A. La protéine BTG2 a été décrite comme interagissant avec et stimulant une déadénylase Caf1. Ce complexe serait impliqué dans la régulation de la quantité d'ARNm ainsi que de leur traduction. L'action de la protéine HOXB9 dans ces mécanismes reste à élucider.

### 2.2.2. Réplication de l'ADN

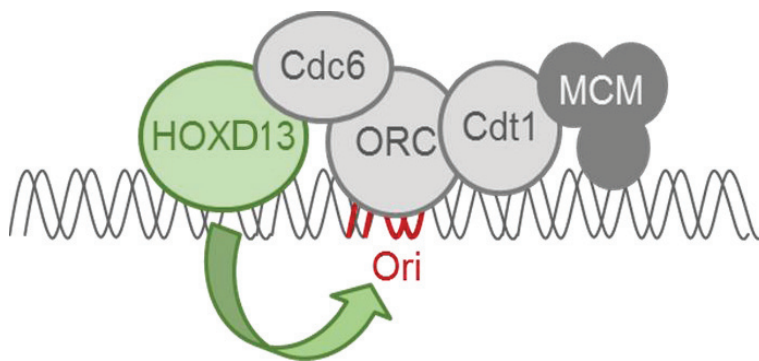
Le processus de réplication de l'ADN est bien décrit. Lors d'une première étape, le complexe de reconnaissance de l'origine de réplication (ORC) vient se fixer sur l'origine de réplication. Suivent alors deux autres protéines, Cdc6 et Cdt1 qui vont permettre l'association avec le complexe hélicase MCM (mini chromosome maintenance) : le complexe de pré-réplication (pre-RC) est formé. Au cours du cycle cellulaire, lors de la transition G1-S, le pré-RC est modifié afin de former les fourches de réplication. Un des régulateurs les mieux décrits de ce processus est la protéine Geminin (GMNN). Elle forme des homodimères qui vont s'associer avec la protéine Cdt1 et ainsi empêcher le pré-RC de se former (**Saxena *et al.*, 2004**).

L'expression de GMNN est donc étroitement régulée au cours du cycle cellulaire : elle est maintenue aux phases G2 et M, puis elle est dégradée pour permettre la réplication de l'ADN.

Afin de mieux comprendre le rôle de GMNN lors de l'embryogénèse, Luo *et al.* ont mené un crible double-hybride entre GMNN et des ADNs préparés à partir d'embryons de souris. Ils ont ainsi dévoilé des interactions directes entre GMNN et Hoxd10 et Hoxa11. Grâce à une confirmation par GST-pulldown, ils ont pu ajouter d'autres protéines HOX à ce jeu d'interactions (Luo *et al.*, 2004). La principale question émanant de cette interaction est de savoir si la protéine GMNN va adopter un rôle de co-facteur de la protéine HOX afin de réguler la transcription, ou, si la protéine HOX va participer au processus de régulation de la réplication de l'ADN médié par GMNN. Nous allons ici décrire certaines des protéines HOX qui ont été étudiées.

#### 2.2.2.1. Initiation de la réplication

La protéine HoxD13 interagit avec les origines de réplication de l'ADN (Salsi *et al.*, 2009) au cours de la phase de réplication G1 et cette interaction n'est jamais retrouvée en dehors de cette phase du cycle cellulaire. Lors de son interaction, la protéine HoxD13 va promouvoir la formation du pre-RC via son interaction avec la protéine Cdc6 (Salsi *et al.*, 2009) et ainsi favoriser l'initiation de la réplication de l'ADN (Figure 2.2.2.1).



**Figure 2.2.2.1 :** La protéine HoxD13 recrute le complexe pré-RC et favorise l'initiation de la réplication. La formation de ce complexe bloque aussi l'interaction entre Cdt1 et GMNN.

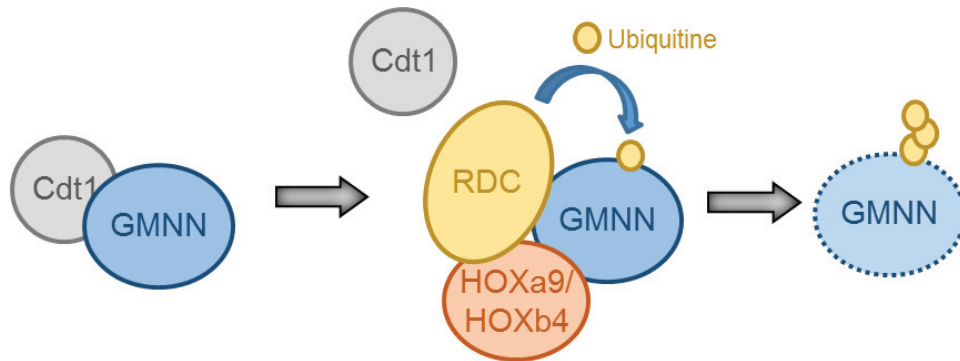
Lors de la phase S, une forte expression de GMNN est nécessaire pour réguler la réplication. Cette protéine va donc interagir à la fois avec Cdt1 et avec la protéine HOX, pour séquestrer les deux protéines et ainsi arrêter le processus de réplication. Il est enfin envisageable que le complexe HOX-GMNN puisse réguler des gènes cibles sur l'ADN, se comportant alors à nouveau comme un complexe de FTs.

Ces propriétés ont aussi été découvertes chez une autre protéine HOX, HoxC13 (Comelli *et al.*, 2009) qui serait capable d'interagir avec CDC6 via son HD, ainsi qu'avec ORC1 et ORC2 (Marchetti *et al.*, 2010).

#### 2.2.2.2. Régulation de la réplication par ubiquitination

Les protéines Hoxb4 et Hoxa9 sont aussi capable de réguler la réplication de l'ADN mais par une toute autre voie. En effet, Ohno *et al.* montrent dans les cellules hématopoïétiques que les protéines HOX vont s'intégrer dans un complexe multiprotéique impliqué dans l'ubiquitination des protéines (Roc1-Ddb1-Cul4a). Elles sont aussi capables d'interagir avec la GMNN par leur HD (Ohno *et al.*, 2013). L'ensemble de ces interactions conduit à l'ubiquitination de la GMNN et donc à sa dégradation par le protéasome (Figure 2.2.2.2). Elle n'est plus capable d'inhiber la réplication. Indirectement donc, les protéines HOX activent la

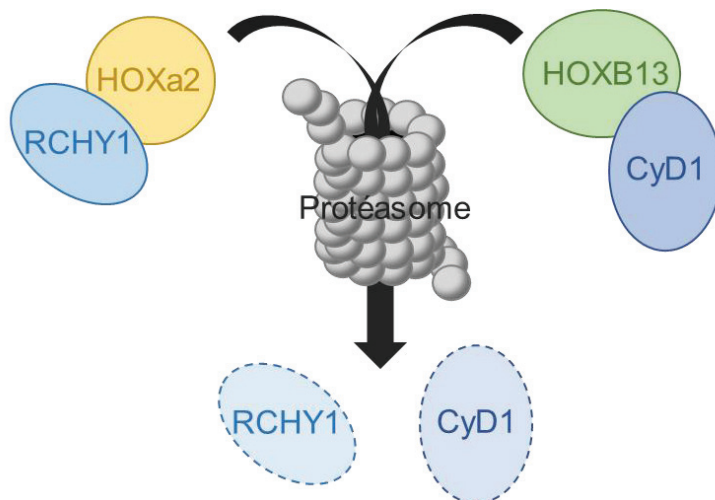
prolifération des cellules souches hématopoïétiques et des progéniteurs du sang, alors que la GMNN maintient la quiescence de ces cellules (**Ohno *et al.*, 2013**).



**Figure 2.2.2.2 :** Les protéines Hoxa9 et Hoxb4 favorisent la dégradation de la GMNN par ubiquitination. Les protéines HOX interagissent avec les protéines du complexe d'ubiquitination Roc1, Dbd1 et Cul4a (RDC) ainsi qu'avec la GMNN.

### 2.2.3. Dégradation protéique

Comme nous venons de le voir, certaines protéines HOX sont capables de recruter les complexes d'ubiquitination pour entraîner la dégradation de protéines telles que la GMNN (**Ohno *et al.*, 2013 ; Ohno *et al.*, 2010**) et influencer ainsi le cycle cellulaire. D'autres protéines HOX sont aussi capables de promouvoir cette dégradation.



**Figure 2.2.3 :** Les protéines Hoxa2 et HoxB13 conduisent les protéines RCHY1 et CyD1 à leur dégradation par le protéasome.

#### 2.2.3.1. Dégradation de l'ubiquitine ligase RCHY1

La protéine RCHY1 a été identifiée comme interacteur potentiel de la protéine Hoxa2 (**Bergiers *et al.*, 2013**). C'est une protéine E3 ubiquitine ligase qui est impliquée dans la dégradation de protéines cibles importantes telles que p53 ou c-MYC qui sont impliquées dans la régulation du cycle cellulaire et dans la mort cellulaire.

Dans une étude récente, Bridoux *et al.* décrivent le rôle d'Hoxa2 dans la dégradation de la protéine RCHY1 (**Bridoux *et al.*, 2015**). En effet, ils montrent que l'interaction HoxA2-RCHY1 se passe dans le noyau et qu'elle dépend de deux domaines importants de la protéine HOX (son HD ainsi qu'une autre région). Cette interaction va ensuite conduire à la dégradation de RCHY1 via les sous unités 19S et 20S du protéasome (**Figure 2.2.3**). Ce phénomène est

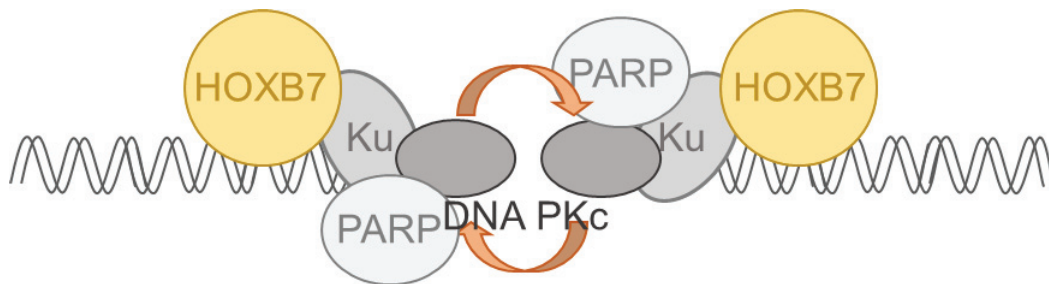
conservé chez les vertébrés et présent au cours de l'embryogénèse, certainement afin de réguler les dégradations induites par l'E3 ligase.

#### 2.2.3.2. Dégradation de la cycline D1

Très décrite dans le cancer de la prostate, la protéine HoxB13 voit son expression fluctuer au cours du cycle cellulaire. Elle est fortement exprimée pendant les phases de croissance cellulaire puis disparaît peu à peu avant la mitose. Il a été montré qu'elle est capable d'interagir avec la cycline D1 et de favoriser ainsi l'ubiquitination de cette dernière (**Hamid et al., 2013**) et sa dégradation par le protéasome (**Figure 2.2.3**). De plus, le rôle de suppresseur de tumeur de la protéine HoxB13 a été démontré (**Jung et al., 2004**) ce qui corrobore son implication dans la régulation du cycle cellulaire.

#### 2.2.4. Réparation de l'ADN

La protéine HoxB7 interagit avec quatre protéines (Ku70, Ku80, DNA-PKc et PARP1) impliquées dans la réparation des lésions double brin de l'ADN (**Rubin et al., 2007**). Ces réparations se produisent par NHEJ (non homologous end joining). En effet, les protéines Ku vont reconnaître la brèche, recruter les DNA PKc puis d'autres enzymes vont intervenir pour compléter la brèche et la refermer. La protéine PARP1 va stimuler les protéines kinases sur la brèche. Rubin *et al.* montrent que la protéine HoxB7 favorise ce processus (**Figure 2.2.4**) et que sa diminution d'expression entraîne une moins bonne réparation des lésions de l'ADN (**Rubin et al., 2007**).



**Figure 2.2.4 : La protéine HoxB7 participe à la réparation des lésions double brin sur l'ADN.**

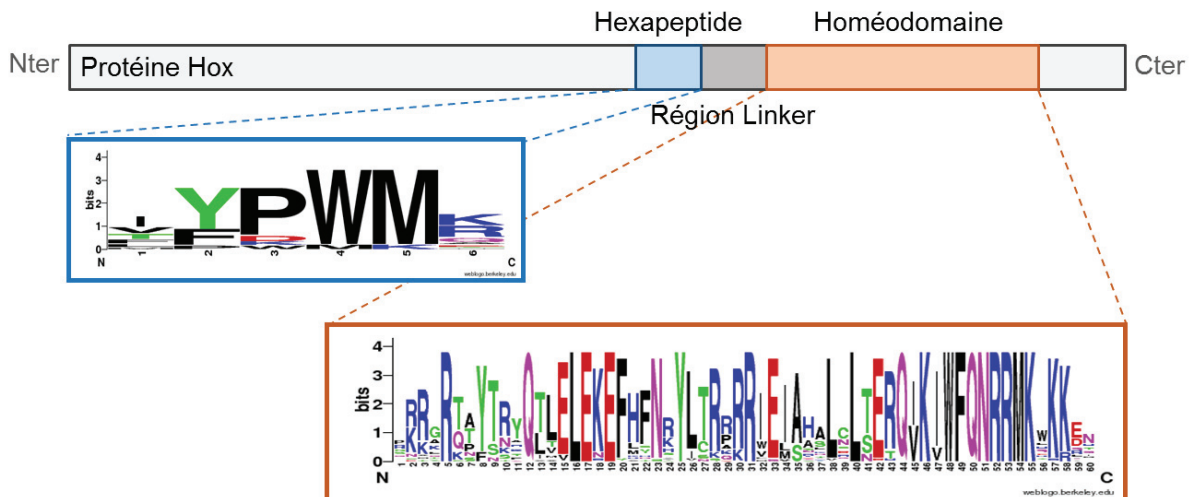
#### 2.2.5. Les protéines HOX pourraient avoir encore d'autres fonctions moléculaires

Certains des partenaires putatifs issus des cribles menés contre diverses protéines HOX laissent supposer que l'ensemble des fonctions non transcriptionnelles des protéines HOX n'a pas encore été élucidé (transduction du signal, adhésion cellulaire, transport vésiculaire ou communication cellule-cellule). En effet, il a été montré que des protéines telles que PLSCR1, TRIP6 et RBPMS sont capables d'interagir avec les protéines HoxA1, HoxA9, HoxB6 ou HoxB9 (**Lambert et al., 2012 ; Tarminiau et al., 2016**).

### 3. STRUCTURE DES PROTEINES HOX

#### 3.1. STRUCTURE « GLOBALE » DES PROTEINES HOX

L'analyse des séquences des protéines HOX de nombreux organismes met en évidence deux régions extrêmement conservées : le domaine de liaison à l'ADN l'**homéodomaine** (HD) et un motif localisé en amont appelé **hexapeptide** (HX) (**Figure 3.1**). Les protéines HOX possèdent d'autres motifs qui sont moins bien conservés que le motif HX (voir plus loin) ainsi que de grandes régions désorganisées, la plupart localisées dans la partie N terminale.

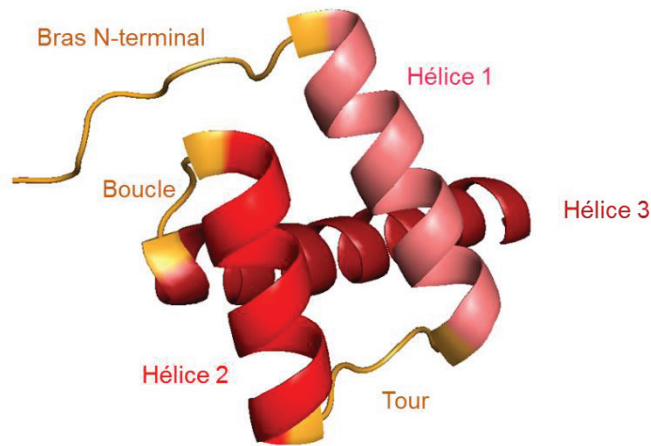


**Figure 3.1. :** Séquences des motifs conservés des protéines HOX. Séquence weblogo du motif hexapeptide (HX) et de l'homéodomaine (HD). La conservation du motif YPWM ou FPWM est mise en évidence pour le motif HX. La longueur de la région linker entre le motif HX et le domaine HD est dépendante de la protéine Hox considérée. Les régions les mieux conservées de l'HD, telles que la fin de la troisième hélice, apparaissent aussi.

##### 3.1.1. Homéodomaine (HD)

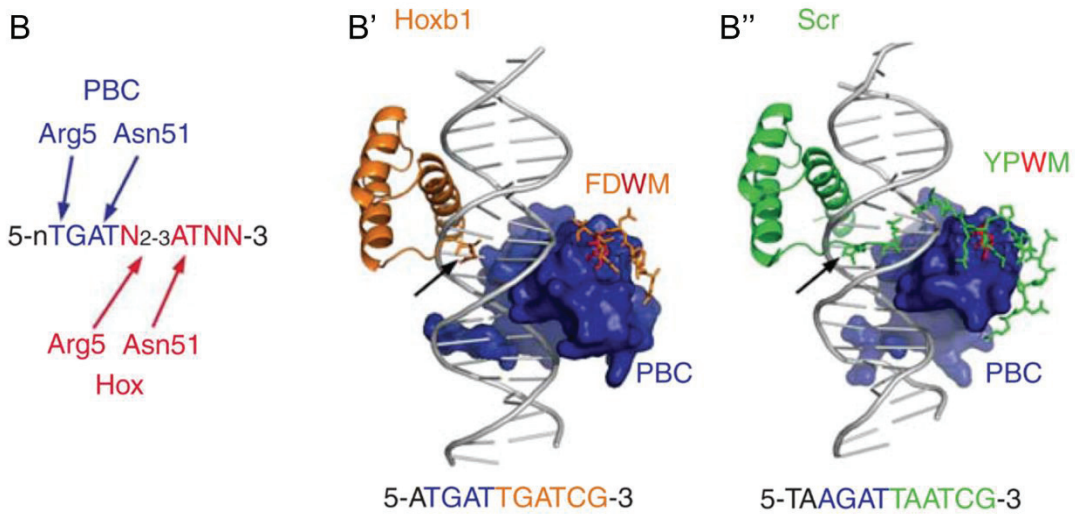
L'homéodomaine (HD) n'est pas un domaine de liaison à l'ADN spécifique des protéines HOX. On le retrouve en effet dans de nombreux autres FTs (**Bürklin et Affolter, 2015**) tels que ceux de la famille des TALE (**Bürklin, 2011**), de la famille des CUT (**Takatori et Saiga, 2008**), de la famille des Paired-Like (**Zhong et Holland, 2011**), et d'autres. Depuis qu'il est étudié, il a été décrit aussi bien chez les animaux (**Mc Ginnis et al., 1984**) que chez les plantes (**Langdale, 1994**) ou les levures (**Shepherd et al., 1984**). Son extrême conservation évolutive signe un rôle primordial et un mécanisme robuste de régulation de l'expression des gènes.

L'homéodomaine est un domaine protéique de 60 acides aminés dans les protéines HOX (ce nombre peut varier pour d'autres protéines à HD). Sa structure est constituée de trois hélices alpha espacées par une boucle puis un tour (**Figure 3.1.1.A**) et précédées par un bras N-terminal flexible. La troisième hélice de l'HD va se loger dans le sillon majeur de l'ADN et le bras N-terminal dans le sillon mineur permettant la régulation de la transcription de gènes cibles. Certains résidus ont été montrés comme nécessaires à cette interaction : c'est le cas de l'arginine en position 5 (dans le bras N-terminal) ainsi que de l'asparagine en position 51 dans la troisième hélice alpha de l'HD (**Figure 3.1.1.B ; Mann et al., 2009**).

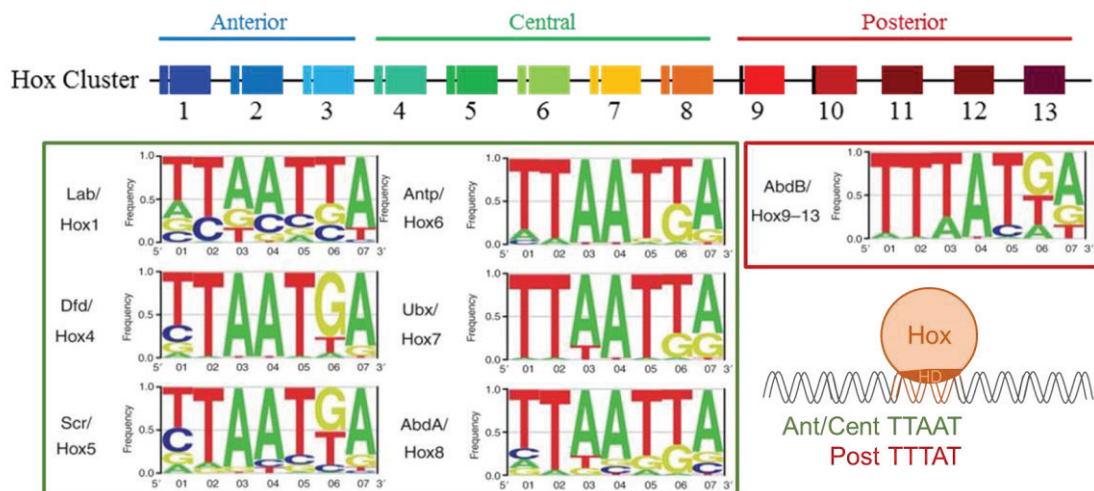


**Figure 3.1.1.A :** Structure de l'HD de la protéine HoxA9. Obtenue à partir de la base de données PDB.

**Figure 3.1.1.B :** Quelques exemples de complexes Hox-PBC-ADN à partir de Mann *et al.*, 2009. (B) Site consensus de liaison à l'ADN du complexe Hox-PBC. L'arg5 contacte le sillon mineur, et l'Asn51 le sillon majeur. (B' et B'') Deux exemples de structures des HD de la protéine Hoxb1 (souris) et Scr (drosophile) en complexe avec l'HD des PBC sur l'ADN



Les protéines HOX vont ainsi reconnaître une séquence sur l'ADN riche en nucléotides AT de 5 à 7 paires de bases (Noyes *et al.*, 2008) avec un cœur de type **TAAT** pour les protéines antérieures et centrales et **TTAT** pour les protéines postérieures (Figure 3.1.1C). Cependant ce domaine à lui seul ne permet pas d'expliquer la spécificité de chaque protéine HOX, capable de réguler des gènes à la fois spécifiques et variés de façon extrêmement fine.

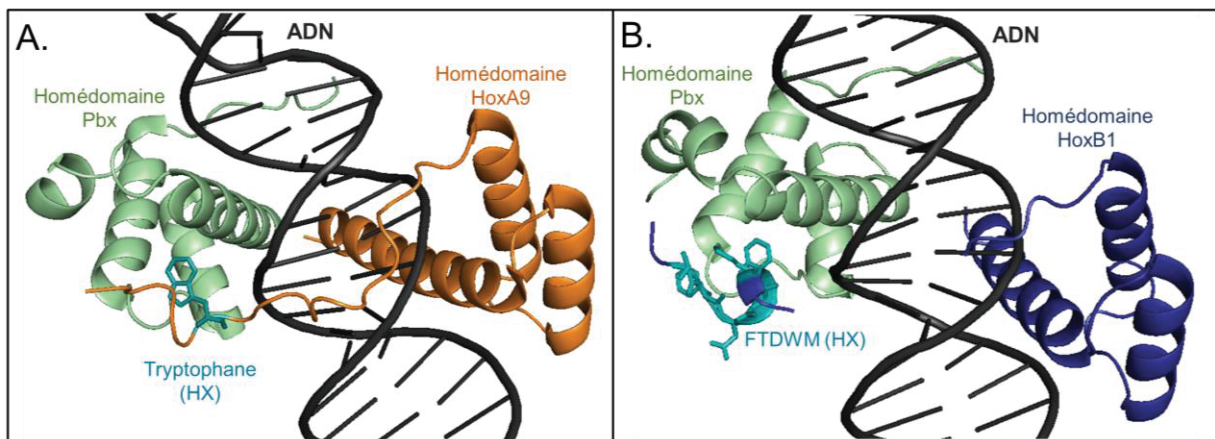


**Figure 3.1.1.C :** Séquences des sites de reconnaissances des différentes protéines HOX.

### 3.1.2. Le motif Hexapeptide (HX)

Lorsque l'on cherche des signatures spécifiques aux protéines HOX (**Merabet et al., 2009**) il se dessine en plus de l'HD un petit motif situé en amont de celui-ci : l'hexapeptide (HX). Ce motif contient une séquence centrale de type YPWM ou FPWM bien conservée dans toutes les protéines HOX antérieures et centrales (**Figure 3.1**). Dans les protéines HOX postérieures des groupes 9 et 10 (AbdB chez la drosophile) ce motif se limite au résidu tryptophane (W). Il est important de noter que les protéines HOX des groupes postérieurs 11 à 13 chez les vertébrés n'ont pas de motif HX ou de résidu W conservé en amont de leur HD. Le motif HX est plus largement appelé PID pour PBC interaction domain (**Morgan et al., 2000**). Le motif PID est caractérisé par un résidu W dans environnement hydrophobe avec des résidus basiques en positions +2 et +5. Il n'est cependant pas spécifique aux protéines HOX puisqu'on le retrouve aussi dans d'autres FTs à HD tels que Tlx1 ou MyoD (**In Der Rieden et al., 2004**).

Ce motif a longtemps été décrit comme étant le siège de l'interaction entre les protéines HOX et leurs partenaires PBC (**Morgan et al., 2000**). Les protéines PBC sont des FTs à HD mais nous décrirons plus en détails ces partenaires dans un autre chapitre (**Chapitre IV.2**). La plupart des résultats *in vitro* démontrent que la mutation de ce motif entraîne la perte d'interaction avec les partenaires PBC (**Knoepfler et al., 1995 ; Neuteboom et al., 1995**).



**Figure 3.1.2 : Structure de l'interaction Hox/PBC sur l'ADN : exemple de deux protéines HOX antérieure et postérieure.** (A) Structure de la partie C-terminale de la protéine HoxA9 (orange) et de l'HD de Pbx1 (vert) sur la double hélice d'ADN. Le tryptophane conservé de l'HX est indiqué en turquoise. (B) Structure de la partie C-terminale de la protéine HoxB1 (bleu) et de l'HD de Pbx1 (vert) sur la double hélice d'ADN. Le motif HX est indiqué en turquoise. La région Linker entre l'HD et l'HX de la protéine HoxB1 n'est pas visualisable sur la structure.

Afin de mieux comprendre les mécanismes régissant l'interaction entre les protéines HOX et leurs partenaires PBC, des analyses structurales ont été réalisées en utilisant des protéines tronquées ne contenant pas toute la partie N-terminale (**Passner et al., 1999 ; Piper et al., 1999 ; LaRonde Leblanc et Wolberger, 2003 ; Mann et al., 2009**). Il est ainsi possible de visualiser cette interaction au niveau de l'ADN pour les protéines HoxA9 ou HoxB1 par exemple (**Figure 3.1.2**). Quelle que soit la protéine Hox, le cristal montre une insertion du motif HX (en particulier du résidu W) dans la poche hydrophobe générée par certains résidus de l'HD de Pbx1 (membre de la famille des PBC).

Enfin, il est important de noter que la distance entre le motif HX et l'HD des protéines HOX, appelée région linker, est étroitement liée au groupe de paralogie. En effet les protéines

HOX antérieures ont une région linker très longue (>50 pour la protéine Labial) alors que celle-ci est très courte pour les protéines HOX postérieures (<5 pour la protéine Hoxa9). La présence de cette région va affecter l'interaction des protéines HOX avec l'ADN de façon spécifique à chaque groupe de paralogie. Elle suppose aussi la possible interaction des protéines HOX avec des partenaires spécifiques supplémentaires via cette région (**Merabet et al., 2003**). La structure d'un complexe entre les protéines Sex Combs reduced (Scr) et Extradenticle (Exd) sur un site non plus artificiel et de forte affinité mais un site physiologique de faible affinité (*fkh250*) montre des interactions supplémentaires qui sont expliquées par la présence de résidus spécifiques dans cette région linker (**Joshi et al., 2007**).

### 3.2. AUTRES MOTIFS DES PROTEINES HOX

#### 3.2.1. Qu'est-ce qu'un « Short Linear Motif » ?

Ces dernières années, le concept de petit motif linéaire (SLiM) capable de réguler les interactions protéine-protéine (IPPs) commence à apparaître de plus en plus dans la littérature (**Neduva et al., 2005 ; Van Roey et al., 2014 ; Tompa et al., 2014**). Ces motifs sont appelés « SLiM » ou « ELM » pour Short Linear interaction Motif ou Eucaryotic Linear Motif respectivement.

Différents types de SLiMs sont répertoriés dans les banques de données (<http://els.eu.org/>) en fonction de leurs propriétés d'interaction protéine-protéine. Le motif HX est ainsi répertorié comme un SLiM de type LIG (**source ELM**) car il agit pour recruter/lier un autre partenaire protéique de la régulation transcriptionnelle (**Tableau 3.2.1**).

Lors de l'étude de la régulation des IPPs, de plus en plus de recherches ont montré que les régions désorganisées des protéines avaient un rôle aussi, voire plus, important que les domaines globulaires ou structurés dans ce mécanisme (**Wright et al., 1999 ; Dunker et al., 2001 ; Tompa, 2012**). Les SLiMs sont décrits pour leur implication dans de nombreux processus cellulaires tels que : le contrôle du cycle cellulaire, l'étiquetage des protéines à destination du protéasome, la stabilisation de complexes, ou encore l'adressage de protéines à différentes localisations sub-cellulaires (**Dyson et al., 2005 ; Diella et al., 2008**).

	Nature de l'ELM
LIG	Site de liaison général d'un ligand
MOD	Site de modification post traductionnelle
TRG	Sites de cible sub-cellulaire
DEG	Site de liaison entraînant l'ubiquitination et la dégradation par le protéasome (degron)
DOC	Site de recrutement d'une enzyme (docking)
CLV	Site de clivage protéolytique

**Tableau 3.2.1 : Les différents types de SLiMs (Short Linear Motifs) ou ELM (Eucaryotic Linear Motif).**

Les SLiMs sont donc de petits motifs protéiques qui forment des interfaces d'IPPs. Ils ont une taille réduite (inférieure à dix acides aminés) et se localisent surtout dans les régions désordonnées des protéines. Grâce à leur taille réduite ils permettent une grande variabilité et plasticité d'interaction. De plus il a été montré qu'ils étaient importants dans la régulation

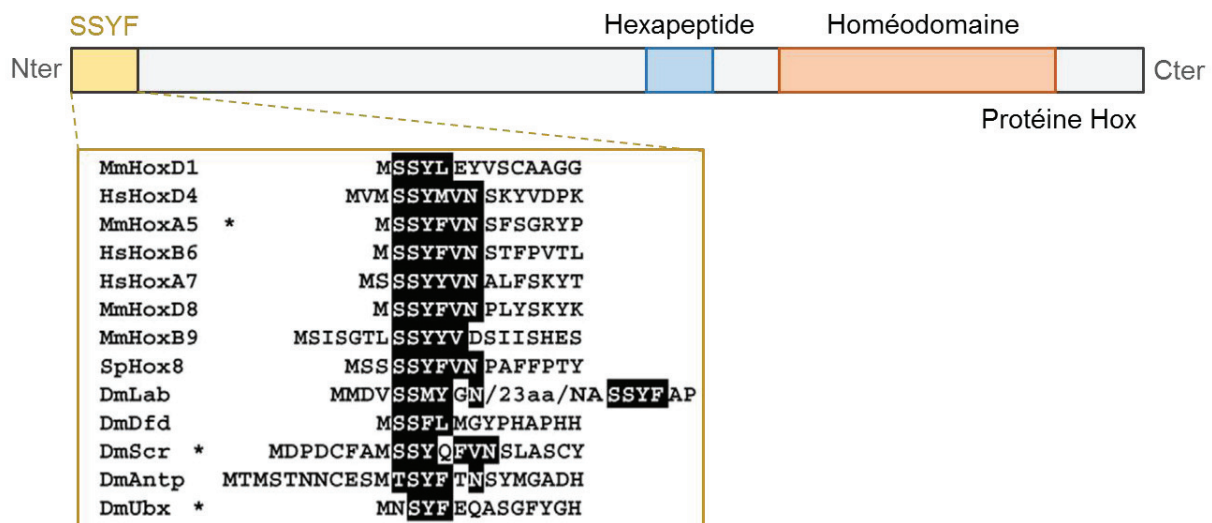


tissus-spécifique des réseaux d'interactions (**Buljan et al., 2013**). Cependant, ils sont moins bien décrits dans la régulation des IPPs entre FTs (**Dinkel et al., 2014**) ce qui suggère que ce type d'interactions est plus difficile à capturer et identifier que les autres.

Dans les protéines HOX seuls quelques SLiMs sont connus pour participer à des interactions spécifiques et à leur régulation (**Sivanantharajah et Percival-Smith, 2014**). L'identification d'autres SLiMs paraît alors essentielle afin de mieux comprendre la spécificité et la diversité de l'activité de régulation de transcription des protéines HOX. Dans la suite de ce chapitre, j décrirai les rares SLiMs dont la fonction génétique et moléculaire a été caractérisée au sein des protéines HOX.

### 3.2.2. Le motif SSYF

Lors d'une étude sur le rôle d'activateur ou de répresseur transcriptionnel de la protéine HOX Ultrabithorax (Ubx) dans la drosophile, **Tour et al.** ont généré des formes tronquées de Ubx (**Tour et al., 2005**). Ils ont ainsi pu constater qu'une région située en position N-terminale de la protéine et conservée chez la plupart des protéines HOX jouait un rôle important dans la régulation de gènes cibles (**Figure 3.2.2.**). Ce motif contient les résidus SSYF (pour sérine-sérine-tyrosine-phenylalanine) et est en effet indispensable aux protéines Ubx et Scr (Sex combs reduced) pour activer des gènes tels que *tsh* (*tea-shirt*) ou *dpp* (*decapentaplegic*) dans des tissus différents.



**Figure 3.2.2. : Le motif SSYF.** Le motif situé en position N terminale de la protéine Hox est très bien conservé. Alignement issu de **Tour et al., 2005** avec les séquences à partir de la Méthionine initiale de protéines HOX de souris (Mm *Mus musculus*), d'Homme (Hs *Homo sapiens*), d'oursin (Sp *Strongylocentrotus purpuratus*) et de mouche (Dm *Drosophila melanogaster*).

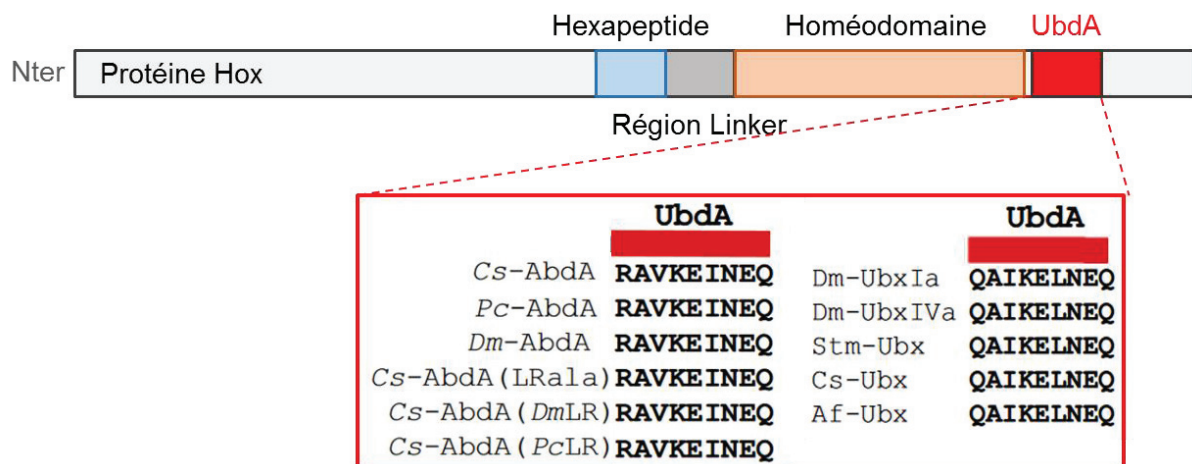
D'autres études avaient déjà porté sur la partie N-terminale de la protéine HoxA5 de souris (**Zhao et al., 1996**) et elles ont montré qu'elle était nécessaire à l'activation d'un gène rapporteur sous le promoteur de *forkhead*.

Le mécanisme impliqué dans cette régulation n'est pas encore décrit mais on peut supposer que ce SLiM est capable de réguler les interactions entre la protéine HOX et des partenaires encore à découvrir, influençant ainsi l'activité de la protéine HOX.

Un co-facteur potentiel capable d'interagir avec le motif SSYF est l'histone acetyltransferase CBP (CREB-binding protein) (Chan *et al.*, 2001). En effet, CBP interagit directement avec la partie N-terminale (les 18 premiers acides aminés) de la protéine HoxB7 et cette interaction augmente l'activation transcriptionnelle médiée par HoxB7 dans des cellules de cancer du sein (Chariot *et al.*, 1999).

### 3.2.3. Le motif UbdA

Le motif UbdA est un petit motif de huit acides aminés que l'on ne retrouve que dans les protéines HOX centrales AbdA et Ubx chez les protostomes (Balavoine *et al.*, 2002). Il est situé en aval de l'HD de ces deux protéines HOX (Figure 3.2.3). Il favorise l'interaction des deux protéines HOX avec le partenaire PBC (Merabet *et al.*, 2007 ; Hudry *et al.*, 2012 ; Foos *et al.*, 2015). La redondance entre les motifs HX et UbdA est régulée par une autre région de la protéine HOX : la région linker qui se situe entre l'HX et l'HD de celle-ci (Figure 3.2.3). La plasticité de cette interaction via ces deux motifs a une influence directe sur l'activité du complexe. En effet, l'utilisation du motif UbdA plutôt que HX va entraîner une répression plus forte du gène cible *Dll* (Saadaoui *et al.*, 2011). Le rôle du motif UbdA met en évidence l'existence de différents modes d'interaction entre les protéines HOX et leurs cofacteur PBC (voir Chapitre IV ; Foos *et al.*, 2015).



**Figure 3.2.3 : Le motif UbdA.** On retrouve ce motif conservé chez les protostomes.

### 3.2.4. Le motif TDWM

Le motif TDWM a une séquence très similaire à celle du motif HX (YPWM). Il est retrouvé spécifiquement dans les protéines AbdA d'insectes (Merabet *et al.*, 2011). Son tryptophane conservé (W) pourrait agir de la même façon que celui du motif HX dans l'interaction de la protéine Hox AbdA avec le partenaire Exd (Hudry *et al.*, 2012 ; Chapitre IV).

### 3.2.5. Utilisation complexe des SLiMs au sein d'une protéine Hox.

En s'intéressant aux motifs UbdA, HX et TDWM, Merabet *et al.* (2011) mettent en évidence une utilisation complexe des SLiMs au sein de la protéine Hox AbdA *in vivo*. En effet ils montrent que ces motifs sont utilisés de manière différentielle en fonction du tissu considéré. Plus particulièrement, ces motifs présentent une activité redondante ou opposée, sans logique apparente au niveau du tissu ou de la ressemblance du motif. Les motifs Hx et TDWM étant très similaires, ils remplissent cependant des fonctions opposées dans certains cas et vis-versa pour les motifs HX et UbdA (Merabet *et al.*, 2011A et B). Cette analyse démontre l'extrême complexité d'utilisation fonctionnelle des petits motifs au sein d'une protéine Hox, complexité qui se traduit certainement par une diversité moléculaire au niveau des IPPs.

## 4. LES PARTENAIRES DES PROTEINES HOX

Le paradoxe des protéines HOX naît de leur extrême conservation moléculaire qui est en contradiction avec leurs rôles aussi spécifiques que variés. Même si certaines protéines HOX montrent une redondance fonctionnelle lorsqu'elles sont exprimées ectopiquement, la plupart d'entre elles ont des gènes cibles bien spécifiques dans un contexte cellulaire lui aussi spécifique (Papagiannouli *et al.*, 2014). Ni leur seul domaine de reconnaissance à l'ADN ni les deux familles de cofacteurs Pbx/Meis ne sont suffisants pour donner une explication complète à ce paradoxe. L'existence d'autres protéines agissant en complexe avec ces FTs a été postulée. Nous allons ici développer en premier lieu les techniques qui ont permis de trouver ces partenaires, puis présenter les co-facteurs décrits dans la littérature.

### 4.1. LES TECHNIQUES DE VISUALISATION DES IPPS

Afin d'isoler de nouveaux partenaires des protéines HOX, il a fallu mener de larges cribles. En effet, le peu de connaissances à ce sujet a nécessité l'utilisation de techniques de détection des IPPs pouvant mettre en évidence un grand nombre de partenaires et ceci de façon assez rapide.

#### 4.1.1. Quelques techniques classiques

Des cribles génétiques chez la drosophile (Kennison et Tamkun, 1988 ; Harding *et al.*, 1995, Florence *et al.*, 1998 ; Merabet *et al.*, 2002) ainsi que des cribles en double-hybride chez la levure (Bondos *et al.*, 2006 ; Lambert *et al.*, 2012) ont mis en évidence un certain nombre de cofacteurs candidats. Cependant, la plupart nécessitent encore d'être caractérisés/validés fonctionnellement et/ou moléculairement.

Dans le cas des cribles en double-hybride chez la levure, le principal biais de la technique est le contexte cellulaire : la levure ne contient pas de gènes *HOX* et elle est très éloignée des tissus animaux dans lesquelles ces gènes s'expriment. Chaque interaction ainsi découverte doit systématiquement être confirmée et vérifiée dans une cellule animale pour s'assurer qu'elle n'est pas un simple artefact du système (Bondos *et al.*, 2006). C'est le cas par exemple de la protéine Hoxa1 de souris pour laquelle cinquante-neuf interacteurs potentiels ont été identifiés en double-hybride chez la levure, dont quarante-cinq nouvelles interactions ont pu être confirmées par la suite grâce à une co-purification en cellules animales, puis

quarante et une par complémentation de fluorescence bimoléculaire, une technique que nous développerons ci-après (**Lambert et al., 2012**). Il est à remarquer que ces techniques sont particulièrement lourdes et fastidieuses.

Les cribles génétiques quant à eux, bien qu'ils soient réalisés en contexte *in vivo*, ne permettent pas d'affirmer que les cofacteurs candidats sont des interacteurs directs de la protéine Hox. Cette validation nécessite donc une confirmation sur le plan moléculaire par des approches biochimiques de type GST pull down ou de co-immunoprécipitation.

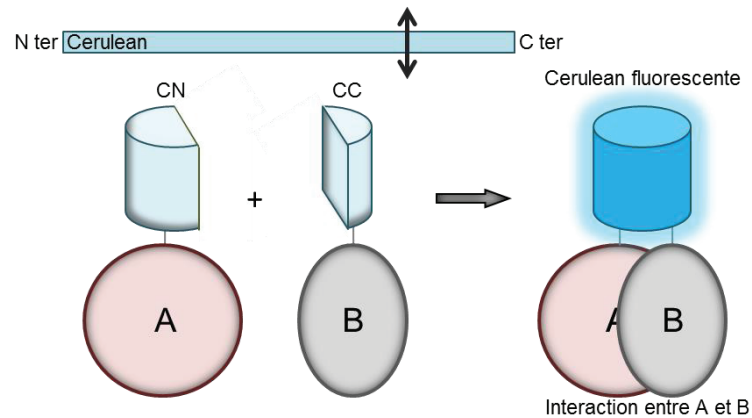
En conclusion, les approches « classiques » ne permettent pas d'isoler assez de nouveaux partenaires des protéines HOX, soit par leur contexte (différent de celui de la protéine endogène et nécessitant alors une vérification) soit par leur niveau de sensibilité trop faible qui ne permet pas de détecter des interactions faibles et dynamiques telles que celles engagées par les SLiMs.

La nécessité d'utiliser une technique capable de détecter des interactions faibles, transitoires, directement *in vivo* est apparue face à ce constat. La complémentation de fluorescence bimoléculaire s'est présentée comme la meilleure des solutions. En effet, cette technique de détection des IPPs est extrêmement sensible *in vivo*, et stabilise les interactions, permettant ainsi de détecter et de conserver des IPPs transitoires et faibles tout en les visualisant dans un contexte cellulaire spécifique. La démocratisation récente d'outils tels que le CRISPR-CAS9 permet d'envisager encore plus facilement des expressions de protéines à des niveaux endogènes.

#### 4.1.2. Principe de la complémentation de fluorescence bimoléculaire (BiFC)

La complémentation de fluorescence bimoléculaire repose sur une propriété intrinsèque des protéines fluorescentes monomériques : lorsqu'elles sont scindées en deux parties non fluorescentes, elles sont capables de se compléter par rapprochement spatial et ainsi reconstituer la fluorescence initiale de la protéine (**Kerppola, 2009**). La complémentation est aussi utilisée avec des enzymes comme la luciférase, la betagalactosidase ou plus récemment la Dam méthylase (**Hass et al., 2015**).

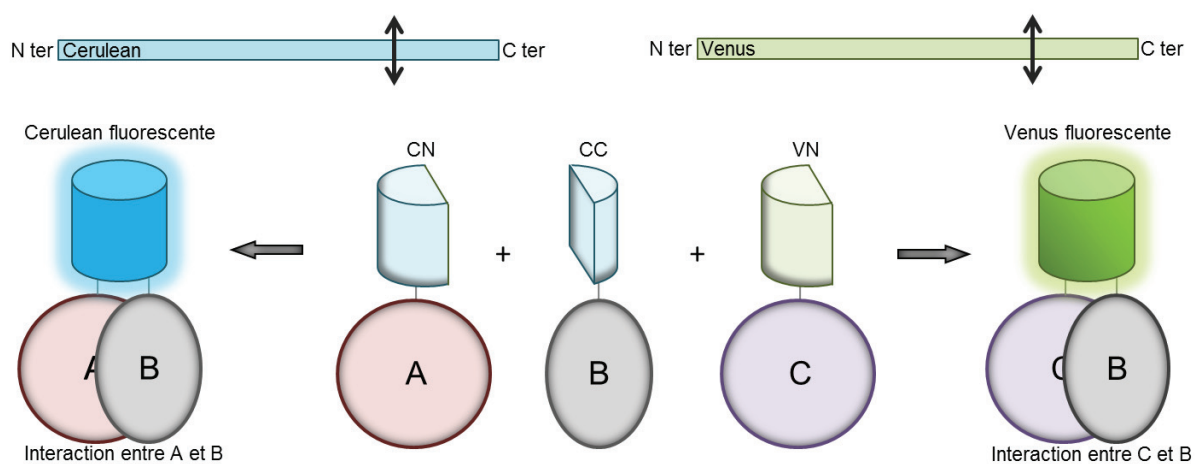
Cette propriété peut ainsi être utilisée pour fusionner deux protéines A et B à la partie N-terminale (CN) ou la partie C-terminale (CC) d'une protéine fluorescente bleue Cerulean. Si les deux protéines A et B interagissent *in vivo* au sein d'un même, alors les deux fragments seront assez proches spatialement pour reconstituer une protéine fluorescente Cerulean fonctionnelle capable d'émettre une fluorescence bleue sous une longueur d'excitation spécifique (**Figure 4.1.2.A**). La complémentation entre les deux moitiés de la protéine fluorescente induit la formation des liaisons covalentes ce qui stabilise le complexe protéique. Ainsi, des protéines exprimées faiblement et/ou à des taux endogènes, et des interactions très brèves peuvent être facilement mises en évidence car elles s'accumulent dans les tissus suite à leur stabilisation.



**Figure 4.1.2.A : Complémentation de Fluorescence Bimoléculaire avec la protéine Cerulean.** L'interaction entre les protéines A et B permet le rapprochement spatial des moitiés de la protéine fluorescente permettant la reconstitution de la Cerulean et donc l'émission d'un signal fluorescent.

On retrouve différentes protéines fluorescentes monomériques dont les propriétés de complémentation (vitesse, intensité de la fluorescence) vont varier d'une protéine à l'autre (**Kodama et Hu, 2012**). Par exemple, la protéine sfGFP (pour superfolder GFP) est décrite pour compléter beaucoup plus vite, et avoir un signal plus fort, mais elle nécessite quelques mutations afin de diminuer les interactions fortuites et donc les faux positifs (**Zhou et al., 2011**).

La protéine Cerulean présente de plus une particularité intéressante : lorsqu'elle est scindée en deux parties, sa partie C-terminale (la plus courte) est capable de compléter avec sa partie N-terminale (et faire un signal bleu) mais aussi avec la partie N-terminale d'une autre protéine fluorescente appelée Venus et émettre un signal vert (**Figure 4.1.2.B**). On parle alors de BiFC multicolore (**Kerppola, 2013**).



**Figure 4.1.2.B : Complémentation de Fluorescence Bimoléculaire Multicolore.** La partie C-terminale de la Cerulean (CC) peut aussi bien compléter avec la partie N-terminale de la Cerulean qu'avec celle de la Venus. La couleur du signal est déterminée par la partie N-terminale de la protéine fluorescente ainsi reconstituée. On peut mener un crible contre deux protéines différentes A et C et trouver des partenaires de l'une ou de l'autre en fonction de la couleur du signal émis par les complexes.

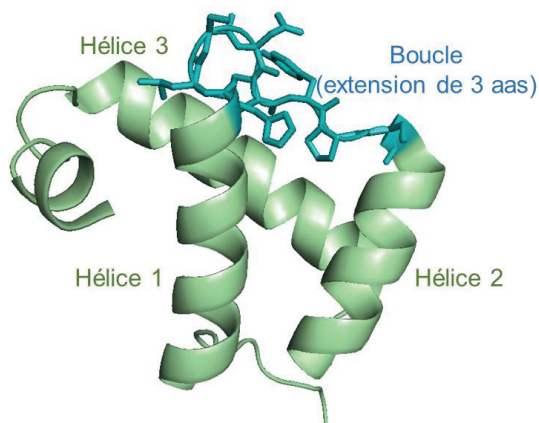
#### 4.1.3. Utilisations de la BiFC pour la recherche de nouveaux partenaires

La BiFC a trouvé de nombreuses utilisations pour des protéines très variées et dans différents organismes. Elle a par exemple été utilisée dans des cribles à l'échelle génomique. En effet, en 2006 Ding *et al.* conduisent un crible contre la protéine AKT-1 pour identifier de nouveaux partenaires de celle-ci dans des cellules HeLa (Ding *et al.*, 2006). Ils utilisent une protéine Venus (désignée dans cet article par IFP pour Intense Fluorescent Protein) dont la partie N terminale est fusionnée à la protéine AKT-1 et exprimée de façon stable dans la lignée cellulaire HeLa. Puis un retrovirus contenant la cassette C terminale de la Venus est mobilisé pour des insertions aléatoires en amont des gènes. L'insertion dans un gène codant pour un cofacteur de la protéine Akt1 permettra l'expression d'une protéine de fusion et entraîne l'apparition d'un signal fluorescent Venus. Les cellules fluorescentes sont alors triées et l'insertion est déterminée par séquençage avec des amorces spécifiques (Figure 4.1.2.2.B). Cette approche a permis d'identifier 24 nouvelles interactions. Plus récemment, cette même technique a été utilisée pour trouver des partenaires de la protéine LMP1 impliquée dans les infections par le virus Epstein-Barr (Talaty *et al.*, 2012).

La BiFC a plus récemment été utilisée pour analyser les interactions Hox-PBC et trouver de nouveaux cofacteurs des protéines HOX *in vivo*. Plus particulièrement la technique a été établie dans l'embryon vivant de drosophile, permettant d'analyser les interactions Hox-cofacteurs avec des niveaux normaux d'expression des protéines HOX (Hudry *et al.*, 2011). La méthode a ensuite été utilisée pour trouver de nouveaux partenaires transcriptionnels des protéines HOX (Hudry *et al.*, 2012). L'approche de type gènes candidats avec des FTs a permis de mettre en évidence que les protéines HOX pouvaient interagir avec des FTs très divers (Baeza, Viala, Heim *et al.*, 2015 : inclus dans le manuscrit au Chapitre II des Résultats).

#### 4.2. LES COFACTEURS A HD DE LA FAMILLE TALE

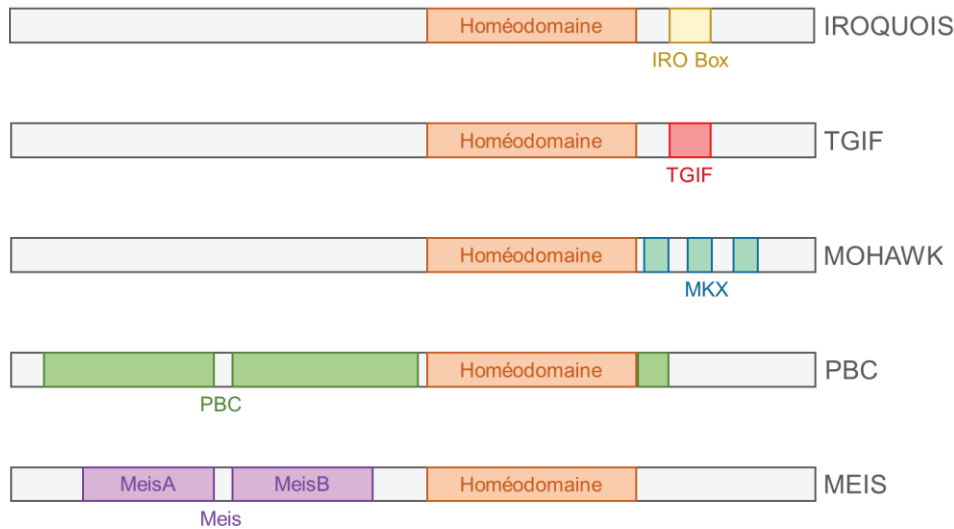
Les cofacteurs les mieux caractérisés des protéines HOX au niveau moléculaire et génétique appartiennent à la sous-classe des protéines TALE (pour three amino acid loop extension) des FTs à HD. Nous allons dans ce chapitre faire une description succincte de ces dernières. D'une façon générale, ce sont des FTs dont l'HD contient une extension de trois acides aminés entre les hélices 2 et 3 (Figure 4.2 ; Bertolino *et al.*, 1995).



**Figure 4.2 : Structure de l'homéodomaine de la protéine Pbx1.** Structure issue de la banque de données de la PDB. .

## 4.2.1. La famille TALE

Cette famille de protéines à HD « atypique » (**Bürglin et Affolter, 2016**) est retrouvée chez les plantes et les animaux (**Bharathan et al., 1997 ; Derelle et al., 2007**) démontrant ainsi une forte conservation évolutive. Chez les plantes, cette famille est divisée en deux classes, les BEL et les KNOX capables de s'hétérodimériser (**Lee et al., 2008**). Chez les animaux elle est divisée en cinq classes de protéines : PBC, MEIS, TGIF, IRO et MKX (**Figure 4.2.1.**) en fonction des autres domaines présents sur ces protéines (**Mukherjee et al., 2009**).



**Figure 4.2.1 :** Les membres de la famille TALE. Les cinq classes de facteurs à HD de la famille TALE sont ici schématisés avec leurs différents domaines spécifiques.

La classe **IRO** (IROQUOIS), définie par un motif IRO box d'environ quinze résidus en C-terminal est représentée par trois gènes dans la Drosophile (*arauca*, *caup*, et *mirror*) et par les gènes IRX dans les vertébrés (**Büglin, 1997**). On en trouve sept (IRX1 à 7) issus de duplications chez les tétrapodes (**Mukherjee et Bürglin, 2007**).

La classe **TGIF** n'est représentée que par deux gènes dans la Drosophile (*achinta* et *vismay*) et la famille des gènes TGIF chez les vertébrés dont le nombre est variable (**Hyman et al., 2003**). Ces protéines contiennent un motif TGIF box de 12 résidus en aval de l'HD riche en résidus hydrophobes et donc susceptibles de favoriser des IPPs (**Mukherjee et Bürglin, 2007**).

La classe **MKX** (MOHAWK) contient trois domaines MKX A, B et C en aval de l'HD aussi riches en résidus hydrophobes et donc susceptibles de favoriser des IPPs. Cette classe n'est représentée que par un seul gène dans l'ensemble des espèces à l'exception du Zebrafish qui en possède une duplication, et de *Caenorhabditis* qui l'a en revanche perdu. (**Mukherjee et Bürglin, 2007**).

La classe **PBC** (Pré-B cell leukemia transcription factor class) est très conservée et contient un large domaine PBC en deux parties en amont de l'HD et un plus petit motif en aval de celui-ci (**Bürglin et Ruvkun, 1992**). Au sein de cette famille, le FT Exd chez la Drosophile est très bien décrit pour interagir avec les protéines HOX ainsi que les Pbx 1 à 4 chez les vertébrés (**chapitre 4.2.2.**).

Pour finir, la classe **MEIS** (Myeloid Ecotropic Integration Site) comprend les membres des sous-classes MEIS et PREP (PBX Regulatory Protein) qui ont toutes deux un domaine MEIS en deux parties A et B (**Bürglin, 1997**) dont la partie A est plus courte dans les Prep que

dans les Meis. Les protéines Meis sont bien conservées, on n'en trouve qu'une copie dans les protostome et trois dans les tetrapodes. Les gènes des protéines Prep sont au nombre de deux chez les mammifères.

Nous allons nous intéresser surtout aux deux dernières classes car elles sont largement décrites dans les interactions avec les protéines HOX. Afin d'exercer leur rôle de FT, les protéines de la famille PBC sont transloquées dans le noyau grâce à leur interaction avec leurs partenaires Meis cytoplasmiques via le domaine PBC-A (**Rieckhof et al., 1997**). La localisation de ces dimères est dépendante de la balance entre l'import et l'export nucléaire médié par les NLS et NES (signal de localisation nucléaire, NLS ; Signal d'export nucléaire, NES) des protéines (**Stevens et Mann, 2007**).

#### 4.2.2. La formation des complexes Hox/Pbx/Meis

Les protéines HOX sont capables de lier une séquence nucléotidique courte de type TAAT ou TTAT. La dimérisation de ces protéines avec les protéines de la famille PBC permet d'augmenter la taille du site de reconnaissance à l'ADN et donc sa spécificité. Les deux HD vont reconnaître de manière coopérative une séquence octanucléotidique TGATXXAT (**Gehring et al., 1994 ; Chan et al., 1996 ; Piper et al., 1999**). La formation des dimères augmente donc l'affinité des deux protéines pour l'ADN.

Comme nous l'avons mentionné plus haut, par des techniques de cristallographie la surface d'interaction entre ces protéines au contact de l'ADN a été établie et le motif HX semble être le siège de celle-ci (**Joshi et al., 2007 ; Mann et al., 2009**). En effet, la plupart des résultats *in vitro* démontrent que la mutation de ce motif entraîne la perte d'interaction avec les partenaires PBC (**Knoepfler et al., 1995 ; Neuteboom et al., 1995**).

Plus récemment il a été mis en évidence *in vitro* qu'il était possible de sauver la perte de l'interaction entre les protéines HOX et leurs partenaires PBC induite par la mutation de l'HX en ajoutant le partenaire MEIS (**Hudry et al., 2012**). En effet, pour la plupart des protéines HOX de *Drosophila* ou de souris mutées dans le motif HX, la présence de Meis restaure la formation de complexes protéiques sur l'ADN suggérant que d'autres régions puissent remplir le rôle d'interaction avec les cofacteurs TALEs dans les protéines HOX. Seuls deux motifs additionnels ont pu être identifiés pour leur rôle dans l'interaction avec les partenaires TALEs, les motifs TDWM et UbdA. Des données cristallographiques ont permis d'avoir les détails moléculaires de l'interaction dans le cas du motif UbdA (**Foos et al., 2015**). En effet, ce motif intervient comme une extension de la troisième hélice de l'HD qui va contacter le partenaire PBC par le côté opposé à celui utilisé par le motif HX sur l'ADN, démontrant ainsi la flexibilité de ces interactions.

De plus, au-delà de leur rôle pour l'adressage et le transport des protéines PBC, les protéines MEIS sont aussi capable d'interagir avec les protéines HOX (**Mann et al., 2009 ; Amin et al., 2015 ; Merabet et Lohmann, 2015**). La formation de complexes Hox-Meis sur l'ADN n'a lieu que pour les protéines HOX des groupes postérieurs mais la pertinence fonctionnelle de ces interactions reste à démontrer. Ces observations soulignent la complexité moléculaire sous-jacente aux interactions Hox-TALE *in vivo*. Il a été suggéré que cette complexité pourrait être à la base des fonctions développementales spécifiques des complexes Hox-TALEs (**Merabet et al., 2009**) et de l'apparition du système Hox-TALEs de patterning le long de l'axe antéro-postérieur chez les animaux (**Hudry et al., 2014**).



### 4.3. AUTRES PARTENAIRES DES PROTEINES HOX

Le paradoxe entre les rôles spécifiques de chaque protéine Hox et le fait qu'elles partagent une même courte séquence de reconnaissance à l'ADN a poussé les biologistes à tenter d'identifier de nouveaux partenaires des protéines HOX qui pourraient avoir un rôle dans la régulation de la transcription (**Rezsohazy et al., 2015**). Nous allons ci-après discuter de certains d'entre eux.

#### 4.3.1. Des facteurs de transcription

Les protéines HOX sont capables de réguler la mise en place ainsi que le devenir de certains tissus sans la présence de leur partenaires PBC (**Galant et al., 2002 ; Perrin et al., 2004 ; Penkov et al., 2013**) ce qui suppose qu'elles possèdent d'autres partenaires pour la régulation de la transcription. Des expériences de CHIP-seq ont mis en évidence de nombreux sites de liaison consensus de FTs connus à proximité des sites de liaison des protéines HOX. C'est le cas par exemple des protéines CREB pour *Hoxa2* (**Donaldson et al., 2012**) ou RUNX1 pour *Hoxb4* (**Oshima et al., 2011**) dont les interactions et les rôles restent à confirmer. D'autres partenaires sont issus de cribles génétiques (**Kennison et Tamkun, 1988 ; Harding et al., 1995, Florence et al., 1998 ; Merabet et al., 2002**) ou de cribles en double-hybride chez la levure (**Bondos et al., 2006 ; Lambert et al., 2012**) et restent aussi à confirmer sur des plans génétiques ou moléculaires.

Nous avons soulevé dans une revue publiée en 2014 (**Merabet et Dard, 2014**) la nécessité de mener des cribles dans des contextes biologiques spécifiques et propres à l'expression des protéines HOX afin de découvrir de nouveaux partenaires transcriptionnels spécifiques. En effet, seule une faible proportion des protéines référencées dans ces banques correspond à des FTs. De plus, la plupart de ceux identifiés tels que des FTs à doigt de zinc (**Suzuki et al., 2003 ; Robertson et al., 2004 ; Mahaffey, 2005**), des effecteurs nucléaires des voies de signalisation (**Merabet et al., 2002 ; Bondos et al., 2006 ; Walsh et Carroll, 2007**) et des FTs spécifiques (**Gebelein et al., 2004 ; Witt et al., 2010**) sont décrit respectivement pour distinguer, régionaliser ou spécifier les activités transcriptionnelles des protéines HOX. Cependant seule une faible proportion d'études s'est intéressée à ces interactions sur leurs gènes cibles (**Sorge et al., 2015**). Une liste de ces partenaires transcriptionnels a été dressée par Rezsohazy et al. (**2015**).

#### 4.3.2. Des régulateurs de la chromatine

La précipitation de la chromatine (CHIP-seq) associée aux protéines Ubx ou *Hoxa9* a mis en évidence un enrichissement en site de liaisons de régulateurs de la chromatine sur l'ADN. Ainsi par exemple, de nombreux motifs de liaison des protéines du groupe Polycomb (**Agrawal et al., 2011**) ou d'histone acétyl transférases telles que CBP/p300 (**Huang et al., 2012**) sont retrouvées à proximité des sites de liaison des protéines HOX.

#### 4.3.3. Autres ?

De nombreux partenaires ont été découverts pour les protéines HOX hors du contexte de la régulation de la transcription. Nous les avons développés plus haut lorsque nous discutons des rôles de protéines HOX. En effet, ces rôles ont été découverts et compris pour certains grâce à la mise en évidence de ces nouveaux partenaires inattendus (**chapitre II.2**).

Il est important de noter qu'une grande partie des IPPs référencées dans les banques telles que DroID (**Murali *et al.*, 2011**), BioGRID (**Chatr-Aryamontri *et al.*, 2015**), BIND, DIP, MINT ou IntAct (**Orchard, 2012**) ne sont pas encore bien décrites. En effet, elles sont pour la plupart issues de cribles à large échelle, à la mode depuis quelques années, mais n'ont encore aucune description fonctionnelle. De plus, ces IPPs peuvent être issues de cribles protéomiques ou de cribles génétiques, questionnant alors sur la pertinence de certaines d'entre elles (**Rezsöhazi, 2014**).

Cette richesse nouvelle des partenaires des protéines HOX ainsi que la plasticité d'interaction de celles-ci associée aux SLiMs suggèrent un nombre étourdissant de possibilités qui restent encore à décrire et à comprendre. Mon travail de thèse tente modestement d'apporter sa pierre à cet édifice rempli de parad'hox.



# RESULTATS

---

## 1. PREMIER CHAPITRE : AUX ORIGINES DES COMPLEXES HOX-TALE

MOLECULAR INSIGHTS INTO THE ORIGIN OF THE HOX-TALE PATTERNING SYSTEM.

Hudry B, Thomas-Chollier M, Volovik Y, Duffraisse M, **Dard A**, Frank D, Technau U, Merabet S.

*Elife. 2014*

## 2. DEUXIEME CHAPITRE : DES PETITS MOTIFS POUR REGULER LE POTENTIEL D'INTERACTION DES PROTEINES HOX

INHIBITORY ACTIVITIES OF SHORT LINEAR MOTIFS UNDERLIE HOX INTERACTOME SPECIFICITY IN VIVO

Baëza M, Viala S, Heim M, **Dard A**, Hudry B, Duffraisse M, Rogulja-Ortmann A, Brun C, Merabet S.

*Elife. 2015*

## 3. TROISIEME CHAPITRE : IDENTIFIER DE NOUVELLES INTERFACES POUR LES COFACTEURS TALE AU SEIN DES PROTEINES HOX HUMAINES

SYSTEMATIC ANALYSIS OF MOLECULAR COMPLEX FORMATION BETWEEN HOX AND TALE PROTEINS REVEALS PARALOG-SPECIFIC TALE INTERACTION DOMAINS IN HUMAN HOX PROTEINS

**Amélie Dard**, Françoise Bleicher, Marilyne Duffraisse, Catherine Lavau and Samir Merabet

4. **QUATRIEME CHAPITRE : TROUVER DE NOUVEAUX PARTENAIRES  
DES PROTEINES HOX SPECIFIQUES D'UN CONTEXTE CANCEREUX**

TRACKING CONTEXT-SPECIFIC TRANSCRIPTION FACTORS REGULATING HOX  
ACTIVITY.

Merabet S, **Dard A.**

*Dev Dyn. 2014*

UN CRIBLE LARGE ECHELLE EN BiFC POUR TROUVER DES PARTENAIRES DES  
PROTEINES HOX DANS DES LIGNEES CELLULAIRES DE CANCER HUMAIN.

MERABET Samir, MANGEOT Philippe, **DARD Amélie**, BLEICHER Françoise, HUGUES  
Sandrine, GILLET Benjamin

## 1. PREMIER CHAPITRE : AUX ORIGINES DES COMPLEXES HOX-TALE

La majorité des animaux (de l'Homme à la mouche en passant par les poissons, les oiseaux, etc.) possède une symétrie bilatérale, c'est-à-dire une symétrie organisée selon un axe longitudinal de la tête à la queue. Il est largement décrit que les protéines HOX et leurs partenaires TALE participent à la mise en place de l'axe antéro-postérieur et ont joué un rôle majeur pour la diversification des formes le long de cet axe chez les bilatériens. Pourtant, on retrouve ces gènes dans le génome des cnidaires, qui se caractérisent par une symétrie radiale, comme chez l'anémone de mer *Nematostella vectensis*. Nous avons pu montrer que les protéines HOX et TALE de *Nematostella* étaient capables d'interagir selon des modes complexes comme leurs homologues de bilatériens. De plus elles sont aussi capables de se substituer aux protéines HOX de drosophile, montrant une étroite conservation à la fois moléculaire et fonctionnelle de ces complexes.

L'ensemble des résultats présentés dans l'article du chapitre I met en évidence comment les familles Hox et TALE ont pu établir un partenariat pendant l'évolution des eucaryotes afin de constituer un des principaux systèmes de spécification des axes longitudinaux chez les Eumetazoaires.



# Molecular insights into the origin of the Hox-TALE patterning system

Bruno Hudry<sup>1\*</sup>, Morgane Thomas-Chollier<sup>2</sup>, Yael Volovik<sup>3</sup>, Marilyne Duffraisse<sup>4</sup>, Amélie Dard<sup>4</sup>, Dale Frank<sup>3</sup>, Ulrich Technau<sup>5</sup>, Samir Merabet<sup>4\*</sup>

<sup>1</sup>MRC Clinical Sciences Centre, Faculty of Medicine, Imperial College London, London, United Kingdom; <sup>2</sup>Unité Mixte de Recherche (UMR) 8197, INSERM U1024, Institut de Biologie de l'ENS (IBENS), Ecole Normale Supérieure (ENS), Centre National de Recherche Scientifique (CNRS), Institut National de la Santé et de la Recherche Scientifique (INSERM), Paris, France; <sup>3</sup>Department of Biochemistry, The Rappaport Family Institute for Research in the Medical Sciences, Faculty of Medicine, Technion, Israel Institute of Technology, Haifa, Israel; <sup>4</sup>Unité Mixte de Recherche (UMR) 5242, Institut de Génomique Fonctionnelle de Lyon (IGFL), Ecole Normale Supérieure (ENS) de Lyon, Centre National de Recherche Scientifique (CNRS), Lyon, France; <sup>5</sup>Department für Molekulare Evolution und Entwicklung, University of Vienna, Vienna, Austria

**Abstract** Despite tremendous body form diversity in nature, bilaterian animals share common sets of developmental genes that display conserved expression patterns in the embryo. Among them are the Hox genes, which define different identities along the anterior–posterior axis. Hox proteins exert their function by interaction with TALE transcription factors. Hox and TALE members are also present in some but not all non-bilaterian phyla, raising the question of how Hox–TALE interactions evolved to provide positional information. By using proteins from unicellular and multicellular lineages, we showed that these networks emerged from an ancestral generic motif present in Hox and other related protein families. Interestingly, Hox-TALE networks experienced additional and extensive molecular innovations that were likely crucial for differentiating Hox functions along body plans. Together our results highlight how homeobox gene families evolved during eukaryote evolution to eventually constitute a major patterning system in Eumetazoans.

DOI: [10.7554/eLife.01939.001](https://doi.org/10.7554/eLife.01939.001)

\*For correspondence: [petithudry@gmail.com](mailto:petithudry@gmail.com) (BH); [samir.merabet@ens-lyon.fr](mailto:samir.merabet@ens-lyon.fr) (SM)

**Competing interests:** The authors declare that no competing interests exist.


**Funding:** See page 21

**Received:** 09 December 2013

**Accepted:** 01 February 2014

**Published:** 18 March 2014

**Reviewing editor:** Diethard Tautz, Max Planck Institute for Evolutionary Biology, Germany

 Copyright Hudry et al. This article is distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use and redistribution provided that the original author and source are credited.

## Introduction

‘What is an animal?’ In 1993, Slack et al. proposed to define an animal by the zootype (*Slack et al., 1993*), a concept illustrating the strong conservation of embryonic expression profiles of developmental genes observed in different bilaterian phyla at that time. Since then, it was found that developmental genes could also display highly dissimilar expression patterns or even be absent in non-bilaterian lineages, showing that the genetic mechanisms underlying embryonic development are not universal.

One major class of developmental genes that was historically considered as highly conserved in the animal kingdom is the Hox genes. Hox genes are expressed along the anteroposterior axis of all bilaterian animals, providing positional information during embryogenesis. Given their important patterning roles, Hox genes are thought to have strongly contributed to morphological diversification of bilaterian organisms during evolution. Accordingly, numerous examples have shown that modifications in Hox genes number, expression, and/or activity could correlate to morphological variations across bilaterian lineages (*Heffer and Pick, 2013*).

The role of Hox genes suffers more ambiguity outside Bilateria, and in particular in the sister Cnidaria group. Cnidarians do contain Hox genes and have undergone a wide range of morphological



**eLife digest** Any animal with a body that is symmetric about an imaginary line that runs from its head to its tail is known as a bilaterian. Humans and most animals are bilateral, whereas jellyfish and starfish are not. Bilateral symmetry can take many forms—as demonstrated by the differences between flies, frogs and humans—but all bilaterians express many of the same genes during development.

One of these groups of genes is known as the Hox family. The expression of specific Hox genes at specific times instructs cells in the developing embryo to adopt different fates according to their position along the anterior–posterior (head to tail) axis. The patterning function of Hox genes relies on the presence of two additional cofactors that belong to the so-called TALE family. Although both Hox and TALE proteins were present early on during animal evolution, it is unclear how and when the interactions between them first began to generate symmetrical body plans.

Now, Hudry et al. have provided insights into the origin of the Hox-TALE network by analysing the expression and molecular properties of Hox and TALE proteins from various multicellular and unicellular organisms. These experiments revealed that Hox and TALE proteins of the sea anemone *Nematostella*, which belongs to a group of animals called cnidarians that have radial rather than bilateral symmetry, interact with one another in a similar manner to the interactions seen in bilaterians.

Hudry et al. then showed that two *Nematostella* Hox genes were able to substitute for their bilaterian equivalents in fruit flies, and that a *Nematostella* TALE gene was able to take over neuronal functions of its equivalent in *Xenopus* frogs. This striking conservation of function between species suggests that Hox and TALE genes were already working together in the common ancestor of all bilaterian and cnidarian animals.

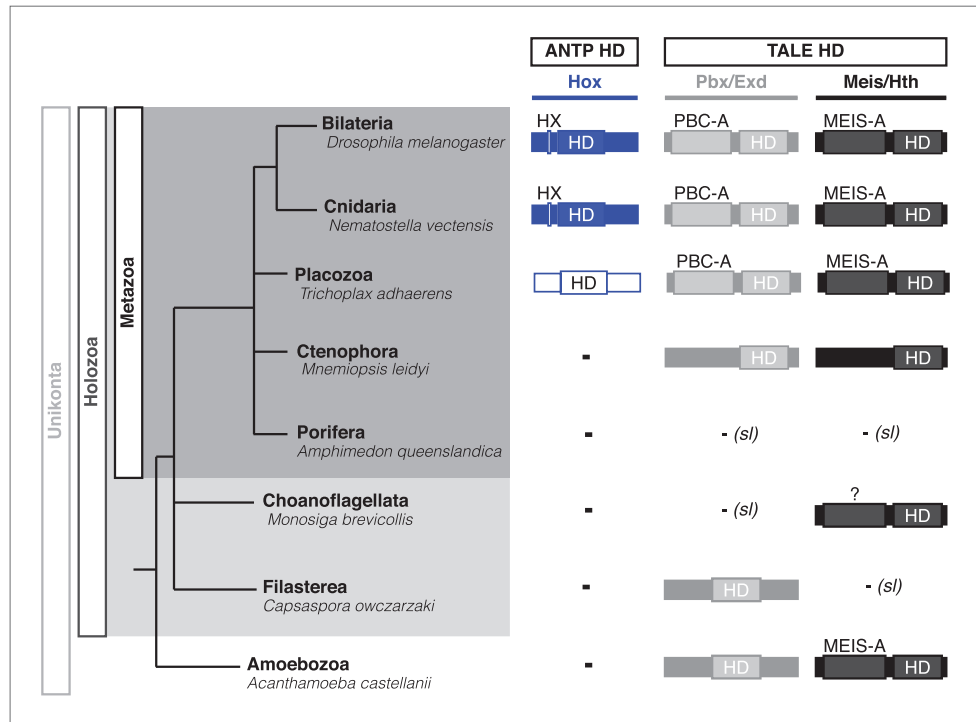
By contrast, TALE members from a unicellular amoeba were unable to interact with Hox proteins, suggesting that Hox–TALE interactions first emerged in multicellular animals. In addition to increasing our knowledge of highly conserved Hox signalling, these data provide insight into the molecular mechanisms that gave rise to the symmetrical body plan that has been adopted, and adapted, by the majority of animals since.

DOI: [10.7554/eLife.01939.002](https://doi.org/10.7554/eLife.01939.002)

radiation during evolution. However, Hox genes have neither a clear collinear nor conserved expression profile along the primary (oral–aboral) axis of the planula larva among different cnidarian lineages (Finnerty et al., 2004; Kamm et al., 2006), which renders their orthology status difficult to assign (specially for posterior/non-anterior cnidarian Hox genes). A role of Hox genes in axis patterning was proposed in the hydrozoan *Eleuthera dichotoma* (Jakob and Schierwater, 2007), but these results are limited by the fact that functional analyses were performed in medusa, a particular developmental stage that is not shared by all cnidarian species. The role of cnidarian Hox genes during early larval stages is, however, currently unclear.

Interestingly, Cnidaria is the only non-bilaterian phylum, which has a bona fide Hox repertoire, whereas others, including ctenophores, sponges, and placozoans, lack Hox genes (Figure 1). This raises the important question of how the Hox gene family acquired its crucial axial patterning functions during metazoan evolution. In Bilateria, Hox patterning functions rely on the presence of the PBC and Meis proteins, which are also present in non-bilaterian phyla (Figure 1). We thus assessed whether a Hox/PBC/Meis network could exist outside Bilateria, and if so whether or not it would rely on identical molecular rules as observed in Bilateria.

Hox and PBC/Meis proteins belong to the ANTP (Antennapedia) and TALE (Three amino acids loop extension) class of homeodomain (HD)-containing TFs, respectively (Saina et al., 2009). The Hox/PBC/Meis network relies on interactions between PBC and Meis proteins on one side, and on interactions between Hox and PBC proteins on the other side (Mann et al., 2009). Some posterior vertebrate Hox members do form dimeric complexes with Meis (Shen et al., 1997; Williams et al., 2005), but these interactions constitute a vertebrate innovation rather than a general rule in Bilateria. Interactions between PBC and Meis occur through conserved regions localized upstream of the HD of both proteins (called PBC-A and MEIS-A domains) and which are thought to derive from a common ancestor domain (Burglin, 1998). These interactions allow the nuclear translocation and stability of



**Figure 1.** Phylogeny of Hox (ANTP superclass) and PBC/Meis (TALE superclass) proteins across eukaryote evolution. Protein motifs required for Hox/PBC/Meis network are indicated when present: homeodomain (HD), hexapeptide (HX), PBC-A, and MEIS-A. Absence of the member in a given group is considered as resulting from a secondary loss (sl), when the ortholog is present more ancestrally. A question mark is indicated for Meis of *Monosiga brevicollis* because of incomplete sequence. The protein indicated in Placozoa is not coloured in blue since it is not a true Hox protein (see main text for details). Examples are provided for a representative species of each group. Members of PBC and Meis classes are called Pbx or Extradenticle (Exd) and Meis or Homothorax (Hth), respectively.

DOI: 10.7554/eLife.01939.003

PBC (*Abu-Shaar et al., 1999*). Interactions between Hox and PBC involve a short conserved motif in Hox proteins, called the hexapeptide (HX), which folds within the hydrophobic pocket formed in part by the extra-three residues of the HD of PBC (*Merabet et al., 2009*). This motif contains a conserved core sequence of four residues in all but AbdB-group Hox proteins, which retain a single tryptophan residue. Additional specific signatures can also be found in the HX of some paralog groups (*Merabet et al., 2009*). Recent data showed that Hox proteins could also interact with PBC partners through other more specific motifs. These alternative interaction modes can be induced by the DNA-binding of the Meis partner, eventually leading to different three-dimensional conformations that could be important for paralog-specific functions (*Galant et al., 2002; Merabet et al., 2007; Hudry et al., 2012*).

PBC and Meis representatives are found from unicellular Amoebozoa and Filasterea groups to metazoan lineages including Ctenophora and Placozoa (*Figure 1*). Sequence analysis shows that only one PBC or Meis representative is present, or that the protein does not contain the PBC-A or MEIS-A interaction domain in most of these lineages (*Figure 1, Figure 2—figure supplements 1 and 2*). Interestingly, all protein features required for PBC/Meis partnership appeared concomitantly with the presence of Hox or Hox-like proteins in Metazoa (*Figure 1*). Of note, the representative species of Placozoa, *Trichoplax adhaerens* (*Ta*), contains a protein that was classified as a ProtoHox (*Schierwater and Kuhn, 1998*) or ParaHox (*Mendivil Ramos et al., 2012*) member (*Figure 1*). However, the absence of any HX motif (*Schierwater and Kuhn, 1998*) suggests that this protein could not interact with PBC/Meis, which is confirmed later (see last section of 'Results').

Protein sequence analysis indicates that a Hox/PBC/Meis network could first be present in Cnidaria. To test this hypothesis, we dissected the molecular properties underlying the formation of

the Hox/PBC/Meis interaction network of the sea anemone *Nematostella vectensis* (Nv), a cnidarian species exhibiting an internal symmetry organized along oral–aboral (primary) and directive (secondary) body axes (**Figure 2A**). This analysis was performed in vitro and in vivo and completed by heterologous functional assays in vertebrate and invertebrate species. Since other members of the ANTP superclass are described to interact with PBC in Bilateria, we also searched for the molecular mechanisms that allowed the emergence of the Hox–TALE network during evolution of homeobox gene families.

Our results show that Hox and TALE proteins from *Nematostella* form interaction networks and perform similar functions to their bilaterian counterparts. Although these networks rely on intricate molecular properties, they originated from an ancestral generic mode of interaction that was kept in other homeobox gene families. Overall our study describes how the molecular cues underlying the Hox–TALE patterning system in Bilateria was established stepwise during eukaryote evolution.

## Results

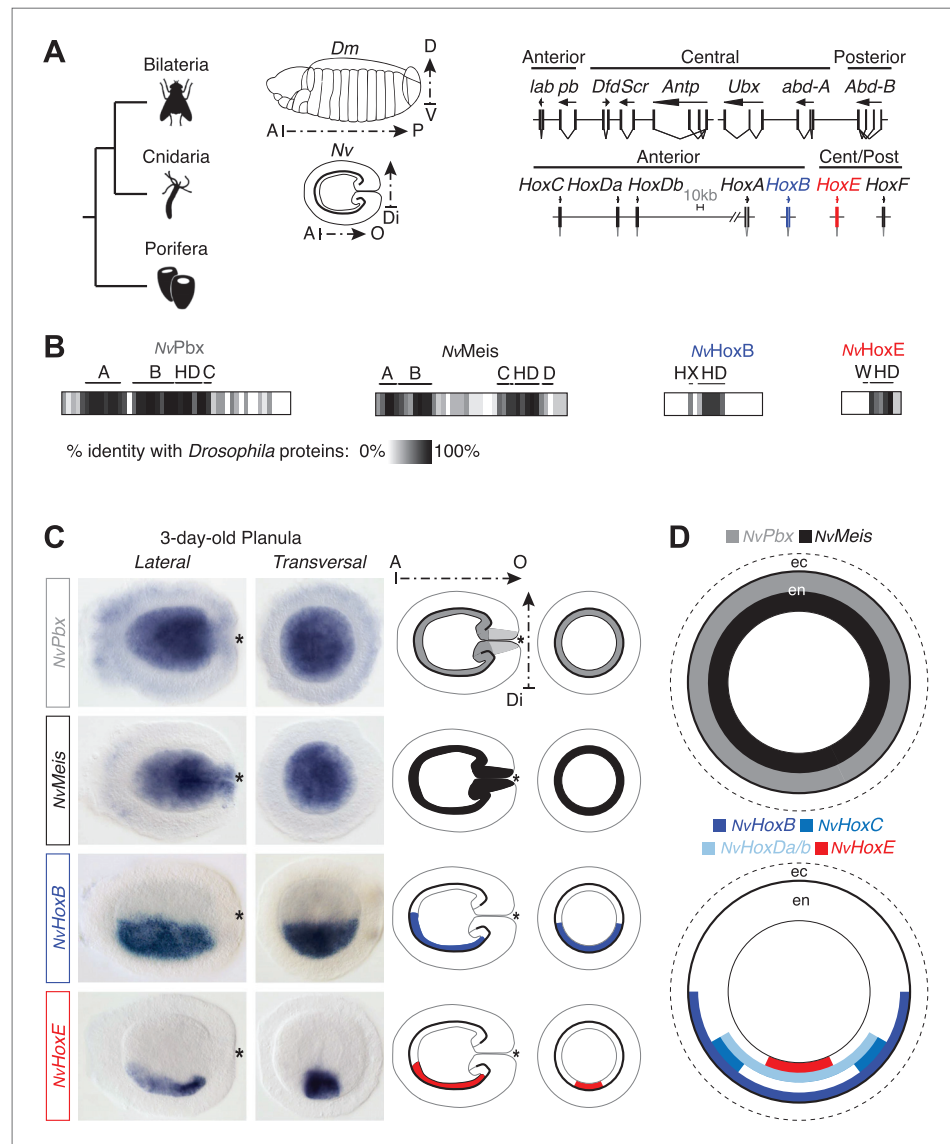
### Hox and TALE members are co-expressed and form protein complexes in the *Nematostella* embryo

The *Nematostella* genome contains seven Hox genes and one representative of the PBC (*NvPbx*) or Meis (*NvMeis*) class (**Figure 2A**). *NvPbx* and *NvMeis* proteins show a high level of sequence conservation with their bilaterian counterparts, especially in the regions encompassing the HD, PBC-A, and Meis-A domains (**Figure 2B**, **Figure 2—figure supplements 1 and 2**). In contrast, sequence similarity between cnidarian Hox proteins and their bilaterian homologs is restricted to the region encompassing the HD, as exemplified for *NvHoxB* (**Figure 2B**, **Figure 2—figure supplement 3**) and *NvHoxE* (**Figure 2B**, **Figure 2—figure supplement 4**). Some *NvHox* proteins do also contain a HX motif, as noticed in *NvHoxB* or *NvHoxE* (**Figure 2B**, **Figure 2—figure supplements 3 and 4**). The HX of *NvHoxE* is more divergent, corresponding to a single tryptophan residue as found in bilaterian Hox posterior paralog groups. Still, the identity of *NvHoxE* (as well as those of *NvHoxF*) remains controversial, being classified as a cnidarian-specific (**Chourrout et al., 2006; Ryan et al., 2006**), posterior (**Gauchat et al., 2000; Ryan et al., 2007**), or central Hox gene (**Thomas-Chollier et al., 2010**). This is not the case for *NvHoxC*, *NvHoxDa*, *NvHoxDb*, *NvHoxA* and *NvHoxB*, which are unambiguously assigned as anterior Hox genes (**Chourrout et al., 2006; Ryan et al., 2006, 2007; Thomas-Chollier et al., 2010; Figure 2A**).

To first verify that Hox, PBC, and Meis products from *Nematostella* could form an interaction network in vivo, we performed in situ hybridization experiments using *NvPbx* (**Matus et al., 2006**) and *NvMeis* probes. Results showed that the transcripts of these two genes are co-expressed in the entire endoderm of the larva (**Figure 2C–D**). Interestingly, several Hox genes were previously described to be expressed in staggered domains along the directive axis in the same tissue (**Finnerty et al., 2004; Matus et al., 2006; Ryan et al., 2007**), which was here confirmed for *NvHoxB* and *NvHoxE* (**Figure 2C–D**).

The relationship between *NvPbx* and *NvMeis* was then analysed by expressing a tagged version of *NvPbx*, alone or with *NvMeis*. Analyses were performed in a 24-hr-old embryo, at a stage when endogenous *NvMeis* is not yet expressed. We observed that the nuclear accumulation of mCherry-*NvPbx* fusion protein is contingent upon co-injection with *NvMeis* (**Figure 3A**). Thus, *NvMeis* is able to stabilize *NvPbx*, eventually triggering its nuclear accumulation. This observation is reminiscent of Pbx/Meis relationships in bilaterians (**Abu-Shaar et al., 1999; Saleh et al., 2000**). We further confirmed this result by visualizing *NvPbx*/*NvMeis* complexes directly in live *Nematostella* embryos through BiFC (Bimolecular Complementation Fluorescence). BiFC relies on the property of fluorescent proteins to be reconstituted when their two non-fluorescent sub-fragments are close enough in space. This method has been developed in several animal model systems to validate interactions between two candidate partners in vivo (**Kodama and Hu, 2012**). In this study, we co-expressed *NvPbx* and *NvMeis* fused respectively to the N-terminal (VN) or C-terminal (VC) fragment of Venus. This resulted in fluorescent signals in the cytoplasm (where interaction occurs first) and nuclei of embryonic cells (**Figure 3B**). No BiFC was obtained between a fusion construct and the complementary isolated VC or VN fragment, highlighting that the interaction between *NvPbx* and *NvMeis* fusion proteins was not artificially induced by the inherent affinity of the VN and VC fragments (**Figure 3B**).

BiFC was also used to visualize interactions between *NvPbx* and *NvHoxE* (**Figure 3C**). The specificity of this interaction was validated by the absence of BiFC between *NvHoxE* and a DNA-binding



**Figure 2.** Hox and TALE members are co-expressed in the endoderm of the *Nematostella* embryo. **(A)** Genomic organisation of Hox genes in *Nematostella vectensis* (*Nv*) and *Drosophila melanogaster* (*Dm*), two representative species of Cnidaria and Bilateria, respectively. Embryos at the planula stage are schematized; A–P: anterior–posterior; D–V: dorsal–ventral; A–O: aboral–oral; Di: directive axis. The *Nematostella* embryo is oriented according to recent findings (*Sinigaglia et al., 2013*). The nomenclature is calqued on (*Chourrou et al., 2006*) to avoid confusions with bilaterian Hox paralogs: *NvHoxC* (*antHox7*), *NvHoxDa* (*antHox8*), *NvHoxDb* (*antHox8a*), *NvHoxA* (*antHox6*), *NvHoxB* (*antHox6a*), *NvHoxE* (*antHox1a*), *NvHoxF* (*antHox1*). The two *Nematostella* Hox genes under study, *NvHoxB* and *NvHoxE*, are highlighted in blue and red respectively. Note that the same colour code is used in other figures. **(B)** Sequence identity between *Nematostella* and *Drosophila* proteins. *NvHoxB* and *NvHoxE* are compared to Labial (*Lab*) and Ultrabithorax (*Ubx*) respectively. The percentage of identity is represented by a grayscale gradient. Conserved domains in bilaterian TALE proteins are indicated (**A**, **B**, **C**, **D**). HX: hexapeptide. HD: homeodomain. See also **Figure 2—figure supplements 1–4**. **(C)** In situ hybridization of *NvPbx*, *NvMeis*, *NvHoxB* and *NvHoxE* in a three-day-old *Nematostella* planula. These four genes are expressed in the endoderm (en), as illustrated in **(D)**. Ec: ectoderm. *NvPbx* and *NvMeis* are illustrated in grey and black, respectively. This colour code is used in other figures. DOI: 10.7554/eLife.01939.004

The following figure supplements are available for figure 2:

**Figure supplement 1.** Protein sequence alignment of PBC members from representative species of Unikonta. DOI: 10.7554/eLife.01939.005

Figure 2. Continued on next page

Figure 2. Continued

**Figure supplement 2.** Protein sequence alignment of Meis members from representative species of Unikonta.

DOI: [10.7554/eLife.01939.006](https://doi.org/10.7554/eLife.01939.006)

**Figure supplement 3.** Protein sequence alignment between NvHoxB and the Labial (Lab) protein from *Drosophila melanogaster*.

DOI: [10.7554/eLife.01939.007](https://doi.org/10.7554/eLife.01939.007)

**Figure supplement 4.** Protein sequence alignment between NvHoxE and the Ultrabithorax (Ubx) or AbdominalB (AbdB) protein from *Drosophila melanogaster*.

DOI: [10.7554/eLife.01939.008](https://doi.org/10.7554/eLife.01939.008)

---

deficient form of NvPbx (**Figure 3C**), showing that the formation of the cnidarian Hox/Pbx complex is DNA-binding dependent, as previously noticed in bilaterians (**Hudry et al., 2011, 2012**). Altogether, these results show that NvPbx and NvMeis are co-expressed with several NvHox genes in the endoderm and that NvHox and NvTALE proteins can constitute an interaction network in vivo.

### Interaction properties between *Nematostella* and bilaterian Hox/TALE complexes are highly similar in vitro

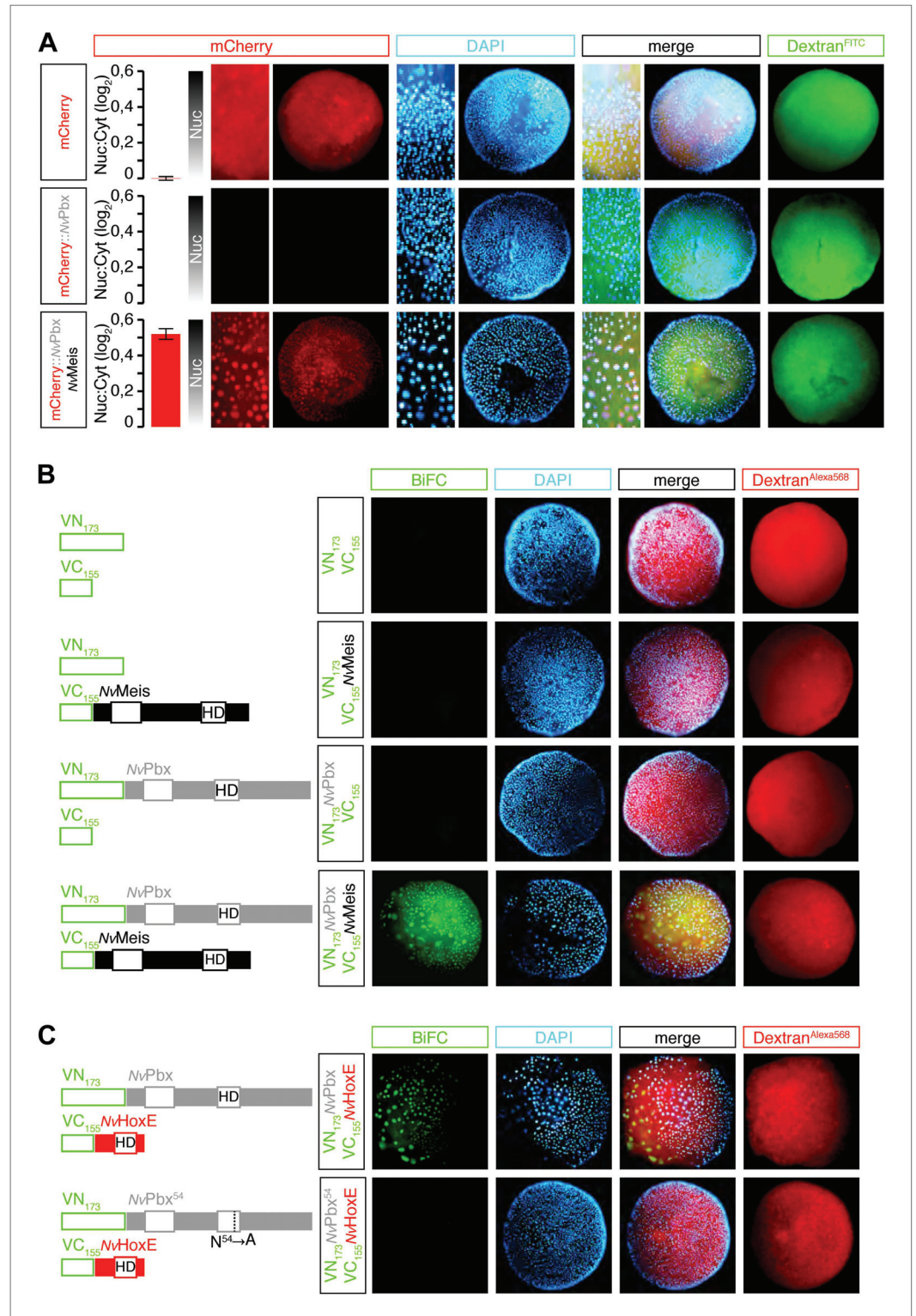
Next, we analysed the molecular properties underlying the assembly of *Nematostella* Hox/Pbx/Meis complexes in vitro. NvPbx and NvMeis proteins were previously shown to interact with bilaterian Hox proteins (**Hudry et al., 2012**). Here, we conducted binding assays with NvHoxB and NvHoxE proteins and assembly properties of protein complexes were measured by electromobility shift assays (EMSAs) on three different DNA probes. DNA probes differ by one nucleotide in the Hox/Pbx binding site (**Figure 4A**), each one corresponding to the preferential DNA-binding sites of previously defined anterior, central, and posterior Hox/Pbx complexes with vertebrate and invertebrate proteins (**Shen et al., 1996; Slattery et al., 2011**).

We observed that NvHoxB/NvPbx and NvHoxE/NvPbx complexes display anterior and central DNA-binding preferences, respectively (**Figure 4B–B'**). The addition of a consensus Meis binding site in a topology found in known Hox target enhancers (**Mann et al., 2009**) confirmed that NvHoxB and NvHoxE do form trimeric complexes with the TALE partners on DNA (**Figure 4C**). Interestingly, the DNA-binding of NvMeis is also sufficient for rescuing the loss of NvHox/NvPbx complex formation upon the HX mutation (**Figure 4C**), suggesting that NvMeis is able to remodel NvHox–NvPbx interactions. Given the sequence divergence between the two NvHox proteins (**Figure 4—figure supplement 1**), these alternative interaction modes are presumably paralog-specific, as previously suggested in bilaterians (**Merabet et al., 2007; Hudry et al., 2012**).

### Genomic binding sites for NvHox and NvTALE proteins are preferentially localized in the promoter region of genes expressed in the endoderm and allow the assembly of Hox/TALE complexes in vitro

Our band shift assays were performed on consensus binding sites previously defined with bilaterian Hox and TALE proteins. To know whether such sites could be used in the context of *Nematostella* development, we searched for their presence in the *Nematostella* genome. We predicted that these DNA-binding sites should be found in the promoter region of genes expressed in the endoderm, where Hox and TALE products are present together. By comparison, the promoter region of genes expressed in the ectoderm should not be enriched in Pbx/Meis-binding sites since the TALE partners are absent in this tissue. Therefore, we performed an in silico analysis based on 76 genes displaying a characterized developmental expression pattern during *Nematostella* embryogenesis. The choice of working with a limited number of genes was motivated by the fact that we did not find any significant enrichment of Hox/PBC/Meis binding sites in a genome-wide search, neither in *Nematostella* nor in other metazoan species (**Figure 5—figure supplement 1**). In addition, by limiting the search in non-coding regions (of at least 60 base pairs) conserved in 12 *Drosophila* genomes, we could find a higher density of Hox/PBC/Meis clusters (**Figure 5—figure supplement 1**). These results highlight that the signal to noise ratio is too low when considering Hox/PBC/Meis binding sites in all non-coding regions, and that the space search needs to be restricted to observe a significant enrichment.

Our in silico analysis revealed a significant enrichment of Hox/Pbx binding sites in the promoter region of genes expressed in the endoderm compared to the ectoderm (**Figure 5A,B**, 'Materials and methods'). Several of these genes also contain a consensus Meis binding sequence within the



**Figure 3.** Hox and TALE members form protein complexes in vivo. **(A)** NvPbx interacts with NvMeis in vivo. The nuclear localisation of a fusion mCherry-NvPbx protein was only observed upon co-injection with NvMeis. Graphs on the left are quantifications of the ratio between nuclear and cytoplasmic fluorescent signals (log<sub>2</sub>). Note that mCherry-NvPbx alone did not lead to any signal, suggesting that the fusion protein is not stable in the absence of NvMeis. **(B)** BIFC assays show that NvPbx interacts with VC<sub>155</sub> and VC<sub>155</sub> NvMeis. **(C)** BIFC assays show that NvPbx interacts with NvHoxE and NvPbx<sup>S4</sup>. *Continued on next page*

Figure 3. Continued

Meis, as noticed in bilaterians. **(B)** BiFC between NvPbx and NvMeis in the *Nematostella* embryo. Fusion constructs are schematized on the left. VN: N-terminal fragment of Venus; VC: C-terminal fragment of Venus. Specificity of BiFC is verified by the absence of fluorescent signals upon the injection of isolated VN or VC fragments, together or with the complementary VC-NvMeis or VN-NvPbx fusion proteins, as indicated. Interaction between NvPbx and NvMeis occurs both in the cytoplasm and nucleus (see text for details). **(C)** BiFC between NvHoxE and NvPbx in the *Nematostella* embryo. Interaction occurs only in the nucleus. Mutation of the residue 54 in the homeodomain (HD) of NvPbx abolishes DNA-binding and BiFC with NvHoxE. In all panels, Dapi (cyan) stains nuclei and Dextran (red) is a control of injection conditions.

DOI: [10.7554/eLife.01939.009](https://doi.org/10.7554/eLife.01939.009)

40 nucleotides surrounding the Hox/Pbx binding site (**Figure 5C**, 'Materials and methods'). Among the candidate target genes containing a putative Hox/Pbx/Meis binding site in their promoter region, we found *NvHoxB* and *NvHoxC*. Thus, auto/cross-regulatory loops may occur between *NvHox* genes, as observed in bilaterians. Since sequences surrounding Hox/TALE binding sites can also strongly influence the protein complex formation (**Ebner et al., 2005; Hudry et al., 2012**), we verified that the binding sites found in proximity of *NvHoxB* and *NvHoxC* could indeed allow the assembly of Hox/TALE complexes in vitro. We confirmed that *NvHoxB* and *NvHoxE* could form dimeric or trimeric complexes on these putative binding sites. Interestingly, *NvHoxB* and *NvHoxE* displayed distinct DNA-binding preferences on these two sites (**Figure 5D–D'**). Again, the assembly of *Nematostella* Hox/Pbx complexes was HX-dependent on both probes, except in the presence of *NvMeis*, highlighting that alternative interaction modes between *Nematostella* Hox and TALE proteins can occur on various DNA-binding sites (**Figure 5—figure supplement 2**). Altogether these results show that the molecular properties underlying the Hox-TALE system are conserved between Cnidaria and Bilateria.

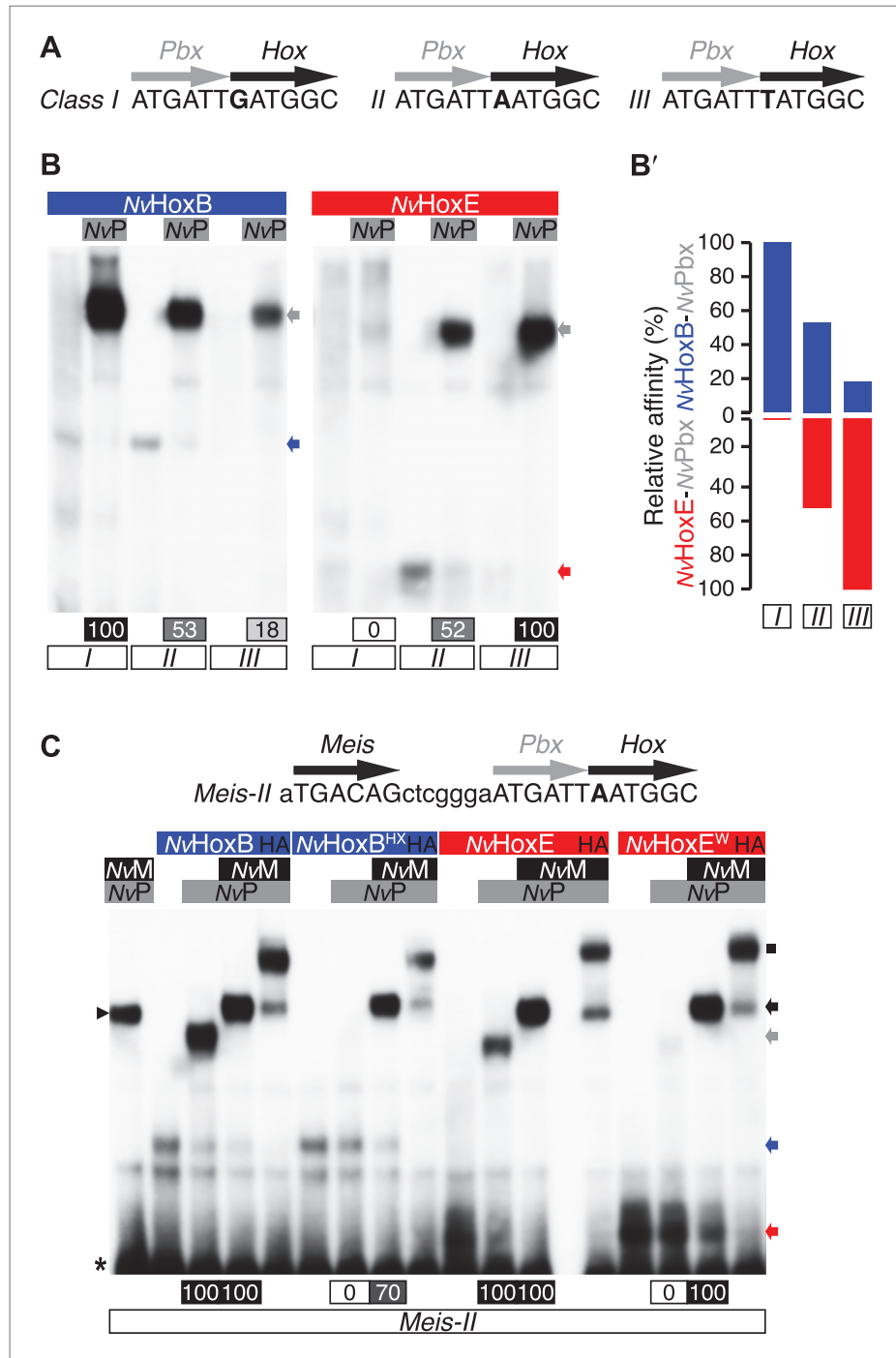
### ***Nematostella* Hox and TALE proteins can execute generic functions in invertebrate and vertebrate species**

To assess whether the *Nematostella* Hox/TALE system could have any conserved biological function, we examined the activity of *NvHox* and *NvTALE* proteins in two different bilaterian organisms, the fly *Drosophila melanogaster* and the frog *Xenopus laevis*.

For Hox assays, series of EMSAs previously confirmed that *NvHoxB* and *NvHoxE* are able to associate with the *Drosophila* Pbx (Extradenticle, Exd) and Meis (Homothorax, Hth) cofactors on different *Drosophila* Hox target enhancers in vitro (**Figure 6—figure supplement 1**). Of note, *NvHoxB* and *NvHoxE* again display anterior or central-like DNA-binding preferences on those physiological target sites, respectively. Furthermore, BiFC validated that *NvHox* proteins could interact with Exd in vivo (**Figure 6—figure supplement 1**). Again, the specificity of BiFC was confirmed with a DNA-binding deficient form of Exd (**Figure 6—figure supplement 1**).

The activity of *NvHox* proteins in *Drosophila* was then measured in two generic Hox assays: the antenna-to-leg transformation in adult (**Casares et al., 1996; Yao et al., 1999**), and the rescue of the Hox *labial* (*lab*) mutant phenotype in a particular structure of the central nervous system called the tritocerebrum (**Hirth et al., 1998**). We found that *NvHoxB* and *NvHoxE* were able to successfully function as their *Drosophila* homologs in both assays (**Figure 6A,B**). The antenna-to-leg transformation by *NvHox* proteins was shown to rely on the assembly of a repressive trimeric complex with *Drosophila* TALE cofactors on cis-regulatory sequences of the *spineless* (*ss*) target gene (**Figure 6A, Figure 6—figure supplement 2; Duncan et al., 2010**). Moreover, *NvHoxB* and *NvHoxE* behave like central and anterior paralogs in the tritocerebrum, respectively (**Figure 6B, Figure 6—figure supplement 2; Hirth et al., 1998**). Finally, the activity of both *NvHox* proteins in the antenna and tritocerebrum appears to be dependent on the integrity of the HX motif (**Figure 6—figure supplement 2**).

The functional conservation of *NvPbx* was addressed by analysing its potential to rescue zygotic *exd* mutant phenotypes in *Drosophila*. As for *NvHox* proteins, we previously verified that *NvPbx* is able to form a protein complex with *Drosophila* Hox and Meis proteins in vitro and in vivo (**Figure 6—figure supplement 3**). We observed that *NvPbx* could rescue the *exd* mutant phenotype in the *Drosophila* larva cuticle (**Figure 6C**). Providing *NvPbx* in this mutant background was also sufficient to rescue the A1 transforming activity of *Ubx* (**Figure 6C'**), which is also known to depend on the integrity of the HX (**Galant et al., 2002**). These results highlight that *DmExd* and *NvPbx* are functionally equivalent, at least for specification functions in the *Drosophila* epidermis.



**Figure 4.** Interaction properties between NvHox and NvTALE proteins in vitro. **(A)** Nucleotide sequence of the different classes of Hox/Pbx binding sites used in band shift experiments. The nucleotide that distinguishes each Hox/Pbx binding site is bolded. **(B–B')** Band shift experiments between NvHoxB or NvHoxE and NvPbx on the three different classes of binding sites, as indicated. Coloured and grey arrows point to monomer or dimer binding, respectively. Graph on the right (**B'**) depicts the relative affinity of each dimeric complex on the three different binding sites, as deduced from the direct quantification on the gel (values are indicated at the bottom). **(C)** Band shift experiments between wild-type or HX-mutated NvHox proteins and NvTALE cofactors, as indicated. Colour codes and annotations are as in **(B)**. Black arrow indicates trimeric NvHox/NvPbx/NvMeis complexes. Other bands are not specific (proteins of the lysate). Black scare highlights the supershift band resulting from the addition of an

Figure 4. Continued on next page



Figure 4. Continued

antibody against the HA tag of NvHox proteins. Asterisk shows the free probe. Note that the loss of dimeric NvHox/NvPbx complex upon the HX mutation is rescued in the presence of NvMeis.

DOI: [10.7554/eLife.01939.010](https://doi.org/10.7554/eLife.01939.010)

The following figure supplements are available for figure 4:

**Figure supplement 1.** Protein sequence alignment between NvHoxE (upper sequence) and NvHoxB (lower sequence).

DOI: [10.7554/eLife.01939.011](https://doi.org/10.7554/eLife.01939.011)

Finally, the functional conservation of NvMeis was measured in the *Xenopus* embryo, which constitutes a well-established developmental model system for assessing Meis activities. In particular, XMeis proteins are known to be required for the specification of posterior cell fates along the AP axis of the central nervous system (Dibner et al., 2001), a function that also involves Pbx1 (Maeda et al., 2001). Accordingly, ectopic expression of XMeis proteins causes anterior neural truncations with a concomitant expansion of hindbrain and spinal cord (Salzberg et al., 1999). This phenotype is reproduced with the fly or mouse Meis proteins, demonstrating that it can constitute a generic assay for assessing Meis function. We observed that the injection of NvMeis in animal caps of *Xenopus* embryos was also able to robustly induce the expression of several posterior spinal cord marker genes (like *Hoxa7*, *cdx1* or *cdx2*), although to a lesser extent than XMeis3 (Figure 6—figure supplement 4). Thus, NvMeis displays a striking functional similarity with its bilaterian homologs.

In sum, our assays highlight a striking functional conservation between *Nematostella* and bilaterian Hox and TALE proteins, suggesting that the Hox/TALE network was already at work in the common ancestor of Cnidaria and Bilateria.

### Genesis of Hox–TALE interaction networks across Metazoa

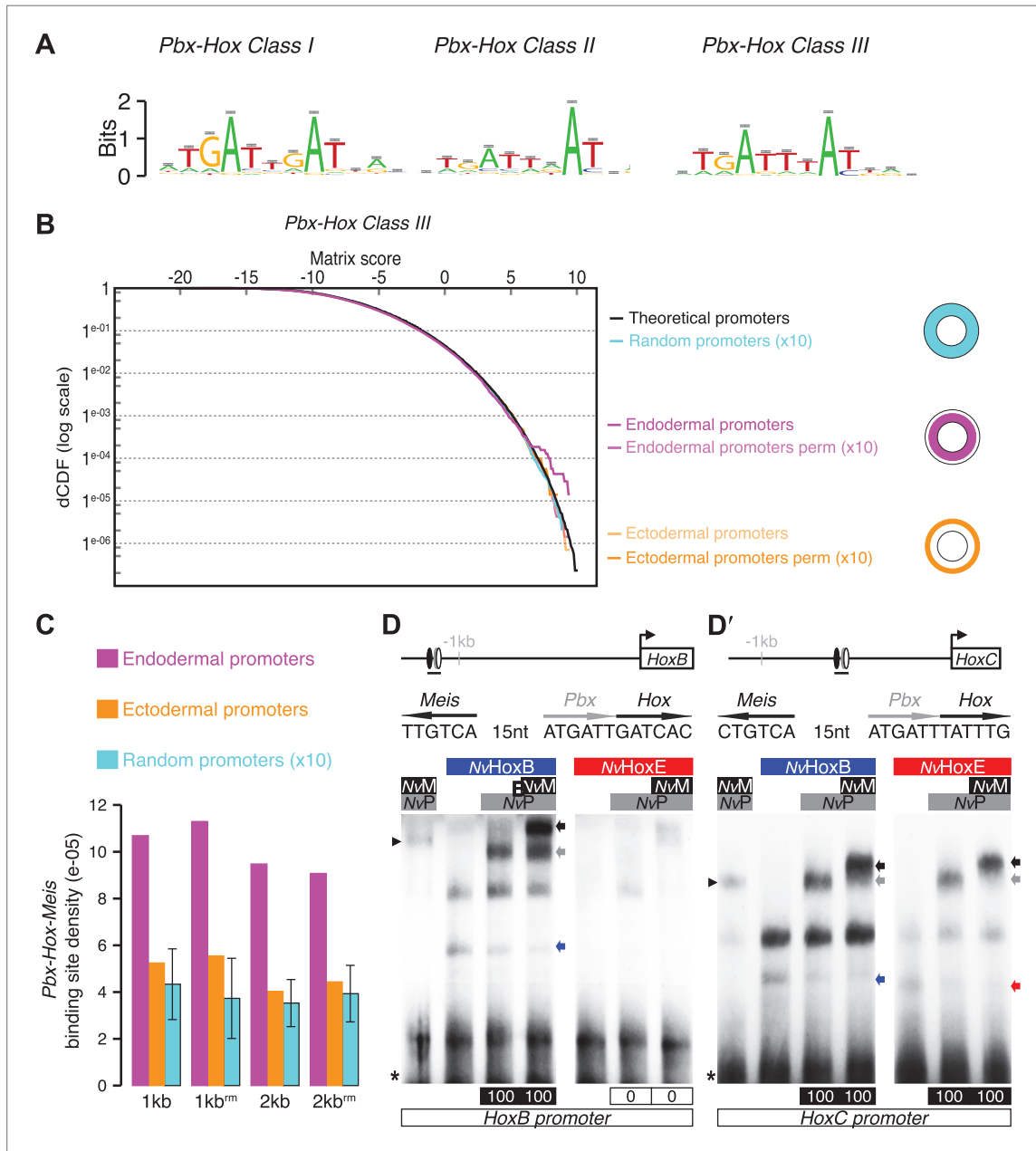
Among the ANTP class, we can distinguish three main subclasses that originally derived from a ProtoANTP ancestor (Saina et al., 2009): Hox/ParaHox, NK and extended-Hox (Figure 7). Consistent with their close evolutionary relationships, Hox, NK and extended-Hox members share common protein features including the presence of an HX motif upstream of the HD (Figure 7). These observations raise the question of the molecular mechanisms that led to the emergence of the Hox-TALE network and more generally to interaction networks between TALE proteins and other members of the ANTP class across metazoan evolution. We postulated that the HX motif could have constituted a major protein scaffold that provided an ancestrally conserved TALE interaction potential to different ANTP members.

To test this hypothesis, we started by analysing the interaction properties between NK and TALE proteins. Previous works showed that some vertebrate NK members could interact with PBC (Brendolan et al., 2005) or PBC/Meis (Rhee et al., 2004), but the role of the HX was not addressed in these interactions. Here, we analysed the molecular properties underlying complex assembly between the *Nematostella* NK representative NvMx and the two NvTALE partners. We observed that NvMx is not able to interact with NvPbx except in the presence of NvMeis (Figure 8A). Trimeric complex formation is however not as strong as with NvHox proteins and is also fully dependent on the integrity of the HX motif (Figure 8A).

Our results suggest that protein region(s) in NvMx could mask the interaction with Pbx in absence of Meis. We confirmed this hypothesis by testing a series of truncated and chimeric proteins generated from NvMx and NvHoxB. We found that deleting the N- and C-terminal parts of NvMx allowed dimeric complex formation with NvPbx (Figure 8B,C). Conversely, the N- and C-terminal regions of NvMx are sufficient to alleviate the interaction between a minimal NvHoxB protein and NvPbx (Figure 8B,C).

Together these results show that the NK-TALE and Hox-TALE interaction networks rely on different molecular properties, in particular with a role of Meis in promoting HX-dependent or HX-independent interaction modes, respectively (Figure 8D).

We next analysed the interaction properties between TALE proteins and the *Drosophila* Engrailed (En) protein, a member of the extended-Hox family, which was recently described to form cooperative DNA-binding complexes with Exd and Hth on physiological target sequences (Fujioka et al., 2012). We observed that En could form dimeric or trimeric complexes with Exd or Exd/Hth respectively, but in all cases these complexes were lost upon the mutation of the HX (Figure 8—figure supplement 1).



**Figure 5.** Genes expressed in the endoderm are enriched in Hox-TALE binding sites in their promoter region. **(A)** Hox/Pbx binding motifs represented as logos. The three motifs represent the binding specificity of the Hox/Pbx complex for sites of class I, II, or III. Matrix was determined by Selex with the *Drosophila* proteins (Slattery et al., 2011). **(B)** Score distributions of the Hox/Pbx Class III matrix. The Y-axis is shown in logarithmic scale to highlight the relevant range of p values (small values). The separation of the pink curve (endoderm) from the black one (theoretical distribution) indicates an enrichment of the Hox/Pbx putative binding sites in the promoter of genes expressed in the endoderm. On the contrary, there is no enrichment in the promoter of genes expressed in the ectoderm, as the orange curve follows the black one. All negative controls also show no enrichment: random sets of gene promoters (cyan), promoter regions randomized by matrix column permutations for the endoderm (light pink) and ectoderm (light orange). **(C)** In silico analysis of Hox/Pbx/Meis binding sites in the promoter region (1 kb or 2 kbs upstream of the transcription start site) of genes expressed in the endoderm (pink), ectoderm (orange), or randomly chosen (cyan). The graph illustrates the preferential enrichment of Hox/Pbx/Meis binding sites in the promoter region of endodermal genes. Rm: repeat masked. **(D–D')** Band shift experiments between NvHox and NvTALE proteins on binding sites found in the promoter region of *NvHoxB* and *NvHoxC* genes. Sequence and genomic position of each binding site are shown above the gel. Colour code and annotations are as in Figure 3. Note the distinct DNA-binding preferences of NvHoxB and NvHoxE on these two different target sites. See also Figure 5—figure supplements 1 and 2.

DOI: 10.7554/eLife.01939.012

Figure 5. Continued on next page

Figure 5. Continued

The following figure supplements are available for figure 5:

**Figure supplement 1.** Genome-wide analysis cannot reveal significant enrichment of Hox/PBC/Meis binding sites.

DOI: [10.7554/eLife.01939.013](https://doi.org/10.7554/eLife.01939.013)

**Figure supplement 2.** NvMeis promotes HX-independent interaction modes on DNA-binding sites found in the promoter region of *NvHoxB* (A) and *NvHoxC* (B).

DOI: [10.7554/eLife.01939.014](https://doi.org/10.7554/eLife.01939.014)

In conclusion, although NvMsx and DmEn do not exhibit identical cooperative DNA-binding properties with PBC and Meis, they both require the HX to interact with the TALE partners (**Figure 8D**). Thus, the acquisition of the HX motif during evolution was likely a key molecular event for the emergence of ANTP-TALE interaction networks. Along the same line, we observed that the Trox2/Gsx protein from *Trichoplax adhaerens*, which does not contain any obvious HX-like sequence (**Schierwater and Kuhn, 1998**), is not able to form any dimeric or trimeric complex with PBC or PBC/Meis, respectively (**Figure 9A**).

To further confirm that ANTP-TALE networks are a metazoan innovation, we analysed the interaction properties of PBC and Meis proteins of *Acanthamoeba castellanii* (Ac), a unicellular organism from the Amoebozoa group (**Figure 1**). Interestingly, AcMeis possesses a MEIS-A domain and displays a high level of sequence similarity with mouse or fly Meis proteins in the HD (**Figure 2—figure supplement 2**). In contrast, AcPbx lacks any PBC-A domain and has a strongly divergent HD when compared to other Pbx proteins (**Figure 2—figure supplement 1**). As expected, we observed that AcPbx could neither bind on a consensus PBC-binding site, nor stimulates the binding of AcMeis on the same probe, suggesting that both proteins could not interact in vitro (**Figure 9B**). This result was confirmed by using central or posterior Hox/PBC/Meis binding sites, on which no dimeric or trimeric complex with Hox proteins could be formed (**Figure 9C,D**). To assess whether absence of protein complexes could be explained by the strong sequence divergence of AcPbx, we repeated experiments with the mouse Pbx1 protein. Under these heterologous partnership conditions, we observed that the binding of AcMeis could be strongly enhanced in the presence of the Pbx partner, suggesting that the two proteins could make interactions (**Figure 9B**). However, neither dimeric nor trimeric complexes could be formed on central and posterior Hox/PBC/Meis nucleotide probes, demonstrating that AcMeis lacks protein feature(s) for making cooperative DNA-binding complexes with Hox and PBC proteins (**Figure 9C–D'**). Not surprisingly, AcMeis is also not able to rescue Hox/PBC complex formation upon the HX mutation (**Figure 9—figure supplement 1**).

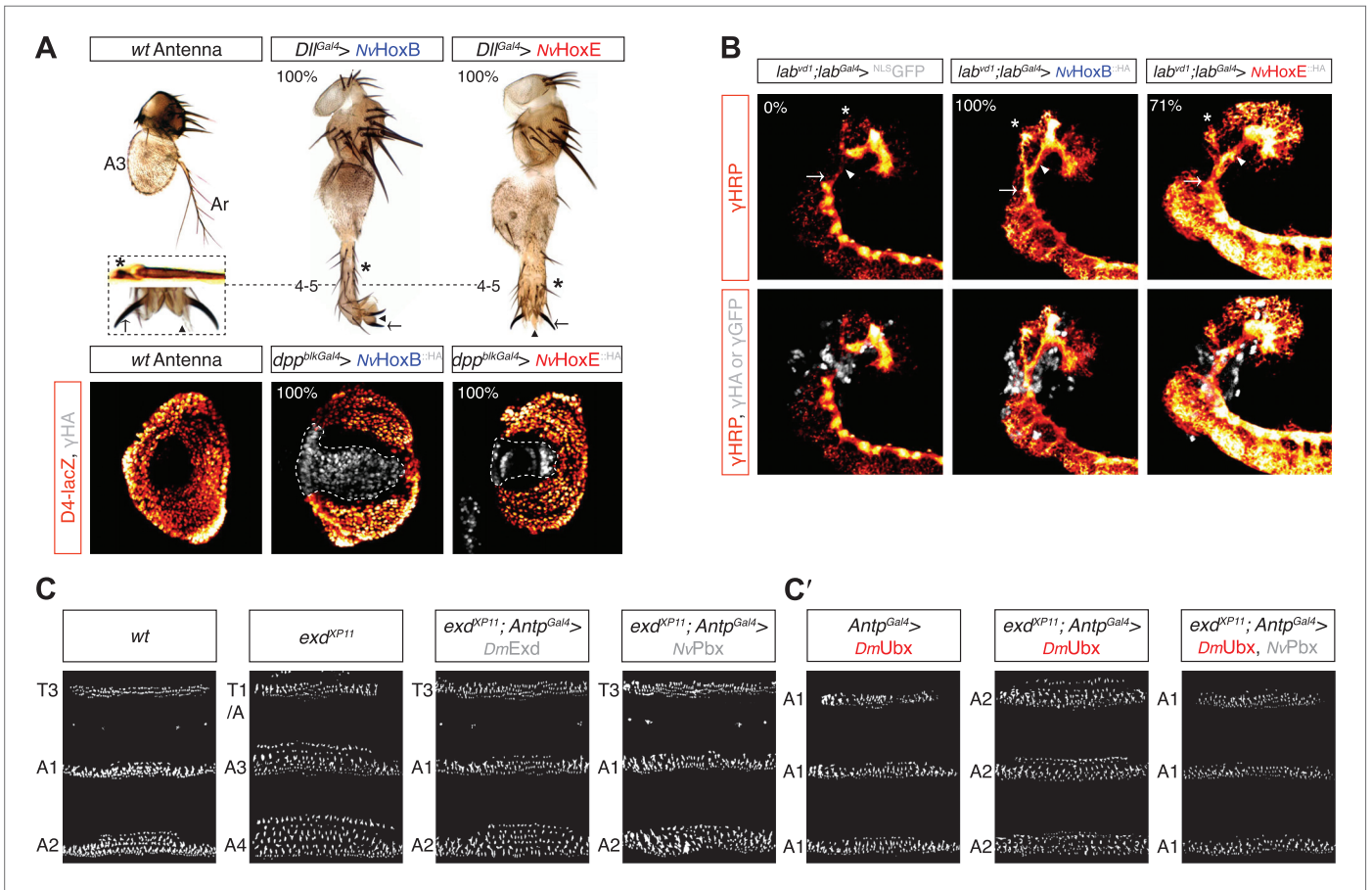
Taken together, our results show that the use of other protein motifs than the HX for interaction with TALE partners is a peculiar property of Hox proteins among the ANTP class. They also emphasize that the PBC/Meis partnership likely evolved concomitantly with the apparition of the HX in the ANTP class. In this context, our work with TALE proteins of *Acanthamoeba* underlines that the evolution of the TALE partners enabled the interaction network with Hox proteins and hence new functions to emerge during eukaryote evolution.

## Discussion

### Molecular evolution of the HX motif

We have shown that identical molecular rules and conserved functions characterize Hox-TALE interaction networks in Cnidaria and Bilateria. The presence of the HX motif in several NvHox members and its requirement for generic Hox/TALE functions strongly suggest that this motif had a pivotal role for the evolution of an active Hox-TALE system in early metazoan lineages. In addition, the observation that a number of cnidarian Hox proteins do not have any HX (**Figure 10, Figure 10—figure supplement 1**) highlights that the ancestral molecular properties and hence functions of the Hox-TALE network could have considerably diverged among different cnidarian lineages.

To date, no HX motif can be found in any member of the ANTP class in Porifera, Ctenophora, or Placozoa group. This motif is present in different ANTP subclasses, including NK, Hox/ParaHox and extended-Hox, specifically in Bilateria and Cnidaria. Different scenarios can be proposed for explaining the evolutionary history of the HX among different ANTP members. These scenarios are hypothetical and diverge according to the putative evolutionary history of the ANTP class homeobox



**Figure 6.** Functional analysis of NvHox and NvPbx proteins in *Drosophila*. **(A)** Antenna-to-leg transforming activities of NvHoxB and NvHoxE. NvHox proteins were expressed in the antenna with the *Distalless* (*Dll*)-*Gal4* driver. Asterisk depicts leg-specific bracted bristles. 4–5 shows the transformation of the arista in two tarsal segments. Arrow and arrowhead in the enlargement indicate the formation of the leg-specific terminal claw and its associated sensory pad respectively. The antenna-to-leg transformation by NvHox proteins (grey) is achieved through the repression of the *spineless* (*ss*) target gene, as observed by the repression of the *ss* enhancer *D4* activity on *lacZ* reporter gene expression (orange). See also **Figure 6—figure supplements 1 and 2**. **(B)** Rescue of the *labial* (*lab*) mutant phenotype in the tritocerebrum by NvHox proteins. The central nervous system is stained with an anti-HRP (orange). Hox or GFP (as a control) proteins (grey) are expressed in the tritocerebrum with a *lab*-*Gal4* driver. Frontal connectives (asterisk), longitudinal connectives (arrowhead) and tritocerebral commissure (arrow) are indicated. In *lab* mutant background, longitudinal connectives are reduced, frontal connectives project ectopically and the tritocerebral commissure is missing (*Hirth et al., 1998*). Expression of NvHoxB or NvHoxE in this mutant context leads to a complete or strong rescue of this phenotype, respectively. See also **Figure 6—figure supplement 2**. **(C–C')** NvPbx can rescue zygotic *exd* mutant phenotypes in the *Drosophila* larva cuticle. **(C)** Larvae homozygous for the zygotic *exd<sup>KP11</sup>* mutation have T3 and A1 segments that resemble to a T1/abdominal or A3 segment, respectively. Thoracic expression of either *DmExd* or *NvPbx* in this mutant background (through the UAS/*Gal4* system, with the *Antennapedia* (*Antp*)-*Gal4* driver) is sufficient to restore the correct specification of T3 and A1, as assessed by the shape and arrangement of denticle belts. **(C')** Ubx normally specifies the A1 segment. Ectopic expression of Ubx with *Antp*-*Gal4* induces A1-like segments in the thorax. In absence of Exd, Ubx produces A2-like segments. Providing back *NvPbx* in this genetic background is sufficient to restore the normal A1-inducing activity of Ubx. See also **Figure 6—figure supplements 3 and 4**.

DOI: [10.7554/eLife.01939.015](https://doi.org/10.7554/eLife.01939.015)

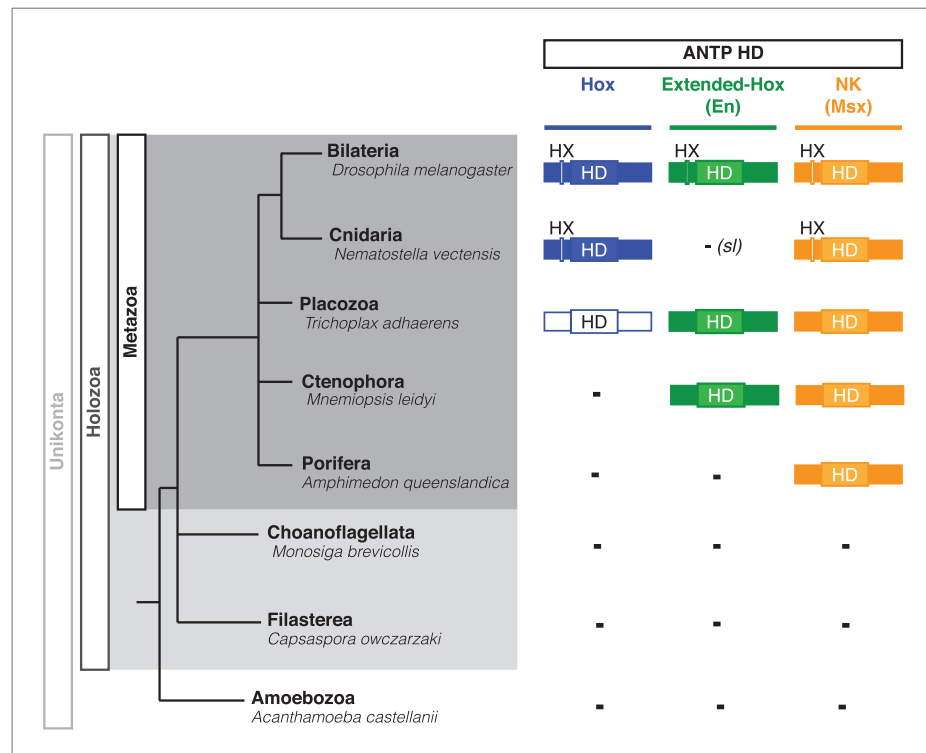
The following figure supplements are available for figure 6:

**Figure supplement 1.** NvHoxB and NvHoxE interact with the *Drosophila* TALE cofactors Extradenticle (Exd) and Homothorax (Hth) in vitro and in vivo. DOI: [10.7554/eLife.01939.016](https://doi.org/10.7554/eLife.01939.016)

**Figure supplement 2.** Role of the HX of NvHox proteins in generic *Drosophila* Hox assays. DOI: [10.7554/eLife.01939.017](https://doi.org/10.7554/eLife.01939.017)

**Figure supplement 3.** NvPbx interacts with the *Drosophila* Ultrabithorax (Ubx) and Homothorax (Hth) proteins in vitro and in vivo. DOI: [10.7554/eLife.01939.018](https://doi.org/10.7554/eLife.01939.018)

**Figure supplement 4.** NvMeis reproduces generic bilaterian Meis activities in the *Xenopus* embryo. DOI: [10.7554/eLife.01939.019](https://doi.org/10.7554/eLife.01939.019)



**Figure 7.** Phylogeny of Hox, Extended-Hox (Engrailed, En) and NK (Msx) proteins during eukaryote evolution. Nomenclature is as in **Figure 1**. The conserved features of any HX motif correspond to a sequence containing an invariant tryptophan residue in a hydrophobic context, with a lysine or arginine residue at position +2 to +5, as previously defined (*In der Rieden et al., 2004*).

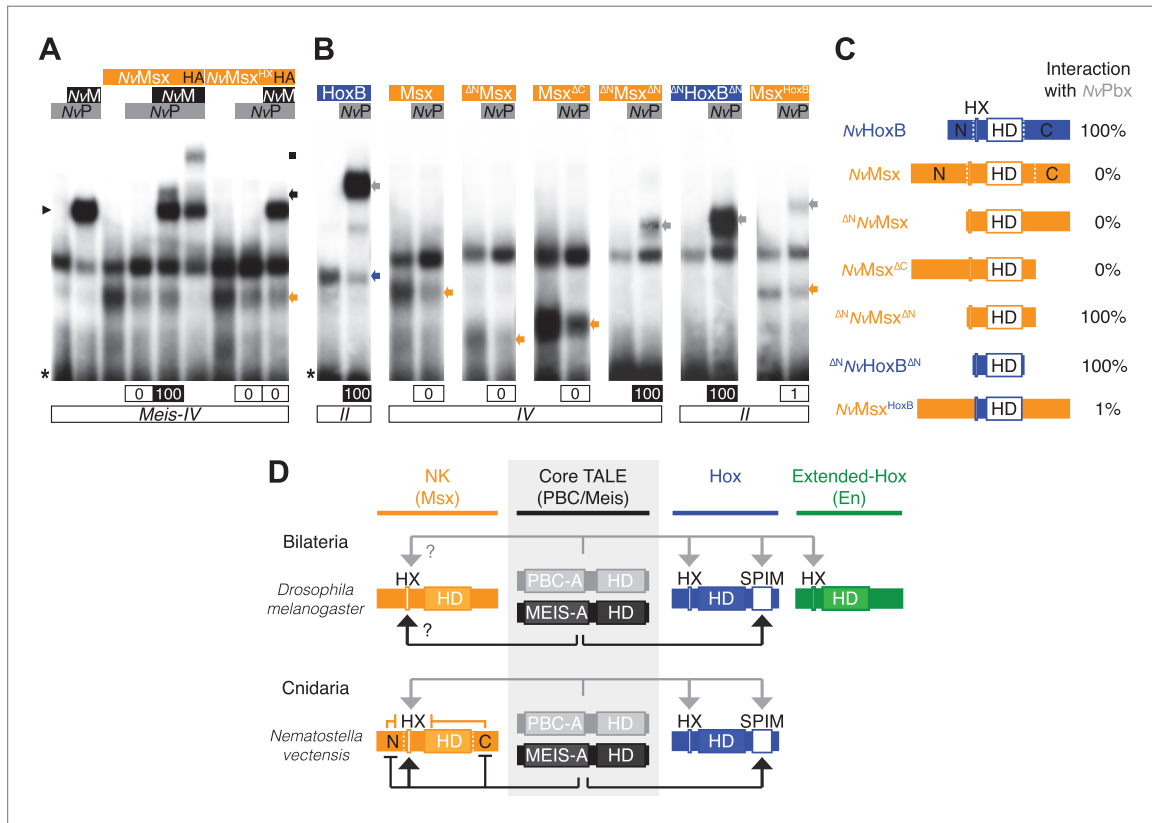
DOI: [10.7554/eLife.01939.020](https://doi.org/10.7554/eLife.01939.020)

genes. In one scenario, HX-containing NK, Hox and extended-Hox members could have emerged from a common HX-containing NK protein, which itself appeared from duplications of an ancestral HX-deficient NK cluster, at the basis of Eumetazoa (Bilateria+Cnidaria, **Figure 11A**). Alternatively, it was recently proposed that Hox/ParaHox, NK and extended-Hox evolved before the origin of poriferans (*Mendivil Ramos et al., 2012*), being already present in the last common ancestor of animals (Urmetazoa). In this scenario, we do not favour the hypothesis of independent acquisitions of the HX in the three homeobox gene families. Indeed, although the HX is a short motif, it is always located upstream and at a reasonable distance of the HD. This invariant position is probably critical to ensure interactions with PBC proteins. Moreover, the distance between the HX and the HD could have played important roles for the acquisition of new functions, as noticed in certain bilaterian Hox proteins (*Prince et al., 2008; Saadaoui et al., 2011*). Thus, the HX could already have been present in the ProtoANTP protein of the Urmetazoa ancestor, being secondarily lost in Porifera, Ctenophora and Placozoa during evolution (**Figure 11B**).

### Hox/TALE functions in the sea anemone *Nematostella vectensis*

Our results showed that *Nematostella* Hox and TALE proteins shared conserved functions with their bilaterian counterparts. Interestingly, *NvHoxB* and *NvHoxE* display anterior-like or central-like properties, respectively. These preferential activities were observed in vitro, at the level of DNA-binding site recognition, but also in the generic rescue assay of the *Drosophila Hox*-mutant tritocerebrum structure. Thus, although *NvHoxB* and *NvHoxE* are not organized in a genomic cluster, they display differential expression profiles and activities, suggesting that ancestral colinearity rules are at least kept for two asymmetrically expressed Hox genes in *Nematostella*.

Considering our in silico data, and given the expression profile of Hox and TALE members in the *Nematostella* embryo, we propose that Hox-TALE interaction networks could be used for regulating gene expression in the endoderm, likely for positioning and specifying the formation of the different



**Figure 8.** Genesis of Hox–TALE interaction networks in Metazoa. **(A)** Band shift experiments between NvMsx and NvTALE cofactors, as indicated. Colour code and annotations are as in previous figures. Note that no dimeric NvMsx/NvPbx complex is formed. Binding reactions with NvMsx proteins were performed on a consensus *Msx/Pbx* binding site derived from vertebrates and containing an additional *Meis* binding site in a topology similar to the Hox probe ('Materials and methods'). **(B)** Band shift experiments between wild-type, truncated or chimeric NvMsx and/or NvHoxB proteins, and NvPbx, as indicated. **(C)** Scheme of the diverse protein constructs and their corresponding interaction affinity level with NvPbx, as assessed from quantification of each band shift. Quantifications with truncated NvMsx proteins were deduced by comparison with the trimeric NvMsx/NvPbx/NvMeis in **(A)**. **(D)** Molecular rules underlying interaction properties between NK, Hox, or extended-Hox members and TALE cofactors in Eumetazoans. In this model, Meis is able to promote HX-dependent interactions between NvMsx and NvPbx by masking inhibitory interaction domains in NvMsx. Whether a similar role could exist in Bilateria remains to be determined. We noticed that the *Drosophila* Msx protein contains an HX and forms trimeric but not dimeric complexes with the *Drosophila* TALE partners (not shown). HX-dependency in those interactions remains to be determined (question marks). Interaction between *DmEn* and PBC is also HX-dependent but does not require the presence of Meis to occur. In contrast to the NK or extended-Hox families, most members of the Hox family have retained a HX motif. This motif is required for generic Hox/Pbx functions. The additional presence of Meis allows revealing specific Pbx interaction motifs (SPIMs), which could be important for distinguishing and/or diversifying the embryonic activities of each Hox paralog group member. See also **Figure 8—figure supplement 1**.

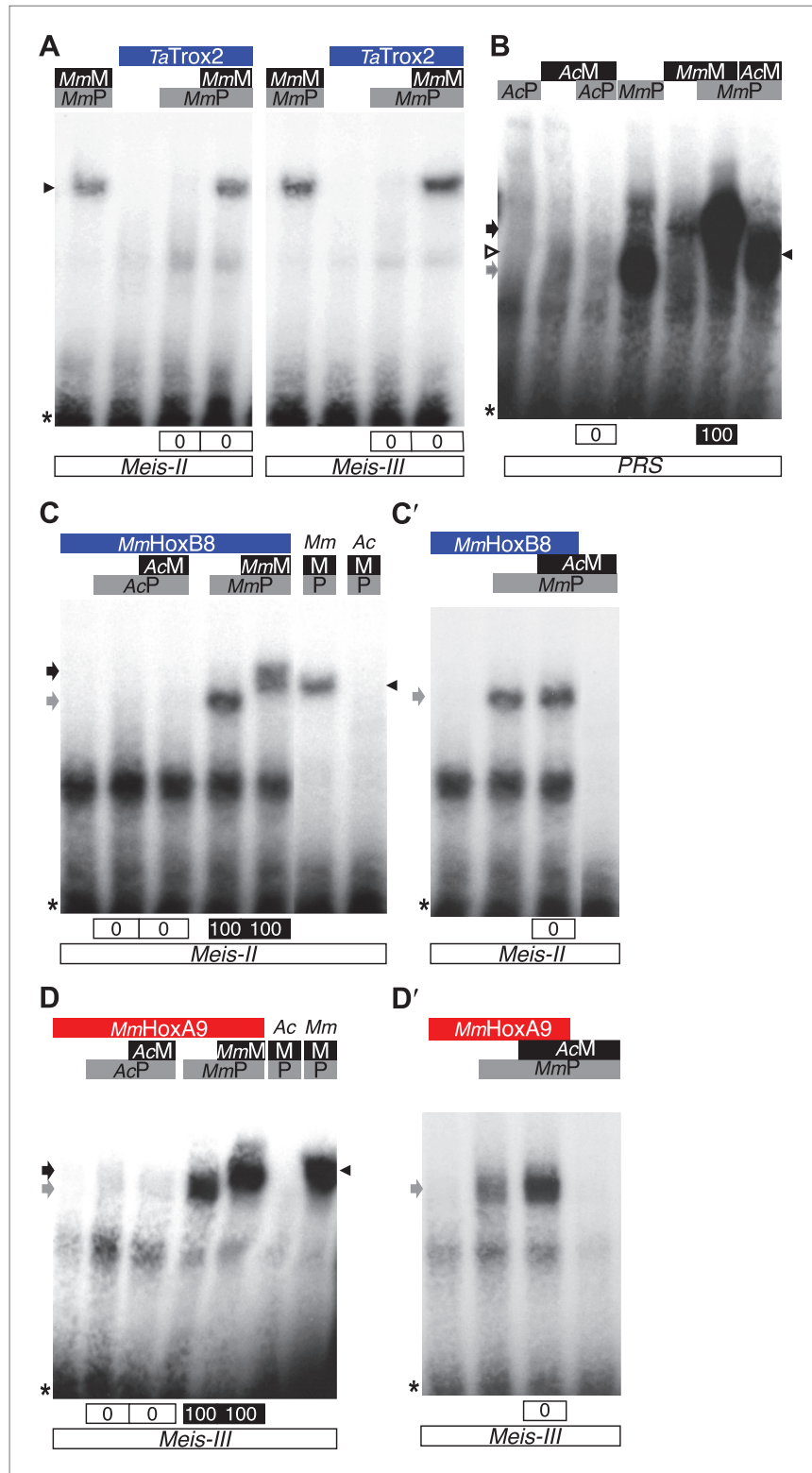
DOI: 10.7554/eLife.01939.021

The following figure supplements are available for figure 8:

**Figure supplement 1.** The *Drosophila* Engrailed (*DmEn*) protein forms HX-dependent DNA-binding complexes with Exd (*DmE*) and Hth (H) on physiological target sites.

DOI: 10.7554/eLife.01939.022

mesenteries along the directive axis. Homology between cnidarian and bilaterian axes is a long-standing and still controversial question. It was first proposed that the oral–aboral axis of *Nematostella* could be orthologous to the bilaterian AP axis, with an anterior-like (*NvHoxA/anthox6*) and central/posterior (*NvHoxF/Anthox1*) Hox gene being expressed at the oral or aboral tip respectively (Finnerty et al., 2004). This expression profile is however not conserved in other cnidarian species (Kamm et al., 2006) and there is compelling evidence that the aboral pole could rather correspond to the anterior end of bilaterians (Sinigaglia et al., 2013). Therefore, expression of *NvHox* genes along the primary body axis could correspond to individual morphogenetic and not positional patterning functions, as recently shown for *NvHoxF* (Sinigaglia et al., 2013). In this context, other HD-containing determinants could



**Figure 9.** Interaction properties of Hox and TALE proteins from *Trichoplax adhaerens* and *Acanthamoeba castellanii*. (A) The ProtoHox/ParaHox Trox2 protein from *Trichoplax adhaerens* does not form DNA-binding complexes with PBC or PBC/Meis in vitro. Band shift experiments are performed with mouse Pbx (MmP) and Meis (MmM) proteins on central (*Meis-II*) and posterior (*Meis-III*) Hox/PBC/Meis binding sites as indicated. Black Figure 9. Continued on next page

Figure 9. Continued

arrowhead shows dimeric Pbx/Meis complexes. (B–D') PBC (AcP) and Meis (AcM) proteins from the unicellular *Acanthamoeba castellanii* organism cannot form protein complexes between each other or with Hox proteins. (B) Band shift experiment on a consensus PBC binding site (PRS, Chang et al., 1995). AcP does not bind DNA, neither as a monomer nor with AcM. A weak monomer DNA-binding of AcM is observed (white arrowhead). This monomer binding is strongly enhanced in the presence of mouse Pbx1 (MmP, black arrowhead). In comparison, MmP binds strongly (grey arrow), and the monomer binding of Meis1 (MmM) is also strongly enhanced in the presence of Pbx1 (black arrow). (C–C') Band shift experiments with mouse HoxB8 (MmHoxB8) and mouse or *Acanthamoeba* TALE cofactors on the central (*Meis-I*) Hox consensus binding probe as indicated. (D–D') Band shift experiments with mouse HoxA9 (MmHoxA9) and mouse or *Acanthamoeba* TALE cofactors on the posterior (*Meis-III*) Hox consensus binding probe as indicated. Complexes with Hox proteins are observed only with mouse Pbx (grey arrows) and Pbx/Meis partners (black arrows) on both probes. AcPbx and AcMeis proteins are not able to form dimeric complexes on these probes, unlike mouse TALE proteins (black arrowheads). See also **Figure 9—figure supplement 1**.

DOI: [10.7554/eLife.01939.023](https://doi.org/10.7554/eLife.01939.023)

The following figure supplements are available for figure 9:

**Figure supplement 1.** PBC and Meis proteins from the unicellular *Acanthamoeba castellanii* organism do not form protein complexes with mouse HX-mutated Hox proteins.

DOI: [10.7554/eLife.01939.024](https://doi.org/10.7554/eLife.01939.024)

play important patterning roles (de Jong et al., 2006), as demonstrated for Six transcription factors in the early specification of the aboral pole (Sinigaglia et al., 2013).

Along the same line, the directive axis of *Nematostella* was proposed to be homologous to the bilaterian DV axis (Matus et al., 2006), but functional analyses revealed that DV patterning genes are also required in the endoderm and ectoderm along the primary axis (Matus et al., 2006; Saina et al., 2009).

Together, these observations highlight that patterning molecules can be used along different longitudinal axes during animal evolution, which renders difficult the comparison between Bilateria and Cnidaria. Here, we propose that the ancestral molecular cues underlying the Hox patterning system along the cnidarian directive axis could have been recruited for AP patterning in Bilateria.

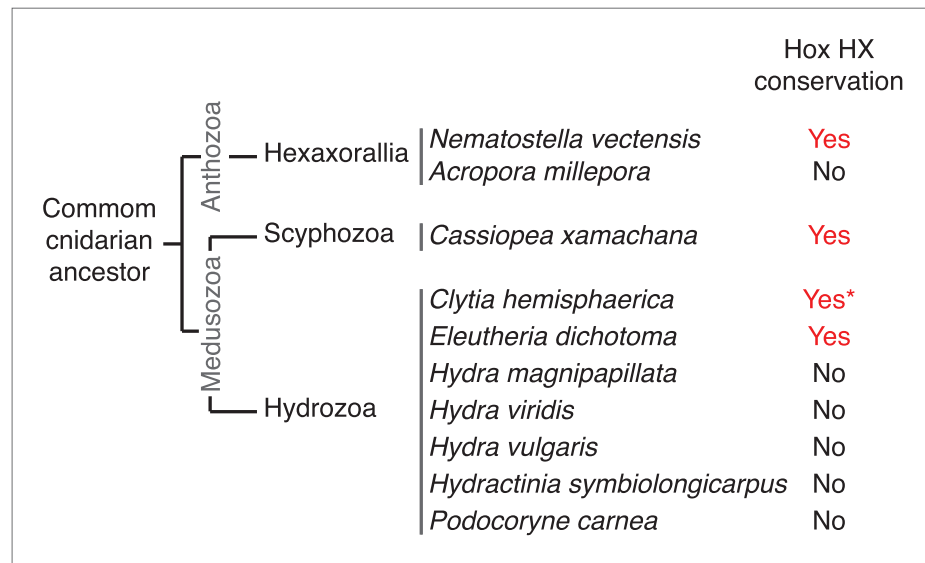
### TALE proteins and the evolution of interaction networks in eukaryotes

Non-TALE and TALE representatives of the HD superfamily were probably already present in first eukaryotes (Derelle et al., 2007; Larroux et al., 2008; Ryan et al., 2010), and it was proposed that interactions between these two classes of TFs could have existed in the common ancestor of plants, fungi, and metazoans (Burglin, 1998). TALE proteins originate from a putative ancestor that contained a MEINOX domain and gave rise to conserved N-terminal interaction domains in different TALE members (Burglin, 1997). Interestingly, although PBC and Meis are not present outside Unikonta, other TALE members are known to interact with each other or with other protein families in plants (Bellaoui et al., 2001; Hackbusch et al., 2005; Kanrar et al., 2006; Hay and Tsiantis, 2010) and fungi (Keleher et al., 1989; Stark and Johnson, 1994; Li et al., 1998; Carr et al., 2004), suggesting that partnership with TALE proteins is a common and ancient feature in eukaryotes (Figure 11C; Burglin, 1998).

Here, we revealed the existence of interactions between typical HD (ANTP) and TALE (PBC/Meis) members in Cnidaria, suggesting that this network, which is also present in Bilateria, was already at work in the Eumetazoa ancestor, before the Cnidaria/Bilateria split. We showed that more ancient TALE proteins, like those from the unicellular *Acanthamoeba* organism could neither interact between each other nor form complexes with Hox proteins. Heterologous interaction assays between AcMeis and mouse Hox and PBC proteins further exemplified that the PBC/Meis partnership is critical for the formation of Hox-TALE networks. This partnership probably appeared with the PBC-A domain in PBC, and concomitantly with the HX motif during eukaryote evolution. Although the MEIS-A domain was more ancient, the Meis partner also clearly acquired additional protein features, allowing the formation and therefore diversifying the activity of Hox/PBC/Meis networks in Metazoa.

Interestingly, the HX is not only conserved in the Hox family but also in different NK and extended-Hox members. Still, the role and the importance of TALE partners in these additional networks might not be equivalent. We propose that apparition of Hox, NK and extended-Hox members was accompanied by functional sub-specialisations that could in part result from divergent molecular interaction properties with TALE cofactors. For example, only two NK sub-family members (Msx and Tlx) have





**Figure 10.** Phylogeny of the HX in Hox proteins of main cnidarian lineages.

DOI: [10.7554/eLife.01939.025](https://doi.org/10.7554/eLife.01939.025)

The following figure supplements are available for figure 10:

**Figure supplement 1.** Protein sequence alignment of the region encompassing the HX (highlighted in red) and HD (highlighted in yellow) of cnidarian Hox members used in **Figure 10—figure supplement 1**.

DOI: [10.7554/eLife.01939.026](https://doi.org/10.7554/eLife.01939.026)

retained a HX motif, suggesting that interaction between NK and TALE proteins is not a general rule. The same rationale applies to extended-Hox members. On the contrary, the HX is present in almost all Hox paralog groups, which coincides with a general requirement of TALE cofactors in Hox functions. In this context, the role of Meis for revealing additional and more specific Pbx interaction motifs (SPIMs) in Hox proteins (Hudry *et al.*, 2012) was also likely an ancestral feature of the Hox/TALE system for distinguishing functions between different Hox paralog members.

In conclusion, we propose that Hox-TALE networks constituted an ancestral regulatory module that was later on exploited for patterning functions in Bilateria. This network was effective as soon as different interaction modes could exist with duplicated Hox family members, allowing diversifying patterning functions along the body axis. This original molecular system was subsequently co-opted by the various contexts of embryogenesis in different eumetazoan phyla, for axis or tissue (see e.g., Di-Poi *et al.*, 2007) patterning, illustrating its remarkable adaptability throughout animal evolution.

## Materials and methods

### Cloning

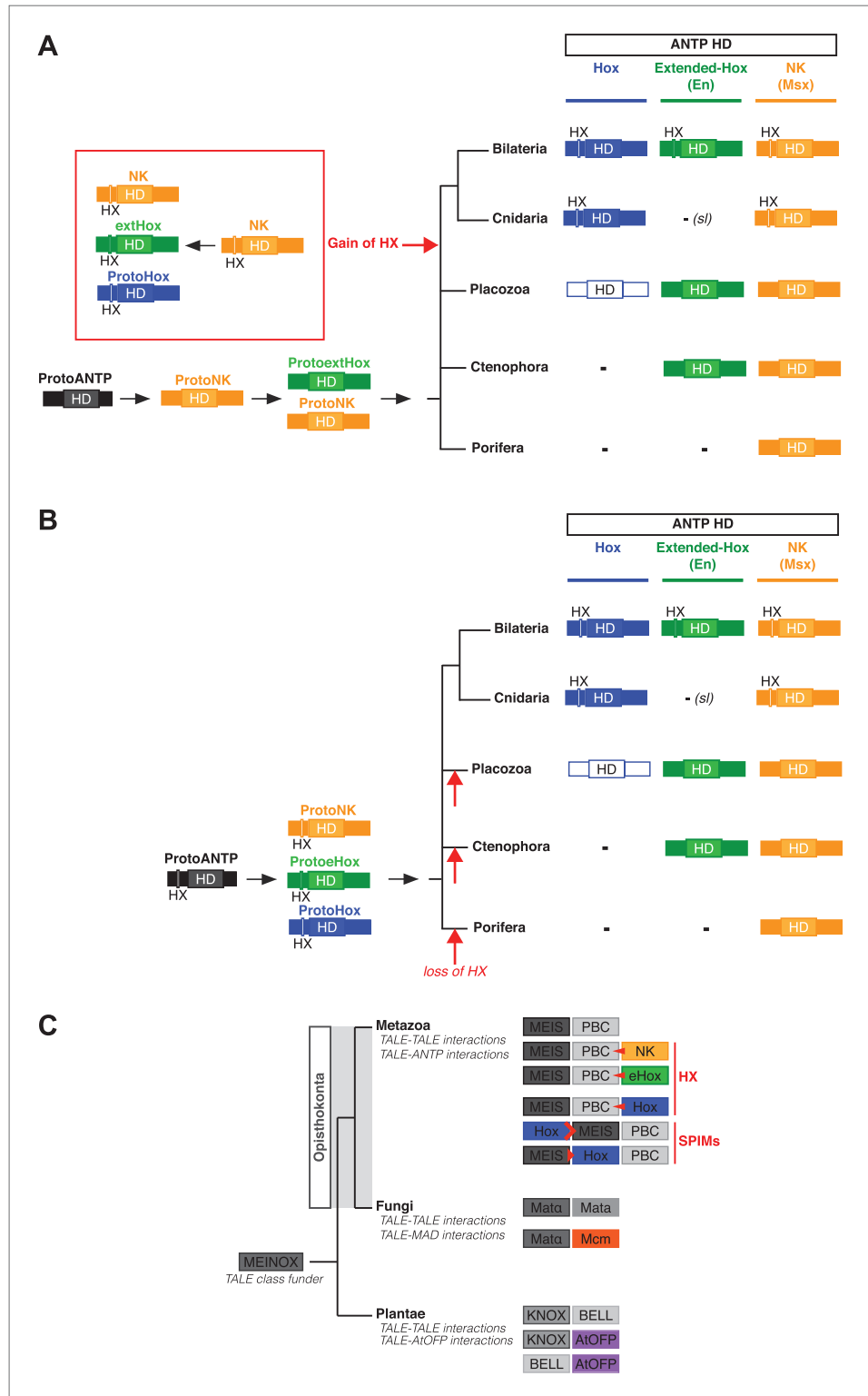
Clones were generated by PCR from full-length complementary DNAs and restriction-cloned in the appropriate vector (see also **Supplementary file 1** for a complete list of all the constructs). Primers used are listed in **Supplementary file 1**. All constructs were sequence-verified before using.

### Fly stocks and transgenic lines

Transgenic lines were established either by the PhiC-31 integrase system (with the pUASTattB vector [Venken *et al.*, 2006; Bischof *et al.*, 2007]) or by classical P-element (with P-AST vector) mediated germ line transformation. Unless otherwise indicated, fly stocks were obtained from the Bloomington *Drosophila* Stock Center. Gal4 drivers used are: *en*-Gal4, *Dll*-Gal4, *Ubx*-Gal4<sup>M1</sup> (kindly provided by Ernesto Sanchez-Herrero), *Antp*-Gal4 (Michel Crozatier), and *dpp*<sup>blink</sup>-Gal4. *lab*<sup>VD1</sup>; *lab*-Gal4 line was provided by Frank Hirth, and *D4*-LacZ line by Ian Duncan.

### Immunostainings, cuticle preparation and in situ hybridization

Immunodetections in *Drosophila* embryos, imaginal discs, and cuticle preparations were performed according to standard procedures. *Nematostella* in situ hybridizations were performed as described



**Figure 11.** Evolution of molecular signatures underlying interaction networks with TALE proteins. **(A)** Model for the apparition of the HX in ANTP members of Cnidaria and Bilateria. In this scenario, the HX appeared in a duplicated NK protein that gave rise to other HX-containing ANTP members in the common ancestor of Bilateria and Cnidaria. Duplications in other lineages occurred without the apparition of the HX. **(B)** Second model for the apparition of Figure 11. Continued on next page

Figure 11. Continued

the HX in ANTP members of Cnidaria and Bilateria. This model is based on a recent work postulating the existence of Hox/ParaHox, NK, and extended-Hox clusters in the common ancestor of animals (**Mendivil Ramos et al., 2012**). In this scenario, it is unlikely that the HX appeared after duplications of the ProtoANTP ancestor independently and at the same place in the ProtoHox, ProtoNK and ProtoExtended(e)Hox families. Thus, the HX motif was probably already present in the ProtoANTP ancestor. This motif was secondarily lost in ANTP members of Porifera, Placozoa and Ctenophora during evolution. See also 'Discussion'. (C) TALE proteins and the generation of interaction networks across major multicellular branches of the eukaryote evolutionary tree. TALE proteins are indicated in grey-filled boxes, with different grey tones corresponding to different TALE members. Other colours depict non-TALE members. Interactions can occur between different TALE members, or between TALE and non-TALE members. These interactions involve different proteins in each major multicellular branch, as indicated. Red signs in Metazoa symbolise interaction modes involving the HX or specific PBC interaction motifs (SPIMs) between TALE (PBC/Meis) and ANTP (NK, extended(e)Hox, Hox) members.

DOI: [10.7554/eLife.01939.027](https://doi.org/10.7554/eLife.01939.027)

(**Genikhovich and Technau, 2009**). The antibodies used were: rat anti-HA (1/500; Molecular probe, Invitrogen, CA, USA), mouse anti- $\beta$ -galactosidase (1/500; Molecular Probe), chicken anti-GFP (1/500; Promega, WI, USA), guinea pig anti-Homothorax (Natalia Azpiazu) and rabbit anti-HRP (1/100; FITC-conjugated, Jackson Immunoresearch, PA, US).

### BiFC analysis in *Nematostella* and *Drosophila* embryos

BiFC analysis in *Drosophila* embryos was performed as previously described (**Hudry et al., 2011**). BiFC in *Nematostella* was achieved by injecting in vitro synthesized mRNAs. mRNA synthesis was performed on a template produced by PCR using the mMessage mMachine T7 Kit (Ambion, Invitrogen, CA, USA). Embryos were co-injected with two BiFC vectors (50 ng/ $\mu$ l each) and fluorophore coupled-dextran. Embryos were allowed to develop for 24 hr and were processed for visualisation.

### Electro-mobility shift assays

Constructs for EMSAs were cloned in the pCDNA3 vector and sequence-verified before using. Proteins were produced with the TNT-T7-coupled in vitro transcription/translation system (Promega, WI, USA). Production yields of wild-type and mutated counterpart proteins were estimated by  $^{35}$ S-methionine labelling. EMSAs were performed as described previously. We used the following double strands radiolabelled probes: *Class I* 5'-ACGCGGGAATGATTGATGGCCCAAATA, *Meis-II* 5'-ATGACAGCTCGGAATGATTAATGGCCCAAATA, *Meis -IV* 5'- ATGACAGCTCGGAATGATTAATACCCCAAATA, *a1a promoter* 5'-TAATATTGTCAGTCAGATTGCAAATGATGATTGATCACTATAG, *a7 promoter* 5'-TAGGTCTGTCAGGCTGCTCTTTCACGATGATTTATGCCTCAC, *box2'* from the *tsh* epidermal enhancer *tsh* 5'-TCATGGACTGAAAACCATAAATTTGATAATTGACTTTCCAC (**McCormick et al., 1995**), *DlIR* 5'-TATTTGGAAATTAATCATTCCCGCGGACAGTT (**Gebelein et al., 2002**), *D4<sup>1-62</sup>* 5'-AGTTTACCA-TTAAATCCCATTTAGGCTGTCAATCATTTGCGCTAATTTTTCTTGCGGCTT (**Duncan et al., 2010**), *class IV* 5'- ATGACAGCTCGGAATGATTAATACCCCAAATA, *lab<sup>48/95</sup>* 5'-AAATTGATGGATTGCCCGCGCCGACTGTCACCG (**Ryoo et al., 1999**), and *modC site I* 5'-CCTCGTCCCACAGCTATAATGATTAATGAACGCGCCGCC (**Joshi et al., 2010**). The sequence of all other probes (*probes class II and III*) is indicated in the corresponding figures. 1  $\mu$ l of rat anti-HA (1/50; Molecular probe, Invitrogen, CA, USA) was used for the 'super-shift' experiments. Quantifications of shifted bands was performed using the Analyze>Gels function of the ImageJ software.

### Bioinformatic analysis

By screening the literature two sets of 38 genes exclusively expressed in the endoderm or the ectoderm of *Nematostella* embryo could be identified. The density of the Pbx-Hox and Meis motifs in the *cis*-regulatory regions of these 72 genes was determined, using a previously established matrix (**Slattery et al., 2011**). For more details, see the following website: <http://www.bigre.ulb.ac.be/Users/morgane/bruno/result.html>.

### *Xenopus* assays

One-cell stage *Xenopus laevis* embryos were microinjected with RNA encoding *Nematostella vectensis* Meis (NvMeis, 0.8 ng) protein. Animal Cap (AC) explants were removed from control and

injected embryos at blastula stage and cultured to neurula stage 16. Total RNA was isolated from five control embryos (CE) and eighteen ACs from the control or injected groups. Semi-quantitative (sq) RT-PCR analysis was performed to posterior neural markers as described (**Elkouby et al., 2010**). EF1alpha controls for RNA levels in each sample. RT-PCR was performed on total RNA from embryos.

## Acknowledgements

We warmly thank Peter Holland, Michael Manuel, Michalis Averof, Vincent Laudet, and Abderrahman Khila for critical comments and Johanne Burden for proofreading the manuscript. We thank Jun Aruga, Ian Duncan, Natalia Azpiazu, Bruno Bello and the Bloomington stock centre for reagents and fly lines. We thank Yacine Graba for allowing us performing part of this work in his laboratory. Research in the author's laboratory is supported by the CNRS, Ecole Normale Supérieure (ENS) of Lyon, EMBO short-term fellow (for B Hudry) and grants from ARC (Association de Recherche contre le Cancer) and FRM (Fondation pour la Recherche Médicale).

---

## Additional information

### Funding

Funder	Author
Association pour la Recherche contre le Cancer	Bruno Hudry, Morgane Thomas-Chollier, Yael Volovik, Marilynne Duffraisse, Amélie Dard, Dale Frank, Ulrich Technau, Samir Merabet
Fondation pour la Recherche Médicale	Bruno Hudry, Morgane Thomas-Chollier, Yael Volovik, Marilynne Duffraisse, Amélie Dard, Dale Frank, Ulrich Technau, Samir Merabet
EMBO short term fellow	Bruno Hudry

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

### Author contributions

BH, SM, Conception and design, Acquisition of data, Analysis and interpretation of data, Drafting or revising the article; MT-C, Acquisition of data, Analysis and interpretation of data, Drafting or revising the article; YV, MD, AD, DF, Acquisition of data, Analysis and interpretation of data; UT, Analysis and interpretation of data, Drafting or revising the article

---

## Additional files

### Supplementary files

- Supplementary file 1. Constructs generated for BiFC, functional and shift assays.

DOI: [10.7554/eLife.01939.028](https://doi.org/10.7554/eLife.01939.028)

---

## References

- Abu-Shaar M**, Ryoo HD, Mann RS. 1999. Control of the nuclear localization of Extradenticle by competing nuclear import and export signals. *Genes & Development* **13**:935–945. doi: [10.1101/gad.13.8.935](https://doi.org/10.1101/gad.13.8.935).
- Bellaoui M**, Pidkowich MS, Samach A, Kushalappa K, Kohalmi SE, Modrusan Z, Crosby WL, Haughn GW. 2001. The Arabidopsis BELL1 and KNOX TALE homeodomain proteins interact through a domain conserved between plants and animals. *The Plant Cell* **13**:2455–2470.
- Bischof J**, Maeda RK, Hediger M, Karch F, Basler K. 2007. An optimized transgenesis system for Drosophila using germ-line-specific phiC31 integrases. *Proceedings of the National Academy of Sciences of the United States of America* **104**:3312–3317. doi: [10.1073/pnas.0611511104](https://doi.org/10.1073/pnas.0611511104).

- Brendolan A**, Ferretti E, Salsi V, Moses K, Quaggin S, Blasi F, Cleary ML, Selleri L. 2005. A Pbx1-dependent genetic and transcriptional network regulates spleen ontogeny. *Development* **132**:3113–3126. doi: [10.1242/dev.01884](https://doi.org/10.1242/dev.01884).
- Burglin TR**. 1997. Analysis of TALE superclass homeobox genes (MEIS, PBC, KNOX, Iroquois, TGIF) reveals a novel domain conserved between plants and animals. *Nucleic Acids Research* **25**:4173–4180. doi: [10.1093/nar/25.21.4173](https://doi.org/10.1093/nar/25.21.4173).
- Burglin TR**. 1998. The PBC domain contains a MEINOX domain: coevolution of Hox and TALE homeobox genes? *Development Genes and Evolution* **208**:113–116. doi: [10.1007/s004270050161](https://doi.org/10.1007/s004270050161).
- Carr EA**, Mead J, Vershon AK. 2004. Alpha1-induced DNA bending is required for transcriptional activation by the Mcm1-alpha1 complex. *Nucleic Acids Research* **32**:2298–2305. doi: [10.1093/nar/gkh560](https://doi.org/10.1093/nar/gkh560).
- Casares F**, Calleja M, Sanchez-Herrero E. 1996. Functional similarity in appendage specification by the Ultrabithorax and abdominal-A Drosophila HOX genes. *The EMBO Journal* **15**:3934–3942.
- Chang CP**, Shen WF, Rozenfeld S, Lawrence HJ, Largman C, Cleary ML. 1995. Pbx proteins display hexapeptide-dependent cooperative DNA binding with a subset of Hox proteins. *Genes & Development* **9**:663–674. doi: [10.1101/gad.9.6.663](https://doi.org/10.1101/gad.9.6.663).
- Chourrout D**, Delsuc F, Chourrout P, Edvardsen RB, Rentzsch F, Renfer E, Jensen MF, Zhu B, de Jong P, Steele RE, Technau U. 2006. Minimal ProtoHox cluster inferred from bilaterian and cnidarian Hox complements. *Nature* **442**:684–687. doi: [10.1038/nature04863](https://doi.org/10.1038/nature04863).
- de Jong DM**, Hislop NR, Hayward DC, Reece-Hoyes JS, Pontynen PC, Ball EE, Miller DJ. 2006. Components of both major axial patterning systems of the Bilateria are differentially expressed along the primary axis of a 'radiate' animal, the anthozoan cnidarian *Acropora millepora*. *Developmental Biology* **298**:632–643. doi: [10.1016/j.ydbio.2006.07.034](https://doi.org/10.1016/j.ydbio.2006.07.034).
- Derelle R**, Lopez P, Le Guyader H, Manuel M. 2007. Homeodomain proteins belong to the ancestral molecular toolkit of eukaryotes. *Evolution & Development* **9**:212–219. doi: [10.1111/j.1525-142X.2007.00153.x](https://doi.org/10.1111/j.1525-142X.2007.00153.x).
- Dibner C**, Elias S, Frank D. 2001. XMeis3 protein activity is required for proper hindbrain patterning in *Xenopus laevis* embryos. *Development* **128**:3415–3426.
- Di-Poi N**, Zakany J, Duboule D. 2007. Distinct roles and regulations for HoxD genes in metanephric kidney development. *PLOS Genetics* **3**:e232. doi: [10.1371/journal.pgen.0030232](https://doi.org/10.1371/journal.pgen.0030232).
- Duncan D**, Kiefel P, Duncan I. 2010. Control of the spineless antennal enhancer: direct repression of antennal target genes by Antennapedia. *Developmental Biology* **347**:82–91. doi: [10.1016/j.ydbio.2010.08.012](https://doi.org/10.1016/j.ydbio.2010.08.012).
- Ebner A**, Cabernard C, Affolter M, Merabet S. 2005. Recognition of distinct target sites by a unique Labial/Extradenticle/Homothorax complex. *Development* **132**:1591–1600. doi: [10.1242/dev.01721](https://doi.org/10.1242/dev.01721).
- Elkouby YM**, Elias S, Casey ES, Blythe SA, Tsabar N, Klein PS, Root H, Liu KJ, Frank D. 2010. Mesodermal Wnt signaling organizes the neural plate via Meis3. *Development* **137**:1531–1541. doi: [10.1242/dev.044750](https://doi.org/10.1242/dev.044750).
- Finnerty JR**, Pang K, Burton P, Paulson D, Martindale MQ. 2004. Origins of bilateral symmetry: Hox and dpp expression in a sea anemone. *Science* **304**:1335–1337. doi: [10.1126/science.1091946](https://doi.org/10.1126/science.1091946).
- Fujioka M**, Gebelein B, Cofer ZC, Mann RS, Jaynes JB. 2012. Engrailed cooperates directly with Extradenticle and Homothorax on a distinct class of homeodomain binding sites to repress sloppy paired. *Developmental Biology* **366**:382–392. doi: [10.1016/j.ydbio.2012.04.004](https://doi.org/10.1016/j.ydbio.2012.04.004).
- Galant R**, Walsh CM, Carroll SB. 2002. Hox repression of a target gene: extradenticle-independent, additive action through multiple monomer binding sites. *Development* **129**:3115–3126.
- Gauchat D**, Mazet F, Berney C, Schummer M, Kreger S, Pawlowski J, Galliot B. 2000. Evolution of Antp-class genes and differential expression of Hydra Hox/paraHox genes in anterior patterning. *Proceedings of the National Academy of Sciences of the United States of America* **97**:4493–4498. doi: [10.1073/pnas.97.9.4493](https://doi.org/10.1073/pnas.97.9.4493).
- Gebelein B**, Culi J, Ryoo HD, Zhang W, Mann RS. 2002. Specificity of Distalless repression and limb primordia development by abdominal Hox proteins. *Developmental Cell* **3**:487–498. doi: [10.1016/S1534-5807\(02\)00257-5](https://doi.org/10.1016/S1534-5807(02)00257-5).
- Genikhovich G**, Technau U. 2009. In situ hybridization of starlet sea anemone (*Nematostella vectensis*) embryos, larvae, and polyps. *Cold Spring Harbour Protocols* **2009**:pdb prot5282. doi: [10.1101/pdb.prot5282](https://doi.org/10.1101/pdb.prot5282).
- Hackbusch J**, Richter K, Muller J, Salamini F, Uhrig JF. 2005. A central role of *Arabidopsis thaliana* ovate family proteins in networking and subcellular localization of 3-aa loop extension homeodomain proteins. *Proceedings of the National Academy of Sciences of the United States of America* **102**:4908–4912. doi: [10.1073/pnas.0501181102](https://doi.org/10.1073/pnas.0501181102).
- Hay A**, Tsiantis M. 2010. KNOX genes: versatile regulators of plant development and diversity. *Development* **137**:3153–3165. doi: [10.1242/dev.030049](https://doi.org/10.1242/dev.030049).
- Heffer A**, Pick L. 2013. Conservation and variation in Hox genes: how insect models pioneered the evo-devo field. *Annual Review of Entomology* **58**:161–179. doi: [10.1146/annurev-ento-120811-153601](https://doi.org/10.1146/annurev-ento-120811-153601).
- Hirth F**, Hartmann B, Reichert H. 1998. Homeotic gene action in embryonic brain development of *Drosophila*. *Development* **125**:1579–1589.
- Hudry B**, Viala S, Graba Y, Merabet S. 2011. Visualization of protein interactions in living *Drosophila* embryos by the bimolecular fluorescence complementation assay. *BMC Biology* **9**:5. doi: [10.1186/1741-7007-9-5](https://doi.org/10.1186/1741-7007-9-5).
- Hudry B**, Remacle S, Delfini MC, Rezsosahay R, Graba Y, Merabet S. 2012. Hox proteins display a common and ancestral ability to diversify their interaction mode with the PBC class cofactors. *PLOS Biology* **10**:e1001351. doi: [10.1371/journal.pbio.1001351](https://doi.org/10.1371/journal.pbio.1001351).
- In der Rieden PM**, Mainguy G, Woltering JM, Durston AJ. 2004. Homeodomain to hexapeptide or PBC-interaction-domain distance: size apparently matters. *Trends in Genetics* **20**:76–79. doi: [10.1016/j.tig.2003.12.001](https://doi.org/10.1016/j.tig.2003.12.001).

- Irimia M, Tena JJ, Alexis MS, Fernandez-Minan A, Maeso I, Bogdanovic O, de la Calle-Mustienes E, Roy SW, Gómez-Skarmeta JL, Fraser HB. 2012. Extensive conservation of ancient microsynteny across metazoans due to cis-regulatory constraints. *Genome Research* **22**:2356–2367. doi: [10.1101/gr.139725.112](https://doi.org/10.1101/gr.139725.112).
- Jakob W, Schierwater B. 2007. Changing hydrozoan bauplans by silencing Hox-like genes. *PLOS ONE* **2**:e694. doi: [10.1371/journal.pone.0000694](https://doi.org/10.1371/journal.pone.0000694).
- Joshi R, Sun L, Mann R. 2010. Dissecting the functional specificities of two Hox proteins. *Genes & Development* **24**:1533–1545. doi: [10.1101/gad.1936910](https://doi.org/10.1101/gad.1936910).
- Kamm K, Schierwater B, Jakob W, Dellaporta SL, Miller DJ. 2006. Axial patterning and diversification in the cnidaria predate the Hox system. *Current Biology* **16**:920–926. doi: [10.1016/j.cub.2006.03.036](https://doi.org/10.1016/j.cub.2006.03.036).
- Kanrar S, Onguka O, Smith HM. 2006. Arabidopsis inflorescence architecture requires the activities of KNOX-BELL homeodomain heterodimers. *Planta* **224**:1163–1173. doi: [10.1007/s00425-006-0298-9](https://doi.org/10.1007/s00425-006-0298-9).
- Keleher CA, Passmore S, Johnson AD. 1989. Yeast repressor alpha 2 binds to its operator cooperatively with yeast protein Mcm1. *Molecular and Cellular Biology* **9**:5228–5230.
- Kodama Y, Hu CD. 2012. Bimolecular fluorescence complementation (BiFC): a 5-year update and future perspectives. *Biotechniques* **53**:285–298. doi: [10.2144/000113943](https://doi.org/10.2144/000113943).
- Larroux C, Luke GN, Koopman P, Rokhsar DS, Shimeld SM, Degnan BM. 2008. Genesis and expansion of metazoan transcription factor gene classes. *Molecular Biology and Evolution* **25**:980–996. doi: [10.1093/molbev/msn047](https://doi.org/10.1093/molbev/msn047).
- Li T, Jin Y, Vershon AK, Wolberger C. 1998. Crystal structure of the MATA1/MATalpha2 homeodomain heterodimer in complex with DNA containing an A-tract. *Nucleic Acids Research* **26**:5707–5718. doi: [10.1093/nar/26.24.5707](https://doi.org/10.1093/nar/26.24.5707).
- Maeda R, Mood K, Jones TL, Aruga J, Buchberg AM, Daar IO. 2001. Xmeis1, a protooncogene involved in specifying neural crest cell fate in *Xenopus* embryos. *Oncogene* **20**:1329–1342. doi: [10.1038/sj.onc.1204250](https://doi.org/10.1038/sj.onc.1204250).
- Mann RS, Lelli KM, Joshi R. 2009. Hox specificity unique roles for cofactors and collaborators. *Current Topics In Developmental Biology* **88**:63–101. doi: [10.1016/S0070-2153\(09\)88003-4](https://doi.org/10.1016/S0070-2153(09)88003-4).
- Matus DQ, Pang K, Marlow H, Dunn CW, Thomsen GH, Martindale MQ. 2006. Molecular evidence for deep evolutionary roots of bilaterality in animal development. *Proceedings of the National Academy of Sciences of the United States of America* **103**:11195–11200. doi: [10.1073/pnas.0601257103](https://doi.org/10.1073/pnas.0601257103).
- Matus DQ, Thomsen GH, Martindale MQ. 2006. Dorsal/ventral genes are asymmetrically expressed and involved in germ-layer demarcation during cnidarian gastrulation. *Current Biology* **16**:499–505. doi: [10.1016/j.cub.2006.01.052](https://doi.org/10.1016/j.cub.2006.01.052).
- McCormick A, Core N, Kerridge S, Scott MP. 1995. Homeotic response elements are tightly linked to tissue-specific elements in a transcriptional enhancer of the teashirt gene. *Development* **121**:2799–2812.
- Mendivil Ramos O, Barker D, Ferrier DE. 2012. Ghost loci imply Hox and ParaHox existence in the last common ancestor of animals. *Current Biology* **22**:1951–1956. doi: [10.1016/j.cub.2012.08.023](https://doi.org/10.1016/j.cub.2012.08.023).
- Merabet S, Saadaoui M, Sambrani N, Hudry B, Pradel J, Affolter M, Graba Y. 2007. A unique Extradenticle recruitment mode in the *Drosophila* Hox protein Ultrabithorax. *Proceedings of the National Academy of Sciences of the United States of America* **104**:16946–16951. doi: [10.1073/pnas.0705832104](https://doi.org/10.1073/pnas.0705832104).
- Merabet S, Hudry B, Saadaoui M, Graba Y. 2009. Classification of sequence signatures: a guide to Hox protein function. *Bioessays* **31**:500–511. doi: [10.1002/bies.200800229](https://doi.org/10.1002/bies.200800229).
- Prince F, Katsuyama T, Oshima Y, Plaza S, Resendez-Perez D, Berry M, Kurata S, Gehring WJ. 2008. The YPWM motif links Antennapedia to the basal transcriptional machinery. *Development* **135**:1669–1679. doi: [10.1242/dev.018028](https://doi.org/10.1242/dev.018028).
- Rhee JW, Arata A, Selleri L, Jacobs Y, Arata S, Onimaru H, Cleary ML. 2004. Pbx3 deficiency results in central hypoventilation. *The American Journal of Pathology* **165**:1343–1350. doi: [10.1016/S0002-9440\(10\)63392-5](https://doi.org/10.1016/S0002-9440(10)63392-5).
- Ryan JF, Burton PM, Mazza ME, Kwong GK, Mullikin JC, Finnerty JR. 2006. The cnidarian-bilaterian ancestor possessed at least 56 homeoboxes: evidence from the starlet sea anemone, *Nematostella vectensis*. *Genome Biology* **7**:R64. doi: [10.1186/gb-2006-7-7-r64](https://doi.org/10.1186/gb-2006-7-7-r64).
- Ryan JF, Mazza ME, Pang K, Matus DQ, Baxeveanis AD, Martindale MQ, Finnerty JR. 2007. Pre-bilaterian origins of the Hox cluster and the Hox code: evidence from the sea anemone, *Nematostella vectensis*. *PLOS ONE* **2**:e153. doi: [10.1371/journal.pone.0000153](https://doi.org/10.1371/journal.pone.0000153).
- Ryan JF, Pang K, NISC Comparative Sequencing Program, Mullikin JC, Martindale MQ, Baxeveanis AD. 2010. The homeodomain complement of the ctenophore *Mnemiopsis leidyi* suggests that Ctenophora and Porifera diverged prior to the ParaHoxozoa. *EvoDevo* **1**:9. doi: [10.1186/2041-9139-1-9](https://doi.org/10.1186/2041-9139-1-9).
- Ryoo HD, Marty T, Casares F, Affolter M, Mann RS. 1999. Regulation of Hox target genes by a DNA bound Homothorax/Hox/Extradenticle complex. *Development* **126**:5137–5148.
- Saadaoui M, Merabet S, Litim-Mecheri I, Arbeille E, Sambrani N, Damen W, Brena C, Pradel J, Graba Y. 2011. Selection of distinct Hox-Extradenticle interaction modes fine-tunes Hox protein activity. *Proceedings of the National Academy of Sciences of the United States of America* **108**:2276–2281. doi: [10.1073/pnas.1006964108](https://doi.org/10.1073/pnas.1006964108).
- Saina M, Genikhovich G, Renfer E, Technau U. 2009. BMPs and chordin regulate patterning of the directive axis in a sea anemone. *Proceedings of the National Academy of Sciences of the United States of America* **106**:18592–18597. doi: [10.1073/pnas.0900151106](https://doi.org/10.1073/pnas.0900151106).
- Saleh M, Huang H, Green NC, Featherstone MS. 2000. A conformational change in PBX1A is necessary for its nuclear localization. *Experimental Cell Research* **260**:105–115. doi: [10.1006/excr.2000.5010](https://doi.org/10.1006/excr.2000.5010).
- Salzberg A, Elias S, Nachaliel N, Bonstein L, Henig C, Frank D. 1999. A Meis family protein caudalizes neural cell fates in *Xenopus*. *Mechanisms of Development* **80**:3–13. doi: [10.1016/S0925-4773\(98\)00187-7](https://doi.org/10.1016/S0925-4773(98)00187-7).
- Schierwater B, Kuhn K. 1998. Homology of Hox genes and the zootype concept in early metazoan evolution. *Molecular Phylogenetics and Evolution* **9**:375–381. doi: [10.1006/mpev.1998.0489](https://doi.org/10.1006/mpev.1998.0489).

- Shen WF**, Chang CP, Rozenfeld S, Sauvageau G, Humphries RK, Lu M, Lawrence HJ, Cleary ML, Largman C. 1996. Hox homeodomain proteins exhibit selective complex stabilities with Pbx and DNA. *Nucleic Acids Research* **24**:898–906. doi: [10.1093/nar/24.5.898](https://doi.org/10.1093/nar/24.5.898).
- Shen WF**, Montgomery JC, Rozenfeld S, Moskow JJ, Lawrence HJ, Buchberg AM, Largman C. 1997. AbdB-like Hox proteins stabilize DNA binding by the Meis1 homeodomain proteins. *Molecular and Cellular Biology* **17**:6448–6458.
- Sinigaglia C**, Busengdal H, Leclere L, Technau U, Rentsch F. 2013. The bilaterian head patterning gene six3/6 controls aboral domain development in a cnidarian. *PLOS Biology* **11**:e1001488. doi: [10.1371/journal.pbio.1001488](https://doi.org/10.1371/journal.pbio.1001488).
- Slack JM**, Holland PW, Graham CF. 1993. The zootype and the phylotypic stage. *Nature* **361**:490–492. doi: [10.1038/361490a0](https://doi.org/10.1038/361490a0).
- Slattery M**, Riley T, Liu P, Abe N, Gomez-Alcala P, Dror I, Zhou T, Rohs R, Honig B, Bussemaker HJ, Mann RS. 2011. Cofactor binding evokes latent differences in DNA binding specificity between Hox proteins. *Cell* **147**:1270–1282. doi: [10.1016/j.cell.2011.10.053](https://doi.org/10.1016/j.cell.2011.10.053).
- Stark MR**, Johnson AD. 1994. Interaction between two homeodomain proteins is specified by a short C-terminal tail. *Nature* **371**:429–432. doi: [10.1038/371429a0](https://doi.org/10.1038/371429a0).
- Thomas-Chollier M**, Ledent V, Leyns L, Vervoort M. 2010. A non-tree-based comprehensive study of metazoan Hox and ParaHox genes prompts new insights into their origin and evolution. *BMC Evolutionary Biology* **10**:73. doi: [10.1186/1471-2148-10-73](https://doi.org/10.1186/1471-2148-10-73).
- Venken KJ**, He Y, Hoskins RA, Bellen HJ. 2006. P[acman]: a BAC transgenic platform for targeted insertion of large DNA fragments in *D. melanogaster*. *Science* **314**:1747–1751. doi: [10.1126/science.1134426](https://doi.org/10.1126/science.1134426).
- Williams TM**, Williams ME, Innis JW. 2005. Range of HOX/TALE superclass associations and protein domain requirements for HOXA13:MEIS interaction. *Developmental Biology* **277**:457–471. doi: [10.1016/j.ydbio.2004.10.004](https://doi.org/10.1016/j.ydbio.2004.10.004).
- Yao LC**, Liaw GJ, Pai CY, Sun YH. 1999. A common mechanism for antenna-to-Leg transformation in *Drosophila*: suppression of homothorax transcription by four HOM-C genes. *Developmental Biology* **211**:268–276. doi: [10.1006/dbio.1999.9309](https://doi.org/10.1006/dbio.1999.9309).

## 2. DEUXIEME CHAPITRE : DES PETITS MOTIFS POUR REGULER LE POTENTIEL D'INTERACTION DES PROTEINES HOX

Afin de mieux comprendre les rôles des protéines HOX dans le développement embryonnaire, nous avons mis au point une technique de visualisation des interactions protéine-protéine (IPPs) directement dans l'embryon vivant de drosophile par BiFC. Nous avons conduit un crible avec des gènes candidats, choisis pour être des facteurs de transcription impliqués dans le développement embryonnaire et dont le rôle et le profil d'expression sont connus.

Au cours de ce crible plusieurs informations clés sont ressorties. D'une part que les protéines HOX ont un « potentiel d'interaction » très important, leur permettant de vastes possibilités d'interaction. Ensuite, qu'un nombre important d'interactions sont communes au sein des cinq protéines Hox testées. Enfin, qu'une part importante de la spécificité des interactomes pouvait reposer sur une inhibition des interactions non spécifiques pouvant en particulier être établies par les régions désorganisées des protéines Hox. Nous montrons que cette inhibition des IPPs est réalisée par certains petits motifs des protéines Hox, comme les motifs HX et UbdA.

Ce travail révèle l'importance des petits motifs et de leur rôle inhibiteur des IPPs pour la spécificité des interactomes des protéines Hox.







# Inhibitory activities of short linear motifs underlie Hox interactome specificity *in vivo*

Manon Baëza<sup>1,2†</sup>, Séverine Viala<sup>1,2†</sup>, Marjorie Heim<sup>1,2†</sup>, Amélie Dard<sup>1,2</sup>, Bruno Hudry<sup>3</sup>, Marilyne Duffraisse<sup>1,2</sup>, Ana Rogulja-Ortmann<sup>4</sup>, Christine Brun<sup>5,6</sup>, Samir Merabet<sup>1,2\*</sup>

<sup>1</sup>Institut de génomique fonctionnelle de Lyon, Centre National de Recherche Scientifique, Lyon, France; <sup>2</sup>Institut de génomique fonctionnelle de Lyon, École normale supérieure de Lyon, Lyon, France; <sup>3</sup>MRC Clinical Sciences Centre, Faculty of Medicine, Imperial College London, Hammersmith Hospital Campus, London, United Kingdom; <sup>4</sup>Institut of Genetics, University of Mainz, Mainz, Germany; <sup>5</sup>Technological Advances for Genomics and clinics, Institut national de la santé et de la recherche médicale, University Aix-Marseille, Parc Scientifique de Luminy, Marseille, France; <sup>6</sup>Centre National de la Recherche Scientifique, Marseille, France

**Abstract** Hox proteins are well-established developmental regulators that coordinate cell fate and morphogenesis throughout embryogenesis. In contrast, our knowledge of their specific molecular modes of action is limited to the interaction with few cofactors. Here, we show that Hox proteins are able to interact with a wide range of transcription factors in the live *Drosophila* embryo. In this context, specificity relies on a versatile usage of conserved short linear motifs (SLiMs), which, surprisingly, often restrains the interaction potential of Hox proteins. This novel buffering activity of SLiMs was observed in different tissues and found in Hox proteins from cnidarian to mouse species. Although these interactions remain to be analysed in the context of endogenous Hox regulatory activities, our observations challenge the traditional role assigned to SLiMs and provide an alternative concept to explain how Hox interactome specificity could be achieved during the embryonic development.

DOI: [10.7554/eLife.06034.001](https://doi.org/10.7554/eLife.06034.001)

\*For correspondence: samir.merabet@ens-lyon.fr

†These authors contributed equally to this work

**Competing interests:** The authors declare that no competing interests exist.


**Funding:** See page 24

**Received:** 15 December 2014

**Accepted:** 16 March 2015

**Published:** 14 April 2015

**Reviewing editor:** Diethard Tautz, Max Planck Institute for Evolutionary Biology, Germany

 Copyright Baëza et al. This article is distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use and redistribution provided that the original author and source are credited.

## Introduction

There is mounting evidence that many protein–protein interactions (PPIs) are mediated by small peptide motifs called linear motifs (LMs) or eukaryotic/short linear motifs (ELMs/SLiMs) (Neduvu and Russell, 2005; Van Roey et al., 2012, 2014; Tompa et al., 2014). These compact interaction interfaces are typically less than 10 residues in length and are often located within intrinsically disordered regions of highly connected proteins. Due to their small size, SLiMs exhibit high evolutionary plasticity and mediate interactions with many different types of proteins. Moreover, SLiMs are known to be important for the rewiring of interaction networks, being the subject of tissue-specific regulatory mechanisms (Buljan et al., 2013).

The contribution of SLiMs to the functional diversification and specification of key regulatory TFs throughout development and evolution remains poorly understood. For example, very few SLiMs listed in the current databases relate to the regulatory interaction between transcription factors (TFs) (Dinkel et al., 2014), suggesting that this particular type of functional interaction is more difficult to capture than others. More generally, classic large-scale screening methods based on affinity purification followed by mass spectrometry (AP-MS) or yeast two-hybrid are more efficient for

**eLife digest** In all animals, it is important that cells are correctly organised into tissues and organs. This organisation starts in the embryo, and cells are instructed to perform different roles depending on their position within the body.

A family of proteins called the Hox proteins coordinates the organisation of the cells in the animal embryo by binding to and controlling the expression of specific genes. To properly control their target genes, Hox proteins need to interact with other proteins called transcription factors that can also bind to the genes. However, only a few of these transcription factors have been identified so far, and it is not clear how Hox proteins are able to interact with them.

Here, Baëza, Viala, Heim et al. identified several more transcription factors that can bind to the Hox proteins in fruit fly embryos. The experiments show that Hox proteins are able to bind to many transcription factors that are very different from each other. Baëza, Viala, Heim et al. also show that two short sections within the Hox proteins known as short linear motifs are important for controlling these interactions. A fly Hox protein that was missing these motifs was able to interact with new transcription factors. This inhibitory role was found in Hox proteins from mice and sea anemones, suggesting that these motifs may play the same role in all animals.

Baëza, Viala, Heim et al.'s findings challenge the traditional view of the role of the short linear motifs in interactions between proteins. Also, the findings provide an alternative explanation for how the Hox proteins are only able to interact with particular transcription factors in animal embryos. The next step will be to find out whether the inhibitory role of short linear motifs could more generally apply to many other protein families.

DOI: [10.7554/eLife.06034.002](https://doi.org/10.7554/eLife.06034.002)

detecting stable interactions between structured domains than for revealing transient interactions involving SLiMs (Landry et al., 2013). Therefore, alternative approaches are needed to decipher SLiM-mediated interactions and functions within the context of developmental regulatory networks in vivo.

Here, we tackle this issue by using Hox proteins as a case study. Hox proteins are homeodomain (HD)-containing TFs present in all cnidarian and bilaterian species (Finnerty, 2003). They are required throughout the embryogenesis for controlling specific cell fates and structures along different axes and in territories as different as the limb bud (Zakany and Duboule, 2007), cardiac outflow tract (Bertrand et al., 2011), and female genital disc (Foronda et al., 2005). The specific functions of Hox proteins in vivo contrast with their ability to recognize closely similar DNA-binding sites as monomers in vitro (Berger et al., 2008; Noyes et al., 2008). This so-called Hox paradox strongly suggests that additional cofactors are required for helping Hox proteins to elicit their diverse and specific transcriptional programs in vivo.

To date, only one type of cofactors is described to specify Hox functions at the molecular level. These cofactors are collectively referred to as the PBC class of HD-containing TFs, and correspond to the Pbx1-4 and Extradenticle (Exd) proteins in mammals and *Drosophila melanogaster*, respectively (Mukherjee and Bürglin, 2007). Biochemical studies have shown that the interaction between Hox and PBC proteins relies on a highly conserved motif of Hox proteins, the hexapeptide (HX) (Mann et al., 2009). The HX motif constitutes the generic signature of Hox proteins after key positions within the HD (Merabet et al., 2009). It belongs to the LIG type of SLiM according to the ELM database (<http://elm.eu.org/>) and contains a core YF/PWM sequence conserved in all but posterior Hox paralogs throughout animal evolution (Merabet et al., 2009). More generally, the HX motif has been defined as corresponding to an invariant Tryptophan residue located in a hydrophobic environment and followed by basic residues from +2 to +5 (In der Rieden et al., 2004). Crystal structures with truncated proteins emphasized the importance of the Tryptophan residue in establishing strong interactions with specific residues of the PBC partner (Mann et al., 2009). However, in vivo analyses showed that mutations within the HX motif (including that of the key Tryptophan residue) did not systematically abolish PBC-dependent functions of Hox proteins (Galant et al., 2002; Merabet et al., 2003). In addition, Hox-PBC interactions are influenced by the promoter environment and can occur in absence of the HX motif in several cnidarian and bilaterian Hox proteins (Hudry et al., 2012, 2014). Reciprocally, the HX is also required for PBC-independent functions (Merabet et al., 2011) and for

interacting with Bip2, a TATA-binding protein associated factor in *Drosophila* (Prince et al., 2008). Together these observations highlight that the HX motif is neither a unique nor an obligatory Hox protein interface for recruiting the PBC cofactor, suggesting that Hox-PBC interactions could rely on the presence of other specific SLiM(s). Notably, another motif called UbdA and present in Ultrabithorax (Ubx) and AbdominalA (AbdA) proteins of protostome lineages was recently described to be important for the formation and activity of the Ubx/Exd complex in *Drosophila* (Merabet et al., 2007; Hudry et al., 2012; Foos et al., 2015).

In summary, our current knowledge on SLiM-mediated interactions in Hox proteins is limited to only two different types of TFs, the PBC and Bip2 proteins. Given the number of embryonic events controlled by Hox proteins, we hypothesize that Hox SLiMs such as the HX and UbdA motifs could interact with a higher number of TFs. Identifying these TFs represents a major challenge to understand part of the molecular cues underlying Hox transcriptional specificity and diversity in vivo.

Here, we exploited the recently developed bimolecular fluorescence complementation (BiFC) (Hudry et al., 2011) to profile a wide range of Hox protein interactions in the *Drosophila* embryo and investigate whether SLiMs could influence their specificity in vivo. As a first step, we identified the respective sets of BiFC interactors of five *Drosophila* Hox paralogs, showing that each Hox interactome relies on a different combination of TFs. The role of the HX and UbdA motifs was then analysed in several Hox interactomes and in different tissues of the live embryo. Our data establish that the ablation of Hox SLiMs not only prevents several interactions but additionally leads to a number of ectopic interactions. These effects differ depending on the Hox protein and tissue considered, suggesting that SLiM activity could be strongly influenced by the protein environment. Furthermore, results obtained with mouse and cnidarian Hox proteins indicate that the inhibitory activity of SLiMs could be important for restricting the inherent binding potential of intrinsically disordered regions.

Altogether, these findings provide new insights on how Hox transcriptional specificity could be reached in vivo and add to the functional repertoire of SLiMs.

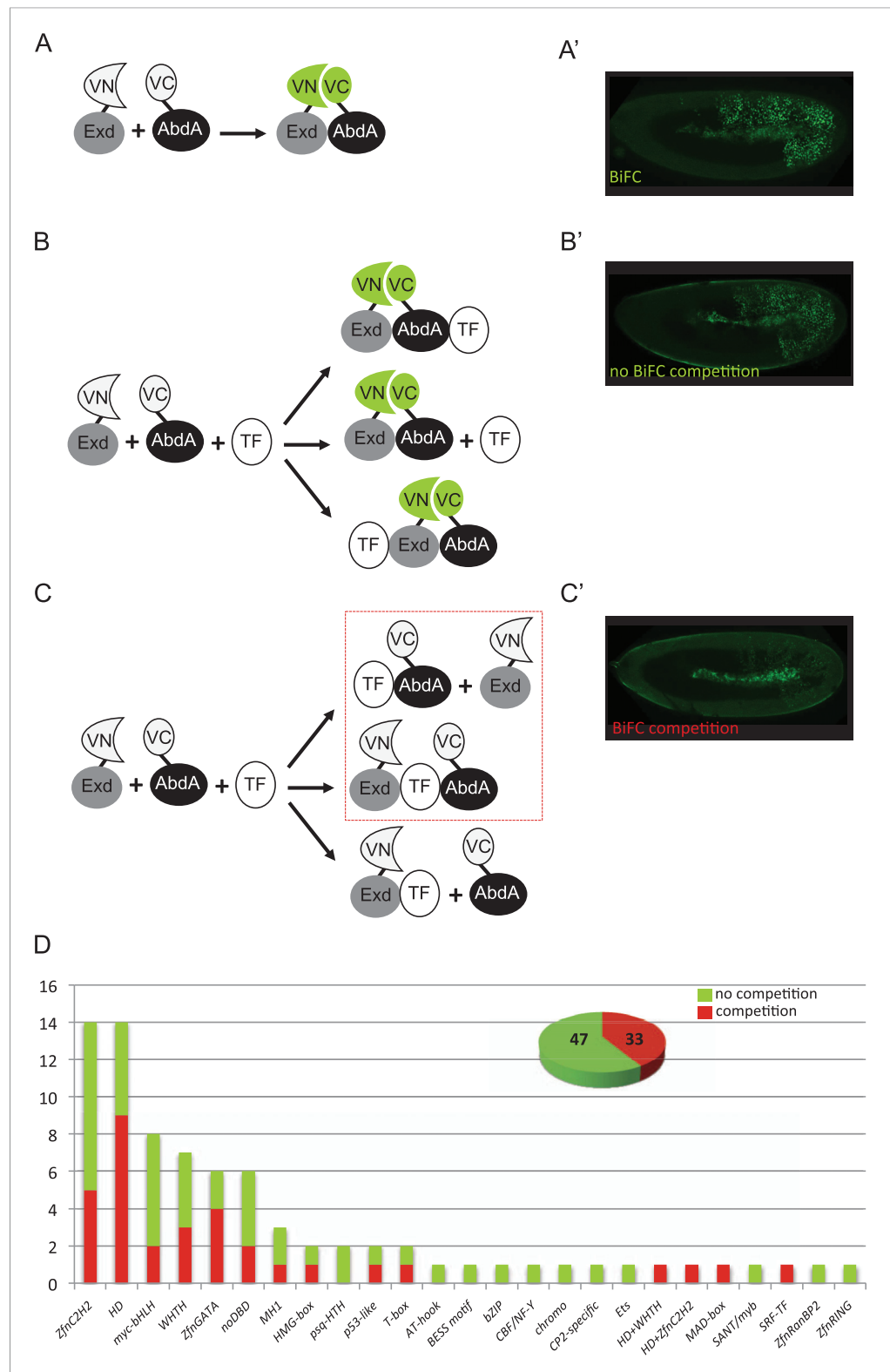
## Results

### A competitive BiFC screen reveals new candidate binding partners of the Hox protein AbdA in the live *Drosophila* embryo

BiFC relies on the property of monomeric fluorescent proteins to be reconstituted from two separate sub-fragments upon spatial rearrangement (Ghosh et al., 2000). This property is used with different types of proteins in various cell and animal model systems to demonstrate the close proximity hence the existence of possible interactions between two putative protein partners (Kerppola, 2008; Kodama and Hu, 2012).

We previously demonstrated that BiFC was sensitive and specific enough for analysing Hox-TF interactions in the live *Drosophila* embryo (Hudry et al., 2011). Experimental parameters were established by using the partnership between AbdA and Exd as a case study. Interaction was visualized by fusing the two partners with complementary fragments of the Venus (yellow: Figure 1A–A'), mCherry (red) or Cerulean (blue) fluorescent protein. Among several controls, we showed that the simultaneous co-expression of a 'cold' AbdA protein (i.e., not fused with a fragment of the fluorescent protein) with AbdA and Exd fusion proteins could induce a titration of the BiFC complex (Hudry et al., 2011). Thus, cold interactions (in this case AbdA–AbdA and AbdA–Exd interactions) could compete against BiFC, leading to a significant decrease of fluorescent signals in the embryo (Hudry et al., 2011).

We reasoned that any protein capable of displacing the AbdA and/or Exd fusion protein from the BiFC complex could lead to a loss of the fluorescence. This readout could thus serve to rapidly identify putative interacting partners of AbdA. In this case, competitive BiFC could not be observed with proteins that are exclusively participating in the AbdA/Exd complex (Figure 1B–B'). In addition, the titration of BiFC signals could only occur when the cold interaction is strong enough to disrupt the assembly between AbdA and Exd (Figure 1C–C'). However, as competitive BiFC could not discriminate between AbdA- and Exd-specific interacting partners (Figure 1C–C'), a second experimental phase will be necessary to confirm the Hox interaction status. Despite these limitations, we decided to test our hypothesis with a reasonable number of candidate TFs and by using a fast genetic approach. To this end, we established a BiFC reporter fly line expressing AbdA and Exd fusion



**Figure 1.** A BiFC competition screen identifies candidate transcription factors (TFs) as potential binding partners of the Hox protein AbdominalA (AbdA). **(A–C)** Principle of the competition test. **(A–A')** Co-expression of Extradenticle (Exd) and AbdA proteins fused to the N-(VN) or C-(VC) terminal fragment of the Venus fluorescent protein leads to BiFC. **(B–B')** Co-expression of Exd, AbdA, and a TF leads to no BiFC competition. **(C–C')** Co-expression of Exd, AbdA, and a TF leads to BiFC competition. **(D)** Summary of the BiFC competition screen results. The number of candidate TFs identified as potential binding partners of AbdA is shown for each TF motif. The total number of TFs identified is 47 (no competition) and 33 (competition). *Figure 1. continued on next page*

Figure 1. Continued

BiFC in the embryo. (B–B') Cases where no BiFC competition could be observed with a cold TF. B' is an illustrative picture of non-competitive BiFC resulting from the simultaneous co-expression of the red fluorescent protein RFP (see also **Figure 1—figure supplement 2**). (C–C') Cases where BiFC competition could be observed with a cold TF. C' is an illustrative picture of competitive BiFC resulting from the simultaneous co-expression of a nuclear-localized form of Exd (see also **Figure 1—figure supplement 1**). Note that AbdA-interacting partners do not obligatory lead to competitive BiFC. TFs that could be validated in the secondary step as AbdA-binding partners (see **Figure 2**) are indicated (dotted-red box). (D) Graph showing the repartition of competitive (red bars) and non-competitive (green bars) TFs with regard to their DNA-binding domain. See also **Supplementary file 1**.

DOI: [10.7554/eLife.06034.003](https://doi.org/10.7554/eLife.06034.003)

The following figure supplements are available for figure 1:

**Figure supplement 1.** Illustrative pictures of competitive TFs.

DOI: [10.7554/eLife.06034.004](https://doi.org/10.7554/eLife.06034.004)

**Figure supplement 2.** Illustrative pictures of non-competitive TFs.

DOI: [10.7554/eLife.06034.005](https://doi.org/10.7554/eLife.06034.005)

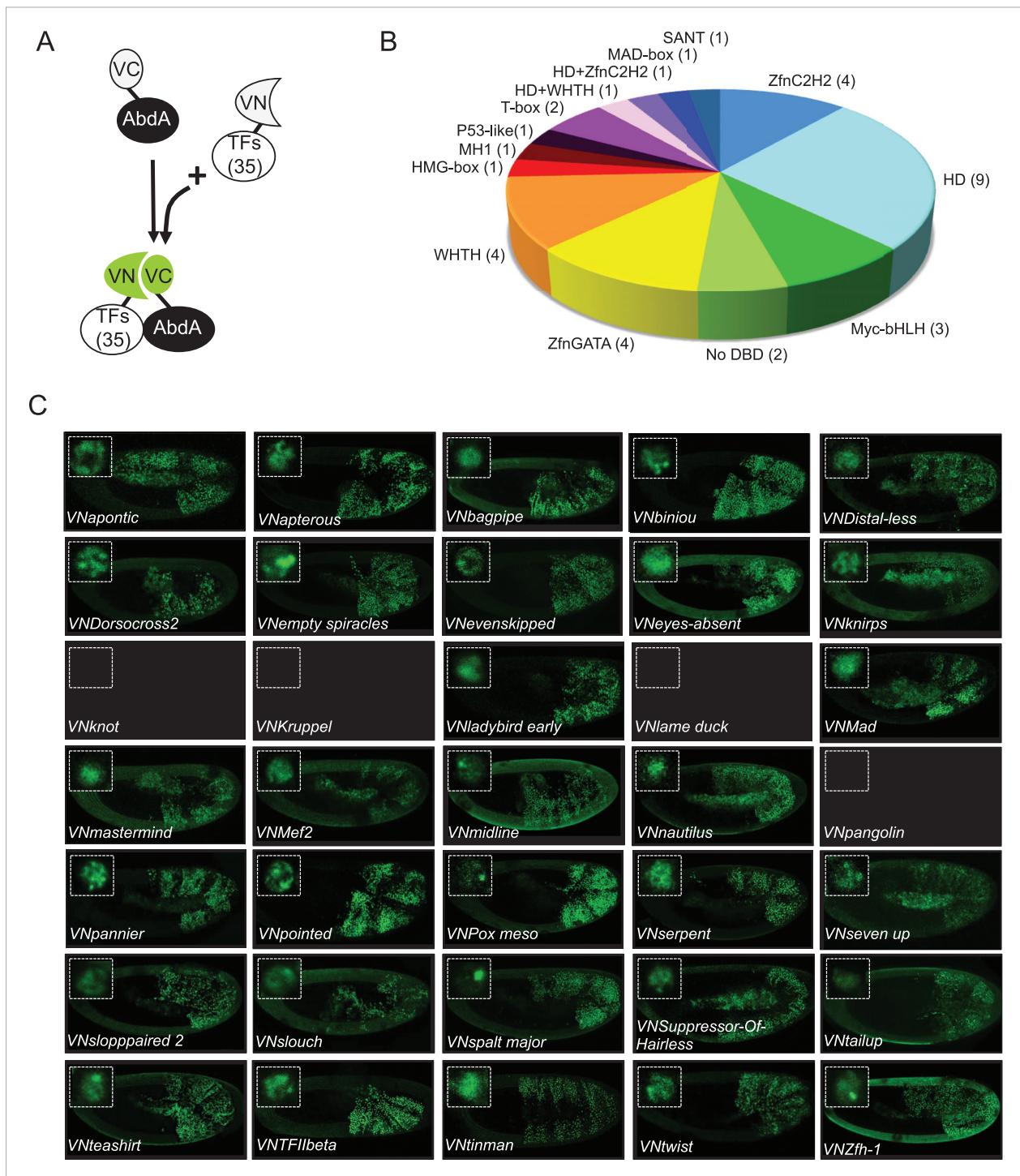
proteins under the control of the *abdA-Gal4* driver ('Materials and methods' and [Hudry et al., 2011]). This fly line can be crossed with individuals containing any UAS-driven cold candidate-binding partner and BiFC could directly be assessed in the embryo progeny. As a control, we verified that the co-expression of a nuclear-localized form of Exd (Kammermeier et al., 2004) could indeed affect BiFC fluorescent signals in the embryo (**Figure 1—figure supplement 1**). In comparison, co-expressing a red fluorescent protein (RFP) under the same condition did not lead to any changes in the BiFC profile (**Figure 1—figure supplement 2**).

Here, we deliberately focused the screen on TFs that could participate in the transcriptional programs of Hox proteins in their various developmental contexts. We chose a starting set of 80 TFs covering different DNA-binding families and displaying distinct expression profiles in the three main germ layers of the embryo (**Supplementary file 1**). We observed that 33 of those TFs could compete against AbdA/Exd assembly (**Figure 1D** and **Figure 1—figure supplements 1, 2**). Among them we found Biniou (Bin) and Mad, previously described to participate in the regulation of Hox target enhancers (Grienenberger et al., 2003; Walsh and Carroll, 2007), and Teashirt (Tsh), known to help Hox proteins in specifying the trunk segments of the embryo (Fasano et al., 1991). Thus, these three TFs validate the competition screen. Interestingly, 65% of the tested HD and GATA TFs (9/14 and 4/6, respectively) were positive in the competition test (**Figure 1D**), representing a strong tendency compared to the other tested TF classes. In total, none of the 33 competitors was previously described as a binding partner of AbdA or Exd, illustrating the efficiency of the competition screen for revealing new candidate cofactors in vivo.

### The majority of AbdA-TF interactions can occur in different cell-contexts

Our genetic competitive approach revealed proteins that could potentially bind to AbdA and/or Exd. We next analysed whether these positive competitors could more specifically interact with AbdA in a complementary BiFC-based approach. To this end, we generated fly lines carrying each corresponding TF as a UAS-driven fusion construct compatible for BiFC (**Figure 2A** and 'Materials and methods'). Two additional TFs (TFIIbeta and Knot [Kn]) were added to the 33 positive competitors, reaching a total of 35 fusion TFs that could be used for BiFC in *Drosophila* (**Figure 2B** and **Supplementary file 1**). TFIIbeta, could not be tested by competition, as no corresponding UAS-driven fly line exists. It is however a good positive candidate as it is described to interact with Ubx in a yeast two-hybrid screen (Bondos et al., 2004). Kn did not compete against BiFC in the first step and was used as a negative candidate interacting partner of the experiment.

BiFC was observed in the epidermis of stage 10 embryos, even for TFs that are not endogenously expressed in this tissue (see 'Materials and methods', **Supplementary file 2** and [Hammonds et al., 2013]). We anticipated that the epidermis was appropriate for the interaction since competition was observed in this tissue. In addition, the epidermis has been shown to tolerate the activation of mesodermal target genes upon the ectopic expression of mesoderm-specific TFs (Cunha et al., 2010), suggesting it is a relatively neutral tissue. Finally, BiFC could also be increased in a heterologous tissue because of the absence of competition by the endogenous gene product, as previously described (Hudry et al., 2011).



**Figure 2.** BiFC validates the AbdA-interaction status of competitive TFs. **(A)** Principle of the BiFC screen between AbdA and the 35 selected TFs. **(B)** Repartition of the 35 selected TFs with regard to their DNA-binding domain. **(C)** Illustrative pictures of BiFC signals obtained between VC-AbdA and the indicated VN-TF in the epidermis of stage 10–12 of live embryos. Fusion constructs are expressed with the *abdA-Gal4* driver. Note that typical nuclear interaction profiles are observed between AbdA and different TFs (white-dotted boxes). See also ‘Materials and methods’, **Supplementary files 2, 3** and **Figure 2—figure supplements 1, 2**. DOI: [10.7554/eLife.06034.006](https://doi.org/10.7554/eLife.06034.006)

The following figure supplements are available for figure 2:

**Figure supplement 1.** Co-immunoprecipitation (co-ip) between AbdA and TFs selected from the set used for BiFC in the *Drosophila* embryo. DOI: [10.7554/eLife.06034.007](https://doi.org/10.7554/eLife.06034.007)

Figure 2. continued on next page

Figure 2. Continued

**Figure supplement 2.** BiFC between mesodermal TFs and AbdA in the mesoderm.

DOI: [10.7554/eLife.06034.008](https://doi.org/10.7554/eLife.06034.008)

Among the 35 TFs tested as fusion constructs, 31 led to BiFC signals with AbdA, including TFIIBeta (**Figure 2C**). These fluorescent signals display homogenous or punctuate distribution within the nucleus, depending on the TF considered (**Figure 2C**). BiFC was negative with Kn as expected, given that no competition was previously observed with this TF. BiFC was however also negative with Krüppel (Kr), Lameduck (Lmd), and Pangolin (Pan), although these three TFs were positive competitors. This discrepancy suggests that the previously observed competition could result from the formation of specific cold complexes with the Exd and not with the AbdA fusion protein. Alternatively, the fusion topologies could forbid the interaction hence BiFC between the three TFs and AbdA. Indeed, the negative influence of fusion topologies on protein–protein interactions was previously described and is hardly predictable (**Hudry et al., 2011**).

To assess the potential influence of fusion topologies on AbdA-TF interactions, we performed co-immunoprecipitation (co-ip) experiments, using an anti-HA antibody recognising a HA-tagged form of AbdA. We reasoned that a small HA epitope (8 residues long) should be more neutral than the Venus fragment (80 residues long) for the interaction with the TF. Practically, the fusion TF was co-expressed with the AbdA-HA construct in S2 cells, and its presence was verified with an anti-GFP antibody recognizing the Venus fragment (see 'Materials and methods'). Experimental parameters were established with the control Exd cofactor. All tested BiFC-positive TFs were found by co-ip (**Supplementary file 3**), highlighting that the S2 cell environment is appropriate for revealing interactions with tissue-specific TFs. Thus, observations from BiFC could be reproduced by co-ip, as previously noticed (**Lee et al., 2011**). Co-ip was also performed with the three positive competitors that did not produce BiFC with AbdA (Kr, Lmd, Pan). We observed that these three TFs could be immuno-precipitated with AbdA (**Supplementary file 3**). This was not the case for Kn, which was negative both in the competition and BiFC tests (**Figure 2—figure supplement 1**). We conclude that inappropriate fusion topologies are likely to be responsible for the absence of BiFC between AbdA and Lmd, Kr or Pan in the epidermis.

Several of the tested TFs are not expressed in the epidermis (**Supplementary file 2**). We thus wondered whether the interaction with those TFs could also be reproduced in their endogenous expression tissue. To this end, we repeated BiFC analyses in the mesoderm, using mesodermal TFs that are not expressed in the epidermis, including Lmd and Kn (**Supplementary file 3**). We observed that TFs interacting with AbdA in the epidermis were also positive in the mesoderm (**Figure 2—figure supplement 2**). Fluorescent signals were however generally weaker than in the epidermis, probably due to the competition by the endogenous gene products. Surprisingly, weak BiFC signals could also be observed with Kn, while Lmd remained negative in the mesoderm (**Figure 2—figure supplement 2**). We concluded that Kn could interact with AbdA but that this interaction is more sensitive to the cell environment in order to occur.

In summary, although our approach was performed with a limited set of TFs (around 12% of all *Drosophila* TFs), it revealed an unexpectedly high number of new binding partners of AbdA. However, whether and how these binding partners could be used in the context of endogenous regulatory activities of AbdA remains to be investigated. This result still illustrates the strong propensity of AbdA to establish interactions with diverse TFs in vivo. In the following, we considered the set of 35 TFs as sufficiently representative for addressing the issue of the molecular mechanisms underlying Hox interactome specificity in vivo.

### Regions outside the homeodomain influence dependent and independent DNA-binding interactions of AbdA with TFs

Our experimental parameters allow quantifying subtle changes in fluorescent signal intensities in the whole *Drosophila* embryo ('Materials and methods' and [**Hudry et al., 2011**]). These variations in the fluorescence intensity can be correlated to differences in interaction affinity. Indeed, high affinity partners lead to fast accumulation, hence strong BiFC signals, and vice versa (**Hudry et al., 2011**). Here, levels of BiFC were used to measure the effects of AbdA mutations on the interaction potential



with each of the 35 TFs. We first focused on the HD, which is responsible for the DNA-binding and the most conserved part of Hox proteins. The HD of Hox proteins is also described to interact with different types of cofactors (Merabet et al., 2009), suggesting it could be involved in several of the observed interactions. More precisely, we asked (i) whether AbdA-TF interactions could depend on the DNA-binding activity of the HD, and (ii) whether the HD could be sufficient for AbdA-TF interactions. BiFC was performed with two corresponding mutant forms of AbdA: one carrying the N51A mutation in the HD, which abolishes the DNA-binding activity of full length AbdA (AbdA51) (Hudry et al., 2011), the other resulting in a truncated version that contains only the HD (AbdAHD) (Boube et al., 2014). The two fusion constructs were inserted on the same genomic locus and expressed at similar levels in the embryo as the wild type AbdA fusion protein ('Materials and methods' and [Hudry et al., 2011; Boube et al., 2014]). BiFC was measured in the epidermis and considered as affected when the fluorescent signal was equal or lower than 50% of the fluorescent level normally obtained with wild type AbdA (see 'Materials and methods' for quantification details).

Among the 31 BiFC-positive interactions, 18 were strongly affected or lost with AbdA51 (compared Figure 3A,B, and Figure 3—figure supplement 1). Still, a significant proportion of the interactions (for 13 TFs) was retained, two of which were even stronger than with wild type AbdA (leading to intensities higher than 120% of the wild type fluorescent signal). Thus, the two corresponding TFs (Pannier and Slouch) preferentially interact with a form of AbdA that is unable to bind DNA. Effects were more drastic with the minimal AbdAHD construct, which kept only eight interactions among the 31 positive TFs (compare Figure 3A,C, and Figure 3—figure supplement 2). One of these interactions (with Serpent, Srp) was also stronger compared to wild type AbdA.

The observation of stronger interactions in some cases prompted us to analyse whether previously negative TFs with AbdA (Kn, Kr, Lmd, and Pan) could produce BiFC with either of the two AbdA mutant forms. We found that Kr could indeed lead to BiFC signals when using the AbdAHD version (Figure 3C and Figure 3—figure supplement 2). Thus, the interaction was strong enough in this particular case to be visualized despite unfavourable fusion topologies, illustrating that HD-surrounding region(s) could inhibit the interaction potential of the HD in vivo.

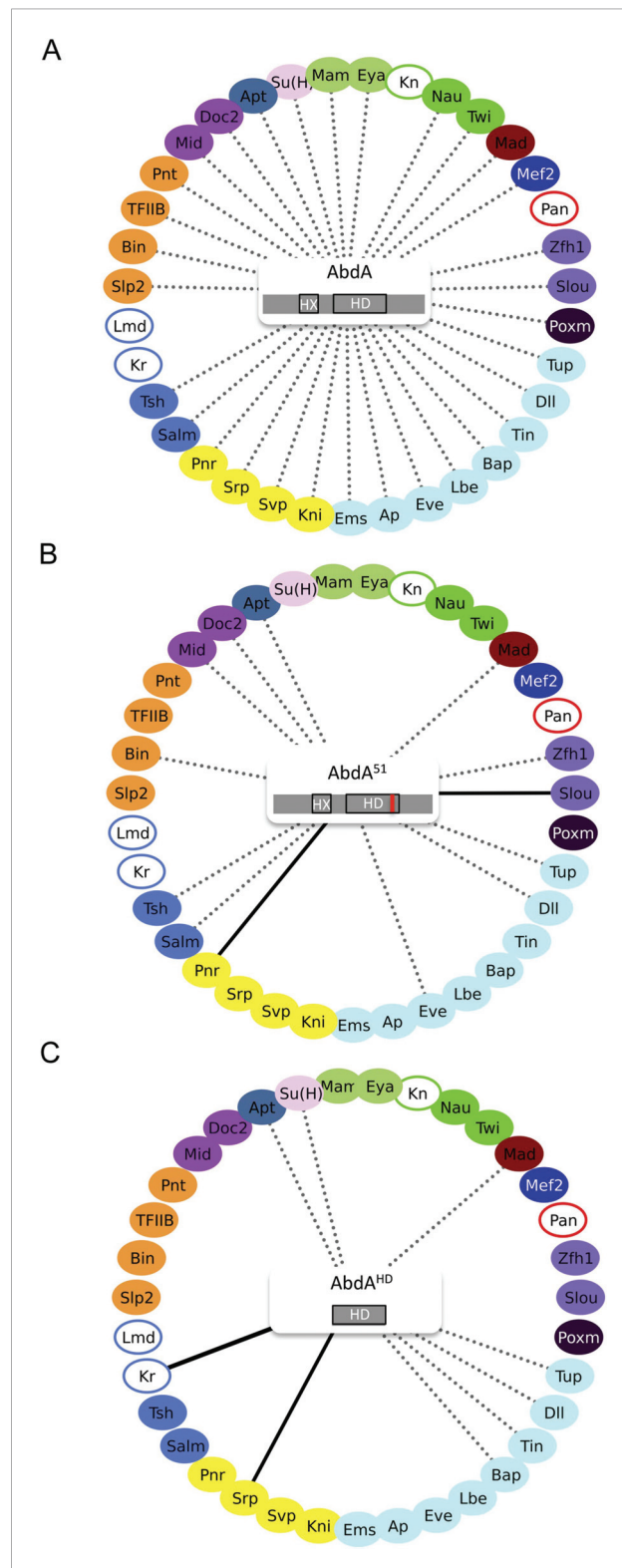
Together these results show that the DNA-binding of AbdA is not systematically required for recruiting the TFs. The HD itself is also not sufficient in most cases, suggesting that surrounding protein region(s) are important for recruiting TFs. Interestingly, these HD-surrounding region(s) can also have an inhibitory role since their absence allows the formation of stronger interactions between the HD and two of the tested TFs. Given that HD-surrounding regions are less conserved in Hox proteins in general, we then asked whether some of the revealed interactions could also be found with other *Drosophila* Hox proteins.

### Specific combinations of TFs underlie the formation of different Hox interactomes in vivo

With the exception of Ubx, *Drosophila* Hox proteins have few redundant functions with AbdA, as reflected at the protein sequence and embryonic expression levels (Figure 4A). We thus wondered whether common vs specific features between Hox proteins could be found in their respective interactomes. To this end, we repeated BiFC between the 35 TFs and four other Hox proteins (see 'Materials and methods'), namely Sex combs reduced (Scr), Antennapedia (Antp), Ultrabithorax (Ubx), and AbdominalB (AbdB).

Overall, we observed an unexpected high proportion of positive interactions with the four additional Hox proteins: from 22 and 18 with Scr and Antp, to 21 and 26 with Ubx and AbdB, respectively (Figure 4B and Figure 4—figure supplements 5–9, Figure 5—figure supplement 1). Because of this high proportion, the majority of interactions are common to several Hox proteins. Despite this number of common interactions, each Hox interactome contains a specific combination of binding partners. Interestingly, the hierarchical clustering of Hox interactomes (see 'Materials and methods') does not reflect the protein sequence similarity between Hox proteins (Figure 4A–B). For example, the interactomes of AbdA and AbdB appear closely similar although AbdB is much more divergent from AbdA than the other Hox proteins.

Taken together these observations show that Hox proteins are able to bind to common and different types of TFs. Still, each Hox interactome contains a distinct combination of TFs, demonstrating a certain



**Figure 3.** Role of the homeodomain (HD) in the AbdA interactome. **(A)** Interactome with wild type AbdA. **(B)** Interactome with the DNA-binding deficient form of AbdA (mutated in the residue 51 of the HD, as illustrated with the red bar). See also **Figure 3—figure supplement 1**. **(C)** Interactome with the HD of AbdA. See also **Figure 3—figure supplement 2**. Each interactome is represented with the 35 TFs. The colour code for TFs **Figure 3. continued on next page**

Figure 3. Continued

corresponds to their type of DNA-binding domain, as shown in the **Figure 2B**. TFs that are not colour-filled correspond to TFs that do not interact with the wild type Hox protein. Those TFs are not connected to the Hox protein. Dotted lines indicate TFs that do interact with the wild type Hox protein. Interactions with Hox variants are depicted as the following: dotted lines indicate unaffected interactions (in between 51% and 119% of the wild type interaction); solid black lines indicate stronger (equal or superior to 120% of the wild type interaction) or novel interactions (with a non-colour-filled TF); absence of the dotted line with a colour-filled TF indicates a partial (equal or below to 50% of the wild type interaction) or complete loss of the interaction. Each Hox variant is schematized in the centre of the interactome. HD: Homeodomain. HX: Hexapeptide. See also 'Materials and methods'.

DOI: [10.7554/eLife.06034.009](https://doi.org/10.7554/eLife.06034.009)

The following figure supplements are available for figure 3:

**Figure supplement 1.** BiFC between the 35 TFs and the homeodomain (HD)-mutated form of AbdA.

DOI: [10.7554/eLife.06034.010](https://doi.org/10.7554/eLife.06034.010)

**Figure supplement 2.** BiFC between the 35 TFs and the homeodomain (HD) of AbdA.

DOI: [10.7554/eLife.06034.011](https://doi.org/10.7554/eLife.06034.011)

level of specificity. Moreover, the hierarchical clustering of Hox interactomes suggests that the recruitment of common cofactors would not obligatorily rely on the same Hox protein interface.

### The common Hexapeptide (HX) motif is differently used within each Hox interactome in vivo

BiFC analyses with AbdA showed that HD-surrounding regions are important for most of the revealed interactions. In addition to the HD, the HX motif represents the second generic signature of Hox proteins. We thus asked whether this common Hox SLiM could be important for recruiting common TFs in vivo. Its role was assessed within each Hox interactome by repeating BiFC between the 35 TFs and the corresponding HX-mutated Hox proteins (see 'Materials and methods' and [Hudry et al., 2012]).

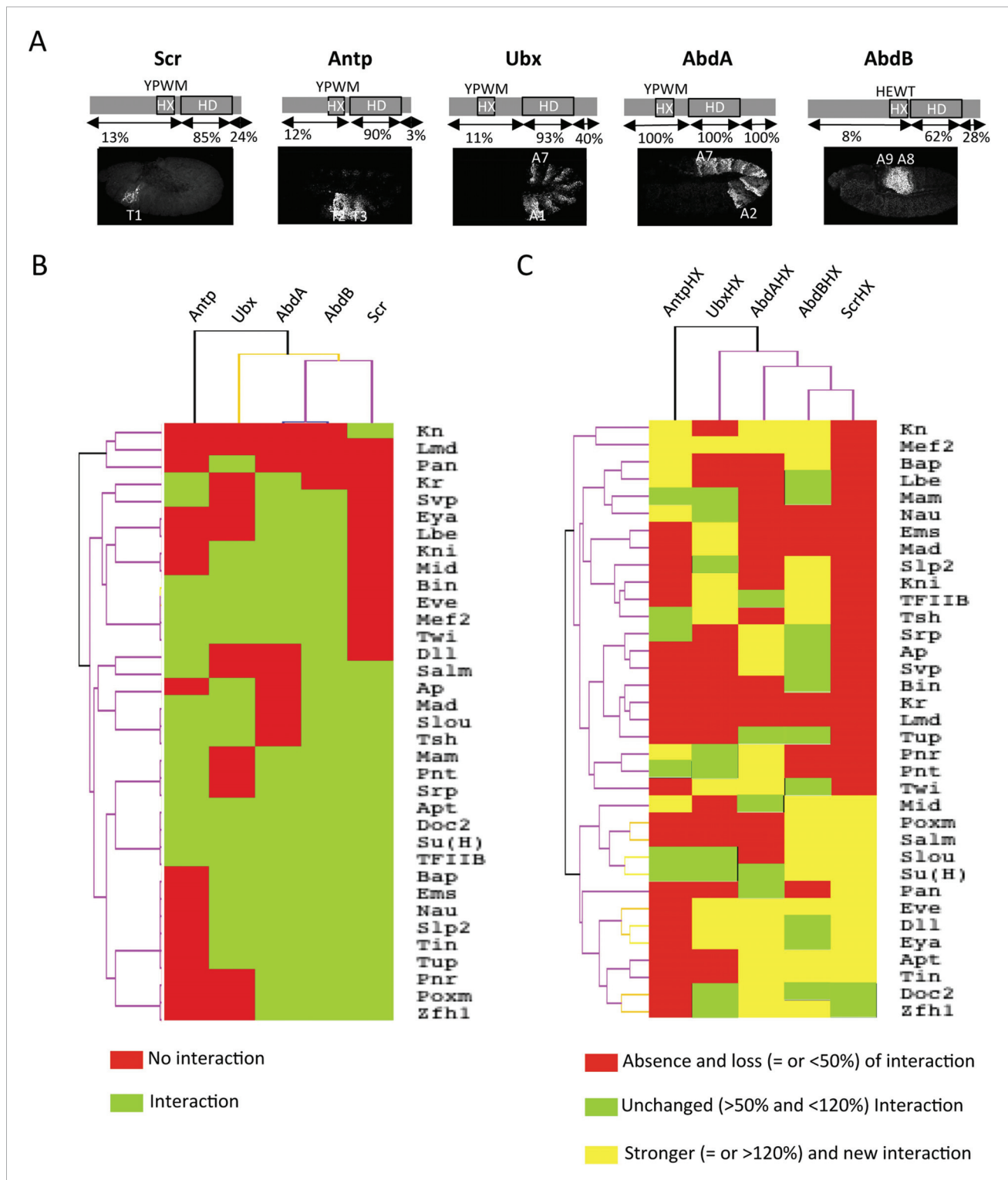
Heatmap representation shows that the HX mutation led to a complete reorganisation of Hox interactomes when compared to the wild type Hox proteins (compare **Figure 4B,C** and **Figure 4—figure supplements 5–9**). More precisely, the HX mutation affects the majority of Hox-TF interactions. Surprisingly, this mutation leads not only to a loss, but also to a gain of the Hox interaction potential, with the appearance of stronger or new interactions (highlighted in yellow in **Figure 4C**). The balance between gain and loss was different depending on the Hox protein: Antp and AbdB were more sensitive to a loss while the reverse was observed with Ubx and AbdA. The HX of Scr was equally responsible for a gain or loss of interactions (**Figure 5** and **Figure 5—figure supplement 1**). Overall, the HX mutation leads more often to a gain than a loss of the Hox interaction potential. Interestingly, the positive or negative influence of the HX on Hox-TF interactions is not identical for each TF. For example, the HX mutation has no effect on AbdA-Twist (Twi) interaction, while it leads to a loss of interaction between Antp and AbdB and the same TF, and to a stronger and novel interaction with Ubx and Scr, respectively (**Figure 5** and **Figure 5—figure supplement 1**). Thus, the role of the HX appears dictated by the Hox protein to which it belongs and not by the interacting TF, therefore reinforcing the fact that HX neighbourhood is important for controlling its interaction properties (Merabet and Hudry, 2011).

In summary, our results show that a short conserved motif, common to all Hox proteins, is specifically used both for promoting and limiting their interaction potential with TFs.

### Common vs paralog-specific Hox protein motifs display different interaction properties with TFs

Since we observed that the HX could inhibit the interaction potential of Hox proteins, we asked whether this property could also be found with other more specific Hox SLiMs. We focused on Ubx and AbdA, in which the HX mutation led to the highest number of gained interactions among the tested Hox proteins. Interestingly, Ubx and AbdA share the UbdA motif, which is conserved in most protostome lineages (Balavoine et al., 2002). This motif was shown to be important for recruiting the Exd cofactor in Ubx (Merabet et al., 2007; Foos et al., 2015) and for tissue-specific activities of AbdA in vivo (Merabet et al., 2011).

As previously done with the HX motif, the role of the UbdA motif was assessed by performing BiFC with UbdA-mutated forms of Ubx and AbdA (**Figure 6A** and [Hudry et al., 2012]). We focused on the



**Figure 4.** Comparison between wild type and Hexapeptide(HX)-mutated Hox interactomes. **(A)** Embryonic expression profile (grey) and protein sequence identity of each of the five *Drosophila* Hox proteins under study. The percentage of sequence identity is given in comparison to AbdA. The conserved core sequence of the HX is also given for each Hox protein. **(B)** Heatmap showing the organisation of wild type Hox interactomes with the 35 TFs. See also **Figure 4—figure supplements 1–4**. **(C)** Heatmap showing the organisation of HX-mutated Hox interactomes with the 35 TFs. See also **Figure 4—figure supplements 5–9**. Interactions are symbolized by a colour code, as indicated. Note that the yellow colour, which corresponds to a gain of the interaction

Figure 4. continued on next page

Figure 4. Continued

potential, appears with the HX mutation in all Hox proteins. Dendrogram branches are coloured according to their bootstrap score: black 100%, grey 90–100%, blue 80–90%, green 70–80%, yellow 60–70%, orange 50–60%, pink 0.1–50%, red 0% support respectively.

DOI: [10.7554/eLife.06034.012](https://doi.org/10.7554/eLife.06034.012)

The following figure supplements are available for figure 4:

**Figure supplement 1.** BiFC between the 35 TFs and Sex combs reduced (Scr).

DOI: [10.7554/eLife.06034.013](https://doi.org/10.7554/eLife.06034.013)

**Figure supplement 2.** BiFC between the 35 TFs and Antennapedia (Antp).

DOI: [10.7554/eLife.06034.014](https://doi.org/10.7554/eLife.06034.014)

**Figure supplement 3.** BiFC between the 35 TFs and Ultrabithorax (Ubx).

DOI: [10.7554/eLife.06034.015](https://doi.org/10.7554/eLife.06034.015)

**Figure supplement 4.** BiFC between the 35 TFs and AbdominalB (AbdB).

DOI: [10.7554/eLife.06034.016](https://doi.org/10.7554/eLife.06034.016)

**Figure supplement 5.** BiFC between the 35 TFs and hexapeptide (HX)-mutated Scr.

DOI: [10.7554/eLife.06034.017](https://doi.org/10.7554/eLife.06034.017)

**Figure supplement 6.** BiFC between the 35 TFs and hexapeptide (HX)-mutated Antp.

DOI: [10.7554/eLife.06034.018](https://doi.org/10.7554/eLife.06034.018)

**Figure supplement 7.** BiFC between the 35 TFs and hexapeptide (HX)-mutated Ubx.

DOI: [10.7554/eLife.06034.019](https://doi.org/10.7554/eLife.06034.019)

**Figure supplement 8.** BiFC between the 35 TFs and hexapeptide (HX)-mutated AbdA.

DOI: [10.7554/eLife.06034.020](https://doi.org/10.7554/eLife.06034.020)

**Figure supplement 9.** BiFC between the 35 TFs and hexapeptide (HX)-mutated AbdB.

DOI: [10.7554/eLife.06034.021](https://doi.org/10.7554/eLife.06034.021)

20 common binding partners of Ubx and AbdA to potentially reveal a common usage mode of the UbdA motif between the two Hox proteins.

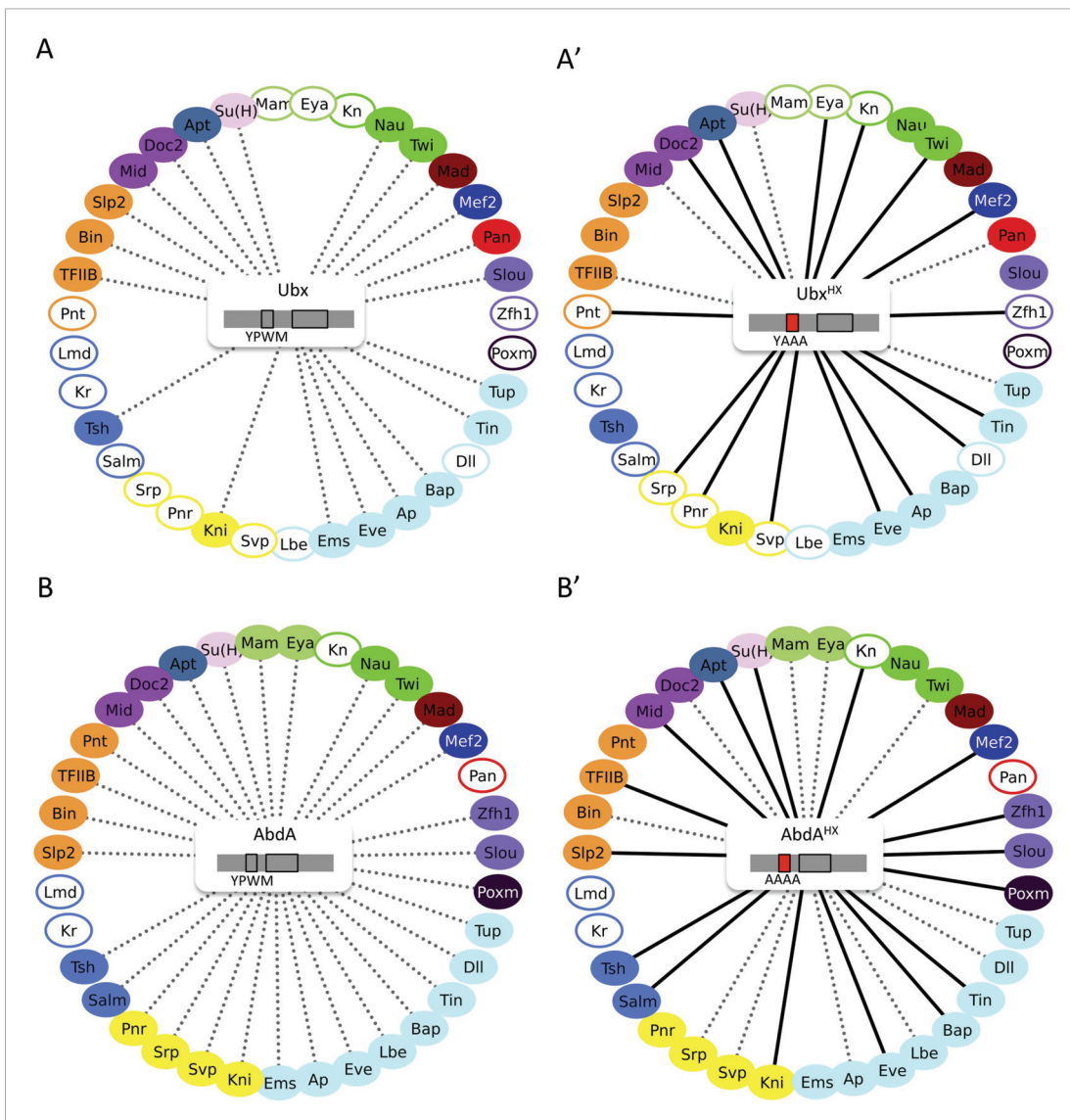
Results show that the UbdA mutation affects the majority of interactions in both Hox proteins, as previously noticed for the HX mutation (**Figure 6B–C** and **Figure 6—figure supplements 1, 2**). Effects can also be categorized as a gain or a loss of the interaction potential. However, the UbdA motif has a more pronounced tendency to be required for the interaction rather than for inhibiting it (**Figure 6B–C**). For example, the UbdA mutation leads to 12 losses and 5 gains among the 20 tested TFs with AbdA. In comparison, 3 losses and 12 gains were induced upon the HX mutation for the same set of interactions. In addition, the HX and UbdA mutations have distinct effects for the majority of interactions established by Ubx or AbdA (**Figure 6B–C**). This result highlights that the two Hox proteins do not use the HX and UbdA motifs similarly. Interestingly, the UbdA mutation often leads to similar effects in Ubx and AbdA compared to the HX mutation (**Figure 6D–E**). More precisely, the UbdA mutation affects the majority (13/20) of the tested TFs in a similar way, which is not the case for the HX mutation (8/20).

In conclusion, the UbdA motif displays preferential Ubx/AbdA-specific interaction properties compared to the HX motif. These interaction properties rely in part on inhibitory activities, highlighting that the negative influence on PPIs is not a specific property of the HX motif. To gain further insights into the molecular property of Hox SLiMs, we next examined whether the regulatory activity of the HX and UbdA motifs could change depending on the embryonic tissue considered.

### The HX and UbdA motifs have tissue-specific interaction properties

Previous work showed that the HX and UbdA motifs have tissue-specific functions in AbdA (**Merabet et al., 2011**), suggesting that their interaction properties with TFs would not be identical in different embryonic tissues. To test this hypothesis, we analysed the interaction potential of HX- or UbdA-mutated AbdA in the mesoderm and nervous system, focusing on TFs that are normally expressed in one and/or both tissues. We also used a set of TFs that were all BiFC-positive with AbdA in the epidermis and all but one sensitive to the HX or UbdA mutation in this tissue (**Figure 7A**).

We observed that the HX and UbdA motifs were less often required in the mesoderm since their mutation affected fewer interactions than in the epidermis (compare **Figure 7A,B**, and **Figure 7—figure supplement 1**). Still, affected TFs again correspond to a loss or a gain of the Hox interaction potential. The inhibitory activity of the HX and UbdA motifs was more pronounced in the



**Figure 5.** The HX mutation increases the interaction potential of Hox proteins with TFs in vivo: example in Ubx and AbdA. **(A–A')** Comparison between wild type and HX-mutated interactomes of Ubx. **(B–B')** Comparison between wild type and HX-mutated interactomes of AbdA. The HX mutation led more frequently to stronger or new interactions than to interaction loses in these two Hox proteins. Colour code and representation are as in **Figure 3**. The HX mutation is indicated and highlighted in red. See also **Figure 5—figure supplement 1**.  
DOI: [10.7554/eLife.06034.022](https://doi.org/10.7554/eLife.06034.022)

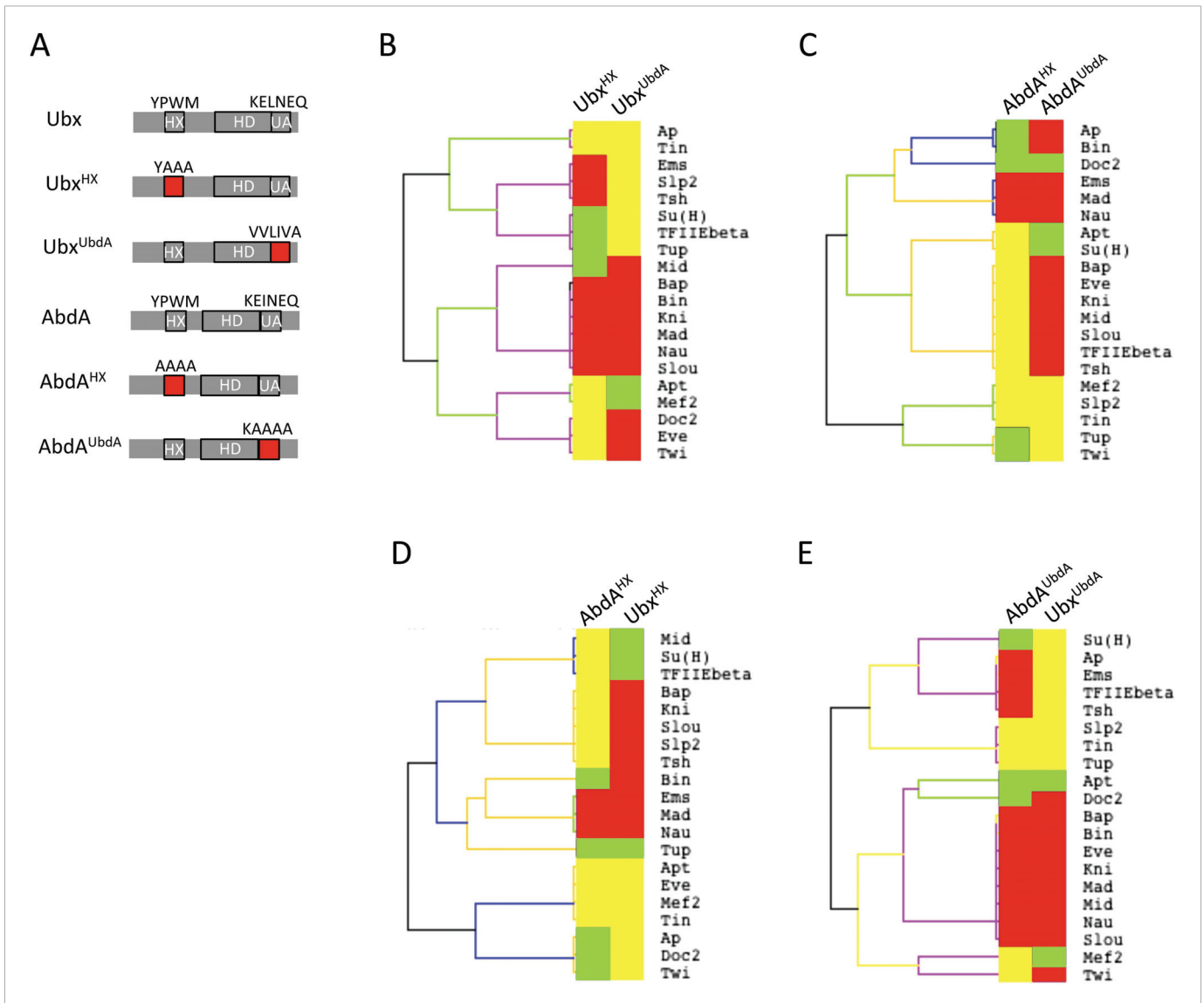
The following figure supplement is available for figure 5:

**Figure supplement 1.** The HX mutation increases the interaction potential of Hox proteins with TFs in vivo: example in Scr, Antp, and AbdB.

DOI: [10.7554/eLife.06034.023](https://doi.org/10.7554/eLife.06034.023)

nervous system, since their mutation led only to stronger interactions with TFs (**Figure 7C** and **Figure 7—figure supplement 2**). This effect was more obvious with the HX mutation, which led to a gain of interaction for all but one TF in this tissue (**Figure 7C**).

In total, all tested TFs were not similarly affected by the HX or UbdA mutation in the three different tissues. Thus, the two motifs are differently used depending on the Hox protein and tissue considered. This specific usage mode is based both on the positive and negative control of PPIs. Given the evolutionary conserved roles of Hox proteins in general, we then tested whether these novel facets of SLiM activity could also be found in Hox proteins from other animal species.



**Figure 6.** Usage mode of the HX and UbdA motifs in Ubx and AbdA proteins. **(A)** Schematic representation of wild type and HX- or UbdA-mutated Ubx and AbdA proteins. **(B)** Heatmap comparing interaction properties of HX- and UbdA-mutated Ubx proteins with a set of 20 TFs. These TFs are common to Ubx and AbdA for BiFC. See also **Figure 6—figure supplement 1**. **(C)** Heatmap comparing interaction properties of HX- and UbdA-mutated AbdA proteins with the same set of TFs. See also **Figure 6—figure supplement 2**. Note that the HX and UbdA mutations have distinct or opposite effects for most of the interactions in Ubx and AbdA. **(D)** Heatmap comparing interaction properties of HX-mutated Ubx and AbdA proteins with the 20 common TFs. **(E)** Heatmap comparing interaction properties of UbdA-mutated Ubx and AbdA proteins with the 20 common TFs. Note that a higher proportion of TFs is similarly affected by the UbdA mutation in Ubx and AbdA when compared to the HX mutation. Colour code is as in **Figure 4**.

DOI: [10.7554/eLife.06034.024](https://doi.org/10.7554/eLife.06034.024)

The following figure supplements are available for figure 6:

**Figure supplement 1.** BiFC with UbdA-mutated Ubx in the epidermis.

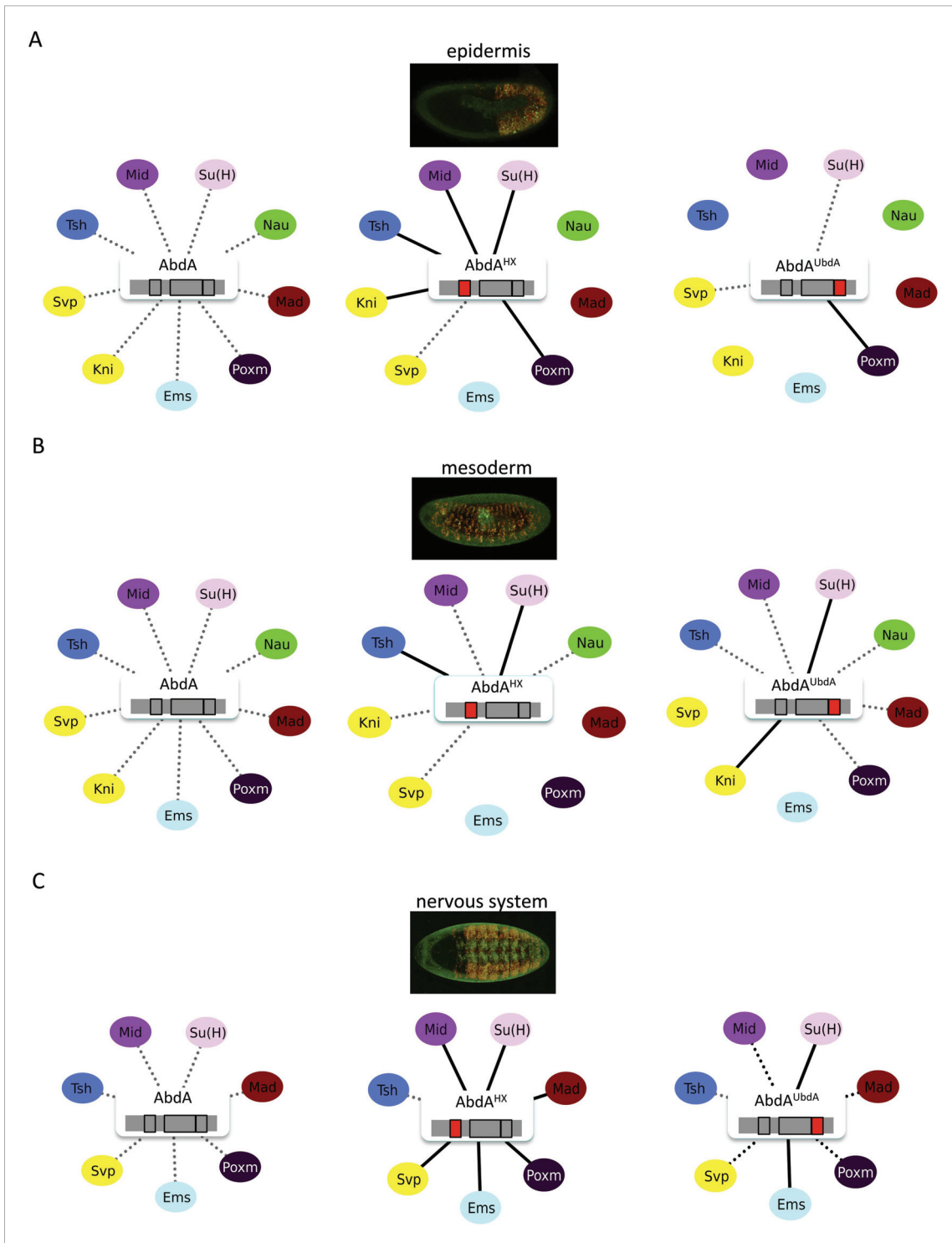
DOI: [10.7554/eLife.06034.025](https://doi.org/10.7554/eLife.06034.025)

**Figure supplement 2.** BiFC with UbdA-mutated AbdA in the epidermis.

DOI: [10.7554/eLife.06034.026](https://doi.org/10.7554/eLife.06034.026)

## Inhibition of PPIs is an evolutionary conserved feature of the HX motif

Our work revealed that the HX could specify *Drosophila* Hox interactomes in part by limiting the interaction potential in a context-dependent manner. Here, we ask whether the inhibitory role of the HX on PPIs could constitute an evolutionary conserved property of Hox proteins. To this end, we used



**Figure 7.** The HX and UbdA motifs of AbdA have different interaction properties in different embryonic tissues. **(A)** Interaction properties of wild type and HX- or UbdA-mutated AbdA in the epidermis. **(B)** Interaction properties of wild type and HX- or UbdA-mutated AbdA in the mesoderm. See also **Figure 7—figure supplement 1**. **(C)** Interaction properties of wild type and HX- or UbdA-mutated AbdA in the nervous system. See **Figure 7. continued on next page**



Figure 7. Continued

also **Figure 7—figure supplement 2**. Picture of an embryo making BiFC (green) and expressing the dsRed fluorescent protein under the control of the Gal4 driver illustrates the tissue of interest in each condition. Interactomes are represented as in **Figure 3**.

DOI: [10.7554/eLife.06034.027](https://doi.org/10.7554/eLife.06034.027)

The following figure supplements are available for figure 7:

**Figure supplement 1.** BiFC with wild type, HX- or UbdA-mutated AbdA in the mesoderm.

DOI: [10.7554/eLife.06034.028](https://doi.org/10.7554/eLife.06034.028)

**Figure supplement 2.** BiFC with wild type, HX- or UbdA-mutated AbdA in the nervous system.

DOI: [10.7554/eLife.06034.029](https://doi.org/10.7554/eLife.06034.029)

the mouse HoxB8 and *Nematostella* HoxE proteins as extreme representatives (**Figure 8A**). HoxB8 is a central Hox protein containing a typical HX motif. HoxE was recently shown to display central-like molecular properties (*Hudry et al., 2014*), although it contains a posterior-like derived HX motif (**Figure 8A**). Overall, HoxB8 and HoxE show little sequence identity with the *Drosophila* AbdA protein outside the HX and HD (**Figure 8A**).

We first addressed whether the HoxB8 and HoxE could interact with the 35 *Drosophila* TFs, as previously done with *Drosophila* Hox proteins (see also 'Materials and methods'). Results show that the mouse and cnidarian Hox proteins interact with a surprisingly large number of TFs, respectively 27 and 22 (**Figure 8B–C** and **Figure 8—figure supplements 1, 2**). Still, each Hox protein interacts with a different set of TFs, underlining the existence of preferential/specific interaction properties when considering the whole interactome.

To directly assess whether the HX could act as an inhibitory motif, we analysed the interaction properties of HX-mutated HoxB8 and HoxE proteins, focusing on TFs that were negative with the corresponding wild type Hox proteins. We observed that more than half of the tested TFs became positive with HX-mutated proteins in both cases (**Figure 8—figure supplements 3, 4**). Thus, the HX is also an inhibitory interaction motif in HoxB8 and HoxE.

To further explore how the HX motif could inhibit PPIs, we considered the HoxE protein, which has a simple organisation in terms of secondary structures. Basically, this protein contains a long intrinsically disordered N-terminal region followed by the ordered HD (**Figure 9A**). Still, this protein establishes a number of common interactions with HoxB8 and AbdA (**Figure 8—figure supplement 5**). Since the HD is unlikely to be sufficient for several of those interactions (as deduced from AbdA: **Figure 3C**), we decided to test the N-terminal region of HoxE. We thus generated fly lines carrying a short HoxE variant, called Nter-HoxE, which corresponds to the residues 1 to 54 and does not contain the HX motif (see 'Materials and methods'). BiFC with the 35 *Drosophila* TFs shows that only five interactions are lost when using this short variant of HoxE (**Figure 9B** and **Figure 9—figure supplement 1**). In contrast, this fragment interacts more strongly or establishes new interactions with seven TFs while 15 other interactions remained unaffected. In total, Nter-HoxE establishes as many interactions as the full length HoxE.

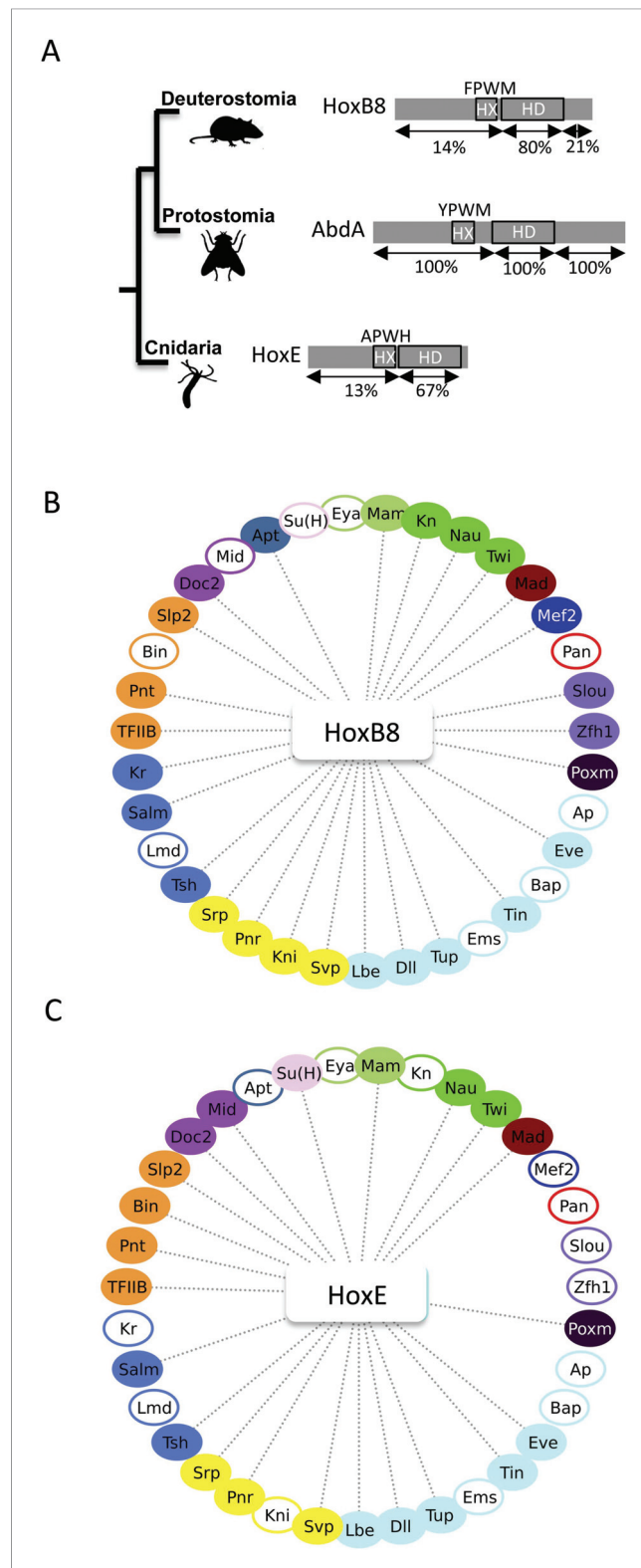
Together these results show that the long disordered region of HoxE is involved in a number of the heterologous interactions observed with *Drosophila* TFs. The observation of seven ectopic interactions also underlines that the interaction potential of the disordered region is tightly controlled in the context of the full-length protein. We propose that SLiMs such as the HX motif are important mediators of this control.

## Discussion

Our work revealed a striking propensity of Hox proteins to interact with different types of TFs. In this context, SLiMs such as the HX are not only important for promoting but also for limiting this strong interaction potential. Effects could vary depending on the Hox protein and the tissue considered, highlighting the adaptability of SLiMs to different environments. Thus, the constraining activity of SLiMs on PPIs is an essential attribute of Hox interactome specificity in vivo.

## A novel set of Hox interacting partners in the fly embryo as revealed by BiFC

Apart from the PBC class, very little is known about the TFs that could help Hox proteins to elicit specific developmental programs in the embryo. As a consequence, the interactome underlying Hox



**Figure 8.** The *Nematostella* HoxE and mouse HoxB8 proteins interact with several *Drosophila* TFs. **(A)** Schematic representations of the Hox proteins and the corresponding animal phylogeny. The percentage of sequence identity is given in comparison to AbdA. **(B)** Interactome between mouse HoxB8 and the 35 *Drosophila* TFs. **(C)** Interactome between mouse HoxE and the 35 *Drosophila* TFs. *Figure 8. continued on next page*

Figure 8. Continued

between *Nematostella* HoxE and the 35 *Drosophila* TFs. Colour code and representation are as in **Figure 3**.

See also **Figure 8—figure supplements 1–5**.

DOI: [10.7554/eLife.06034.030](https://doi.org/10.7554/eLife.06034.030)

The following figure supplements are available for figure 8:

**Figure supplement 1.** BiFC between *Drosophila* TFs and the mouse HoxB8 protein.

DOI: [10.7554/eLife.06034.031](https://doi.org/10.7554/eLife.06034.031)

**Figure supplement 2.** BiFC between *Drosophila* TFs and the *Nematostella* HoxE protein.

DOI: [10.7554/eLife.06034.032](https://doi.org/10.7554/eLife.06034.032)

**Figure supplement 3.** BiFC between *Drosophila* TFs and HX-mutated HoxB8.

DOI: [10.7554/eLife.06034.033](https://doi.org/10.7554/eLife.06034.033)

**Figure supplement 4.** BiFC between *Drosophila* TFs and HX-mutated HoxE.

DOI: [10.7554/eLife.06034.034](https://doi.org/10.7554/eLife.06034.034)

**Figure supplement 5.** *Drosophila* AbdA, mouse HoxB8 and *Nematostella* HoxE interact with several common TFs in vivo.

DOI: [10.7554/eLife.06034.035](https://doi.org/10.7554/eLife.06034.035)

---

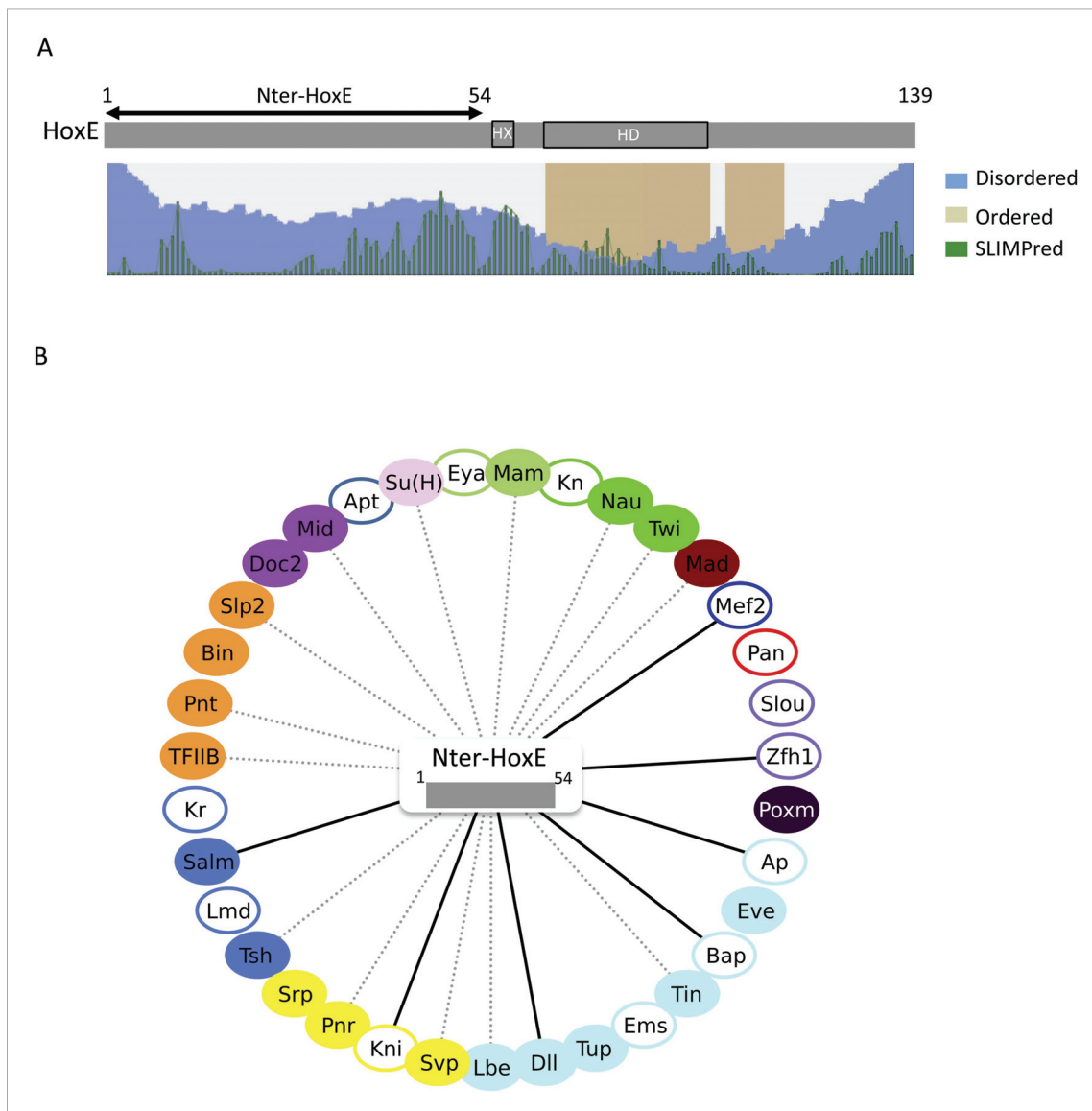
embryonic functions remains largely elusive. This lack of knowledge is explained by the difficulty of identifying the transcriptional partners that could participate in each context-specific activity of Hox proteins. We hypothesize that many of those interactions are too weak and/or too dynamic to be efficiently trapped through classic high throughput approaches. In support of this, a yeast two-hybrid screen with the *Drosophila* Ubx protein led to the characterization of less than 15 TFs as interacting partners (**Bondos et al., 2006**), strongly contrasting with the numerous functions ensured by this Hox protein during embryogenesis. A similar approach with the mouse HoxA1 protein also led to the characterization of less than 20 TFs (**Lambert et al., 2012**).

Our approach relied on the sequential analysis of tandem interactions between Hox proteins and individual TFs. The interaction screen was performed in two complementary steps with a set of TFs covering different DNA-binding families and displaying various expression profiles during *Drosophila* embryogenesis. This set of TFs is expected to be representative of the diverse transcriptional regulatory activities of Hox proteins in vivo.

The competition experiment was performed in the epidermis, even for TFs that are not endogenously expressed in this tissue. Despite this limitation, we found a high proportion (33/80) of positive competitive events, which confirmed the sensitivity of BiFC for this type of approach. A similar strategy was reported in cell culture for identifying drug molecules that could affect the assembly or the localisation of a specific protein complex (**Morell et al., 2008**). Thus, competitive BiFC could certainly be applied more generally in the future for selecting any kind of new interacting molecules upon the screening of subtle variations in fluorescent reporter signals.

BiFC then showed that all but three competitive TFs could interact with AbdA, making a total of 31 TFs as new Hox interacting partners. In comparison, only seven TFs were so far described to interact with AbdA (**Merabet and Dard, 2014**). These results were reproduced in different tissues of the embryo or corroborated by co-ip experiments in S2 cells, illustrating that Hox-TF interactions could occur in different cell contexts upon co-expression. Importantly, the specificity of each interaction was supported by the loss of fluorescent signals when using mutated or truncated Hox variants, and by the observation of typical nuclear interaction profiles with different TFs.

It is important to stress that all these results revealed an interaction potential between Hox proteins and TFs. Whether and how these interactions could be used in the context of the endogenous gene products is an open question. For example, some of the interactions revealed with the fusion Scr protein are unlikely to occur with the endogenous Scr product, which displays a quite restrained expression profile in the embryo compared to the other tested Hox proteins (**Supplementary file 3** and [**Hammonds et al., 2013**]). In addition, the fact that TF-encoding genes were not expressed under the control of their endogenous promoter forbids assessing the role of tissue-specific expression levels in Hox interactome properties. In this context, the recent advent of genetic tools in *Drosophila*, including Mimic elements (**Gnerer et al., 2015**) and the CRISPR/Cas9 system (**Bassett and Liu, 2014**), could certainly add to the functional relevance of BiFC observations in the future. Nevertheless, the high proportion of positive events among a starting set of 80 TFs strongly suggests that AbdA, and



**Figure 9.** The intrinsically disordered region of HoxE establishes a number of interactions with *Drosophila* TFs. (A) Scheme of full length *Nematostella* HoxE with its predicted SLiMs (green bars), and disordered (blue waves) or ordered (brown blocks) regions. Adapted from iupred (<http://iupred.enzim.hu/>). The N-terminal disordered region used for BiFC is indicated (Nter-HoxE). (B) Interactome between Nter-HoxE and the 35 *Drosophila* TFs. Colour code and representation are as in **Figure 3**. See also **Figure 9—figure supplement 1**.

DOI: [10.7554/eLife.06034.036](https://doi.org/10.7554/eLife.06034.036)

The following figure supplement is available for figure 9:

**Figure supplement 1.** BiFC between *Drosophila* TFs and the N-terminal (1-54) fragment of the *Nematostella* HoxE protein.

DOI: [10.7554/eLife.06034.037](https://doi.org/10.7554/eLife.06034.037)

Hox proteins in general, have a strong potential to interact with a number of different TFs in vivo. This assumption could be verified by using a library of 600 TFs compatible for BiFC in *Drosophila* (Bischof et al., 2013).

### Specific vs common cofactors among Hox interactomes

Hox proteins play numerous functions in all embryonic germ layers. These functions can be highly specific (Brodu et al., 2002; Li-kroeger et al., 2008) or common to several (Gebelein et al., 2002; Coiffier et al., 2008) Hox proteins, suggesting they could rely on the interaction with different types of cofactors.

Here, we present the first interactomes of five Hox proteins with a set of 35 TFs. BiFC and co-ip experiments revealed that all tested TFs could interact with two or more Hox proteins. Although this result should be expanded to more TFs, it suggests that the interaction between Hox proteins and TFs is generally not exclusive. As a corollary, we hypothesize that Hox interactome specificity is unlikely to rely solely on the interaction with specific TFs.

Despite a number of common interactions, each Hox interactome contains a different set of positive TFs. AbdA is the Hox protein establishing the highest number of interactions, which is consistent with the fact that it served as a bait protein in the starting competition screen. However, the observation that Hox proteins do not interact systematically with the same set of cofactors shows their specificity. Interestingly, this specificity is not only occurring at the DNA-binding level since the loss of AbdA DNA-binding activity did not affect all interactions (18 interactions of 31 were affected). Our results thus emphasize the need of considering post-DNA-binding mechanisms for understanding Hox functional specificity *in vivo*.

The strong interaction potential of Hox proteins is consistent with their wide spectrum of regulatory activities in the embryo. It probably constitutes an inherent feature of other classes of developmental TFs that intervene in several regulatory processes throughout embryogenesis. Along this line, observations from genome-wide analyses showed that target cis-regulatory sequences allow the assembly of large multi-protein complexes (Moorman *et al.*, 2006; Blaxter, 2010; Kvon *et al.*, 2012). In addition, TFs can generally bind thousands of sites across the genome (Li *et al.*, 2008). Thus, the high number of possibilities in protein–protein and protein–DNA contacts likely reflects the propensity of developmental TFs to regulate target gene expression in many different cell contexts.

Finally, clustering analyses showed that the similarity between the different Hox interactomes does not follow the level of sequence identity between Hox proteins. Thus, common interactions might rely on different Hox protein interfaces. Accordingly, we found that the HD, which contains the highest score of sequence identity between Hox proteins, was not sufficient to ensure the majority of interactions established by AbdA. In addition, the common HX motif is not similarly used in the different Hox interactomes. Interaction properties of the HX motif were also different depending on the tissue considered, highlighting the strong flexibility and adaptability of this motif to the surrounding protein environment. Restricting the analysis to the closely related Ubx and AbdA proteins did not reveal a higher level of similarity in the interaction properties of the HX motif. In contrast, the UbdA motif showed a more frequent similar usage mode between the two Hox proteins. Thus, SLiMs conserved at different evolutionary extents provide different levels of specificity to Hox interaction properties.

### Evolutionary perspective of Hox interactomes

A high interaction potential was not only observed with *Drosophila*, but also with mouse and cnidarian Hox proteins. This result is particularly striking with the cnidarian HoxE protein, which is capable of interacting with TFs that are specific of the Bilateria group, including Biniou, Midline, Pointed or Teashirt. Although these observations are not functionally informative, they indicate that the strong interaction potential of Hox proteins is an ancestral feature that was probably present before its full exploitation in bilaterian lineages.

In addition, the observation that highly divergent Hox proteins could interact with the same set of cofactors questions the role of conserved and non-conserved regions in Hox functions. Interestingly, long intrinsic disordered regions characterize Hox proteins in general (Merabet and Dard, 2014), and a recent study showed that they could serve in Ubx to bind different partners in a competitive or cooperative way (Hsiao *et al.*, 2014). Results obtained with the cnidarian HoxE protein confirm the important role of a long disordered region in mediating interactions with different TFs. We suggest that the acquisition of long intrinsic disordered regions was a key for providing functional diversity to Hox proteins during animal evolution.

### A revised view of SLiMs in mediating protein–protein interactions

Our work provides an original experimental strategy for analysing the role of SLiMs in the context of full-length proteins *in vivo*. Results show that the HX mutation affects a number of interactions in all tested Hox proteins. Surprisingly, the absence of the HX motif could lead to a stronger or new interaction potential with TFs. A gain of interaction was observed with *Drosophila*, mouse, and cnidarian Hox proteins, suggesting that this molecular property is evolutionary conserved in the

animal kingdom. In total, the HX motif appears more often involved in limiting rather than in promoting interactions with TFs.

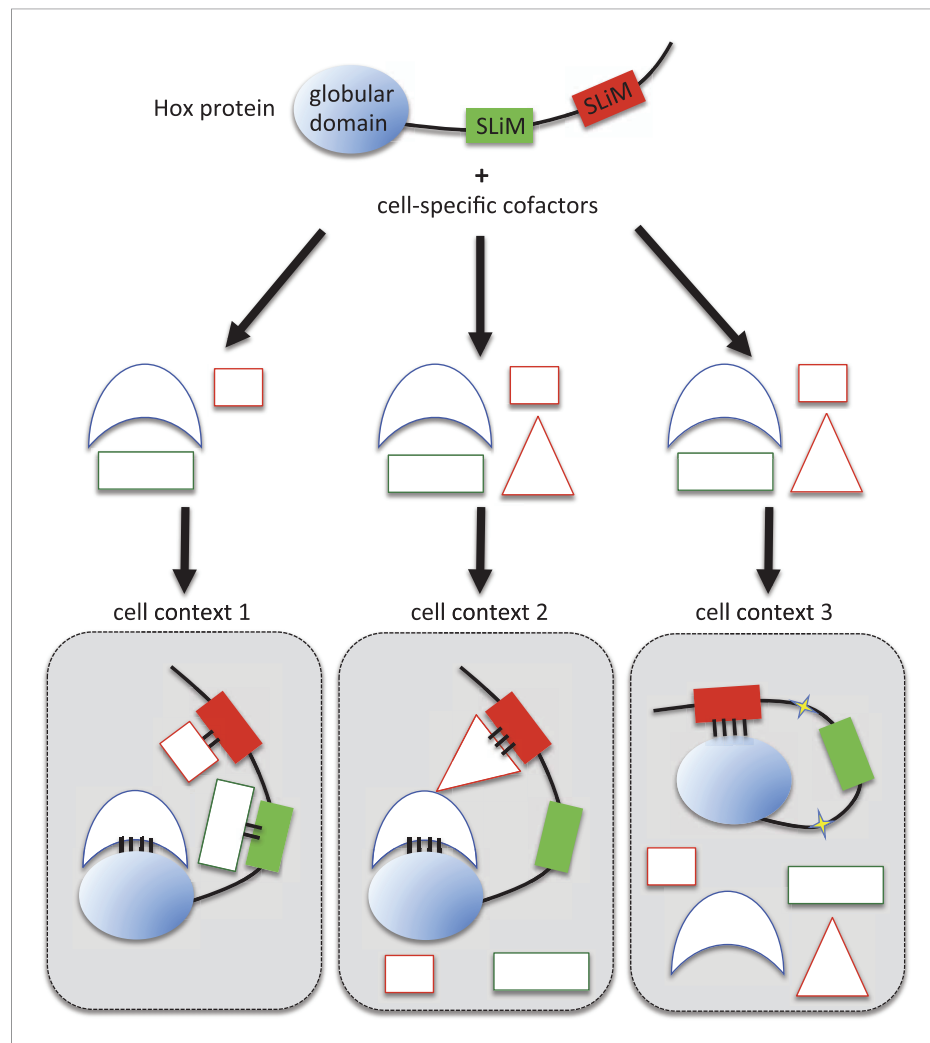
The inhibitory effect of the HX motif on PPIs was most pronounced in Ubx and AbdA, which were also used for the analysis of the UbdA motif. We found that this motif was also required for limiting the interaction potential of the two Hox proteins, although to a lesser extent than the HX motif. Thus, the negative regulation of PPIs is not a specific property of the HX motif.

Inhibitory activity of SLiMs on Hox protein function can be reconsidered in the light of previous functional data. For example, the HX mutation was shown to convert AbdA from a repressor to a strong activator of the *decapentaplegic* (*dpp*) target enhancer (Merabet et al., 2011). This striking transcriptional conversion is difficult to assign to a simple loss of interaction with co-repressor(s). Along the same line, the HX mutation increases the interaction potential of Ubx with Exd in vivo (Hudry et al., 2012), and confers an AbdA-like activity to Ubx for segment specification in the epidermis (Galant et al., 2002). Similarly, the mouse HoxB8 protein was shown to provoke dominant negative phenotypes in absence of its HX motif (Medina-martinez & Rami, 2003), which is also difficult to reconcile with a simple loss of interactions. Finally, the interaction between HoxA11 and Foxo1a in placental mammal lineages is reported to result from the loss of a specific Foxo1a-inhibitory interaction domain and not from the gain of a new binding interface in HoxA11 (Brayer et al., 2011). The gain of interaction between HoxA11 and Foxo1a comprises one of the major regulatory events that led to placentation in mammals, illustrating a so far unexpected role of a PPI inhibitory domain in interactome rewiring during evolution (Lynch et al., 2008).

More generally, protein autoinhibition is described for the regulation of other molecular events, including protein-DNA interactions and actin polymerisation (Pufall, 2002; Lee et al., 2005; Padrick and Rosen, 2010). It relies on the presence of inhibitory modules that were recently found to be enriched in intrinsically disordered regions (Trudeau et al., 2013). SLiMs can also be categorized as hiding motifs (Van Roey et al., 2013), but this role has so far been described for few mammalian proteins in the context of their intracellular transport or post-translational modification.

Here, we demonstrate that two different SLiMs could be used in a cell type-dependent manner for promoting or limiting the Hox interaction potential with TFs. We propose different models for explaining the underlying molecular mechanisms. In the classic situation, SLiMs are positively used as context-specific interaction modules, together with globular domains, for the recruitment of cell-specific cofactors (cell context 1 in Figure 10). Alternatively, SLiMs could also be important for restraining the interaction potential (cell contexts 2 and 3 in Figure 10). In one mechanism, the inhibitory activity would rely on the interaction with a particular partner that will mask or forbid the recruitment of the other SLiM-interacting cofactors (cell context 2 in Figure 10). This mechanism implies that the interaction between the SLiM and the hiding partner will be strong and stable enough to overcome the binding of the other cofactors. In a second mechanism, SLiMs could also directly act as a masking peptide, preventing recognition and/or binding of undesired cofactors (cell context 3 in Figure 10). SLiMs are classically required to make interactions with structured globular domains in *trans* (Stein and Aloy, 2008), but intramolecular contacts following post-translational modifications have also been reported in a few cases (Pawson et al., 2001). Thus, SLiMs could also establish interactions in *cis*, potentially upon cell-specific modifications (i.e., phosphorylations), to eventually inhibit the recruitment of inappropriate cofactors. A similar role was previously described for the HX motif of Labial, which is able to prevent the binding of the HD on the DNA (Chan et al., 1996). Interestingly, this inhibition is relieved upon interaction with Exd, highlighting the influence of the environment on the SLiM inhibitory activity.

In summary, we showed that highly conserved SLiMs are used in a context-dependent manner for constraining the interaction potential of Hox proteins with surrounding TFs. This molecular strategy has certainly been underestimated to date. We propose that the inhibiting interaction properties of SLiMs could apply more generally to the fine-tuning of highly connected interactomes. It is interesting to note that SLiMs can also be produced as individual molecules from short open reading frames (Kondo et al., 2007; Magny et al., 2013) or even from long non coding RNAs (Ruiz-Orera et al., 2014). We anticipate that short peptides could act as buffering molecules, helping hub proteins to discriminate their correct partners among hundreds of possible interactions within the 'messy' (Tawfik, 2010) cell environment.



**Figure 10.** Molecular mechanisms underlying context-dependent activities of SLiMs in protein-protein interactions. The Hox protein is represented as containing a globular structured domain together with two different SLiMs embedded in a disordered region, as indicated. This protein will present different interaction properties with a set of cofactors that could vary depending on the cell context considered. Preferential interactions between cofactors and the protein domain and SLiMs are represented by a colour code. Black bars symbolize the various levels of interaction affinity. In the cell context 1, cofactors are recruited through specific interactions with the globular domain and the two SLiMs. In the cell context 2, there is a supplementary triangular cofactor that displays higher affinity with the red SLiM than the square cofactor. As a consequence, interaction will occur with this triangular (hiding) cofactor, which forbids the interaction with the other SLiM. In this context, the red SLiM behaves as an inhibitory interaction motif. In the cell context 3, post-translational modifications in the disordered region (yellow stars) allow the inhibitory SLiM to establish interactions in *cis* with the globular domain. These intra-molecular contacts forbid the binding of the other cofactors. The last two mechanisms illustrate how the inhibitory activity of SLiMs could help in distinguishing/specifying interactomes with an identical set of cofactors.

DOI: [10.7554/eLife.06034.038](https://doi.org/10.7554/eLife.06034.038)

## Materials and methods

### Fusion protein constructs and transgenic lines

Several VC-Hox fusion constructs were previously generated (Hudry et al., 2012; Boube et al., 2014): these correspond to wild type and mutated/truncated variants of all *Drosophila* Hox proteins and wild type and HX-mutated HoxE. Other Hox fusion constructs (HoxB8, HoxB8<sup>HX</sup> and Nter-HoxE) were generated by PCR and restriction-cloned in the pUAST or pUASTattB vector, in fusion with the C-

terminal (155–238) fragment of Venus (VC) at 3' end (see **Supplementary file 4**). Fusion TFs were also generated by PCR from full-length complementary DNAs and restriction-cloned in fusion in pUAST or pUASTattB vector with the N-terminal (1–173) fragment of Venus (VN) at the 5' or 3' end (see **Supplementary file 4**). Primers used for cloning of each TF are listed in **Supplementary file 5**. For all fusion constructs, a linker of three to five amino acids was added to separate the Venus fragment from the protein. All constructs were sequence-verified before injection. Transgenic lines were established either by the PhiC31 integrase system (with the pUASTattB vector [Venken et al., 2006; Bischof et al., 2007]) or by classic P-element (with PUAST vector) mediated germ line transformation. Expression level of Hox fusion constructs was verified as previously described (Hudry et al., 2011). Briefly, flies were crossed at different temperatures with the *abdA-Gal4* driver and embryos were collected for immunostaining with a chicken anti-GFP antibody (Abcam ab13970, England) recognizing the VN and VC fragments. Fluorescent revelation (with a secondary anti-chicken antibody coupled to AlexaFluor488, Abcam150169) was used to compare the expression level between the different conditions with wild type and mutated Hox fusion constructs. The temperature for each fly cross was adjusted accordingly, allowing comparing BiFC signals with Hox fusion proteins expressed at similar levels. The same anti-GFP antibody was used to verify the correct expression level of each generated VN-TF fly line.

### Fly stocks and genetic crosses

Gal4 drivers used are: *Antp-Gal4*, *Ubx-Gal4*, *abdA-Gal4*, and *AbdB-Gal4* (de Navas et al., 2006; Hudry et al., 2012). Fly lines generating in this work are listed in **Supplementary file 4**.

Competition tests were performed in one generation by crossing each candidate UAS-TF (see **Supplementary file 1** for the type of the UAS fly line) with the BiFC reporter fly line containing the UAS-VC-*abdA* (homozygous on the second chromosome) and UAS-VN-*exd* (homozygous on the fourth chromosome) constructs, together with the *abdA-Gal4* driver balanced over a *TM6tubulineGal80* third balancer (Hudry et al., 2011). Under these conditions, half of the embryo progeny (with homozygous UAS-TF fly line) could display affected BiFC signals in presence of a competitive TF.

Complementation tests were performed by crossing *en mass* virgin females containing the VN-TF and VC-Hox constructs (as non-established fly lines resulting from a previous cross between VN-TF and VC-Hox individuals) with males carrying the corresponding *Hox-Gal4* driver. Over night egg laying was performed at different temperatures, according to the expression level of the VC-Hox variant.

### Immunostaining

Embryo collection, preparation, and immunodetections were performed according to standard procedures. The antibodies used were: chicken anti-GFP (Abcam ab13970, 1/500), mouse anti-Scr (6H4.1, 1/100), mouse anti-*Antp* (4C3, 1/100), mouse anti-*Ubx* (FP3.38, 1/100), rabbit anti-*AbdA* (Dm. Abd-A.1, 1/100), mouse anti-*AbdB*

### Co-immunoprecipitation (Co-ip) experiments and western blot

Co-ips were performed in S2 cells, which were transfected with a HA-tagged form of *AbdA*, together with an actin-*Gal4* plasmid and the corresponding VN-cofactor. *AbdA*-HA construct was generated by PCR, using an oligonucleotide bearing the HA sequence and cloned into the pUASTattB vector. The construct was sequence verified. Transfection was realised using the X-tremeGENE HP DNA Transfection Reagent (Roche). Cells were lysed 48 hr later and nuclear extracts were prepared as classically described. Ip was performed with a polyclonal rabbit anti-HA antibody (Abcam ab9110). Presence of *AbdA*-HA was verified by western blot using a monoclonal anti-HA antibody (HA.11 from Covance). Presence of the associated VN-cofactor was revealed with a chicken anti-GFP antibody recognising the VN fragment (Abcam ab13970).

### BiFC analysis in *Drosophila* embryos

Experimental parameters allowing a comparable expression level between wild type and mutated VC-Hox proteins were previously established for several constructs (Hudry et al., 2012; Boube et al., 2014) or deduced from additional immunostaining experiments for new constructs (*HoxB8*, *HoxB8<sup>HX</sup>*, and *Nter-HoxE*). Fly crosses for BiFC analyses were set up at the defined temperature over night.



After the removal of the flies, the embryos were kept at 4°C for 24 hr before live imaging. Live embryos were dechorionated and mounted in the halocarbon oil 10S (commercialized by VWR, Pennsylvania, USA). Quantification of the BiFC signals was realised by taking unsaturated images at the same desired stage, depending on the tissue considered (epidermis: stage 10, mesoderm: stage 12, nervous system: stage 14). For BiFC analysis in the CNS, embryos were manually aligned on the dorsal face. Observations were performed at least twice (from two different over night egg laying periods) for a same genotype. A minimum of 5 embryos of the correct developmental stage was considered in each case. Pictures were acquired using a LSM780 confocal microscope (Zeiss, Jena, Germany). For Venus fluorescence, filters were adjusted at 500 nm for excitation and 535 nm for emission. Identical parameters of acquisition were applied between the different genotypes. The number and intensity of the all pixels (for each embryo) were measured in the tissue of interest using the histogram function of the ImageJ Software. The quantification of fluorescence complementation is shown for each condition by boxplot representation using R-Software. Boxplot depicts: the smallest value, lower quartile, median (black line), upper quartile, and largest value for each condition.

## Network visualisation and heatmap

Networks were represented using Cytoscape 3.0 (Shannon *et al.*, 2003). A hierarchical clustering algorithm (with Euclidian distance and average linking) was applied to the matrix using the MeV software suite (Saeed *et al.*, 2006). The bootstrap method was used for resampling the data and provides a statistical support for each tree node.

## Acknowledgements

We thank Johannes Bischof and Bart Desplancke for sharing unpublished tools. We thank Muriel Boube, Henri-Marc Bourbon, and Luis Mojica-Vasquez for the *Distalless-VN* fly line, Damien Ahr for his help in fly genetics and BiFC observations and Yacine Graba for freedom during the earliest part of this work in his laboratory. We thank Eileen Furlong, Christophe Jagla, Manfred Frasch, Marc Haenlin, Michael Taylor, Alan M. Michelson, Stefan Thor, James Castelli-Gair, the Bloomington Stock Center and FlyORF for fly lines and cDNAs. We warmly thank François Payre, Andreas Zanzoni, and Johannes Bischof for their helpful comments on the manuscript and Joanne Burden for proofreading. This work was supported by grants from ARC and FRM.

---

## Additional information

### Funding

Funder	Author
Fondation pour la Recherche Médicale	Séverine Viala, Marjorie Heim, Amélie Dard, Bruno Hudry, Marilyne Duffraisse, Ana Rogulja-Ortmann, Christine Brun, Samir Merabet
Association pour la Recherche sur le Cancer	Séverine Viala, Marjorie Heim, Amélie Dard, Bruno Hudry, Marilyne Duffraisse, Ana Rogulja-Ortmann, Christine Brun, Samir Merabet

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

### Author contributions

MB, SV, MH, AD, BH, MD, AR-O, Acquisition of data, Analysis and interpretation of data; CB, Analysis and interpretation of data, Drafting or revising the article; SM, Conception and design, Acquisition of data, Analysis and interpretation of data, Drafting or revising the article, Contributed unpublished essential data or reagents

### Author ORCIDs

Christine Brun,  <http://orcid.org/0000-0002-5563-6765>

## Additional files

### Supplementary files

• Supplementary file 1. List of the 80 candidate transcription factors (TFs) tested as potential cofactors of AbdA. Each TF used in the competition test corresponds either to an artificial UAS construct or to a UAS-containing transposon inserted in the endogenous promoter locus. These constructs were obtained from stock centres or from particular laboratories, as indicated. Red and green-filled boxes indicate positive or negative competitor status, respectively. The expression profile of each TF is provided for the main embryonic germ layers and was compiled according to databases (<http://flybase.org> and [Hammonds et al., 2013]). Yellow and black-filled boxes depict presence or absence of expression, respectively. TFs that were subsequently used for BiFC with AbdA are bolded and annotated with a grey box. Note that TFIIBeta was directly included in the BiFC analysis without doing a preliminary competition test (see main text for details).

DOI: [10.7554/eLife.06034.039](https://doi.org/10.7554/eLife.06034.039)

• Supplementary file 2. Spatiotemporal expression pattern of AbdA and TFs under study as deduced from in situ hybridization experiments (Hammonds et al., 2013).

DOI: [10.7554/eLife.06034.040](https://doi.org/10.7554/eLife.06034.040)

• Supplementary file 3. List of the TFs used for BiFC in the mesoderm. Colour code for tissue expression is as in **Supplementary file 1**. Green or red boxes depict presence or absence of BiFC, respectively. Results obtained by co-ip in S2 cells are also indicated for each TF. Note that only Kn is interacting specifically in the mesoderm among the tested TFs.

DOI: [10.7554/eLife.06034.041](https://doi.org/10.7554/eLife.06034.041)

• Supplementary file 4. List of the constructs and transgenic fly lines generated in this study.

DOI: [10.7554/eLife.06034.042](https://doi.org/10.7554/eLife.06034.042)

• Supplementary file 5. List of the primers used to clone each TF in fusion with the Venus fragment.

DOI: [10.7554/eLife.06034.043](https://doi.org/10.7554/eLife.06034.043)

## References

- Balavoine G, de Rosa R, Adoutte A. 2002. Hox clusters and bilaterian phylogeny. *Molecular Phylogenetics and Evolution* **24**:366–373. doi: [10.1016/S1055-7903\(02\)00237-3](https://doi.org/10.1016/S1055-7903(02)00237-3).
- Bassett AR, Liu JL. 2014. CRISPR/Cas9 and genome editing in *Drosophila*. *Journal of Genetics and Genomics* **41**: 7–19. doi: [10.1016/j.jgg.2013.12.004](https://doi.org/10.1016/j.jgg.2013.12.004).
- Berger MF, Badis G, Gehrke AR, Talukder S, Philippakis AA, Peña-Castillo L, Alleyne TM, Mnaimneh S, Botvinnik OB, Chan ET, Khalid F, Zhang W, Newburger D, Jaeger SA, Morris QD, Bulyk ML, Hughes TR. 2008. Variation in homeodomain DNA binding revealed by high-resolution analysis of sequence preferences. *Cell* **133**:1266–1276. doi: [10.1016/j.cell.2008.05.024](https://doi.org/10.1016/j.cell.2008.05.024).
- Bertrand N, Roux M, Ryckebusch L, Niederreither K, Dollé P, Moon A, Capecchi M, Zaffran S. 2011. Hox genes define distinct progenitor sub-domains within the second heart field. *Developmental Biology* **353**:266–274. doi: [10.1016/j.ydbio.2011.02.029](https://doi.org/10.1016/j.ydbio.2011.02.029).
- Bischof J, Björklund M, Furger E, Schertel C, Taipale J, Basler K. 2013. A versatile platform for creating a comprehensive UAS-ORFeome library in *Drosophila*. *Development* **140**:2434–2442. doi: [10.1242/dev.088757](https://doi.org/10.1242/dev.088757).
- Bischof J, Maeda RK, Hediger M, Karch F, Basler K. 2007. An optimized transgenesis system for *Drosophila* using germ-line-specific phiC31 integrases. *Proceedings of the National Academy of Sciences of USA* **104**:3312–3317. doi: [10.1073/pnas.0611511104](https://doi.org/10.1073/pnas.0611511104).
- Blaxter M. 2010. Genetics. Revealing the dark matter of the genome. *Science* **330**:1758–1759. doi: [10.1126/science.1200700](https://doi.org/10.1126/science.1200700).
- Bondos SE, Catanese DJ Jr, Tan XX, Bicknell A, Li L, Matthews KS. 2004. Hox transcription factor ultrabithorax 1b physically and genetically interacts with disconnected interacting protein 1, a double-stranded RNA-binding protein. *The Journal of Biological Chemistry* **279**:26433–26444. doi: [10.1074/jbc.M312842200](https://doi.org/10.1074/jbc.M312842200).
- Bondos SE, Tan XX, Matthews KS. 2006. Physical and genetic interactions link hox function with diverse transcription factors and cell signaling proteins. *Molecular & Cellular Proteomics* **5**:824–834. doi: [10.1074/mcp.M500256-MCP200](https://doi.org/10.1074/mcp.M500256-MCP200).
- Boube M, Hudry B, Immarigeon C, Carrier Y, Bernat-Fabre S, Merabet S, Graba Y, Bourbon HM, Cribbs DL. 2014. *Drosophila melanogaster* Hox transcription factors access the RNA polymerase II machinery through direct homeodomain binding to a conserved motif of mediator subunit Med19. *PLOS Genetics* **10**:e1004303. doi: [10.1371/journal.pgen.1004303](https://doi.org/10.1371/journal.pgen.1004303).
- Brayer KJ, Lynch VJ, Wagner GP. 2011. Evolution of a derived protein-protein interaction between HoxA11 and Foxo1a in mammals caused by changes in intramolecular regulation. *Proceedings of the National Academy of Sciences of USA* **108**:E414–E420. doi: [10.1073/pnas.1100990108](https://doi.org/10.1073/pnas.1100990108).

- Brodu V**, Elstob PR, Gould AP. 2002. abdominal A specifies one cell type in *Drosophila* by regulating one principal target gene. *Development* **129**:2957–2963.
- Buljan M**, Chalancon G, Dunker AK, Bateman A, Balaji S, Fuxreiter M, Babu MM. 2013. Alternative splicing of intrinsically disordered regions and rewiring of protein interactions. *Current Opinion in Structural Biology* **23**: 443–450. doi: [10.1016/j.sbi.2013.03.006](https://doi.org/10.1016/j.sbi.2013.03.006).
- Chan SK**, Pöpperl H, Krumlauf R, Mann RS. 1996. An extradenticle-induced conformational change in a HOX protein overcomes an inhibitory function of the conserved hexapeptide motif. *The EMBO Journal* **15**: 2476–2487.
- Coiffier D**, Charroux B, 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](https://doi.org/10.1242/dev.009662).
- Cunha PM**, Sandmann T, Gustafson EH, Ciglar L, Eichenlaub MP, Furlong EE. 2010. Combinatorial binding leads to diverse regulatory responses: Lmd is a tissue-specific modulator of Mef2 activity. *PLoS Genetics* **6**:e1001014. doi: [10.1371/journal.pgen.1001014](https://doi.org/10.1371/journal.pgen.1001014).
- de Navas L**, Foronda D, Suzanne M, Sánchez-Herrero E. 2006. 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*. *Mechanisms of Development* **123**:860–867. doi: [10.1016/j.mod.2006.07.010](https://doi.org/10.1016/j.mod.2006.07.010).
- Dinkel H**, Van Roey K, Michael S, Davey NE, Weatheritt RJ, Born D, Speck T, Krüger D, Grebnev G, Kuban M, Strumillo M, Uyar B, Budd A, Altenberg B, Seiler M, Chemes LB, Glavina J, Sánchez IE, Diella F, Gibson TJ. 2014. The eukaryotic linear motif resource ELM: 10 years and counting. *Nucleic Acids Research* **42**:D259–D266. doi: [10.1093/nar/gkt1047](https://doi.org/10.1093/nar/gkt1047).
- Fasano L**, Röder L, Coré N, Alexandre E, Vola C, Jacq B, Kerridge S. 1991. The gene teashirt is required for the development of *Drosophila* embryonic trunk segments and encodes a protein with widely spaced zinc finger motifs. *Cell* **64**:63–79. doi: [10.1016/0092-8674\(91\)90209-H](https://doi.org/10.1016/0092-8674(91)90209-H).
- Finnerty JR**. 2003. The origins of axial patterning in the metazoa: how old is bilateral symmetry? *The International Journal of Developmental Biology* **47**:523–529.
- Foos N**, Maurel-Zaffran C, Maté MJ, Vincentelli R, Hainaut M, Berenger H, Pradel J, Saurin AJ, Ortiz-Lombardía M, Graba Y. 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](https://doi.org/10.1016/j.str.2014.12.011).
- Foronda D**, Estrada B, de Navas L, Sánchez-Herrero E. 2005. Requirement of Abdominal-A and Abdominal-B in the developing genitalia of *Drosophila* breaks the posterior downregulation rule. *Development* **133**:117–127. doi: [10.1242/dev.02173](https://doi.org/10.1242/dev.02173).
- Galant R**, Walsh CM, Carroll SB. 2002. Hox repression of a target gene: extradenticle-independent, additive action through multiple monomer binding sites. *Development* **3126**:3115–3126.
- Gebelein B**, Culi J, Ryoo HD, Zhang W, Mann RS. 2002. Specificity of distalless repression and limb primordia development by abdominal hox proteins. *Developmental Cell* **3**:487–498.
- Ghosh I**, Hamilton AD, Regan L. 2000. Antiparallel leucine zipper-directed protein reassembly: Application to the green fluorescent protein. *Journal of the American Chemical Society* **122**:5658–5659. doi: [10.1021/ja994421w](https://doi.org/10.1021/ja994421w).
- Gnerer JP**, Venken KJT, Dierick HA. 2015. Gene-specific cell labeling using MiMIC transposons. *Nucleic Acids Research* **1**–13.
- Grienenberger A**, Merabet S, Manak J, Iltis I, Fabre A, Bérenger H, Scott MP, Pradel J, Graba Y. 2003. Tgfbeta 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](https://doi.org/10.1242/dev.00760).
- Hammonds AS**, Bristow CA, Fisher WW, Weiszmann R, Wu S, Hartenstein V, Kellis M, Yu B, Frise E, Celniker SE. 2013. Spatial expression of transcription factors in *Drosophila* embryonic organ development. *Genome Biology* **14**:R140. doi: [10.1186/gb-2013-14-12-r140](https://doi.org/10.1186/gb-2013-14-12-r140).
- Hsiao H-C**, Gonzalez KL, Catanese DJ Jr, Jordy KE, Matthews KS, Bondos SE. 2014. The intrinsically disordered regions of the *Drosophila melanogaster* Hox protein ultrabithorax select interacting proteins based on partner topology. *PLoS ONE* **9**:e108217. doi: [10.1371/journal.pone.0108217](https://doi.org/10.1371/journal.pone.0108217).
- Hudry B**, Remacle S, Delfini MC, Rezsóhazy R, Graba Y, Merabet S. 2012. Hox proteins display a common and ancestral ability to diversify their interaction mode with the PBC class cofactors. *PLoS Biology* **10**:e1001351. doi: [10.1371/journal.pbio.1001351](https://doi.org/10.1371/journal.pbio.1001351).
- Hudry B**, Thomas-Chollier M, Volovik Y, Duffraisse M, Dard A, Frank D, Technau U, Merabet S. 2014. Molecular insights into the origin of the Hox-TALE patterning system. *eLife* **3**:e01939. doi: [10.7554/eLife.01939](https://doi.org/10.7554/eLife.01939).
- Hudry B**, Viala S, Graba Y, Merabet S. 2011. Visualization of protein interactions in living *Drosophila* embryos by the bimolecular fluorescence complementation assay. *BMC Biology* **9**:5. doi: [10.1186/1741-7007-9-5](https://doi.org/10.1186/1741-7007-9-5).
- In der Rieden PM**, Mainguy G, Woltering JM, Durston AJ. 2004. Homeodomain to hexapeptide or PBC-interaction-domain distance: size apparently matters. *Trends in Genetics* **20**:76–79. doi: [10.1016/j.tig.2003.12.001](https://doi.org/10.1016/j.tig.2003.12.001).
- Kammermeier L**, Sprecher SG, Müller M, Miller DF, Kaufman TC, Reichert H, Hirth F. 2004. Hox gene cross-regulatory interactions in the embryonic brain of *Drosophila*. *Mechanisms of Development* **121**:527–536.
- Kerppola TK**. 2008. Bimolecular fluorescence complementation (BiFC) analysis as a probe of protein interactions in living cells. *Annual Review of Biophysics* **37**:465–487. doi: [10.1146/annurev.biophys.37.032807.125842](https://doi.org/10.1146/annurev.biophys.37.032807.125842).
- Kodama Y**, Hu CD. 2012. Bimolecular fluorescence complementation (BiFC): a 5-year update and future perspectives. *Biotechniques* **53**:285–298. doi: [10.2144/000113943](https://doi.org/10.2144/000113943).
- Kondo T**, Hashimoto Y, Kato K, Inagaki S, Hayashi S, Kageyama Y. 2007. Small peptide regulators of actin-based cell morphogenesis encoded by a polycistronic mRNA. *Nature Cell Biology* **9**:660–665. doi: [10.1038/ncb1595](https://doi.org/10.1038/ncb1595).

- Kvon EZ**, Stampfel G, Yáñez-Cuna JO, Dickson BJ, Stark A. 2012. HOT regions function as patterned developmental enhancers and have a distinct cis-regulatory signature. *Genes & Development* **26**:908–913. doi: [10.1101/gad.188052.112](https://doi.org/10.1101/gad.188052.112).
- Lambert B**, Vandeputte J, Remacle S, Bergiers I, Simonis N, Twizere JC, Vidal M, Rezsohazy R. 2012. Protein interactions of the transcription factor Hoxa1. *BMC Developmental Biology* **12**:29. doi: [10.1186/1471-213X-12-29](https://doi.org/10.1186/1471-213X-12-29).
- Landry CR**, Levy ED, Abd Rabbo D, Tarassov K, Michnick SW. 2013. Extracting insight from noisy cellular networks. *Cell* **155**:983–989. doi: [10.1016/j.cell.2013.11.003](https://doi.org/10.1016/j.cell.2013.11.003).
- Lee GM**, Donaldson LW, Pufall MA, Kang HS, Pot I, Graves BJ, McIntosh LP. 2005. The structural and dynamic basis of Ets-1 DNA binding autoinhibition. *The Journal of Biological Chemistry* **280**:7088–7099. doi: [10.1074/jbc.M410722200](https://doi.org/10.1074/jbc.M410722200).
- Lee OH**, Kim H, He Q, Baek HJ, Yang D, Chen LY, Liang J, Chae HK, Safari A, Liu D, Songyang Z. 2011. Genome-wide YFP fluorescence complementation screen identifies new regulators for telomere signaling in human cells. *Molecular & Cellular Proteomics* **10**:M110.001628. doi: [10.1074/mcp.M110.001628](https://doi.org/10.1074/mcp.M110.001628).
- Li X**, MacArthur S, Bourgon R, Nix D, Pollard DA, Iyer VN, Hechmer A, Simirenko L, Stapleton M, Luengo Hendriks CL, Chu HC, Ogawa N, Inwood W, Sementchenko V, Beaton A, Weizmann R, Celniker SE, Knowles DW, Gingeras T, Speed TP, Eisen MB, Biggin MD. 2008. Transcription factors bind thousands of active and inactive regions in the *Drosophila* blastoderm. *PLOS Biology* **6**:e27. doi: [10.1371/journal.pbio.0060027](https://doi.org/10.1371/journal.pbio.0060027).
- Li-kroeger D**, Witt LM, Grimes HL, Cook TA, Gebelein B. 2008. Hox and senseless antagonism functions as a molecular switch to regulate EGF secretion in the *Drosophila* PNS. *Developmental Cell* **15**:298–308. doi: [10.1016/j.devcel.2008.06.001](https://doi.org/10.1016/j.devcel.2008.06.001).
- Lynch VJ**, Tanzer A, Wang Y, Leung FC, Gellersen B, Emera D, Wagner GP. 2008. Adaptive changes in the transcription factor HoxA-11 are essential for the evolution of pregnancy in mammals. *Proceedings of the National Academy of Sciences of USA* **105**:14928–14933. doi: [10.1073/pnas.0802355105](https://doi.org/10.1073/pnas.0802355105).
- Magny EG**, Pueyo JI, Pearl FM, Cespedes MA, Niven JE, Bishop SA, Couso JP. 2013. Conserved regulation of cardiac calcium uptake by peptides encoded in small open reading frames. *Science* **341**:1116–1120. doi: [10.1126/science.1238802](https://doi.org/10.1126/science.1238802).
- Mann RS**, Lelli KM, Joshi R. 2009. Hox Specificity unique roles for cofactors and collaborators. *Current Topics in Developmental Biology* **88**:63–101. doi: [10.1016/S0070-2153\(09\)88003-4](https://doi.org/10.1016/S0070-2153(09)88003-4).
- Medina-martinez O**, Ramí 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. *Developmental Biology* **264**:77–90. doi: [10.1016/j.ydbio.2003.07.020](https://doi.org/10.1016/j.ydbio.2003.07.020).
- Merabet S**, Dard A. 2014. Tracking context-specific transcription factors regulating hox activity. *Developmental Dynamics* **243**:16–23. doi: [10.1002/dvdy.24002](https://doi.org/10.1002/dvdy.24002).
- Merabet S**, Hudry B. 2011. On the border of the homeotic function: re-evaluating the controversial role of cofactor-recruiting motifs: the role of cofactor-recruiting motifs in conferring Hox evolutionary flexibility may critically depend on the protein environment. *Bioessays* **33**:499–507. doi: [10.1002/bies.201100019](https://doi.org/10.1002/bies.201100019).
- Merabet S**, Kambris Z, Capovilla M, Bérenger H, Pradel J, Graba Y. 2003. The hexapeptide and linker regions of the AbdA Hox protein regulate its activating and repressive functions. *Developmental Cell* **4**:761–768. doi: [10.1016/S1534-5807\(03\)00126-6](https://doi.org/10.1016/S1534-5807(03)00126-6).
- Merabet S**, Saadaoui M, Sambrani N, Hudry B, Pradel J, Affolter M, Graba Y. 2007. A unique Extradenticle recruitment mode in the *Drosophila* Hox protein Ultrabithorax. *Proceedings of the National Academy of Sciences of USA* **104**:16946–16951. doi: [10.1073/pnas.0705832104](https://doi.org/10.1073/pnas.0705832104).
- Merabet S**, Hudry B, Saadaoui M, Graba Y. 2009. Classification of sequence signatures: a guide to Hox protein function. *Bioessays* **31**:500–511. doi: [10.1002/bies.200800229](https://doi.org/10.1002/bies.200800229).
- Merabet S**, Litim-Mecheri I, Karlsson D, Dixit R, Saadaoui M, Monier B, Brun C, Thor S, Vijayraghavan K, Perrin L, Pradel J, Graba Y. 2011. Insights into Hox protein function from a large scale combinatorial analysis of protein domains. *PLOS Genetics* **7**:e1002302. doi: [10.1371/journal.pgen.1002302](https://doi.org/10.1371/journal.pgen.1002302).
- Moorman C**, Sun LV, Wang J, de Wit E, Talhout W, Ward LD, Greil F, Lu XJ, White KP, Bussemaker HJ, van Steensel B. 2006. Hotspots of transcription factor colocalization in the genome of *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences of USA* **103**:12027–12032. doi: [10.1073/pnas.0605003103](https://doi.org/10.1073/pnas.0605003103).
- Morell M**, Czihal P, Hoffmann R, Otvos L, Avilés FX, Ventura S. 2008. Monitoring the interference of protein-protein interactions in vivo by bimolecular fluorescence complementation: the DnaK case. *Proteomics* **8**:3433–3442. doi: [10.1002/pmic.200700739](https://doi.org/10.1002/pmic.200700739).
- Mukherjee K**, Bürglin TR. 2007. Comprehensive analysis of animal TALE homeobox genes: new conserved motifs and cases of accelerated evolution. *Journal of Molecular Evolution* **65**:137–153. doi: [10.1007/s00239-006-0023-0](https://doi.org/10.1007/s00239-006-0023-0).
- Neduva V**, Russell RB. 2005. Linear motifs: evolutionary interaction switches. *FEBS Letters* **579**:3342–3345. doi: [10.1016/j.febslet.2005.04.005](https://doi.org/10.1016/j.febslet.2005.04.005).
- Noyes MB**, Christensen RG, Wakabayashi A, Stormo GD, Brodsky MH, Wolfe SA. 2008. Analysis of homeodomain specificities allows the family-wide prediction of preferred recognition sites. *Cell* **133**:1277–1289. doi: [10.1016/j.cell.2008.05.023](https://doi.org/10.1016/j.cell.2008.05.023).
- Padrick SB**, Rosen MK. 2010. Physical mechanisms of signal integration by WASP family proteins. *Annual Review of Biochemistry* **79**:707–735. doi: [10.1146/annurev.biochem.77.060407.135452](https://doi.org/10.1146/annurev.biochem.77.060407.135452).
- Pawson T**, Gish GD, Nash P. 2001. SH2 domains, interaction modules and cellular wiring. *Trends in Cell Biology* **11**:504–511. doi: [10.1016/S0962-8924\(01\)02154-7](https://doi.org/10.1016/S0962-8924(01)02154-7).
- Prince F**, Katsuyama T, Oshima Y, Plaza S, Resendez-Perez D, Berry M, Kurata S, Gehring WJ. 2008. The YPWM motif links Antennapedia to the basal transcriptional machinery. *Development* **135**:1669–1679. doi: [10.1242/dev.018028](https://doi.org/10.1242/dev.018028).

- Pufall MA**, Graves BJ. 2002. Autoinhibitory domains: modular effectors of cellular regulation. *Annual Review of Cell and Developmental Biology* **18**:421–462. doi: [10.1146/annurev.cellbio.18.031502.133614](https://doi.org/10.1146/annurev.cellbio.18.031502.133614).
- Ruiz-Orera J**, Messeguer X, Subirana JA, Alba MM. 2014. Long non-coding RNAs as a source of new peptides. *eLife* **3**:e03523. doi: [10.7554/eLife.03523](https://doi.org/10.7554/eLife.03523).
- Saeed AI**, Bhagabati NK, Braisted JC, Liang W, Sharov V, Howe EA, Li J, Thiagarajan M, White JA, Quackenbush J. 2006. TM4 microarray software suite. *Methods in Enzymology* **411**:134–193. doi: [10.1016/S0076-6879\(06\)11009-5](https://doi.org/10.1016/S0076-6879(06)11009-5).
- Shannon P**, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Research* **13**:2498–2504. doi: [10.1101/gr.1239303](https://doi.org/10.1101/gr.1239303).
- Stein A**, Aloy P. 2008. Contextual specificity in peptide-mediated protein interactions. *PLOS ONE* **3**:e2524. doi: [10.1371/journal.pone.0002524](https://doi.org/10.1371/journal.pone.0002524).
- Tawfik DS**. 2010. Messy biology and the origins of evolutionary innovations. *Nature Chemical Biology* **6**:692–696. doi: [10.1038/nchembio.441](https://doi.org/10.1038/nchembio.441).
- Tompa P**, Davey NE, Gibson TJ, Babu MM. 2014. Million peptide motifs for the molecular biologist. *Molecular Cell* **55**:161–169. doi: [10.1016/j.molcel.2014.05.032](https://doi.org/10.1016/j.molcel.2014.05.032).
- Trudeau T**, Nassar R, Cumberworth A, Wong ET, Woollard G, Gsponer J. 2013. Structure and intrinsic disorder in protein autoinhibition. *Structure* **21**:332–341. doi: [10.1016/j.str.2012.12.013](https://doi.org/10.1016/j.str.2012.12.013).
- Van Roey K**, Dinkel H, Weatheritt RJ, Gibson TJ, Davey NE. 2013. The switches.ELM resource: a compendium of conditional regulatory interaction interfaces. *Science Signaling* **6**:rs7. doi: [10.1126/scisignal.2003345](https://doi.org/10.1126/scisignal.2003345).
- Van Roey K**, Gibson TJ, Davey NE. 2012. Motif switches: decision-making in cell regulation. *Current Opinion in Structural Biology* **22**:378–385. doi: [10.1016/j.sbi.2012.03.004](https://doi.org/10.1016/j.sbi.2012.03.004).
- Van Roey K**, Uyar B, Weatheritt RJ, Dinkel H, Seiler M, Budd A, Gibson TJ, Davey NE. 2014. Short linear motifs: Ubiquitous and functionally diverse protein interaction modules directing cell regulation. *Chemical Reviews* **114**:6733–6778. doi: [10.1021/cr400585q](https://doi.org/10.1021/cr400585q).
- Venken KJ**, He Y, Hoskins RA, Bellen HJ. 2006. P[acman]: a BAC transgenic platform for targeted insertion of large DNA fragments in *D. melanogaster*. *Science* **314**:1747–1751. doi: [10.1126/science.1134426](https://doi.org/10.1126/science.1134426).
- Walsh CM**, Carroll SB. 2007. Collaboration between Smads and a Hox protein in target gene repression. *Development* **134**:3585–3592. doi: [10.1242/dev.009522](https://doi.org/10.1242/dev.009522).
- Zakany J**, Duboule D. 2007. The role of Hox genes during vertebrate limb development. *Current Opinion in Genetics & Development* **17**:359–366. doi: [10.1016/j.gde.2007.05.011](https://doi.org/10.1016/j.gde.2007.05.011).

### 3. TROISIEME CHAPITRE : IDENTIFIER DE NOUVELLES INTERFACES POUR LES COFACTEURS TALE AU SEIN DES PROTEINES HOX HUMAINES

Depuis quelques années maintenant, il a été montré que les protéines HOX étaient capables d'interagir avec leurs partenaires TALE par d'autres motifs que le motif HX ancestral. En effet, l'apparition de motifs paralogues-spécifiques tels que les motifs UbdA ou TDWM a permis de se questionner sur la présence d'autres motifs de ce type au sein des protéines HOX.

Nous nous sommes intéressés à l'ensemble des groupes de paralogie des protéines HOX humaines et de souris pour mener un crible plus large sur le rôle du motif HX pour la formation des complexes HOX-TALE. Nous avons pu constater que les protéines HOX peuvent toutes utiliser (à l'exception des groupes de paralogie 1 et 2) une interface alternative au motif HX pour former des complexes avec les partenaires TALE.

La dissection moléculaire de plusieurs complexes a pu mettre en évidence différents types d'interface conservés à différents niveaux chez les mammifères mais toujours de manière paralogue-spécifique.

Ces résultats établissent définitivement le rôle d'interfaces spécifiques à chaque groupe de paralogie pour la formation des complexes Hox-TALE chez les mammifères.



# Systematic analysis of molecular complex formation between Hox and TALE proteins reveals paralog-specific TALE interaction domains in human Hox proteins

Amélie Dard<sup>1</sup>, Françoise Bleicher<sup>1</sup>, Marilynne Duffraisse<sup>1</sup>, Catherine Lavau<sup>2</sup> and Samir Merabet<sup>1,#</sup>.

<sup>1</sup> IGFL, ENS Lyon, UMR5242, 46 Allée d'Italie, 69007 Lyon France.

<sup>2</sup> Pediatric Hematology Oncology, Duke University Medical Center, Research Drive, LSRC building, Room B-365, Box 91001, Durham, NC 27710, USA.

# corresponding author

## Abstract

Hox proteins achieve paralog-specific function with their generic Pbx and Meis cofactors, which belong to the TALE family of homeodomain-containing transcription factors. Cumulative data converge to the preponderant role of a highly conserved Hox protein motif called hexapeptide (HX) for Hox-TALE interactions. Few Hox proteins are also described to associate with the TALE cofactors without the HX motif, but whether this constitutes a general rule among Hox paralog groups in mammals remains unknown. In addition, no alternative TALE interaction domain has been characterized in mammalian Hox proteins so far. Here we systematically analysed TALE interaction properties of human and mouse Hox proteins from all paralog groups *in vitro* and in live cells. We demonstrate that the HX motif is dispensable for Hox-TALE interactions in all except the two most anterior paralog groups. We further identify alternative TALE interaction interfaces that are conserved at different evolutionary extents in human Hox proteins. Our findings definitively establish the general role of paralog-specific signatures for Hox-TALE interactions. These protein signatures could constitute a target of choice for abolishing the activity of specific Hox/TALE complexes in cancer.



## Introduction

Hox genes encode for homeodomain (HD)-containing transcription factors that regulate numerous developmental processes during embryogenesis [1][2]. This evolutionary conserved gene family has also key regulatory functions in the adult organism, in particular for maintaining the pool of stem cells in different cell lineages [3]. The diverse and specific transcriptional activities of Hox proteins often depend on the presence of the Pbx and Meis cofactors, which belong to the TALE (Three Amino acids Loop Extension) family of HD-containing TFs [4]. The role of TALE proteins as Hox cofactors was originally identified from mutations affecting the *Drosophila Pbx* ortholog gene called *extradentile (exd)* [5]. *Pbx* and *Meis* were subsequently found in humans from chromosomal translocations leading to acute leukemia [6][7][8][9]. *Pbx* and *Meis* are by now described to participate with Hox proteins in numerous developmental contexts in vertebrates [10] and invertebrates [11], regulating cell processes as diverse as apoptosis [12], differentiation [13] or proliferation [10][14][15]. Not surprisingly, these cofactors are also not neutral for the oncogenic potential of Hox proteins in several solid cancers [16] and leukaemia [17].

Given its major regulatory impact in development and disease, the partnership between Hox, Pbx and Meis proteins has been the subject of numerous molecular studies [18]. Original studies showed that vertebrate Hox proteins all except posterior paralog groups 11-13 could associate with Pbx on DNA [19][20][21]. In contrast, DNA-bound Hox/Meis complexes are only formed with Hox proteins from posterior paralog groups 9-13 *in vitro* [22], but their functional significance remains to be determined. In addition to these dimeric complexes, Hox proteins also assemble in trimeric complexes with Pbx and Meis [23]. These trimeric complexes are formed with Hox proteins from paralog groups 1 to 10, relying on Hox-Pbx interactions on the one side, and Pbx-Meis interactions on the other side [18]. Interaction with Meis allows Pbx to translocate into the nucleus *in vivo* and depends on evolutionary conserved domains located in the N-terminal part of both proteins [24]. Hox-Pbx interaction involves a short Hox protein motif located upstream of the HD and called hexapeptide (HX) [25][26][27][28][21][19]. This motif is present in all Hox members from paralog groups 1 to 10 and has been defined as corresponding to an invariant Trp residue located within a hydrophobic environment and followed by basic residues at +2 to +5 positions [29]. Crystal structures from vertebrate [30][31] and invertebrate [32][33] Hox/Pbx complexes converge to a preponderant role of the Trp residue for making strong hydrophobic contacts within the HD of Pbx. The absence of a conserved Trp residue upstream of the HD of Hox proteins from posterior paralog groups 11-14 is in accordance with their inability to form

a dimeric complex with Pbx *in vitro*. The role of the HX as a key Pbx-interaction motif is further exemplified by the study of HX-mimicking peptides, which were shown to abolish the formation and activity of Hox/Pbx complexes *in vitro* and in cancer cell cultures [34][35][36][37][38][39] [40].

Functional analyses of HX-mutated Hox proteins *in vivo* are more controversial. In particular, the *Drosophila* Hox proteins Ultrabithorax (Ubx) and AbdominalA (AbdA) were shown to carry out several Pbx-dependent functions in the absence of their HX motif [41][42][43], highlighting that other Hox protein region(s) could be used in addition to the HX motif to recruit the Pbx cofactor *in vivo*. Accordingly, two additional Pbx interaction motifs were found in Ubx and AbdA. One motif called “UbdA” is located downstream of the HD in Ubx and AbdA proteins of protostome lineages [44] and was recently solved in a Ubx/Exd crystal structure [45]. The other motif corresponds to a duplicated HX motif that is specifically conserved in insect AbdA proteins [46].

Additional work revealed that three other *Drosophila* Hox proteins (Sex combs reduced, Antennapedia and Abdominal-B) and five over the 39 mouse Hox proteins (Hoxb6, Hoxb7, Hoxb8, Hoxa9 and Hoxa10) could also interact with Pbx in the absence of the HX motif [46]. This dispensability was often revealed in the presence of Meis, underlining its major impact for uncovering HX-independent interaction modes. This property was also observed with cnidarian Hox proteins [47], highlighting its ancestral origin during animal evolution. Altogether, these results led to the proposition that different interaction modes between Hox and TALE proteins could be important for paralog-specific functions of Hox proteins [48][49][50]. Although very appealing, this model awaits further experimental validation. In particular, neither Hox proteins from all Hox paralog groups have been tested for their capacity to interact with the TALE cofactors in the absence of the HX motif, nor additional TALE interaction domains have been described in non-*Drosophila* Hox proteins. As a consequence, whether Hox-TALE interactions could rely on specific Hox proteins domains and whether this complexity could constitute a general rule among the different Hox paralog groups remain to be established.

Here we describe a systematic analysis of the interaction properties of human and mouse Hox proteins with the Pbx and Meis cofactors, tacking into account all Hox paralog groups (Figure 1). Our work relies on the analysis of Hox/TALE complex formation by using consensus DNA binding sites *in vitro* and Bimolecular Fluorescence complementation in human live cells. Results show that Hox proteins from all except anterior paralog groups 1 and 2 have the ability to interact with the TALE cofactors in the absence of the HX motif.

Molecular dissection of complex formation of four different Hox/TALE complexes further reveals the existence of several paralog-specific TALE interaction domains in human Hox proteins. This work definitively establishes the existence of paralog-specific Hox-TALE interaction modes, which has strong consequences for the development of future therapeutic tools against Hox/TALE-induced cancers and leukaemia.

## Results

### **Hox proteins from paralog groups 3 to 12 interact with the TALE cofactors in the absence of the HX motif *in vitro***

To assess the role of the HX motif for TALE cofactors recruitment *in vitro*, we performed band shift experiments on consensus nucleotide sites that are preferentially recognized by anterior, central or posterior Hox/TALE complexes (see materials and method). Complex formation with the Pbx1 or Pbx1 and Meis1 cofactors was analysed with one or several representative members of all mammalian Hox paralog groups, in the presence or absence of the HX motif (Figure 1).

Results show that the HX mutation strongly affects dimeric and trimeric complex formation for human Hox proteins from paralog groups 1 and 2 only (Figure 2A-B). Hox proteins from paralog groups 3 to 10 are all able to interact with the TALE cofactors in the absence of their HX motif (Figure 2C-L). This dispensability is only observed in the context of the trimeric complex, highlighting that Meis has a critical role for uncovering HX-independent interaction modes, as noticed elsewhere [46]. Surprisingly, the HX mutation can also significantly increase the formation of trimeric complexes *in vitro*, as observed for HoxA5, HoxB6 and HoxD10 (Figure 2F, G, and L). Thus, the HX motif could have a dual activity, promoting or inhibiting the interaction with TALE cofactors in the context of a dimeric or trimeric complex, respectively.

Finally, complex formation with the TALE cofactors was analysed with the posterior HoxA11, HoxD12 and HoxB13 proteins, which are naturally devoid of any HX-like signature upstream of the HD. These posterior Hox proteins do not form dimeric complexes with Pbx1, but do form a trimeric complex with Pbx1 and Meis1 in the case of HoxA11 and HoxD12 (Figure 2M-O). Taken together, results obtained from *in vitro* band shift experiments demonstrate a common dispensability of the HX motif for trimeric complex formation with the TALE cofactors among mammalian paralog groups.

### **Hox proteins from paralog groups 3 to 13 interact with the TALE cofactors in the absence of the HX motif in live cells**

Observations raised from band shift experiments are limited to a single DNA-binding event *in vitro*. We next assessed whether these observations could be reproduced at a larger scale level in live cells. To this end, we used the Bimolecular Fluorescence Complementation (BiFC) method, which allows analysing protein-protein interactions *in vivo*. BiFC relies on the property of monomeric fluorescent proteins to be reconstituted from two separate sub-fragments upon spatial proximity [51]. This property has been used with different types of proteins and in different cell and animal contexts [52][53]. In particular, the sensitivity and specificity of BiFC has previously been established for analysing Hox-TALE interactions in mouse COS7 cells [46].

Here we fused each Hox protein to the N-terminal fragment of the Venus fluorescent protein (making VN-Hox fusion proteins) and analysed their interaction property with Pbx1 fused to the complementary C-terminal fragment of Venus (VC-Pbx1; see materials and methods). BiFC was performed in live HEK293T cells, which derive from human embryonic kidney. These cells endogenously express Meis1 but not Pbx1 [54], allowing doing BiFC in the context of trimeric complexes. Fusion proteins were systematically co-expressed with a red fluorescent marker to assess for the efficiency of the transfection and to normalize the BiFC fluorescent levels between the different conditions (see materials and methods). Finally, BiFC was quantified in conditions where the fluorescence level was never at saturation with wild type Hox proteins (see materials and methods).

Under these experimental conditions, the loss of interaction was never as strong as in band shift experiments, which could in part be explained by the fact that BiFC is stabilizing the protein complex once weak protein-protein interactions are established. Still, BiFC is globally reproducing *in vitro* observations, with a strong decrease of the fluorescent level (50% or more) in the case of the HX mutation in the anterior HoxA1, HoxB1 and HoxB2 proteins (Figure 2A'-C'). Exception was for HoxB13, which produced BiFC signals with Pbx1 although no complex could be observed *in vitro* (Figure 2O-O'). This discrepancy suggests that HoxB13 and TALE proteins have the potential to interact together in live cells and that the nucleotide probe used for *in vitro* binding assay was potentially not appropriate for their trimeric association. The other Hox proteins display little (less than 30%) or no variation in the fluorescent level when mutated in their HX motif (Figure 2D'-N'). Of note, a stronger BiFC signal resulting from the HX mutation was also observed with Hoxd4 and

HoxB6, highlighting that the HX motif can have a general inhibitory impact on the Hox-TALE interaction in live cells (Figure 2E' and G').

Altogether, BiFC analysis in HEK cells establishes that the HX motif can be largely dispensable for Hox-TALE interactions in all except two paralog groups. Since only two alternative TALE interaction motifs have been described in *Drosophila* Hox proteins so far, we next investigated efforts to determine which part(s) of human Hox proteins could contain additional TALE interaction motifs.

Our approach was first based on the *in silico* prediction of Hox protein structure and motifs by using the SLiMPred program [55]. The analysis shows that paralog groups differ by the overall structure and motif organisation in Hox members. For example, mammalian Hox proteins from paralog groups 2 to 6 have in general four or more predicted motifs (in addition to the HX motif) and no long ordered protein regions (20 residues long or more, in addition to the HD), as exemplified in the human HoxA5 protein (Supplementary Figure 1). In contrast, members from paralog groups 1, 7 and 8 have rarely more than two short conserved motifs, but have two or more long ordered protein regions, as noticed in human HoxA7 (Supplementary Figure 2) and HoxC8 (Supplementary Figure 3) proteins. Finally, Hox proteins from posterior paralog groups 9-13 display an intermediary composition, with the presence of three to four short motifs and two to three long ordered regions, as noticed in the human HoxA9 protein (Supplementary Figure 4). Of note, all Hox proteins contain long disorder protein regions. Given these observations, we chose human Hox proteins that are representative of these three different types of organisation for molecular dissection with TALE cofactors.

### **The human HoxA5 protein contains an alternative TALE-interaction interface in the HD-containing C-terminal part**

The representative paralog member with several motifs and no long ordered region (with the exception of the HD) we tested is the human HoxA5 protein. Two different deleted forms were generated, removing the first 90 or 160 residues at the N-terminus (Figure 3A). The C-terminal part downstream of the HD of HoxA5 was not considered since it contains only 15 residues with no conserved predicted motifs (Supplementary Figure 1).

Results show that none of the N-terminal deletions affects the interaction with the TALE cofactors either *in vitro* or in live cells (Figure 3A'-A''). Removing the first 160 residues increases monomere binding on DNA, which could explain the higher level of complex formation when compare to the dN90 deleted form (Figure 3A'). The additional

mutation of the HX motif in the context of the dN160 deletion has also no significant effect *in vitro* and in live cells (Figure 3A'-A'').

From this analysis, we conclude that none of the predicted short motifs in the N-terminal region of HoxA5 are important for the interaction with TALE cofactors. Given the absence of conserved paralog-specific signatures in the short C-terminal part (Supplementary Figure 1), we suggest that the HD of HoxA5 is likely responsible of the TALE interaction properties observed in the absence of the HX motif. This role could be more particularly assumed by paralog-specific residues, as found in the human HoxA9 protein (see below).

### **The human HoxA7 protein contains an alternative TALE-interaction motif in the N-terminus part**

The representative paralog member with few motifs and long ordered regions we tested is the human HoxA7 protein. As for HoxA5, the C-terminal part downstream of the HD is relatively short, non-ordered and with no relevant predicted protein motif (Supplementary Figure 2). We thus generated two N-terminal deletions that remove one or the two predicted short motifs in the N-terminal part of HoxA5 (Figure 3B).

Results show that the deletion removing the first motif has no effect on the interaction with TALE cofactors (Figure 3B'-B''). By comparison, the deletion removing the second motif affects complex formation, both *in vitro* (with a loss of 50% of trimeric complexes: Figure 3B') and in live cells (with a loss of 80% of BiFC: Figure 3B'').

Effects observed with the large deletion suggest that the second short protein motif could be involved in the interaction with TALE cofactors. This motif contains the GYGAGAGAF sequence that is conserved only in mammalian Hox proteins from the paralog group 7 (Supplementary Figure 2). We thus specifically deleted this sequence and analysed the resulting effects on TALE recruitment, with or without the HX mutation in the context of full length HoxA7. Results show that the micro-deletion of the GA motif leads to a similar loss of complex formation as the large N-terminal deletion *in vitro* (Figure 3B'), and to a significant loss of BiFC in live cells (with 50% loss: Figure 3B''). More surprisingly, the simultaneous mutation of the HX motif has a tendency to rescue the interaction potential of the GA-mutated HoxA7 protein *in vitro* and in live cells (Figure 3B'-B'').

In conclusion, the analysis with HoxA7 reveals the existence of a new TALE-interaction motif that is specific of the paralog group 7 in mammals. The effects observed in the context of the HX mutation suggest that the two TALE interaction motifs are not simply

working in a redundant manner but could have opposite activities against a potential third TALE interaction interface present elsewhere in HoxA7.

### **Interaction between the human HoxC8 protein and TALE cofactors involves a short C-terminal domain in HoxC8**

The second representative paralog member with few motifs and long ordered regions we tested is the human HoxC8 protein. HoxC8 has only two predicted short motifs in addition to the HX motif: one corresponding to the conserved MSSYF Hox activation motif located at the N-terminus [56]. The other motif is Gln-rich and is located in the last 30 residues, downstream of the HD (Supplementary Figure 3). According to these predictions, we generated two different deletions, one removing most of the N-terminal region upstream of the HX motif, and the other removing the entire C-terminal part downstream of the HD (Figure 3C).

Results show that the long N-terminal deletion has a tendency to increase complex formation *in vitro* (Figure 3C'). This tendency is confirmed in live cells with fluorescent signals that are two times stronger than with wild type HoxC8 (Figure 3C''). In contrast, the C-terminal deletion leads to a strong loss of complex formation *in vitro* (with 10% of remaining complexes: Figure 3C') and a significant decrease of BiFC signals in live cells (with 50% loss of fluorescence: Figure 3C'). To assess whether the C-terminal region could act redundantly with the HX motif, we examined the effect of the HX mutation in the context of the C-terminally deleted HoxC8 protein. As observed with HoxA7, the HX mutation is reversing the effect of the motif deletion, with some partial rescue of complex formation *in vitro* and in live cells (Figure 3C'-C'').

These results demonstrate that the predicted coiled-coil domain downstream of the HD of HoxC8 constitutes an additional TALE interaction interface. Interestingly, this domain is only present in HoxC8 and not in the other Hox members of the paralog group 8 in mammals, highlighting a sub-paralog specific level of conservation (Supplementary Figure 3).

### **Interaction between HoxA9 and TALE cofactors involves paralog-specific residues of the HoxA9 HD**

The last Hox protein we dissected for interaction with TALE cofactors is the human HoxA9 protein, which has an intermediate organisation with the presence of four conserved predicted motifs and long ordered regions upstream of the HX motif (Supplementary Figure 4). The C-

terminal part of HoxA9 is very short (seven residues long) with no conserved motif (Supplementary Figure 4), and was therefore not considered in the analysis.

According to this organisation, we generated a long N-terminal deletion removing three of the four predicted motifs and analysed the resulting interaction properties with TALE cofactors (HoxA9dN190: Figure 4A). Results show that this truncated form is still able to interact with the TALE cofactors *in vitro* (Figure 4B) and in live cells (Figure 4B'), pointing to a preponderant role of the HD as an important TALE interaction interface. This role was confirmed by doing band shift experiments (Figure 4B) and BiFC (Figure 4B') with a minimal HoxA9 protein limited to the HD. The role of the HD of HoxA9 as a main TALE interaction interface was further validated by testing a wild type or HX-mutated chimeric HoxA1 protein containing the HD of HoxA9 in place of the HD of HoxA1 (Figure 4A). In this case, the HD of HoxA9 is sufficient to rescue the HX mutation of HoxA1 (Figure 4C-C'). Reversely, placing the HD of HoxA1 in HoxA9 abolishes HX-independent interaction properties of HoxA9 (Figure 4A-D-D'). Altogether these results establish the HD of HoxA9 as a key interaction interface with TALE cofactors. Interestingly, the same swapping experiments between HoxA1 and HoxA9 proved the HD of HoxA9 essential for leukemogenesis [57].

We next analysed the role of residues that are specifically conserved in the paralog group 9 in vertebrates, with typical chemical properties, and in positions that are favourable for external protein-protein interactions (Supplementary Figure 5). This analysis led to the identification of six candidate residues (Figure 4E and Supplementary Figure 5). Three of these residues are located in the N-terminal arm (K4, C6 and P7), one in the loop between helices 1 and 2 (M24), one in the helix 2 (D29), and one in the recognition helix 3 (M56; Figure 6A). Each of these residues was mutated and the effect analysed with or without the HX mutation in the context of full length HoxA9 (Figure 4F-I, 4F'-I' and Supplementary Figure 6). Band shift experiments showed that the interaction was strongly affected when both the HX motif and residue M24 (Supplementary Figure 6), D29 or M56 were mutated (Figure 4G-H). To rule out a possible effect of the N-terminal arm, we analysed a HoxA9 protein mutated in both the C6 and P7 residues: although this double mutation affects complex formation when coupled to the HX motif mutation *in vitro*, no significant loss was observed in live cells (Supplementary Figure 6). We thus conclude that the N-terminal arm could be important on a particular DNA-binding site, but not at a more global scale level. No strong loss of BiFC was obtained with the other mutations (Figure 4G'-H' and Supplementary Figure 6). Still, given the effect of the single D29 and M56 mutation *in vitro*, we tested a



HoxA9 protein mutated in those two residues, in combination or not with the HX motif mutation. Results show that mutating the D29 and M56 residues, together with the HX mutation, strongly affects complex formation with TALE cofactors *in vitro* (with 25% of remaining trimeric complex: Figure 4I) and in live cells (with less than 10% of BiFC signals: Figure 4I'). Thus, the HX motif and residues D29 and M56 appear to have a redundant and additive role for the interaction with TALE cofactors. The emplacement and orientation of these three TALE interaction interfaces suggest that they could be involved in different contacts with the TALE cofactors (Supplementary Figure 7).

## **Discussion**

The role of TALE cofactors for Hox function is well established in numerous developmental and oncogenic contexts. In contrast, most of our knowledge on the molecular cues underlying this partnership is limited to a single and common Hox protein motif, the HX motif. Although this motif was originally described to be crucial for Hox-Pbx interactions, recent data demonstrated that it could be dispensable in some instances, in particular when Meis was present in the binding reaction. However, whether this could constitute a general property among Hox paralog groups remains to be determined. In addition, very few is known about the nature of alternative Hox protein domains/motifs that could be used for Hox-TALE interactions.

Our exhaustive analysis covering all Hox paralog groups with human and mouse Hox proteins definitively establishes that dispensability of the HX motif for Hox-TALE interactions is a general molecular property. Our molecular dissection of Hox/TALE complexes further reveals that the molecular strategies used for the Hox-TALE partnership are highly diverse and paralog-specific.

## **Versatility in Hox-TALE interaction properties**

*In silico* predictions revealed that the global organisation of mammalian Hox proteins was not identical depending on the paralog group considered. Not only could the proportion of long disordered regions but also the number of predicted short motifs and long order regions varies among the different paralog groups. Three different types of organisations could be distinguished: (i) presence of several SLiMs (four or more, in addition to the HX motif) with no long ordered protein regions (with the exception of the HD), as noticed in paralog groups 2 to 6, (ii) presence of two or more long ordered protein regions with few SLiMs (two or less), as noticed in paralog groups 1, 7 and 8, and (iii) presence of several SLiMs (three or more)

with long ordered regions, as noticed in paralog groups 9 to 13. These different types of organisation could reflect different molecular strategies for interacting with TALE cofactors, with the preferential use of SLiMs in paralog groups 2-6 or long ordered regions in paralog groups 9-13, for example.

We however observed that the global organisation of the Hox protein is not predictive of the molecular strategy used for the Hox-TALE interaction. In the case of HoxA5, the alternative mode interaction is not relying on one of the numerous predicted SLiMs, but on the region encompassing the HD and devoid of any predicted SLiM. Reversely, HoxA7 and HoxC8 are using one of their rare predicted motifs to interact with the TALE cofactors. It is interesting to note that these motifs are located either within an ordered (in HoxA7) or disordered (in HoxC8) protein region. Finally, interaction between HoxA9 and TALE proteins involves paralog-specific residues of the HD and not other predicted SLiMs or ordered domains.

The various Hox protein interfaces used for the interaction with TALE cofactors are systematically specific of the paralog group but conserved at different evolutionary extents. In the case of HoxA9, paralog-specific residues of the HD involved in the interaction with TALE cofactors are conserved in all vertebrate proteins. By comparison, the GAGA motif found in the human HoxA7 protein is present in all mammalian Hox members of the paralog group 7, while the E-rich motif of the human HoxC8 protein is also conserved in mammals, but only in the sub-group C of the paralog group 8. Since the mouse Hoxb8 protein is also known to interact with the TALE cofactors in the absence of the HX motif [46], it suggests that mammalian Hox proteins from a same paralog group could use different interfaces with the TALE cofactors.

### **Complex molecular cues underlie the relationship between the HX motif and other TALE interaction interfaces in Hox proteins**

It is naively assumed that different protein motifs involved in the same regulatory function could work in a redundant or additive manner. Previous work established that different motifs of the *Drosophila* AbdA protein could in fact have opposite activities for the same function [43]. Interacting partners were however not known in this case and the antagonistic activities of AbdA protein motifs could therefore potentially rely on different molecular interactions.

Here our work revealed the role of different Hox protein regions in interaction with the same cofactors. Surprisingly, the intra-molecular relationships existing between these different regions are again different depending on the Hox protein considered. Results

obtained with the human HoxA9 protein show that the TALE interaction property of paralog specific residues of the HD is redundant with the TALE interaction property of the HX motif. This level of redundancy is illustrated by the effect of the simultaneous mutation of the HX motif together with the residues D29 and M56 of the HD, which is the only context abolishing TALE interaction *in vitro* and in live cells. The independent contribution of these three different TALE interaction interfaces is exemplified by the analysis of chimeric HoxA1/HoxA9 proteins and by the observation that they are each located on distinct interfaces in 3D modelling. The HX motif can also be completely or almost completely dispensable for Hox-TALE interactions, underlying a preponderant role of the alternative motif. This was the case of HoxA7 with the GAGA motif and HoxC8 with the E-rich motif. Of note, complex formation was never abolished in the context of these two Hox proteins, suggesting the existence of additional TALE interaction interfaces. Finally, the HX motif can also inhibit Hox-TALE interactions since its mutation can significantly increase Hox/TALE complex formation *in vitro* (as noticed for Hoxd4, HoxA5, HoxB6 and HoxD10) and in live cells (as noticed for Hoxd4 and HoxB6). This negative activity was also observed to a less extent in the context of the GAGA- or E-rich motif mutation. Together, these observations highlight the various relationships that could exist between alternative TALE interaction interfaces and the HX motif for the Hox-TALE interaction. Although not addressed in this study, these complex intra-molecular relationships could potentially depend on the DNA-binding site and/or the cell context considered.

How TALE cofactors interact with different Hox protein regions remain to be determined. The fact that Meis is systematically required for alternative interaction modes with mammalian Hox proteins suggest that it could engage direct contacts in place of Pbx in the trimeric complex. Alternatively, interaction with Meis could change Hox interaction properties of Pbx. In any case, molecular contacts established in the context of the trimeric complex await the resolution of crystal structures. Our finding of TALE interaction interfaces in human Hox proteins could help designing minimal Hox protein fragments for such studies in the future.

### **Incidence of versatile Hox-TALE interaction properties for Hox/TALE activity in cancer**

Cumulative data have established that Hox and TALE proteins could cooperate in several solid cancers [16] and leukaemia [17]. This cooperative activities were described in several instances to be abolished when providing an HX-mimicking peptide in the cell culture

medium [34][35][36][37][38][39] [40], and it was proposed that this peptide could constitute a promising therapeutic tool against Hox-Pbx induced cancers. Since this peptide is targeting a common Hox protein motif, it can however not be used to abolish the activity of a specific Hox/TALE complex. This is problematic when considering that numerous Hox genes are dysregulated in the same cancer [58] and can display pro- or anti-oncogenic potential depending on the cell context considered [59]. For example, HoxA1 [60][61] and HoxB7 [62] promote proliferation and invasiveness while HoxA10 [63] has a tumour suppressor function in breast cancer.

Although this remains to be confirmed, our results suggest that the oncogenic activity of Hox/TALE complexes could also potentially rely on alternative TALE interaction motifs that are paralog-specific. Designing peptides targeting these paralog-specific signatures could therefore open new therapeutic perspectives for abolishing the formation hence the oncogenic activity of a specific Hox/TALE complex in a particular cancer cell type.

## **Materials and methods**

### **Protein constructs**

Hox and Pbx variants were generated by PCR from full-length complementary DNAs and restriction-cloned alone or in fusion with the N-terminal (VN) of Venus, or the C-terminal (CC) fragment of Cerulean in the PcDNA3 vector, respectively (see Table S1 for a complete list of all constructs). Complementation between VN and CC produces a Venus-like fluorescent signal, as previously described [64][65]. Primers used are available upon request. For all fusion constructs, a linker of five amino acids was added to separate the Venus fragment from the protein of interest. All constructs were sequence-verified before using.

### **Protein Expression and Electrophoretic Mobility Shift Assays**

Constructs cloned in the PcDNA3 vector were produced with the TNT-T7-coupled *in vitro* transcription/translation system (Promega). Production yields of wild type and mutated counterpart proteins were estimated by 35S-methionine labeling. EMSAs were performed as described previously [66]. Hox proteins from paralog groups 1-5 and 6-13 were respectively tested on the previously defined central (5'-ATGACAGCTCGGGAATGATTAATTGG CCAAATA-3') and posterior (5'-ATGACAGCTCGGGAATGATTTATGGCCCAAATA-3') nucleotide probes. The choice of the nucleotide probe for each paralog group was dictated by the formation of trimeric complexes with wild type Hox proteins. Although these probes

were defined according to the binding properties of mouse Hox/TALE complexes [46], we observed that human Hox proteins from central paralog groups 6 to 8 formed more affine DNA-binding complexes with TALE cofactors on the so-called posterior probe. The quantification of shift complexes between wild type and mutated Hox proteins was done by using Adobe Photoshop CS6 software.

### **BiFC analysis in HEK293T cells**

Human HEK293T cells were maintained in Dulbecco's modified Eagle's medium (1 g/L glucose, Invitrogen) supplemented with 10% foetal bovine serum (Invitrogen), 100 IU/ml penicillin, and 100 mg/ml streptomycin (Invitrogen) at 37°C in a humidified 5% CO<sub>2</sub> atmosphere. 24 h before transfection, 10<sup>5</sup> cells were plated on glass coverslips. Transfections were carried out using the JetPRIME reagent (Polyplus), with a total amount of 1,5 µg of DNA: 500 ng of the VN-Hox fusion vector, 500 ng of the VC-Pbx1 fusion vector, and 500 ng of the pCMV-mCherry vector. Coverslips were taken 20h after transfection and directly mounted on glass slides for image capture under a Zeiss LSM780 confocal microscope. Four to six different fields of cells were acquired under the same confocal parameters at the 20x objective from three independent experiments in each condition. Quantification of green (BiFC) and red (transfection efficiency) fluorescent in all nuclei was realized by using the histogram function of the ImageJ software. A mean ratio was established for each acquisition. Corresponding values are illustrated by a boxplot representation using R Software. Boxplot depicts the smallest value, lower quartile, median, upper quartile, and largest value for each condition.

### **Acknowledgements**

We thank Robert K. Slany and René Rezsöházy for cDNAs. This work was supported by Association pour la Recherche sur le Cancer (ARC), Fondation pour la Recherche Médicale (FRM), Ligue Nationale Contre le Cancer, centre National de Recherche Scientifique (CNRS) and Ecole Normale Supérieure (ENS) de Lyon.

### **Figure legends**

**Figure 1. Representation of Hox proteins with regard to their paralog group affiliation in mammals.** Hox proteins are schematized with their two generic signatures (boxes): the hexapeptide motif, followed by the homeodomain. Note that the HX motif is absent in Hox members of paralog groups 11-13. Hox proteins analysed in this study are surrounded by a

dotted line. They correspond to human Hox proteins except for Hoxd4, Hoxa11 and Hoxd12 that are from mouse. Mouse Hox proteins that have previously been analysed for their interaction properties with the TALE cofactors are shaded (Hudry et al., PlosBiology 2012).

**Figure 2. Interaction properties between Hox proteins and TALE cofactors *in vitro* and live HEK293T cells.**

**A-O.** Band shift experiments with wild type or HX-mutated Hox proteins, as indicated (coloured boxes). Presence of Pbx1 (P) or Meis1 (M) in the binding reaction is respectively indicated by grey and black boxes above the gel. Coloured arrows indicate the monomer binding when present. Grey and black arrows depict dimeric and trimeric complexes with Pbx1 or Pbx1 and Meis1, respectively. Black arrowhead indicates the Pbx1/Meis1 dimer. The percentage of complex formation with HX-mutated Hox proteins compared to wild type Hox proteins is indicated below the gel (see also materials and methods). Supershift experiments with an antibody recognizing the Flag tag of HoxA11, HoxD12 and HoxB13 was performed to verify the presence or not of the Hox proteins in the trimeric complex. **A'-O'.** BiFC between Pbx1 and wild type or HX-mutated Hox proteins, as indicated. An illustrative confocal picture taken from live transfected cells is given in each case. The quantification of BiFC signals (green) with HX-mutated Hox proteins is provided as a percentage of the signals measured with the corresponding wild type Hox protein (see also materials and methods). This quantification takes into account the efficiency of the transfection (as assessed with the red signal). Statistical values are indicated as a boxplot representation.

**Figure 3. Defining alternative TALE interaction signatures in human HoxA5, HoxA7 and HoxC8 proteins.**

**A-A''.** Alternative TALE interaction signature of HoxA5 resides in the HD-encompassing region. **A.** Scheme of the tested constructs. Predicted short protein motifs are illustrated by small boxes (see also Supplementary Figure 1). The HX motif (small box) and HD (large box) are surrounded in black. Mutation of the HX motif is highlighted in red. **A'.** Band shifts experiments with the various HoxA5 constructs, as indicated. Colour code and arrows are as in Figure 2. None of the deletions is affecting complex formation. Quantification with the deleted forms takes into account the smaller size of the complex. **A''.** BiFC in live HEK cells between Pbx1 and the various HoxA5 constructs, as indicated. One illustrative confocal capture is given for each condition. Quantification of the fluorescent signal is as in Figure 2 and represented as a boxplot. None of the deletions affects the fluorescent signal when compared to wild type HoxA5. **B-B''.** Alternative TALE interaction

signature of HoxA7 resides in a N-terminally located GAGAGAF conserved motif. **B.** Scheme of the tested constructs. The two N-terminally predicted motifs are illustrated by small boxes (see also Supplementary Figure 2). Brackets indicate the specific micro deletion of the GAGAGAF motif. The HX motif and HD are represented as in A. **B’.** Band shifts experiments with the various HoxA7 constructs, as indicated. Colour code and arrows are as in Figure 2. Only deletions removing the GAGAGAF motif significantly affect trimeric complex formation with the TALE cofactors. Note that the HX mutation reverses this effect. **B’’. BiFC** in live HEK cells between Pbx1 and the various HoxA7 constructs, as indicated. One illustrative confocal capture is given for each condition. Quantification of the fluorescent signal is as in Figure 2 and represented as a boxplot. Deletions removing the GAGAGAF motif significantly affect the fluorescent signal when compared to wild type HoxA7. This effect is slightly rescued by the additional HX mutation. **C-C’’. Alternative TALE interaction signature of HoxC8 resides in a C-terminally located E-rich conserved motif. C.** Scheme of the tested constructs. Representation of predicted motifs, HX motif and HD is the same as previously. The box corresponding to the C-terminal E-rich motif is larger since this conserved motif covers 17 residues (see Supplementary Figure 3). **C’.** Band shifts experiments with the various HoxC8 constructs, as indicated. Colour code and arrows are as in Figure 2. The C-terminal deletion removing the E-rich motif affects significantly trimeric complex formation with the TALE cofactors. The additional mutation of the HX motif slightly rescues complex formation in the absence of the E-rich motif. **C’’. BiFC** in live HEK cells between Pbx1 and the various HoxC8 constructs, as indicated. One illustrative confocal capture is given for each condition. Quantification of the fluorescent signal is as in Figure 2 and represented as a boxplot. The C-terminal deletions removing the E-rich motif significantly affect the fluorescent signal when compared to wild type HoxA7. This effect is slightly rescued by the HX mutation, as observed *in vitro*.

**Figure 4. Role of paralog specific residues of the HD of the human HoxA9 protein in Hox-TALE interaction. A.** Scheme of the deleted and chimeric proteins tested in B-D’’. See also Supplementary Figure 4. **B-D.** Band shifts experiments with the various HoxA9 constructs, as indicated. Colour code and arrows are as in Figure 2. A supershift against the Flag tag of the HoxA9 HD confirms association with TALE cofactors (last lane in B). **B’-D’.** BiFC in live HEK cells between Pbx1 and the various HoxA9 constructs, as indicated. One illustrative confocal capture is given for each condition. Quantification of the fluorescent signal is as in Figure 2 and represented as a boxplot. **E.** Scheme of the HoxA9 HD positions

that are paralog-specific and accessible for protein-protein interactions (highlighted in light red). See also Supplementary Figure 5. **F-I**. Band shifts experiments with the various HD-mutated forms of HoxA9, as indicated. Colour code and arrows are as in Figure 2. See also Supplementary Figure 6. **F'-I'**. BiFC in live HEK cells between Pbx1 and HD-mutated forms of HoxA9, as indicated. One illustrative confocal capture is given for each condition. Quantification of the fluorescent signal is as in Figure 2 and represented as a boxplot. See also Supplementary Figure 6. Mutation of residues D29 and M56, together with the HX mutation, is the only combination that strongly affects complex formation both *in vitro* and in live cells.

### Supplementary Figure legends

**Supplementary Figure 1.** Prediction and conservation of ordered (brown), disorder (blue) and short linear motifs (SLIMs, green) in the human HoxA5 protein. The level of conservation of each residue among vertebrate and invertebrate species is illustrated with the red bars. The sequence below is from mouse Hoxa5. Figure was obtained from SLiMPred: [http://bioware.ucd.ie/~compass/biowareweb/Server\\_pages/slimpred\\_legacy.php](http://bioware.ucd.ie/~compass/biowareweb/Server_pages/slimpred_legacy.php). Seven conserved short motifs are predicted upstream of the HX motif (green arrows). The hexapeptide motif (HX), homeodomain (HD) and the sites used for N-terminal HoxA5 deletions (black arrows) are indicated.

**Supplementary Figure 2.** A. Prediction and conservation of ordered (brown), disorder (blue) and short linear motifs (SLIMs, green) in the human HoxA7 protein. The level of conservation of each residue among vertebrate and invertebrate species is illustrated with the red bars. The sequence below is from mouse Hoxa7. Figure was obtained from SLiMPred: [http://bioware.ucd.ie/~compass/biowareweb/Server\\_pages/slimpred\\_legacy.php](http://bioware.ucd.ie/~compass/biowareweb/Server_pages/slimpred_legacy.php). Two conserved short motifs are predicted upstream of the HX motif (green arrows). The hexapeptide motif (HX), homeodomain (HD) and the sites used for N-terminal and internal HoxA7 deletions (black arrows) are indicated. B. Sequence alignments of the short predicted GA motif. Alignments were realized by using protein sequences from *Homo sapiens*, *Mus musculus* and *Danio rerio*. The motif is specifically conserved in mammalian Hox proteins from the paralog group 7.

**Supplementary Figure 3.** A. Prediction and conservation of ordered (brown), disorder (blue) and short linear motifs (SLIMs, green) in the human HoxC8 protein. The level of



conservation of each residue among vertebrate and invertebrate species is illustrated with the red bars. The sequence below is from mouse Hoxc8. Figure was obtained from SLiMPred:

[http://bioware.ucd.ie/~compass/biowareweb/Server\\_pages/slimpred\\_legacy.php](http://bioware.ucd.ie/~compass/biowareweb/Server_pages/slimpred_legacy.php).

Two conserved short motifs are predicted in the HoxC8 sequence (green arrows): one upstream of the HX motif and the other downstream of the HD. The hexapeptide motif (HX), homeodomain (HD) and the sites used for N-terminal or C-terminal deletions in HoxC8 (black arrows) are indicated. The C-terminal part contains an E-rich motif. B. Sequence alignments of the short predicted E-rich motif. Alignments were realized by using protein sequences from *Homo sapiens*, *Mus musculus* and *Danio rerio*. The motif is specifically conserved in mammalian HoxC proteins from the paralog group 8.

**Supplementary Figure 4.** Prediction and conservation of ordered (brown), disorder (blue) and short linear motifs (SLiMs, green) in the human HoxA9 protein. The level of conservation of each residue among vertebrate and invertebrate species is illustrated with the red bars. The sequence below is from mouse Hoxa9. Figure was obtained from SLiMPred:

[http://bioware.ucd.ie/~compass/biowareweb/Server\\_pages/slimpred\\_legacy.php](http://bioware.ucd.ie/~compass/biowareweb/Server_pages/slimpred_legacy.php).

Four conserved short motifs are predicted upstream of the HX motif in HoxA9. The hexapeptide motif (HX), homeodomain (HD) and the sites used for the N-terminal deletion in HoxA9 (black arrow) are indicated.

**Supplementary Figure 5.** Sequence alignment of Hox homeodomains in vertebrates.

Sequences are taken from *Homo sapiens*, *Mus musculus* and *Danio rerio*. Typical residues that are specifically conserved in Hox9 proteins are highlighted in red. Positions that are accessible or not for protein-protein interactions are denoted with a blue or white box respectively (according to Merabet *et al.*, *BioEssay* 2008).

**Supplementary Figure 6. Role of paralog specific residues of the HD of the human HoxA9 protein in Hox-TALE interaction.** A. Scheme of the HoxA9 HD positions that are paralog-specific and accessible for protein-protein interactions (highlighted in light red). B-G. Band shifts experiments with the various HD-mutated forms of HoxA9, as indicated. Colour code and arrows are as in Figure 2. B'-G'. BiFC in live HEK cells between Pbx1 and HD-mutated forms of HoxA9, as indicated. One illustrative confocal capture is given for each condition. Quantification of the fluorescent signal is as in Figure 2 and represented as a

boxplot. Only the C6P7 or M24 mutations, together with the HX mutation, affect significantly complex formation *in vitro*. These effects are however not reproduced by BiFC in live cells.

**Supplementary Figure 7. Position and orientation of paralog-specific residues in the human HoxA9 HD.** A. Side chain positioning of paralog-specific residues (highlighted in orange) in two differently rotated HoxA9 HDs. The Trp (W) of the HX motif lying upstream of the HD is indicated. B-C'. Side chain positioning and chemical properties of wild type (B-C) or mutated (B'-C') paralog-specific residues that are important for complex formation with TALE cofactors in the HoxA9 HD (surrounded by a dotted circle). Two different rotations of the HD are given. Non-charged/hydrophobic residues are in white, positively and negatively charged residues are in red or blue, respectively. The position, orientation and chemical properties of the D29 and M56 residues of the HD suggest that they could be involved in different contacts with TALE cofactors.

#### **Supplementary Table S1**

Constructs generated in the study. All constructs were cloned in the PcDNA3 vector. Restrictions sites used for the cloning of full-length cDNAs in fusion or not with fragments of Venus or Cerulean are indicated, as well as amino acids mutations.

#### **References**

- 1 Pearson, J.C. *et al.* (2005) Modulating Hox gene functions during animal body patterning. *Nat. Rev. Genet.* 6, 893–904
- 2 Hueber, S.D. and Lohmann, I. Shaping segments: Hox gene function in the genomic age. *BioEssays*, 30. (2008) , 965–979
- 3 Seifert, A. *et al.* (2015) Role of Hox genes in stem cell differentiation. *World J. Stem Cells* 7, 583–95
- 4 Mukherjee, K. and Bürglin, T.R. (2007) Comprehensive analysis of animal TALE homeobox genes: New conserved motifs and cases of accelerated evolution. *J. Mol. Evol.* 65, 137–153
- 5 Rauskolb, C. *et al.* (1993) Extradenticle, a Regulator of Homeotic Gene Activity, Is a Homolog of the Homeobox-Containing Human Proto-Oncogene Pbx1. *Cell* 74, 1101–1112
- 6 Kamps, M.P. *et al.* (1990) A new homeobox gene contributes the DNA binding domain of the t(1;19) translocation protein in pre-B all. *Cell* 60, 547–555

- 7 Nourse, J. *et al.* (1990) Chromosomal translocation t(1;19) results in synthesis of a homeobox fusion mRNA that codes for a potential chimeric transcription factor. *Cell* 60, 535–545
- 8 Moskow, J.J. *et al.* (1995) Meis1, a PBX1-related homeobox gene involved in myeloid leukemia in BXH-2 mice. *Mol. Cell. Biol.* 15, 5434–5443
- 9 Steelman, S. *et al.* (1997) Identification of a conserved family of Meis1-related homeobox genes. *Genome Res.* 7, 142–156
- 10 Moens, C.B. and Selleri, L. (2006) Hox cofactors in vertebrate development. DOI: 10.1016/j.ydbio.2005.10.032
- 11 Hombría, J.C.-G. and Lovegrove, B. (2003) Beyond homeosis--HOX function in morphogenesis and organogenesis. *Differentiation.* 71, 461–76
- 12 Domsch, K. *et al.* (2015) The HOX-Apoptosis Regulatory Interplay in Development and Disease. In *Current Topics in Developmental Biology* 114pp. 121–158
- 13 Cerda-Esteban, N. and Spagnoli, F.M. Glimpse into Hox and tale regulation of cell differentiation and reprogramming. , *Developmental Dynamics*, 243. (2014) , 76–87
- 14 Argiropoulos, B. and Humphries, R.K. (2007) Hox genes in hematopoiesis and leukemogenesis. *Oncogene* 26, 6766–76
- 15 Chen, J.Y. *et al.* (2016) Hoxb5 marks long-term haematopoietic stem cells and reveals a homogenous perivascular niche. *Nature* 530, 223–227
- 16 Bhatlekar, S. *et al.* (2014) HOX genes and their role in the development of human cancers. *J. Mol. Med. (Berl).* 92, 811–23
- 17 Sitwala, K. V *et al.* (2008) HOX proteins and leukemia. *Int. J. Clin. Exp. Pathol.* 1, 461–474
- 18 Mann, R.S. *et al.* Chapter 3 Hox Specificity. Unique Roles for Cofactors and Collaborators. , *Current Topics in Developmental Biology*, 88. (2009) , Elsevier Inc., 63–101
- 19 Chang, C.P. *et al.* (1995) Pbx proteins display hexapeptide-dependent cooperative DNA binding with a subset of Hox proteins. *Genes Dev.* 9, 663–674
- 20 Shen, W.F. *et al.* (1996) Hox homeodomain proteins exhibit selective complex stabilities with Pbx and DNA. *Nucleic Acids Res.* 24, 898–906
- 21 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. Cell. Biol.* 15, 5811–9
- 22 Shen, W.F. *et al.* (1997) AbdB-like Hox proteins stabilize DNA binding by the Meis1

- homeodomain proteins. *Mol. Cell. Biol.* 17, 6448–58
- 23 Shanmugam, K. *et al.* (1999) PBX and MEIS as non-DNA-binding partners in trimeric complexes with HOX proteins. *Mol. Cell. Biol.* 19, 7577–88
- 24 Berthelsen, J. *et al.* (1999) The subcellular localization of PBX1 and EXD proteins depends on nuclear import and export signals and is modulated by association with PREP1 and HTH. *Genes Dev.* 13, 946–53
- 25 Shanmugam, K. *et al.* (1997) Residues flanking the HOX YPWM motif contribute to cooperative interactions with PBX. *J. Biol. Chem.* 272, 19081–7
- 26 Schnabel, C. a *et al.* (2000) HoxA9-mediated immortalization of myeloid progenitors requires functional interactions with TALE cofactors Pbx and Meis. *Oncogene* 19, 608–16
- 27 Phelan, M.L. *et al.* (1995) Cooperative interactions between HOX and PBX proteins mediated by a conserved peptide motif. *Mol. Cell. Biol.* 15, 3989–97
- 28 Neuteboom, S.T. *et al.* (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
- 29 Rieden, P.M.J.I. der *et al.* (2003) Homeodomain to hexapeptide or PBC-interactiondomain distance: size apparently matters. *Trends Genet.* 20, 76–79
- 30 Piper, D.E. *et al.* (1999) Structure of a HoxB1-Pbx1 heterodimer bound to DNA: role of the hexapeptide and a fourth homeodomain helix in complex formation. *Cell* 96, 587–97
- 31 Laronde-leblanc, N.A. and Wolberger, C. (2003) Structure of HoxA9 and Pbx1 bound to DNA : Hox hexapeptide and DNA recognition anterior to posterior. DOI: 10.1101/gad.1103303.with
- 32 Passner, J.M. *et al.* (1999) Structure of a DNA-bound Ultrabithorax-Extradenticle homeodomain complex. *Nature* 397, 714–719
- 33 Joshi, R. *et al.* (2007) Functional specificity of a Hox protein mediated by the recognition of minor groove structure. *Cell* 131, 530–43
- 34 Aulisa, L. *et al.* (2009) Inhibition of cancer cell proliferation by designed peptide amphiphiles. *Acta Biomater.* 5, 842–53
- 35 Morgan, R. *et al.* (2007) Antagonism of HOX/PBX dimer formation blocks the in vivo proliferation of melanoma. *Cancer Res.* 67, 5806–13
- 36 Plowright, L. *et al.* (2009) HOX transcription factors are potential therapeutic targets in non-small-cell lung cancer (targeting HOX genes in lung cancer). *Br. J. Cancer* 100,

- 37 Morgan, R. *et al.* (2014) Targeting HOX transcription factors in prostate cancer. *BMC Urol.* 14, 1–9
- 38 Morgan, R. *et al.* (2010) Targeting HOX and PBX transcription factors in ovarian cancer. *BMC Cancer* 10, 89
- 39 Morgan Richard, Boxall Angie, Harrington Kevin J., Simpson Guy R., Gillett Cheryl, Michael Agnieszka, P.H.S. (2012) Targeting the HOX/PBX dimer in breast cancer. *Breast Cancer res Treat* 136, 389–398
- 40 Li, Z. *et al.* (2013) PBX3 is an important cofactor of HOXA9 in leukemogenesis. *Blood* 121, 1422–31
- 41 Galant, R. *et al.* (2002) Hox repression of a target gene: extradenticle-independent, additive action through multiple monomer binding sites. *Development* 129, 3115–3126
- 42 Merabet, S. *et al.* (2007) A unique Extradenticle recruitment mode in the Drosophila Hox protein Ultrabithorax. *Proc. Natl. Acad. Sci. U. S. A.* 104, 16946–51
- 43 Merabet, S. *et al.* (2011) Insights into Hox protein function from a large scale combinatorial analysis of protein domains. *PLoS Genet.* 7, e1002302
- 44 Balavoine, G. *et al.* (2002) Hox clusters and bilaterian phylogeny. *Mol. Phylogenet. Evol.* 24, 366–373
- 45 Foos, N. *et al.* (2015) A Flexible Extension of the Drosophila Ultrabithorax Homeodomain Defines a Novel Hox/PBC Interaction Mode. *Structure* 23, 270–279
- 46 Hudry, B. *et al.* (2012) Hox proteins display a common and ancestral ability to diversify their interaction mode with the PBC class cofactors. *PLoS Biol.* 10, e1001351
- 47 Hudry, B. *et al.* (2014) Molecular insights into the origin of the Hox-TALE patterning system. *Elife* 3, e01939
- 48 Merabet, S. and Hudry, B. (2013) Hox transcriptional specificity despite a single class of cofactors: are flexible interaction modes the key? Plasticity in Hox/PBC interaction modes as a common molecular strategy for shaping Hox transcriptional activities. *Bioessays* 35, 88–92
- 49 Merabet, S. and Mann, R.S. (2016) To Be Specific or Not: The Critical Relationship Between Hox And TALE Proteins. *Trends Genet.* xx, 1–14
- 50 Merabet, S. and Galliot, B. (2015) The TALE face of Hox proteins in animal evolution. *Front. Genet.* 6,
- 51 Kerppola, T.K. (2006) Visualization of molecular interactions by fluorescence complementation. *Nat. Rev. Mol. Cell Biol.* 7, 449–56

- 52 Shyu, Y.J. and Hu, C.-D. (2008) Fluorescence complementation: an emerging tool for biological research. *Trends Biotechnol.* 26, 622–30
- 53 Kodama, Y. and Hu, C.-D. (2012) Bimolecular fluorescence complementation (BiFC): a 5-year update and future perspectives. *Biotechniques* 53, 285–98
- 54 Sultan, M. *et al.* (2008) A Global View of Gene Activity and Alternative Splicing by Deep Sequencing of the Human Transcriptome. *Science* (80-. ). 321, 956–960
- 55 Mooney, C. *et al.* (2012) Prediction of short linear protein binding regions. *J. Mol. Biol.* 415, 193–204
- 56 Tour, E. *et al.* (2005) Evolutionarily conserved domains required for activation and repression functions of the Drosophila Hox protein Ultrabithorax. *Development* 132, 5271–5281
- 57 Breitinger, C. *et al.* (2012) The homeodomain region controls the phenotype of HOX-induced murine leukemia. *Blood* 120, 4018–27
- 58 Makiyama, K. *et al.* (2005) Aberrant expression of HOX genes in human invasive breast carcinoma. *Oncol Rep* 13, 673–679
- 59 Shah, N. and Sukumar, S. (2010) The Hox genes and their roles in oncogenesis. *Nat. Rev. Cancer* 10, 361–71
- 60 Zhang, X. *et al.* (2006) HOXA1 is required for E-cadherin-dependent anchorage-independent survival of human mammary carcinoma cells. *J. Biol. Chem.* 281, 6471–81
- 61 Delval, S. *et al.* (2011) The Pbx interaction motif of Hoxa1 is essential for its oncogenic activity. *PLoS One* 6, e25247
- 62 Wu, X. *et al.* (2006) HOXB7, a homeodomain protein, is overexpressed in breast cancer and confers epithelial-mesenchymal transition. *Cancer Res.* 66, 9527–34
- 63 Carrio, M. *et al.* (2005) Homeobox D10 induces phenotypic reversion of breast tumor cells in a three-dimensional culture model. *Cancer Res.* 65, 7177–7185
- 64 Hu, C.-D. and Kerppola, T.K. (2003) Simultaneous visualization of multiple protein interactions in living cells using multicolor fluorescence complementation analysis. *Nat. Biotechnol.* 21, 539–45
- 65 Hudry, B. *et al.* (2011) Visualization of protein interactions in living Drosophila embryos by the bimolecular fluorescence complementation assay. *BMC Biol.* 9, 5
- 66 Pöpperl, H. *et al.* (1995) Segmental expression of Hoxb-1 is controlled by a highly conserved autoregulatory loop dependent upon exd/pbx. *Cell* 81, 1031–42

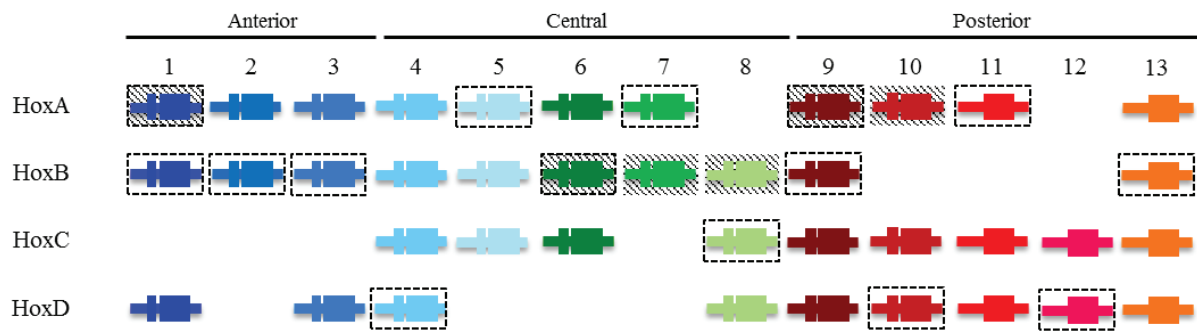


Figure 1

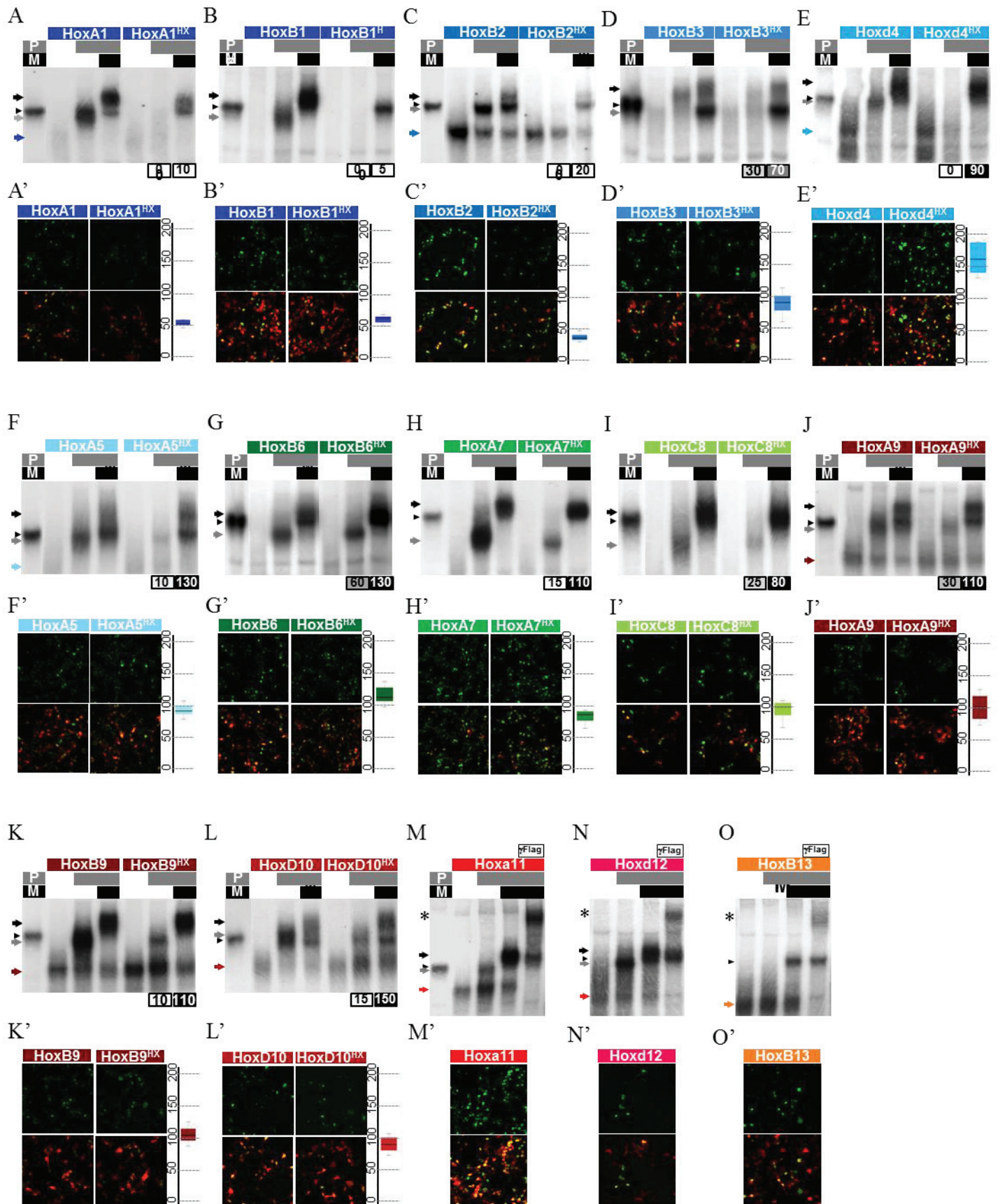


Figure 2



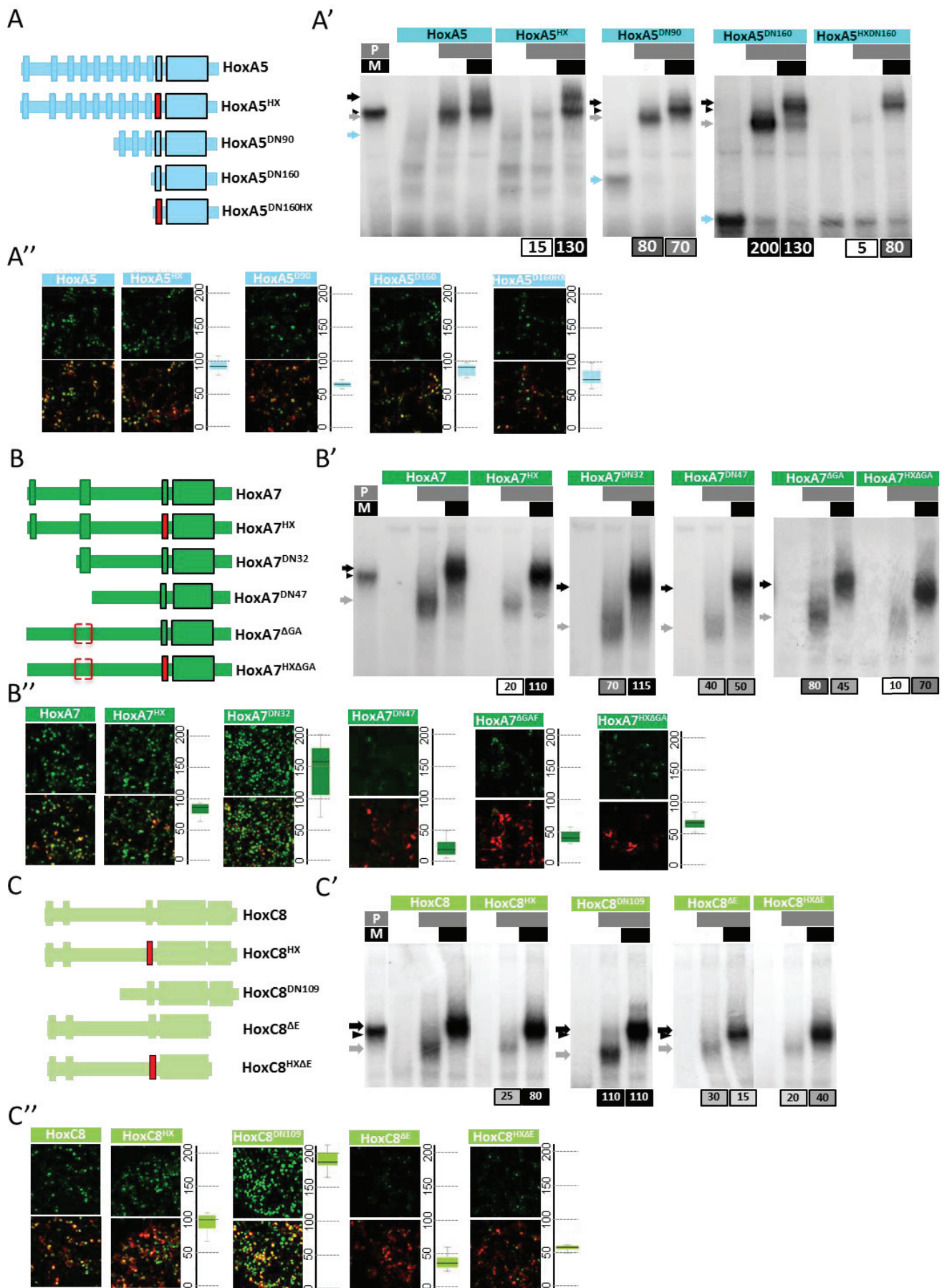


Figure 3

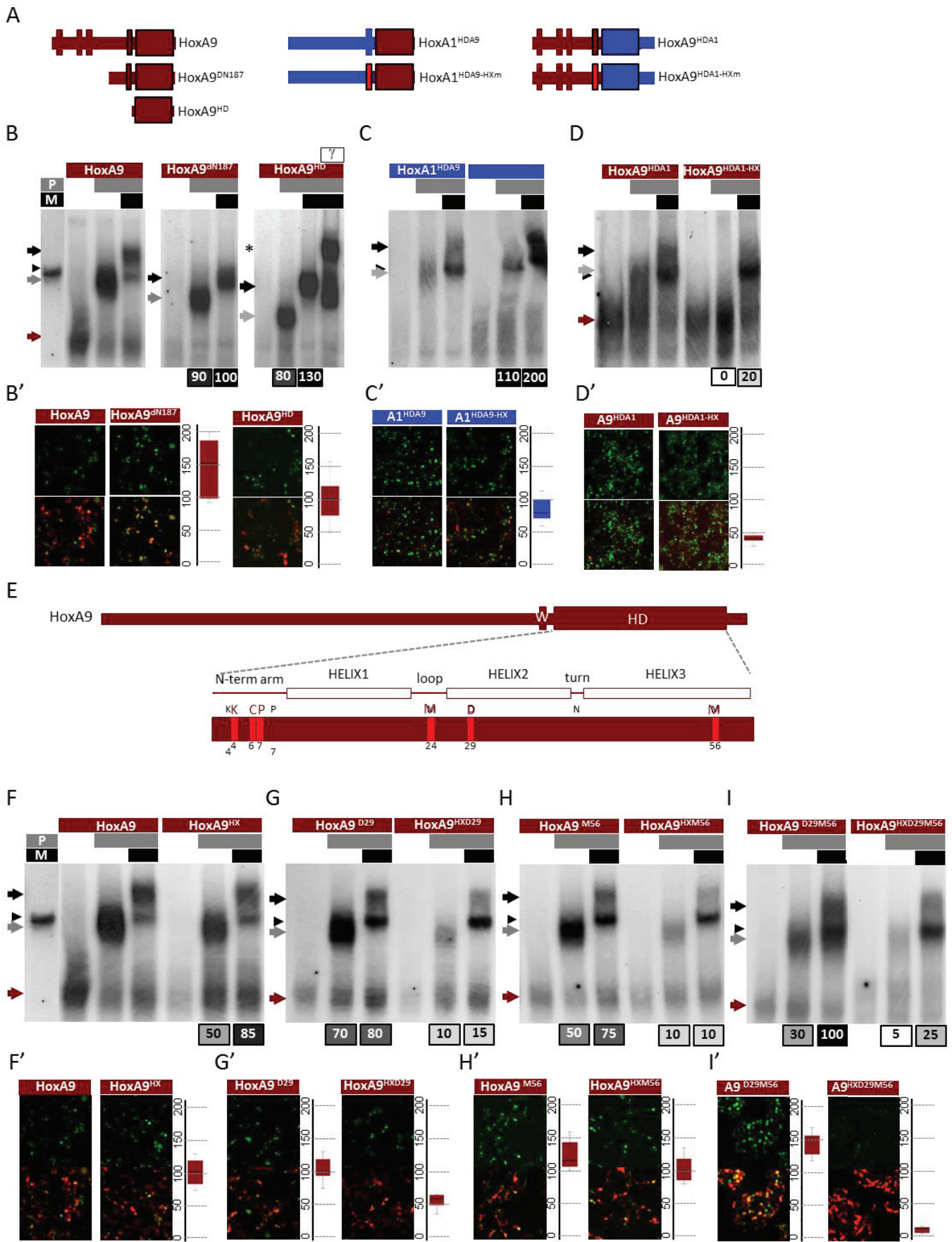
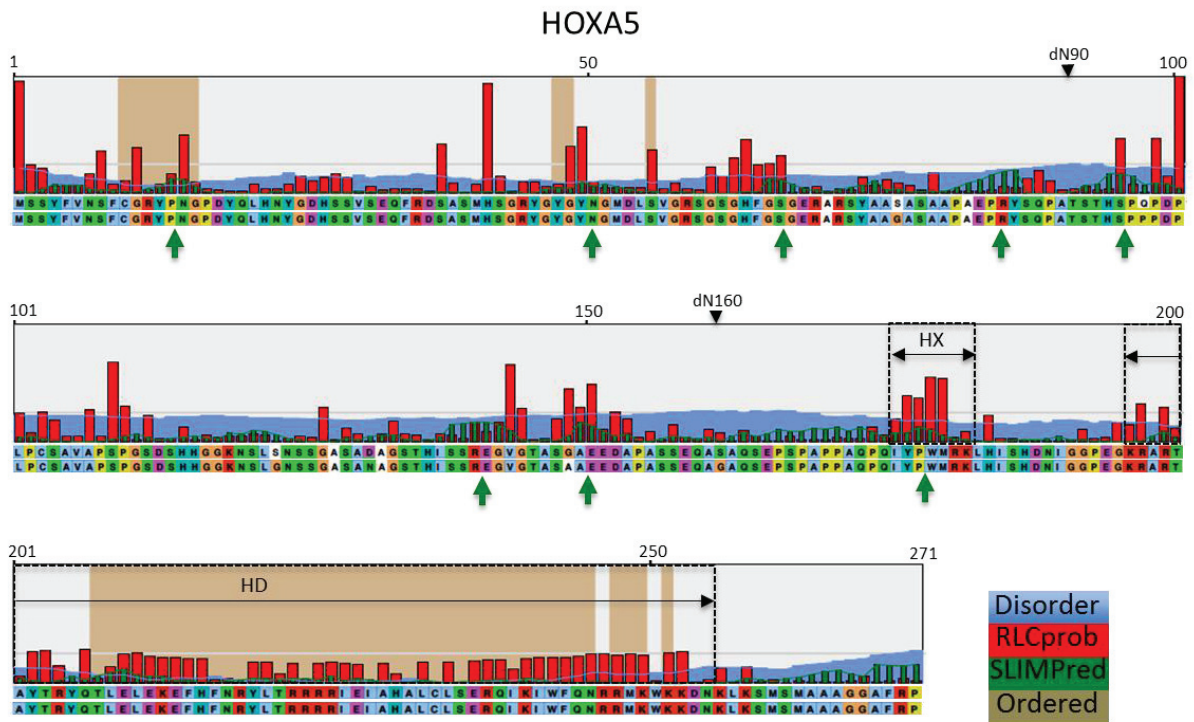
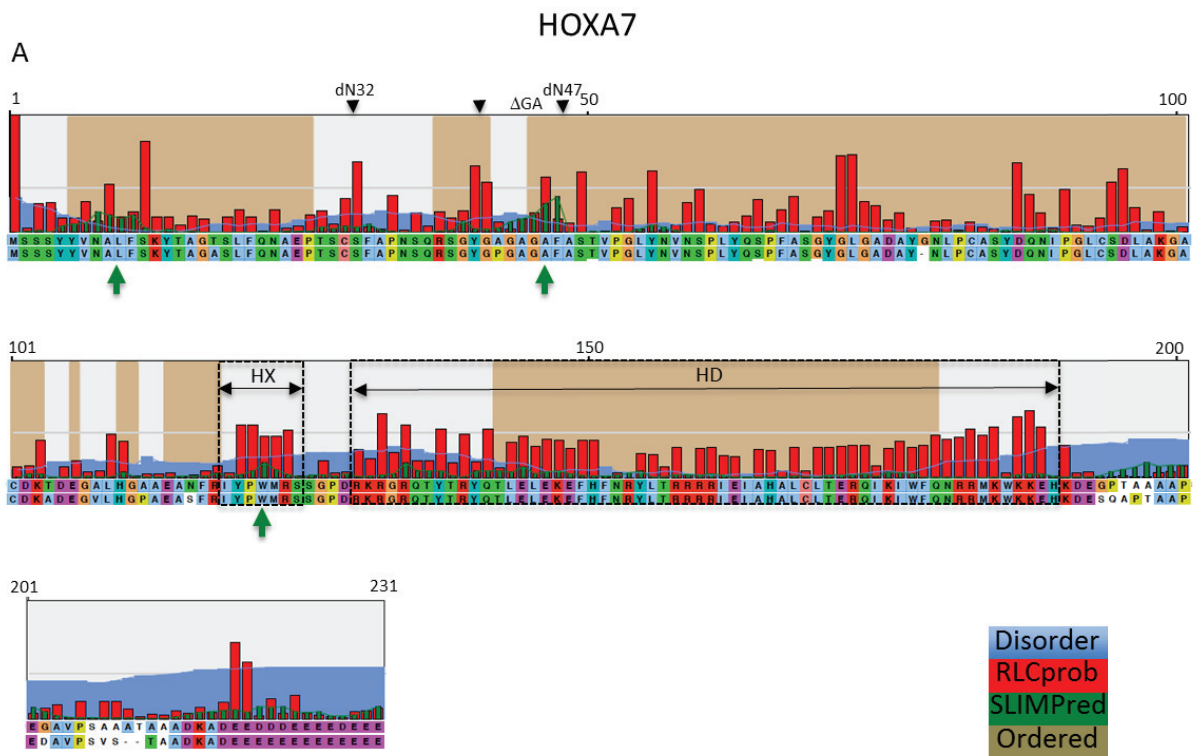


Figure 4



Supplementary Figure 1

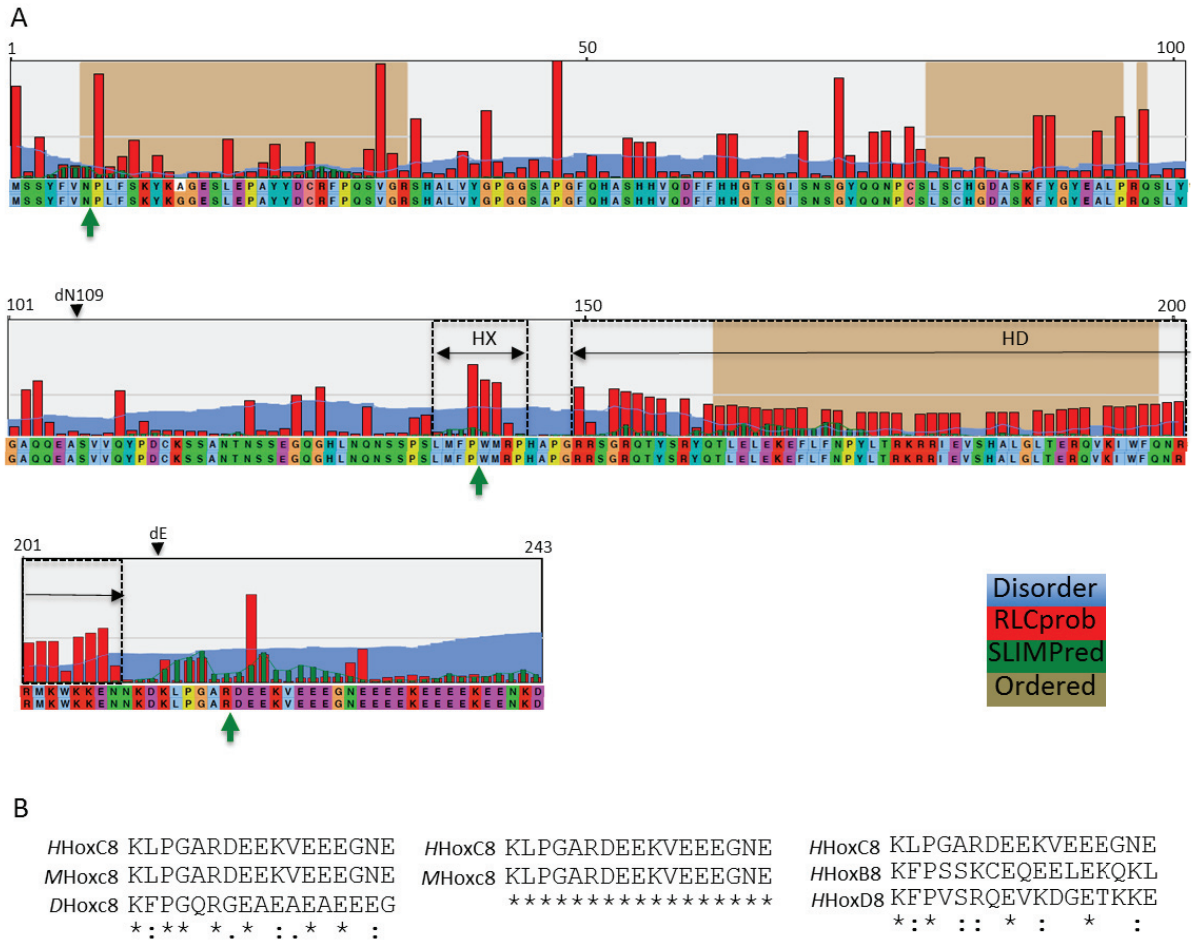


**B**

<i>HHoxA7</i> GYGAGAGAF	<i>HHoxA7</i> GYGAGAGAF	<i>HHoxA7</i> GYGAG-AGAF	<i>HHoxA7</i> GYGAGAGAF
<i>MHoxa7</i> GYGPGAGAF	<i>MHoxa7</i> GYGPGAGAF	<i>HHoxB7</i> GYGAGSGASF	<i>HHoxB8</i> PSSGGSFQH
<i>DHoxa7</i> YPSATFGAT	*** . *****	<i>MHoxb7</i> GYGAGPGAPF	<i>HHoxA6</i> G-ASSLPDK
.. **		***** . . . *	. . . .

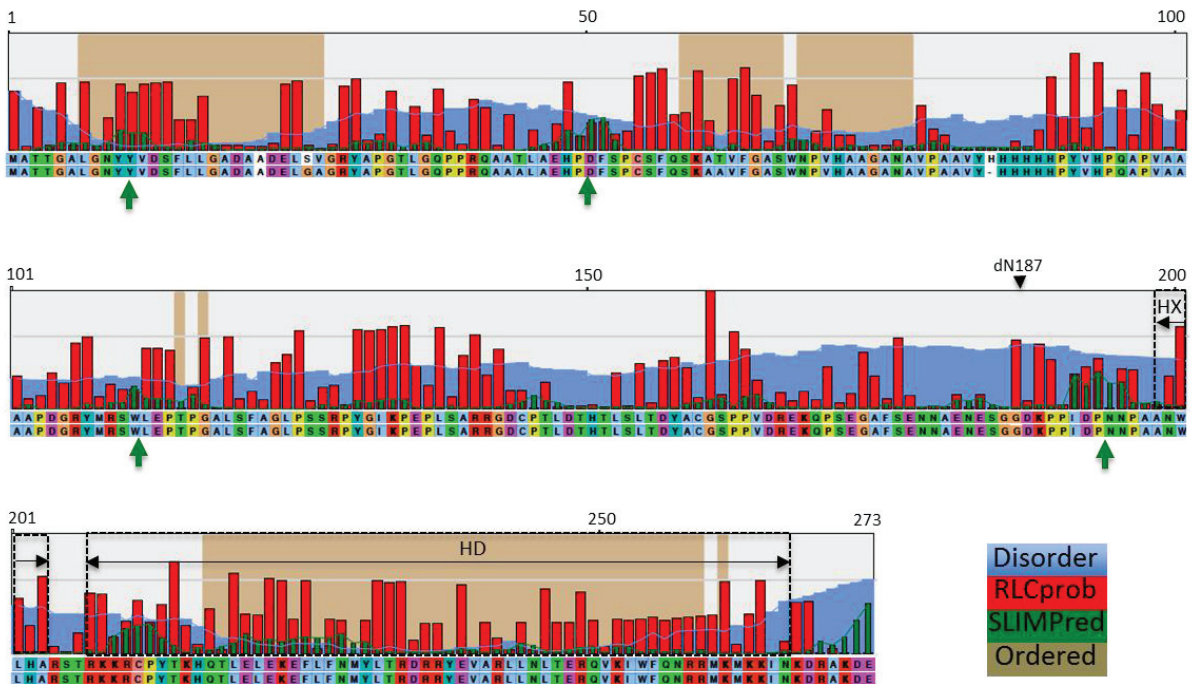
Supplementary Figure 2

### HOXC8

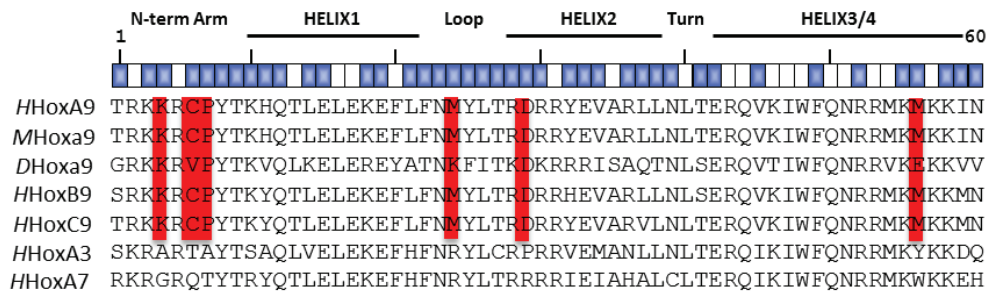


Supplementary Figure 3

### HOXA9

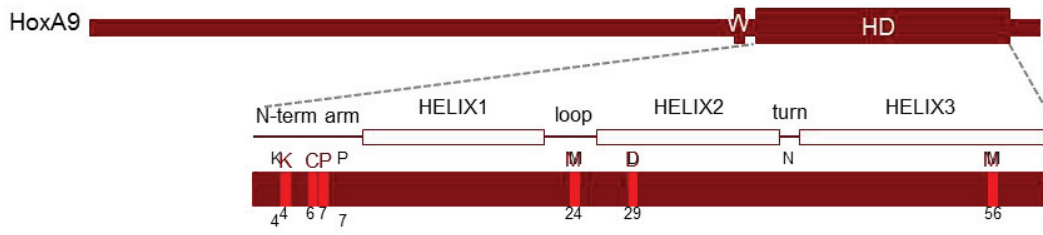


Supplementary Figure 4

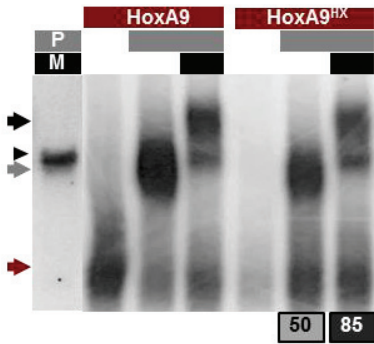


Supplementary Figure 5

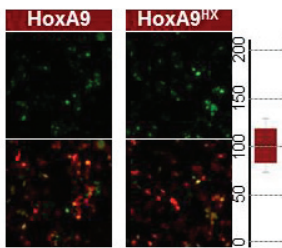
A



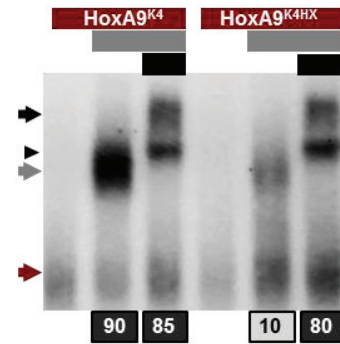
B



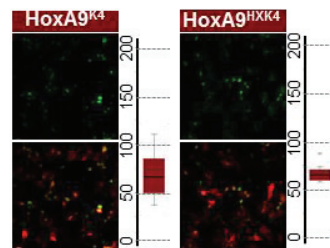
B'



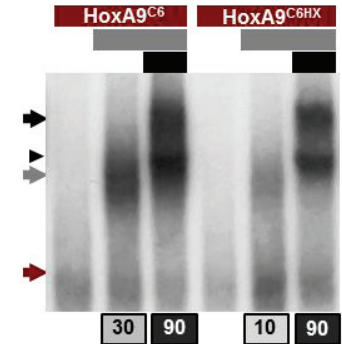
C



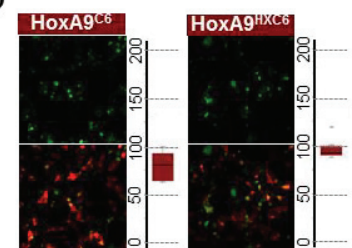
C'



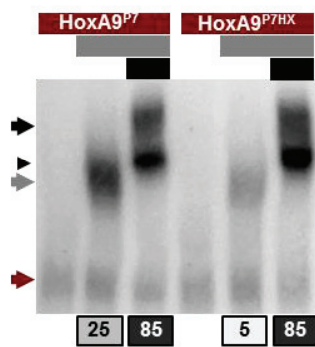
D



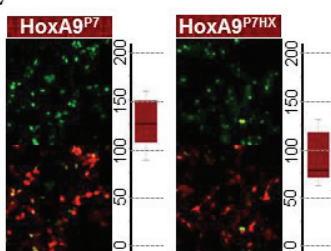
D'



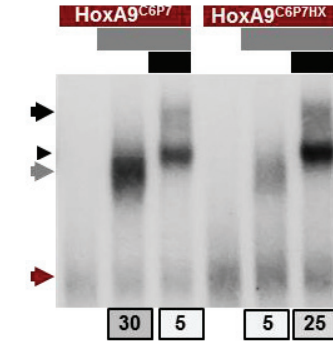
E



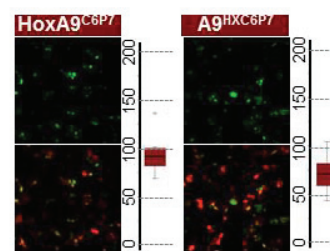
E'



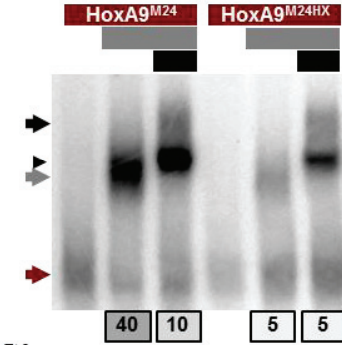
F



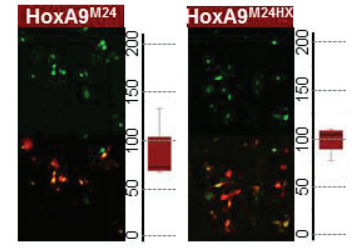
F'



G

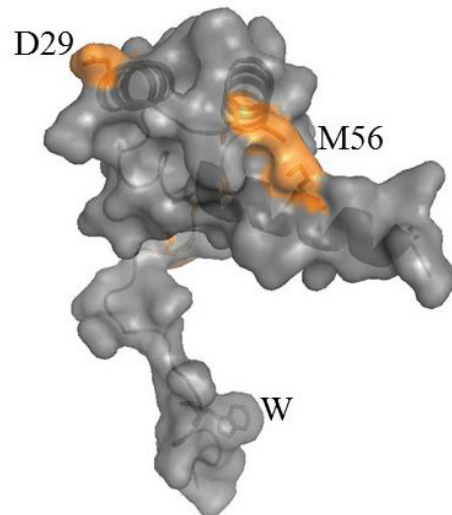
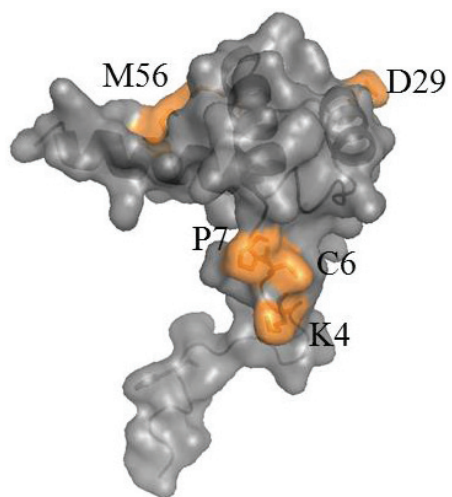


G'

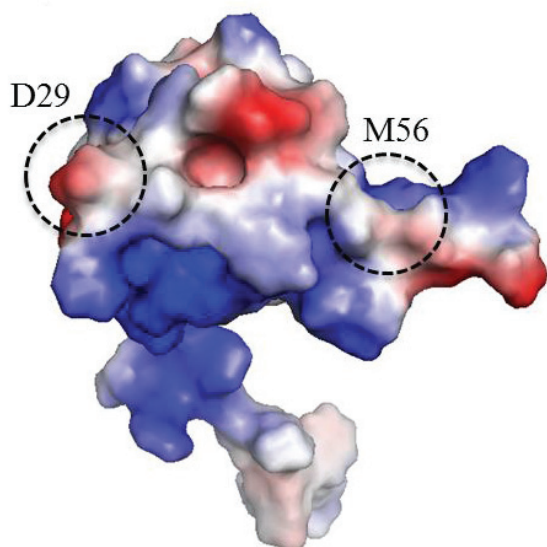


Supplementary Figure 6

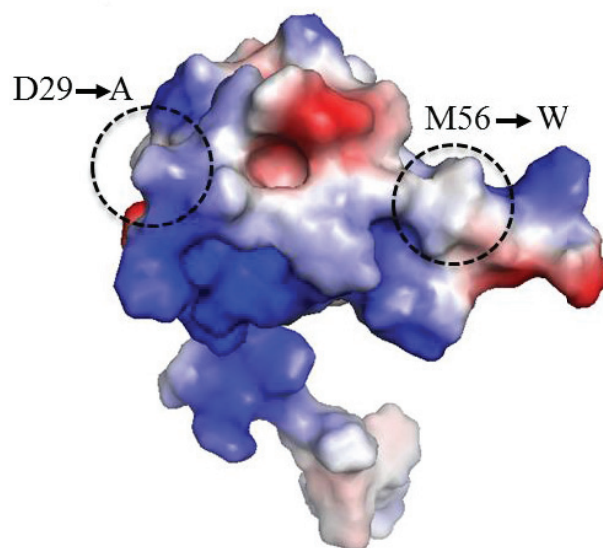
A



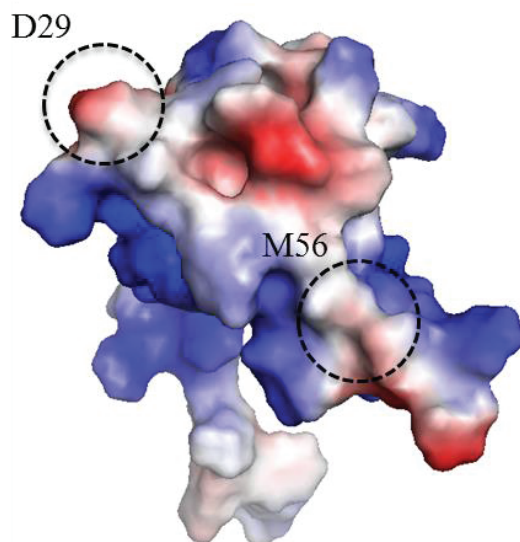
B



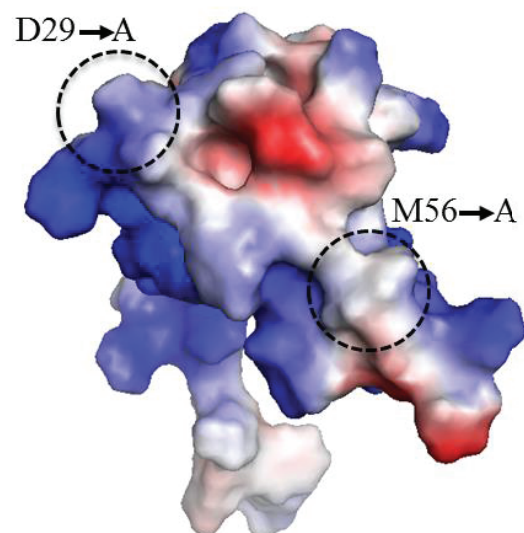
B'



C



C'



Supplementary Figure 7

#### **4. QUATRIEME CHAPITRE : TROUVER DE NOUVEAUX PARTENAIRES DES PROTEINES HOX SPECIFIQUES D'UN CONTEXTE CANCEREUX**

Nous avons séparé ce dernier chapitre en deux parties. La première est une revue discutant de la nécessité d'étudier les protéines HOX et leurs partenaires dans des contextes spécifiques. En effet, la plupart des cribles visant à identifier des partenaires des protéines HOX se sont avérés infructueux car ils se tenaient dans des contextes éloignés de ceux où les protéines HOX sont normalement exprimées. Nous savons maintenant que la plupart des rôles de ces facteurs de transcription peuvent reposer sur des interactions faibles et transitoires et il est alors nécessaire de se placer dans un contexte spécifique avec une méthode de détection sensible pour identifier de nouveaux partenaires de ces protéines.

Dans la deuxième partie, nous présentons le principe d'un crible d'interaction en lignées cellulaires humaines cancéreuses par BiFC, afin d'identifier de nouveaux partenaires des protéines HOX. Il repose sur la constitution de lignées cellulaires ayant intégré une banque d'environ 10 000 gènes humains compatible pour la BiFC. Ces lignées peuvent être transfectées avec n'importe quelle protéine d'intérêt afin de trouver les interacteurs au sein de cette banque de gènes modifiés. Une expérience pilote réalisée avec la protéine humaine HoxA9 a permis de trouver des partenaires spécifiques du contexte cellulaire. Cette approche a fait l'objet d'une demande de brevet dont la déclaration d'invention se trouve dans les annexes de cette thèse.





# Tracking Context-Specific Transcription Factors Regulating Hox Activity

Samir Merabet\* and Amélie Dard

**Background:** Hox proteins are key developmental regulators involved in almost every embryonic tissue for specifying cell fates along longitudinal axes or during organ formation. It is thought that the panoply of Hox activities relies on interactions with tissue-, stage-, and/or cell-specific transcription factors. High-throughput approaches in yeast or cell culture systems have shown that Hox proteins bind to various types of nuclear and cytoplasmic components, illustrating their remarkable potential to influence many different cell regulatory processes. However, these approaches failed to identify a relevant number of context-specific transcriptional partners, suggesting that these interactions are hard to uncover in non-physiological conditions. Here we discuss this problematic. **Results:** In this review, we present intrinsic Hox molecular signatures that are probably involved in multiple (yet specific) interactions with transcriptional partners. We also recapitulate the current knowledge on Hox cofactors, highlighting the difficulty to tracking context-specific cofactors through traditional large-scale approaches. **Conclusion:** We propose experimental approaches that will allow a better characterisation of interaction networks underlying Hox contextual activities in the next future. *Developmental Dynamics* 243:16–23, 2014. © 2013 Wiley Periodicals, Inc.

**Key words:** Hox; TALE; cofactors; transcription; interaction network; embryogenesis

## Key Findings:

- BiFC as a new tool for deciphering new interaction networks in vivo.

Accepted 11 June 2013

## INTRODUCTION

A major advance in biology was the discovery that the embryonic development of multicellular (metazoans) animals is controlled by a relatively small repertoire of conserved molecules. Among these, Hox proteins are unique in specifying distinct cell fates along the anterior-posterior (AP) axis of all bilaterians. The evolutionary conserved function of *Hox* genes has fascinated biologists for decades, starting with the genetic characterisation of the famous homeotic mutant phenotypes observed in nature (Lewis, 1978; Bender et al., 1983). These studies led to the discov-

ery of another striking property, the so-called collinear rule, whereby the genomic organisation of *Hox* genes mirrors their embryonic expression profile along the body plan (Harding et al., 1985; McGinnis and Krumlauf, 1992).

Besides early patterning functions, *Hox* genes are required later on during embryogenesis. For example, *Hox* genes are reiteratively used for the specification of respiratory (Philippidou et al., 2012) and limb-innervating (Lacombe et al., 2013) motor neurons, or for the formation of structures as different as the human brainstem, inner ear, and cardiac outflow tract

(Tischfield et al., 2005). Finally, Hox genes are also important for the homeostasis of several cell lineages in adults (Alharbi et al., 2013). In total, the patterning functions of Hox proteins are likely to represent a minor fraction of their wide spectrum of regulatory activities in the developing embryo (Hombria and Lovegrove, 2003).

Hox proteins are homeodomain (HD)-containing transcription factors (TFs) that preferentially recognize similar DNA-binding sites in vitro (Berger et al., 2008; Noyes et al., 2008). This molecular property contrasts with their highly specific

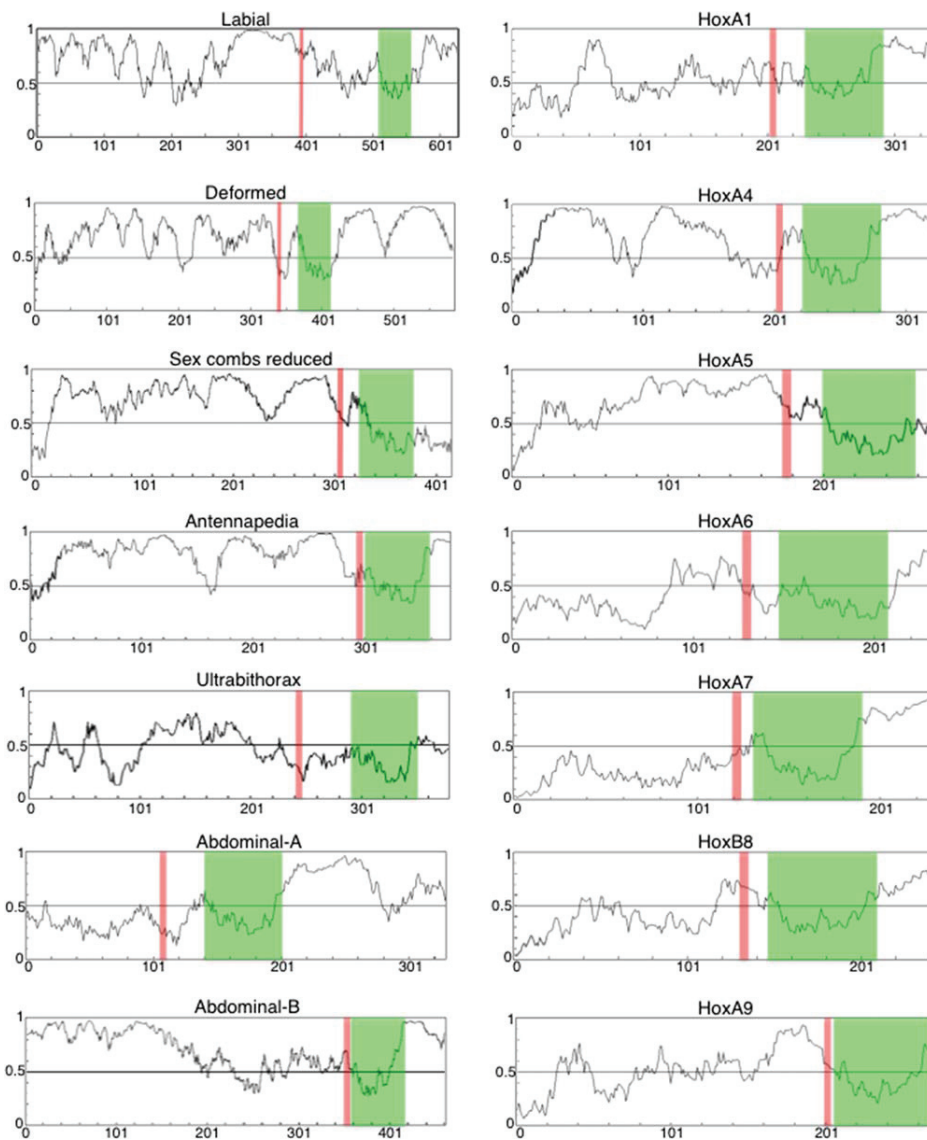
IGFL, UMR5242, CNRS/ENS Lyon, Lyon, France

Grant sponsor: CNRS; Grant sponsor: Ecole Normale Supérieure (ENS) of Lyon; Grant sponsor: Association de Recherche contre le Cancer (ARC); Grant sponsor: Fondation pour la Recherche Médicale (FRM).

\*Correspondence to: Samir Merabet, IGFL, UMR5242, CNRS/ENS Lyon, 46 Allée d'Italie, 69364 Lyon Cedex 07, France.  
E-mail: samir.merabet@ens-lyon.fr

DOI: 10.1002/dvdy.24002

Published online 24 June 2013 in Wiley Online Library (wileyonlinelibrary.com).



**Fig. 1.** Prediction of disordered regions in Hox proteins. Disordered regions in *Drosophila melanogaster* (left) and *Homo sapiens* (right) Hox proteins are predicted by the IUPred algorithm (Dosztanyi et al., 2005) and correspond to values above 0.5 (black line). Green and red boxes enclose the homeodomain and hexapeptide motif in each Hox protein, respectively. Note that long disordered regions are present in all Hox proteins.

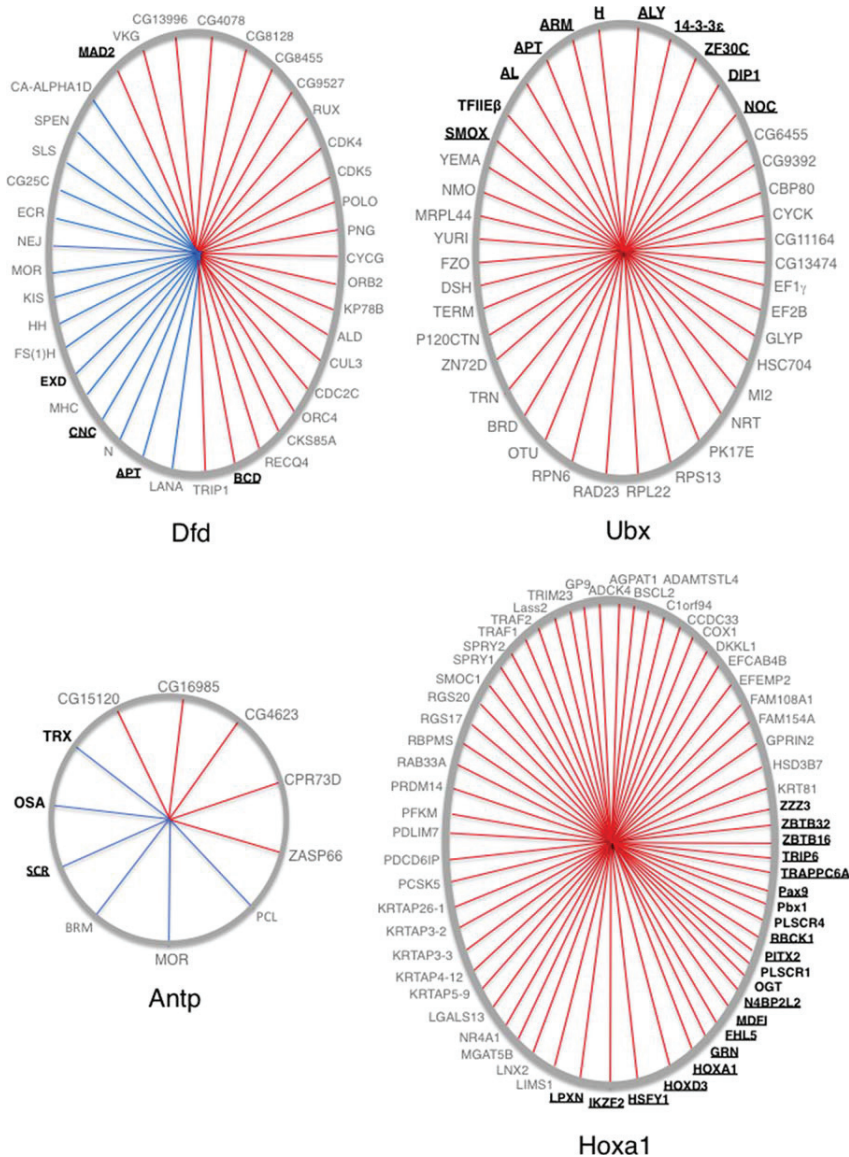
functions in vivo (the famous Hox paradox), implying that additional transcriptional partners, here referred to as context-specific cofactors, intervene for shaping Hox activities in a tissue-, stage-, and/or cell-dependent manner. However, until now only a few of the many postulated context-specific cofactors have been identified using high-throughput proteomic approaches, including protein complex purification from cell culture systems (Rubin et al., 2007; Guruharsha et al., 2011) or two-hybrid assays in mammalian cells (Rual et al., 2005; Lievens et al.,

2009; Ravasi et al., 2010; Lambert et al., 2012) and yeast (Bai et al., 2000; Giot et al., 2003; Stanyon et al., 2004; Formstecher et al., 2005; Bondos et al., 2006). Thus, it is assumed that context-specific Hox-cofactor interactions can only be identified within the physiological environment.

Herein, we discuss this regulatory aspect of Hox proteins, recapitulating the current knowledge on Hox-cofactor interactions and speculating on future experimental approaches that will help to identify context-specific cofactors.

## RECONSIDERING INTRINSICALLY DISORDERED REGIONS IN HOX PROTEINS

Hox family members can be recognized by invariable residues within the HD (Merabet et al., 2009), and by the presence of a short motif containing a conserved Trp residue localized upstream of the HD at variable distances (In der Rieden et al., 2004; Merabet et al., 2009). This motif, called the hexapeptide (HX), is necessary for recruiting a common and crucial class of Hox cofactors collectively



**Fig. 2.** Hox interaction networks obtained with large-scale approaches. Interaction networks result from genetic screens in *Drosophila* (blue lines) or yeast hybrid assays (red lines), and relate to the *Drosophila* Deformed (Dfd), Antennapedia (Antp), Ultrabithorax (Ubx), and mouse Hoxa1 proteins. The representation mode is based on the graphic view of BioGRID (Stark et al., 2011). Transcription factors are highlighted in bold. Potential context-specific transcriptional partners are underlined. Note that the large majority of interacting partners are not transcription factors.

regrouped under the term of PBC proteins (Burglin, 1998). PBC proteins are the only cofactors known to modulate Hox DNA-binding specificities, and their contribution to Hox target gene regulation has been extensively studied at the genetic and molecular levels (for a recent review, see Mann et al., 2009).

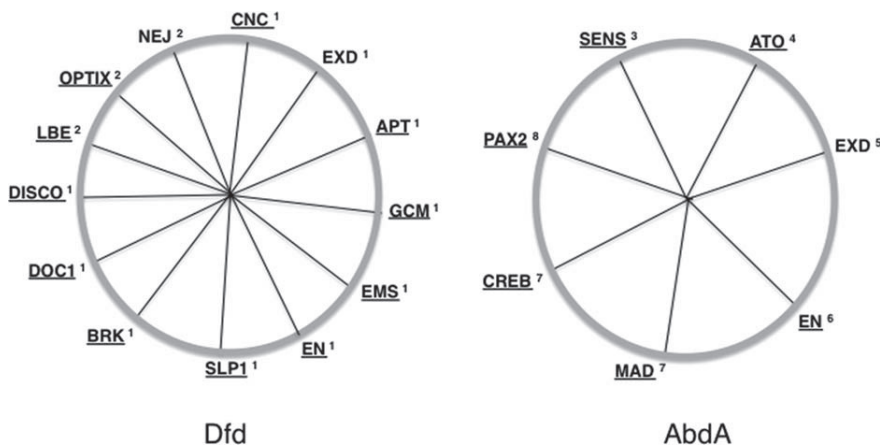
Other short protein motifs are described, which are located in different places and conserved to various

evolutionary extents within Hox family members (Merabet et al., 2009). For example, some are specific for individual Hox proteins in a given animal lineage (like the TDWM motif in insect AbdominalA [AbdA] proteins; Merabet et al., 2011), while others are found in several paralog groups and animal phyla (like the UbdA motif in Ultrabithorax [Ubx] and AbdA proteins of protostomes; Balavoine et al., 2002). Studies in

*Drosophila* have shown that these motifs, including the HX, can display a strong functional plasticity in their usage mode (Galant et al., 2002; Merabet et al., 2003, 2011). First, they can be mutated individually without leading to pleiotropic defects. Second, these protein motifs are functionally interdependent, working in a combined or opposing manner depending on the developmental context considered. And third, Hox function in different tissues might even depend on the same mode of motif usage. For example, the HX and UbdA motifs are both required for the correct regulatory activities of AbdA in contexts as different as *Drosophila* oenocytes and cardiac cells (Merabet et al., 2011). Altogether these observations highlight that multi-functionality is a hallmark of Hox protein motifs, which very likely relies on flexible yet selective interactions with context-specific cofactors.

One possibility to explain the selective plasticity of Hox-cofactor interactions is to take into consideration the role of intrinsically disordered regions. Such regions are often present in DNA-binding proteins (Ward et al., 2004; Lobley et al., 2007; Garza et al., 2009) and have been described in several instances as facilitating binding diversity within protein interaction networks (Dunker et al., 2005; Fong et al., 2009). How could this be achieved? One model is that intrinsically disordered protein regions work as flexible linkers that differentially influence the presentation of protein domains depending on the surrounding environment (Fuxreiter et al., 2007). For example, it has been proposed that the activity of intrinsically disordered regions could be modulated by post-translational modifications such as phosphorylations or acetylations (Solt et al., 2006; Vuzman et al., 2012). Alternatively, intrinsically disordered regions could also establish direct reversible interactions with specific protein partners (Fong et al., 2009).

In line with these hypotheses, large portions of Hox proteins are predicted to be unstructured (Fig. 1) (Liu et al., 2008). Such regions have been described to influence the DNA-binding activity of the *Drosophila* Hox protein Ultrabithorax (Ubx) (Liu



**Fig. 3.** Hox interaction networks resulting from analyses on isolated Hox target enhancers. Data are presented for the *Drosophila* Dfd and AbdominalA (AbdA) proteins. In these studies, a significant proportion of interacting partners correspond to context-specific transcription factors. Nomenclature is the same as in Figure 2. Numbers correspond to the following references, which are also listed in the main text: 1: Stobe et al. (2009); 2: Sorge et al. (2012); 3: Li-Kroeger et al. (2008); 4: Witt et al. (2010); 5: Merabet et al. (2003); 6: Gebelein et al. (2004); 7: Grienemberger et al. (2003); 8: Li-Kroeger et al. (2012).

et al., 2008). In addition, intrinsically disordered regions of Ubx are phosphorylated in a tissue-specific manner (Gavis and Hogness, 1991), suggesting that this mechanism could be used to modulate the binding properties with context-specific protein partners. Considering that the presence of long disordered regions is a conserved feature in Hox proteins (Fig. 1), we postulate that they could have a major impact on Hox functions, possibly by influencing the interaction properties with context-specific cofactors.

### HOX PROTEINS AND THEIR COFACTORS: DIVERSITY AND PAUCITY

The plethora of Hox embryonic functions described in different species and developmental backgrounds contrasts with the overall small number of Hox cofactors characterized so far. This lack of knowledge is best illustrated by the interaction networks depicted in current databases (DPIM) (Guruharsha et al., 2011), DroID (Murali et al., 2011), or BioGRID (Chatr-Aryamontri et al., 2013), where the large majority of Hox proteins appears to contain less than 10 cofactors.

Genetic screens in *Drosophila* (Kennison and Tamkun, 1988; Harding et al., 1995; Gellon et al., 1997; Florence and McGinnis, 1998; Merabet et al., 2002) or yeast hybrid

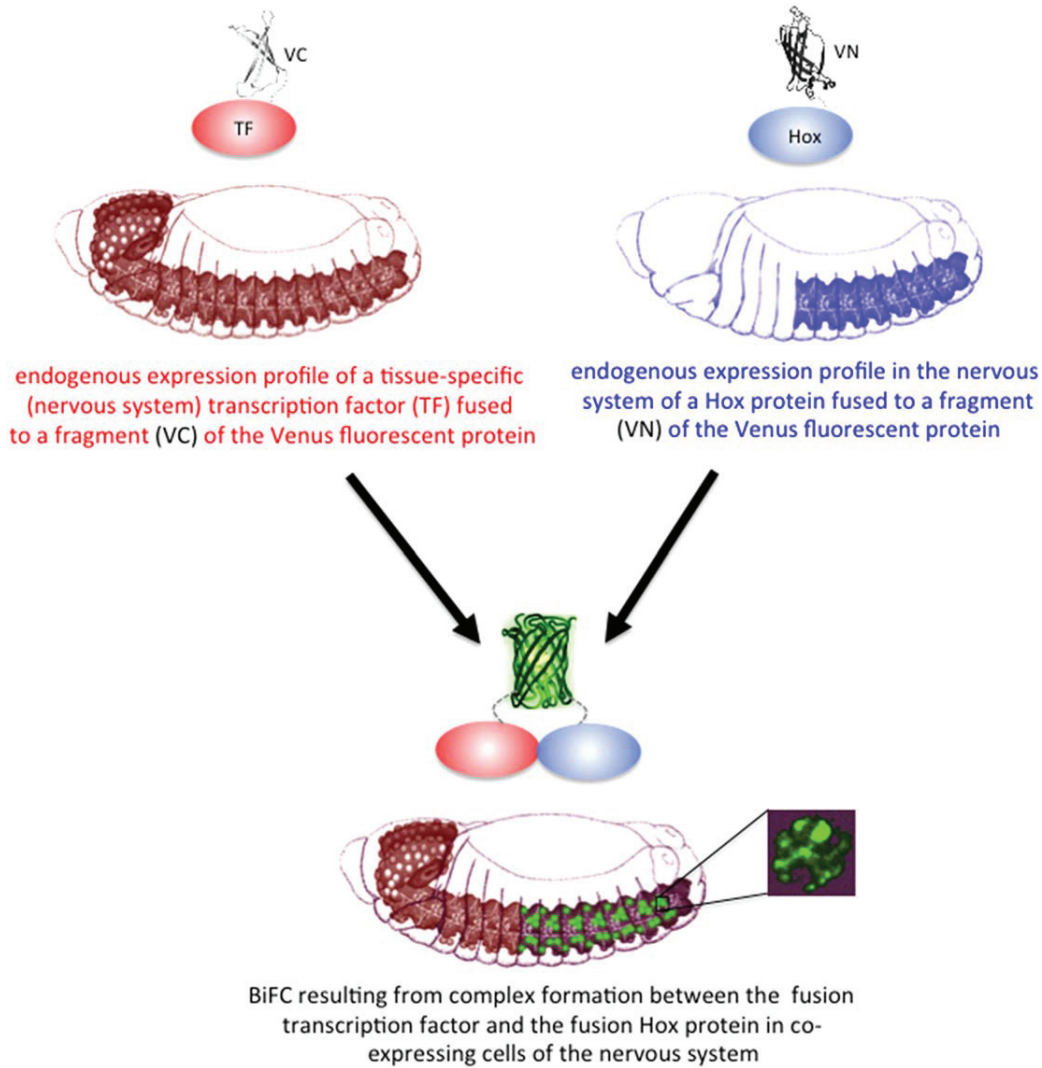
assays (Bondos et al., 2006; Lambert et al., 2012) have identified a significant number of candidate cofactors for a handful of Hox proteins (Fig. 2). Of note, most of the genetic interactions await further molecular characterisation. In the case of high-throughput yeast hybrid screens, interactions could be confirmed in their large majority in live animal cells (Lambert et al., 2012). Other interactions were validated in vivo (Bondos et al., 2006), highlighting that yeast hybrid assays could allow revealing functionally relevant interacting partners of Hox proteins. Importantly, genetic or yeast hybrid screens have considerably enlarged our vision of the molecular mode of action of Hox proteins by showing that their interacting partners were not obligatory nuclear components. In particular, Hox proteins were found to interact with cytoplasmic proteins involved in cell regulatory processes as diverse as signal transduction, mRNA stability (see also Wiellette et al., 1999), post-translational modifications, or endosomal trafficking. Thus, Hox proteins could be required in many different cell regulatory processes, illustrating that their wide range of activities is potentially not limited to gene regulation.

Along the same line, Hox nuclear partners identified by high-throughput approaches are not obligatory transcription factors. These

observations are in accordance with other individual studies describing a molecular link between Hox proteins and nuclear export (Topisirovic et al., 2005), DNA replication (Luo et al., 2004; Salsi et al., 2009; Miotto and Graba, 2010) or histone modification (Shen et al., 2001; Agelopoulos et al., 2012).

Finally, the proportion of TFs, and especially of context-specific TFs, remains extremely small in Hox interaction networks resulting from large-scale approaches (Fig. 2). By comparison, low-throughput analyses focusing on individual Hox activities were more successful in revealing context-specific cofactors, as exemplified with the *Drosophila* AbdA and Deformed (Dfd) proteins (Fig. 3). In these studies, context-specific cofactors such as zinc finger TFs (Suzuki et al., 2003; Robertson et al., 2004; Mahaffey, 2005), nuclear effectors of signalling pathways (Grieder et al., 1997; Mann and Affolter, 1998; Bai et al., 2000; Saleh et al., 2000; Yang et al., 2000; Marty et al., 2001; Merabet et al., 2002; Grienemberger et al., 2003; Bondos, 2006; Li et al., 2006; Walsh and Carroll, 2007), and cell-specific TFs (Gebelein et al., 2004; Gong et al., 2007; Li-Kroeger et al., 2008; Stobe et al., 2009; Witt et al., 2010) were described to distinguish, regionalize, or specify Hox transcriptional activities, respectively. However, considered together, few studies have dissected interactions on physiological target enhancers (Zappavigna et al., 1996; Prevot et al., 2000; Di Rocco et al., 2001; Gebelein et al., 2004; Hersh and Carroll, 2005; Pan et al., 2005; Gong et al., 2007; Taghli-Lamalle et al., 2007; Li-Kroeger et al., 2008; Williams et al., 2008; Stobe et al., 2009; Witt et al., 2010; Sorge et al., 2012). As a consequence, the molecular rules underlying context-specific activities of Hox proteins still remain poorly understood.

Such molecular paucity contrasts with the observation that Hox proteins are co-expressed with numerous context-specific TFs during embryogenesis. In addition, Chip-seq experiments with Hox proteins expressed in different developmental contexts have revealed that close distance arrangement (0 to 10 base pairs) of Hox motifs and binding sites for specific TFs is a common feature among Hox response



**Fig. 4.** Combining BiFC with recombineering methods to analyse interactions between Hox proteins and candidate cofactors in live *Drosophila* embryos. In this theoretical assay, genes are expressed under the control of their endogenous promoter as a fusion construct with the C-terminal (VC) or N-terminal (VN) fragment of the Venus fluorescent protein. Of note, the illustrative expression of the Hox protein in the nervous system is only part of the overall Hox expression profile in the embryo (not depicted). This genetic background could be achieved by replacing a genomic transposon insertion or by new recombineering methods, as discussed in the main text. Endogenous fusion Hox protein (blue) and fusion candidate cofactor (here a context specific transcription factor, red) can form a protein complex in a sub-population of co-expressing cells, leading to BiFC signals (green) in specific nuclei of the ventral nerve cord.

elements (Slattery et al., 2011). The architecture of such binding sites has been proposed to be important for cell-type specific functions of Hox proteins (Sorge et al., 2012).

In conclusion, there is a remarkable gap between the number of expected and the actually identified context-specific transcriptional partners of Hox proteins. One explanation is that Hox proteins collaborate with these factors without necessarily forming a protein complex, as observed on the *reaper* (Stobe et al., 2009) or *spalt* (Walsh and Carroll, 2007) cis-regulatory elements. Alternatively,

context-specific interactions are by definition highly dynamic during development and therefore difficult to trap. Moreover, such interactions very likely depend on other transcriptional partners that act again in a context-specific manner. Altogether, these features of Hox protein activity might explain why high-throughput approaches in yeast or cell culture were unsuccessful in capturing context-dependent Hox-cofactor interactions. In addition, due to their early embryonic and/or pleiotropic functions, standard genetic interaction screens are equally inappropriate to

uncover the identity of Hox cofactors later in development during tissue and organ formation. Thus, probably a high number of context-specific transcriptional partners of Hox proteins remain to be discovered.

## FUTURE DIRECTIONS

Our current knowledge on the identity of potential Hox cofactors is based primarily on experimental methods designed to identify protein-protein interaction partners using *ex vivo* instead of *in vivo* systems. As discussed above, these approaches are

problematic in revealing temporal or tissue-specific interactions. In the last years, high-throughput mass spectrometry methods have been extensively used for purifying high-quality proteomes from various multicellular organisms (Ahrens et al., 2010). Mass spectrometry can be combined with different affinity purification strategies to identify new interactors (see, for example, Li et al., 2011). In particular, a recently established approach in *Drosophila* allowed the isolation of interaction partners of endogenously triple-tagged-proteins (Rees et al., 2011). Although those interactions were obtained from a heterologous population of embryonic cells, recent advances in cell purification methods such as INTACT (Steiner et al., 2012) should allow us to decipher protein interaction networks in a tissue- or stage-specific context in the near future.

Complementary to the highly innovative biochemical technologies, imaging methods have also undergone considerable technical improvements and individual protein-protein interactions can now be traced in a living cell or organism. One of the most promising visualisation methods is BiFC (Bimolecular Fluorescence Complementation), which has recently been used in the *Drosophila* embryo to analyse Hox-PBC interactions under normal levels of protein expression (Hudry et al., 2011). BiFC relies on the property of fluorescent proteins to be reconstituted when their non-fluorescent moieties are close enough in space (Kodama and Hu, 2012). Combined with the genetic tools in *Drosophila*, BiFC has the potential to validate in vivo any cofactor found with the above-mentioned biochemical approaches. For example, endogenous fusion proteins could be generated with new recombineering methods (Bozas et al., 2009; Liu et al., 2012; Cho et al., 2013; Cong et al., 2013; Hwang et al., 2013; Mali et al., 2013) or through replacement of protein-trap transposons (Rees et al., 2011; Venken et al., 2011). Alternatively, a collection of fly lines was also recently established for using BiFC with the UAS/Gal4 system (Bischof et al., 2013). Overall, these recently developed genetic tools allow imaging Hox-cofactor interactions at the tissue

or sub-cellular level in live *Drosophila* embryos (Fig. 4).

## CONCLUSIONS

Deciphering the molecular code of Hox transcriptional activity is a challenging task that requires the elucidation of cell type-specific regulatory networks both at the protein-protein and protein-DNA levels. In this context, the potentially important role of disorganized regions within Hox proteins is an important aspect that needs to be explored further. Thus, despite decades of intense efforts, many facets of Hox contextual activity remain to be clarified. Thanks to the establishment of cell type-specific purification and imaging tools and of elaborate genetic techniques, this question can now be solved in the normal developmental context.

## ACKNOWLEDGMENTS

We thank Ingrid Lohmann for critical comments and Johanne Burden for proofreading the manuscript. We apologize to colleagues whose work was not cited due to space constraints. The authors declare no conflict of interest.

## REFERENCES

- Agelopoulos M, McKay DJ, Mann RS. 2012. Developmental regulation of chromatin conformation by Hox proteins in *Drosophila*. *Cell Rep* 1:350–359.
- Ahrens CH, Brunner E, Qeli E, Basler K, Aebersold R. 2010. Generating and navigating proteome maps using mass spectrometry. *Nat Rev Mol Cell Biol* 11:789–801.
- Alharbi RA, Pettengell R, Pandha HS, Morgan R. 2013. The role of HOX genes in normal hematopoiesis and acute leukemia. *Leukemia* 27:1000–1008.
- Bai S, Shi X, Yang X, Cao X. 2000. Smad6 as a transcriptional corepressor. *J Biol Chem* 275:8267–8270.
- Balavoine G, de Rosa R, Adoutte A. 2002. Hox clusters and bilaterian phylogeny. *Mol Phylogenet Evol* 24:366–373.
- Bender W, Akam M, Karch F, Beachy PA, Peifer M, Spierer P, Lewis EB, Hogness DS. 1983. Molecular genetics of the bithorax complex in *Drosophila melanogaster*. *Science* 221:23–29.
- Berger MF, Badis G, Gehrke AR, Talukder S, Philippakis AA, Pena-Castillo L, Alleyne TM, Mnaimneh S, Botvinnik OB, Chan ET, Khalid F, Zhang W, Newburger D, Jaeger SA, Morris QD, Bulyk ML, Hughes TR. 2008. Variation in homeodomain DNA binding revealed by high-resolution

analysis of sequence preferences. *Cell* 133:1266–1276.

- Bischof J, Bjorklund M, Furger E, Schertel C, Taipale J, Basler K. 2013. A versatile platform for creating a comprehensive UAS-ORFeome library in *Drosophila*. *Development* 140:2434–2442.
- Bondos S. 2006. Variations on a theme: Hox and Wnt combinatorial regulation during animal development. *Sci STKE* 2006:pe38.
- Bondos SE, Tan XX, Matthews KS. 2006. Physical and genetic interactions link hox function with diverse transcription factors and cell signaling proteins. *Mol Cell Proteomics* 5:824–834.
- Bozas A, Beumer KJ, Trautman JK, Carroll D. 2009. Genetic analysis of zinc-finger nuclease-induced gene targeting in *Drosophila*. *Genetics* 182:641–651.
- Burglin TR. 1998. The PBC domain contains a MEINOX domain: coevolution of Hox and TALE homeobox genes? *Dev Genes Evol* 208:113–116.
- Chatr-Aryamontri A, Breitkreutz BJ, Heinicke S, Boucher L, Winter A, Stark C, Nixon J, Ramage L, Kolas N, O'Donnell L, Reguly T, Breitkreutz A, Sellam A, Chen D, Chang C, Rust J, Livstone M, Oughtred R, Dolinski K, Tyers M. 2013. The BioGRID interaction database: 2013 update. *Nucleic Acids Res* 41:D816–823.
- Cho SW, Kim S, Kim JM, Kim JS. 2013. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat Biotechnol* 31:230–232.
- Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F. 2013. Multiplex genome engineering using CRISPR/Cas systems. *Science* 339:819–823.
- Di Rocco G, Gavalas A, Popperl H, Krumlauf R, Mavilio F, Zappavigna V. 2001. The recruitment of SOX/OCT complexes and the differential activity of HOXA1 and HOXB1 modulate the Hoxb1 auto-regulatory enhancer function. *J Biol Chem* 276:20506–20515.
- Dosztanyi Z, Csizmek V, Tompa P, Simon I. 2005. IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics* 21:3433–3434.
- Dunker AK, Cortese MS, Romero P, Iakoucheva LM, Uversky VN. 2005. Flexible nets. The roles of intrinsic disorder in protein interaction networks. *FEBS J* 272:5129–5148.
- Florence B, McGinnis W. 1998. A genetic screen of the *Drosophila* X chromosome for mutations that modify Deformed function. *Genetics* 150:1497–1511.
- Fong JH, Shoemaker BA, Garbuzynskiy SO, Lobanov MY, Galzitskaya OV, Panchenko AR. 2009. Intrinsic disorder in protein interactions: insights from a comprehensive structural analysis. *PLoS Comput Biol* 5:e1000316.

- Formstecher E, Aresta S, Collura V, Hamburger A, Meil A, Trehin A, Reverdy C, Betin V, Maire S, Brun C, Jacq B, Arpin M, Bellaiche Y, Bellusci S, Benaroch P, Bornens M, Chanet R, Chavrier P, Delattre O, Doye V, Fehon R, Faye G, Galli T, Girault JA, Goud B, de Gunzburg J, Johannes L, Junier MP, Mirouse V, Mukherjee A, Papadopoulos D, Perez F, Plessis A, Rosse C, Saule S, Stoppa-Lyonnet D, Vincent A, White M, Legrain P, Wojcik J, Camonis J, Daviet L. 2005. Protein interaction mapping: a *Drosophila* case study. *Genome Res* 15:376–384.
- Fuxreiter M, Tompa P, Simon I. 2007. Local structural disorder imparts plasticity on linear motifs. *Bioinformatics* 23:950–956.
- Galant R, Walsh CM, Carroll SB. 2002. Hox repression of a target gene: extradenticle-independent, additive action through multiple monomer binding sites. *Development* 129:3115–3126.
- Garza AS, Ahmad N, Kumar R. 2009. Role of intrinsically disordered protein regions/domains in transcriptional regulation. *Life Sci* 84:189–193.
- Gavis ER, Hogness DS. 1991. Phosphorylation, expression and function of the Ultrabithorax protein family in *Drosophila melanogaster*. *Development* 112:1077–1093.
- Gebelein B, McKay DJ, Mann RS. 2004. Direct integration of Hox and segmentation gene inputs during *Drosophila* development. *Nature* 431:653–659.
- Gellon G, Harding KW, McGinnis N, Martin MM, McGinnis W. 1997. A genetic screen for modifiers of Deformed homeotic function identifies novel genes required for head development. *Development* 124:3321–3331.
- Giot L, Bader JS, Brouwer C, Chaudhuri A, Kuang B, Li Y, Hao YL, Ooi CE, Godwin B, Vitols E, Vijayadamodar G, Pochart P, Machineni H, Welsh M, Kong Y, Zerhusen B, Malcolm R, Varrone Z, Collis A, Minto M, Burgess S, McDaniel L, Stimpson E, Spriggs F, Williams J, Neurath K, Ioime N, Agee M, Voss E, Furtak K, Renzulli R, Aanensen N, Carrolla S, Bickelhaupt E, Lazovatsky Y, DaSilva A, Zhong J, Stanyon CA, Finley RL, Jr., White KP, Braverman M, Jarvie T, Gold S, Leach M, Knight J, Shimkets RA, McKenna MP, Chant J, Rothberg JM. 2003. A protein interaction map of *Drosophila melanogaster*. *Science* 302:1727–1736.
- Gong KQ, Yallowitz AR, Sun H, Dressler GR, Wellik DM. 2007. A Hox-Eya-Pax complex regulates early kidney developmental gene expression. *Mol Cell Biol* 27:7661–7668.
- Grieder NC, Marty T, Ryoo HD, Mann RS, Affolter M. 1997. Synergistic activation of a *Drosophila* enhancer by HOM/EXD and DPP signaling. *EMBO J* 16:7402–7410.
- Grienerberger A, Merabet S, Manak J, Iltis I, Fabre A, Berenger H, Scott MP, Pradel J, Graba Y. 2003. Tgfbeta signaling acts on a Hox response element to confer specificity and diversity to Hox protein function. *Development* 130:5445–5455.
- Guruharsha KG, Rual JF, Zhai B, Mintseris J, Vaidya P, Vaidya N, Beekman C, Wong C, Rhee DY, Cenaj O, McKillip E, Shah S, Stapleton M, Wan KH, Yu C, Parsa B, Carlson JW, Chen X, Kapadia B, VijayRaghavan K, Gygi SP, Celniker SE, Obar RA, Artavanis-Tsakonas S. 2011. A protein complex network of *Drosophila melanogaster*. *Cell* 147:690–703.
- Harding K, Wedeen C, McGinnis W, Levine M. 1985. Spatially regulated expression of homeotic genes in *Drosophila*. *Science* 229:1236–1242.
- Harding KW, Gellon G, McGinnis N, McGinnis W. 1995. A screen for modifiers of Deformed function in *Drosophila*. *Genetics* 140:1339–1352.
- Hersh BM, Carroll SB. 2005. Direct regulation of knot gene expression by Ultrabithorax and the evolution of cis-regulatory elements in *Drosophila*. *Development* 132:1567–1577.
- Hombria JC, Lovegrove B. 2003. Beyond homeosis: HOX function in morphogenesis and organogenesis. *Differentiation* 71:461–476.
- Hudry B, Viala S, Graba Y, Merabet S. 2011. Visualization of protein interactions in living *Drosophila* embryos by the bimolecular fluorescence complementation assay. *BMC Biol* 9:5.
- Hwang WY, Fu Y, Reyon D, Maeder ML, Tsai SQ, Sander JD, Peterson RT, Yeh JR, Joung JK. 2013. Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat Biotechnol* 31:227–229.
- In der Rieden PM, Mainguy G, Woltering JM, Durston AJ. 2004. Homeodomain to hexapeptide or PBC-interaction-domain distance: size apparently matters. *Trends Genet* 20:76–79.
- Kennison JA, Tamkun JW. 1988. Dosage-dependent modifiers of polycomb and antennapedia mutations in *Drosophila*. *Proc Natl Acad Sci USA* 85:8136–8140.
- Kodama Y, Hu CD. 2012. Bimolecular fluorescence complementation (BiFC): a 5-year update and future perspectives. *Biotechniques* 53:285–298.
- Lacombe J, Hanley O, Jung H, Philippidou P, Surmeli G, Grinstein J, Dasen JS. 2013. Genetic and functional modularity of Hox activities in the specification of limb-innervating motor neurons. *PLoS Genet* 9:e1003184.
- Lambert B, Vandeputte J, Remacle S, Bergiers I, Simonis N, Twizere JC, Vidal M, Rezsöházy R. 2012. Protein interactions of the transcription factor Hoxa1. *BMC Dev Biol* 12:29.
- Lewis EB. 1978. A gene complex controlling segmentation in *Drosophila*. *Nature* 276:565–570.
- Li X, Nie S, Chang C, Qiu T, Cao X. 2006. Smads oppose Hox transcriptional activities. *Exp Cell Res* 312:854–864.
- Li Y, Franklin S, Zhang MJ, Vondriska TM. 2011. Highly efficient purification of protein complexes from mammalian cells using a novel streptavidin-binding peptide and hexahistidine tandem tag system: application to Bruton's tyrosine kinase. *Protein Sci* 20:140–149.
- Li-Kroeger D, Witt LM, Grimes HL, Cook TA, Gebelein B. 2008. Hox and senseless antagonism functions as a molecular switch to regulate EGF secretion in the *Drosophila* PNS. *Dev Cell* 15:298–308.
- Lievens S, Vanderroot N, Van der Heyden J, Gesellchen V, Vidal M, Tavernier J. 2009. Array MAPPIT: high-throughput interactome analysis in mammalian cells. *J Proteome Res* 8:877–886.
- Liu J, Li C, Yu Z, Huang P, Wu H, Wei C, Zhu N, Shen Y, Chen Y, Zhang B, Deng WM, Jiao R. 2012. Efficient and specific modifications of the *Drosophila* genome by means of an easy TALEN strategy. *J Genet Genomics* 39:209–215.
- Liu Y, Matthews KS, Bondos SE. 2008. Multiple intrinsically disordered sequences alter DNA binding by the homeodomain of the *Drosophila* hox protein ultrabithorax. *J Biol Chem* 283:20874–20887.
- Lobley A, Swindells MB, Orengo CA, Jones DT. 2007. Inferring function using patterns of native disorder in proteins. *PLoS Comput Biol* 3:e162.
- Luo L, Yang X, Takihara Y, Knoetgen H, Kessel M. 2004. The cell-cycle regulator geminin inhibits Hox function through direct and polycomb-mediated interactions. *Nature* 427:749–753.
- Mahaffey JW. 2005. Assisting Hox proteins in controlling body form: are there new lessons from flies (and mammals)? *Curr Opin Genet Dev* 15:422–429.
- Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM. 2013. RNA-guided human genome engineering via Cas9. *Science* 339:823–826.
- Mann RS, Affolter M. 1998. Hox proteins meet more partners. *Curr Opin Genet Dev* 8:423–429.
- Mann RS, Lelli KM, Joshi R. 2009. Hox specificity unique roles for cofactors and collaborators. *Curr Top Dev Biol* 88:63–101.
- Marty T, Viganò MA, Ribeiro C, Nussbaumer U, Grieder NC, Affolter M. 2001. A HOX complex, a repressor element and a 50 bp sequence confer regional specificity to a DPP-responsive enhancer. *Development* 128:2833–2845.
- McGinnis W, Krumlauf R. 1992. Homeobox genes and axial patterning. *Cell* 68:283–302.
- Merabet S, Catala F, Pradel J, Graba Y. 2002. A green fluorescent protein reporter genetic screen that identifies modifiers of Hox gene function in the *Drosophila* embryo. *Genetics* 162:189–202.
- Merabet S, Kambris Z, Capovilla M, Berenger H, Pradel J, Graba Y. 2003. The hexapeptide and linker regions of the AbdA Hox protein regulate its activating and repressive functions. *Dev Cell* 4:761–768.
- Merabet S, Hudry B, Saadaoui M, Graba Y. 2009. Classification of sequence signatures: a guide to Hox protein function. *Bioessays* 31:500–511.
- Merabet S, Litim-Mecheri I, Karlsson D, Dixit R, Saadaoui M, Monier B, Brun



- C, Thor S, Vijayraghavan K, Perrin L, Pradel J, Graba Y. 2011. Insights into Hox protein function from a large scale combinatorial analysis of protein domains. *PLoS Genet* 7:e1002302.
- Miotto B, Graba Y. 2010. Control of DNA replication: a new facet of Hox proteins? *Bioessays* 32:800–807.
- Murali T, Pacifico S, Yu J, Guest S, Roberts GG, 3rd, Finley RL, Jr. 2011. DroID 2011: a comprehensive, integrated resource for protein, transcription factor, RNA and gene interactions for *Drosophila*. *Nucleic Acids Res* 39:D736–743.
- Noyes MB, Christensen RG, Wakabayashi A, Stormo GD, Brodsky MH, Wolfe SA. 2008. Analysis of homeodomain specificities allows the family-wide prediction of preferred recognition sites. *Cell* 133:1277–1289.
- Pan L, Glenn ST, Jones CA, Gross KW. 2005. Activation of the rat renin promoter by HOXD10.PBX1b.PREP1, Ets-1, and the intracellular domain of notch. *J Biol Chem* 280:20860–20866.
- Philippidou P, Walsh CM, Aubin J, Jeannotte L, Dasen JS. 2012. Sustained Hox5 gene activity is required for respiratory motor neuron development. *Nat Neurosci* 15:1636–1644.
- Prevot D, Voeltzel T, Birot AM, Morel AP, Rostan MC, Magaud JP, Corbo L. 2000. The leukemia-associated protein Btg1 and the p53-regulated protein Btg2 interact with the homeoprotein Hoxb9 and enhance its transcriptional activation. *J Biol Chem* 275:147–153.
- Ravasi T, Suzuki H, Cannistraci CV, Katayama S, Bajic VB, Tan K, Akalin A, Schmeier S, Kanamori-Katayama M, Bertin N, Carninci P, Daub CO, Forrest AR, Gough J, Grimmond S, Han JH, Hashimoto T, Hide W, Hofmann O, Kamburov A, Kaur M, Kawaji H, Kubosaki A, Lassmann T, van Nimwegen E, MacPherson CR, Ogawa C, Radovanovic A, Schwartz A, Teasdale RD, Tegner J, Lenhard B, Teichmann SA, Arakawa T, Ninomiya N, Murakami K, Tagami M, Fukuda S, Imamura K, Kai C, Ishihara R, Kitazume Y, Kawai J, Hume DA, Ideker T, Hayashizaki Y. 2010. An atlas of combinatorial transcriptional regulation in mouse and man. *Cell* 140:744–752.
- Rees JS, Lowe N, Armean IM, Roote J, Johnson G, Drummond E, Spriggs H, Ryder E, Russell S, St Johnston D, Lilley KS. 2011. In vivo analysis of proteomes and interactomes using Parallel Affinity Capture (iPAC) coupled to mass spectrometry. *Mol Cell Proteomics* 10:M110 002386.
- Robertson LK, Bowling DB, Mahaffey JP, Imiolczyk B, Mahaffey JW. 2004. An interactive network of zinc-finger proteins contributes to regionalization of the *Drosophila* embryo and establishes the domains of HOM-C protein function. *Development* 131:2781–2789.
- Rual JF, Venkatesan K, Hao T, Hirozane-Kishikawa T, Dricot A, Li N, Berriz GF, Gibbons FD, Dreze M, Ayivi-Guedehoussou N, Klitgord N, Simon C, Boxem M, Milstein S, Rosenberg J, Goldberg DS, Zhang LV, Wong SL, Franklin G, Li S, Albala JS, Lim J, Fraughton C, Llamosas E, Cevik S, Bex C, Lamesch P, Sikorski RS, Vandenhaute J, Zoghbi HY, Smolyar A, Bosak S, Sequerra R, Doucette-Stamm L, Cusick ME, Hill DE, Roth FP, Vidal M. 2005. Towards a proteome-scale map of the human protein-protein interaction network. *Nature* 437:1173–1178.
- Rubin E, Wu X, Zhu T, Cheung JC, Chen H, Lorincz A, Pandita RK, Sharma GG, Ha HC, Gasson J, Hanakahi LA, Pandita TK, Sukumar S. 2007. A role for the HOXB7 homeodomain protein in DNA repair. *Cancer Res* 67:1527–1535.
- Saleh M, Rambaldi I, Yang XJ, Featherstone MS. 2000. Cell signaling switches HOX-PBX complexes from repressors to activators of transcription mediated by histone deacetylases and histone acetyltransferases. *Mol Cell Biol* 20:8623–8633.
- Salsi V, Ferrari S, Ferraresi R, Cossarizza A, Grande A, Zappavigna V. 2009. HOXD13 binds DNA replication origins to promote origin licensing and is inhibited by geminin. *Mol Cell Biol* 29:5775–5788.
- Shen WF, Krishnan K, Lawrence HJ, Largam C. 2001. The HOX homeodomain proteins block CBP histone acetyltransferase activity. *Mol Cell Biol* 21:7509–7522.
- Slattery M, Ma L, Negre N, White KP, Mann RS. 2011. Genome-wide tissue-specific occupancy of the Hox protein Ultrabithorax and Hox cofactor Homothorax in *Drosophila*. *PLoS One* 6:e14686.
- Solt I, Magyar C, Simon I, Tompa P, Fuxreiter M. 2006. Phosphorylation-induced transient intrinsic structure in the kinase-inducible domain of CREB facilitates its recognition by the KIX domain of CBP. *Proteins* 64:749–757.
- Sorge S, Ha N, Polychronidou M, Friedrich J, Bezdán D, Kaspar P, Schaefer MH, Ossowski S, Henz SR, Mundorf J, Ratzler J, Papagiannouli F, Lohmann I. 2012. The cis-regulatory code of Hox function in *Drosophila*. *EMBO J* 31:3323–3333.
- Stanyon CA, Liu G, Mangiola BA, Patel N, Giot L, Kuang B, Zhang H, Zhong J, Finley RL, Jr. 2004. A *Drosophila* protein-interaction map centered on cell-cycle regulators. *Genome Biol* 5:R96.
- Stark C, Breitkreutz BJ, Chatr-Aryamontri A, Boucher L, Oughtred R, Livstone MS, Nixon J, Van Auken K, Wang X, Shi X, Reguly T, Rust JM, Winter A, Dolinski K, Tyers M. 2011. The BioGRID Interaction Database: 2011 update. *Nucleic Acids Res* 39:D698–704.
- Steiner FA, Talbert PB, Kasinathan S, Deal RB, Henikoff S. 2012. Cell-type-specific nuclei purification from whole animals for genome-wide expression and chromatin profiling. *Genome Res* 22:766–777.
- Stobe P, Stein MA, Habring-Muller A, Bezdán D, Fuchs AL, Hueber SD, Wu H, Lohmann I. 2009. Multifactorial regulation of a hox target gene. *PLoS Genet* 5:e1000412.
- Suzuki M, Ueno N, Kuroiwa A. 2003. Hox proteins functionally cooperate with the GC box-binding protein system through distinct domains. *J Biol Chem* 278:30148–30156.
- Taghli-Lamalle O, Gallet A, Leroy F, Malapert P, Vola C, Kerridge S, Fasano L. 2007. Direct interaction between Teshirt and Sex combs reduced proteins, via Tsh's acidic domain, is essential for specifying the identity of the prothorax in *Drosophila*. *Dev Biol* 307:142–151.
- Tischfield MA, Bosley TM, Salih MA, Alorainy IA, Sener EC, Nester MJ, Oystreck DT, Chan WM, Andrews C, Erickson RP, Engle EC. 2005. Homozygous HOXA1 mutations disrupt human brainstem, inner ear, cardiovascular and cognitive development. *Nat Genet* 37:1035–1037.
- Topisirovic I, Kentsis A, Perez JM, Guzman ML, Jordan CT, Borden KL. 2005. Eukaryotic translation initiation factor 4E activity is modulated by HOXA9 at multiple levels. *Mol Cell Biol* 25:1100–1112.
- Venken KJ, Schulze KL, Haelterman NA, Pan H, He Y, Evans-Holm M, Carlson JW, Levis RW, Spradling AC, Hoskins RA, Bellen HJ. 2011. MiMIC: a highly versatile transposon insertion resource for engineering *Drosophila melanogaster* genes. *Nat Methods* 8:737–743.
- Vuzman D, Hoffman Y, Levy Y. 2012. Modulating protein-DNA Interactions by post-translational modifications at disordered regions. *Pac Symp Biocomput* 17:188–199.
- Walsh CM, Carroll SB. 2007. Collaboration between Smads and a Hox protein in target gene repression. *Development* 134:3585–3592.
- Ward JJ, Sodhi JS, McGuffin LJ, Buxton BF, Jones DT. 2004. Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *J Mol Biol* 337:635–645.
- Wiellette EL, Harding KW, Mace KA, Ronshaugen MR, Wang FY, McGinnis W. 1999. spen encodes an RNP motif protein that interacts with Hox pathways to repress the development of head-like sclerites in the *Drosophila* trunk. *Development* 126:5373–5385.
- Williams TM, Selegue JE, Werner T, Gompel N, Kopp A, Carroll SB. 2008. The regulation and evolution of a genetic switch controlling sexually dimorphic traits in *Drosophila*. *Cell* 134:610–623.
- Witt LM, Gutzwiller LM, Gresser AL, Li-Kroeger D, Cook TA, Gebelein B. 2010. Atonal, Senseless, and Abdominal-A regulate rhomboid enhancer activity in abdominal sensory organ precursors. *Dev Biol* 344:1060–1070.
- Yang X, Ji X, Shi X, Cao X. 2000. Smad1 domains interacting with Hox-8 induce osteoblast differentiation. *J Biol Chem* 275:1065–1072.
- Zappavigna V, Falciola L, Helmer-Citterich M, Mavilio F, Bianchi ME. 1996. HMG1 interacts with HOX proteins and enhances their DNA binding and transcriptional activation. *EMBO J* 15:4981–4991.

## 4. QUATRIEME CHAPITRE : TROUVER DE NOUVEAUX PARTENAIRES DES PROTEINES HOX SPECIFIQUES D'UN CONTEXTE CANCEREUX

### 4.1. LE CANCER : UN CONTEXTE PROPICE A L'ETUDE DES PROTEINES HOX

Nous ne présenterons ici que les cancers solides sur lesquels nous avons travaillé par la suite. Mais, comme nous l'avons introduit, des dérèglements des protéines Hox sont retrouvés dans de nombreux autres cancers (**Introduction Chapitre II.1**; **Bhatlekar et al., 2014**). On sait que 37 des 39 protéines Hox humaines peuvent être dérégulées dans les cancers, mais nous avons choisi de travailler sur la protéine HoxA9 en particulier pour conduire ce crible, car elle est suffisante pour générer une leucémie myéloïde aigue, et elle est une des protéines Hox dont l'expression est dérégulée dans le plus de cancers solides différents. Afin de mieux comprendre le rôle du motif HX dans un environnement cellulaire donné, nous avons donc identifié des partenaires de la protéine HoxA9 sauvage ou mutée dans son motif HX dans trois différentes lignées cellulaires.

#### 4.1.1. Sein

L'expression des gènes HOX a été étudiée à la fois dans les cellules saines et cancéreuses du sein. Il a été montré que 17 des 39 gènes HOX humains sont toujours exprimés dans le tissu sain de sein adulte, avec majoritairement des protéines des groupes A et C (**Cantile et al., 2003**). Deux études montrent que, la surexpression de certains gènes Hox (**Hur et al., 2014**) et la diminution de l'expression d'autres gènes Hox (**Makiyama et al., 2005**) sont corrélées au cancer du sein.

Si l'on prend l'exemple de la protéine **HoxA9** sur laquelle nous allons ensuite travailler, elle est normalement exprimée dans le tissu sain de sein adulte. On observe une perte de son expression dans le cadre de cancer du sein (**Makiyama et al., 2005**). On la retrouve donc très faiblement exprimée dans la lignée cancéreuses **MDA-MB231** (**Morgan et al., 2012**)

#### 4.1.2. Prostate

La plupart des gènes HOX qui sont surexprimés dans les cancers de la prostate sont des gènes du groupe C (**HOXC4, HOXC5, HOXC6, HOXC8**; **Miller et al., 2003**). La mieux décrite est la surexpression du gène **HOXC8** qui entraîne une perte de la capacité de différenciation des cellules de cancer de la prostate. La protéine HoxC8 joue donc un rôle dans les capacités invasives et métastatique des cellules cancéreuses (**Waltregny et al., 2002**). Les protéines HoxB13 (**Jung et al., 2004**) HoxC6, (**Ramachandran et al., 2005**), HoxB3 (**Chen et al., 2013**) ont aussi été décrites pour augmenter les propriétés prolifératives et/ou migratoires des cellules cancéreuses de la prostate.

La protéine **HoxA9** a quant à elle été montrée comme critique lors du développement embryonnaire de la glande prostatique ainsi que leur de la carcinogénèse de la prostate (**Javed et Langley, 2013**). Afin d'étudier les effets des complexes Hox/Pbx dans les cellules cancéreuses pancréatiques, **Morgan et al. (2014)** ont étudié les différents niveaux d'expression

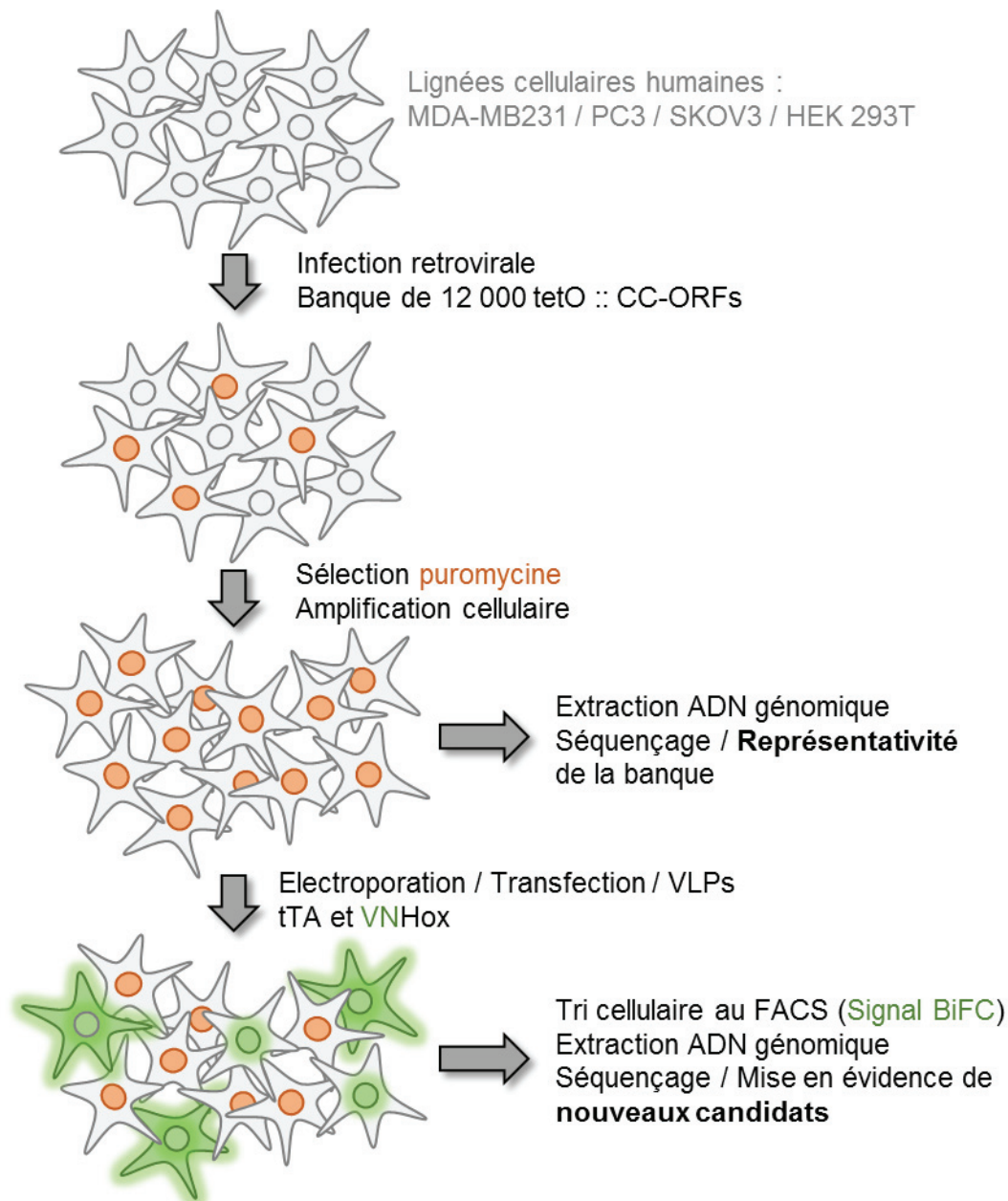
des protéines Hox dans différentes lignées cellulaires cancéreuses pancréatiques parmi lesquelles la lignée **PC3** que nous allons utiliser plus tard dans laquelle le niveau d'expression de la protéine HoxA9 est particulièrement élevé.

#### 4.1.3. Ovaire

Les gènes HOX sont impliqués dans la différenciation du canal de Müller lors du développement et ne sont plus exprimés dans l'épithélium normal ovarien adulte. Cependant, les protéines HoxA9, HoxA10, HoxA11 (**Cheng et al., 2005**), HoxA7 (**Naora et al., 2001**), HoxB7, HoxB13 (**Yamashita et al., 2006**) et d'autres sont retrouvées surexprimées dans ces cancers de l'ovaire, induisant l'hyperprolifération, la migration, l'invasion, des cellules cancéreuses ovariennes. Dans le cadre de la lignée cancéreuse **SKOV3** que nous allons utiliser pour notre crible, nous trouvons une surexpression de différentes protéines Hox (**Morgan et al., 2010**) dont notre protéine d'intérêt **HoxA9**.

#### 4.2. LE MODELE :

Les mécanismes d'action des protéines Hox dans les cancers sont encore mal connus. Des gènes cibles ont d'ores et déjà été identifiés par des analyses de type CHIP-seq dans différents contextes cancéreux, complétées par des analyses transcriptomiques. Nous nous sommes intéressé à identifier de nouveaux partenaires des protéines Hox, qui pourraient être spécifiques d'un contexte cancéreux et donc expliquer leur rôle dans ces maladies. Afin de répondre à cette question, nous avons mis au point un crible permettant d'identifier des partenaires des protéines HOX dont le principe repose sur la technique de la BiFC (**Figure 4.2**). En effet, notre protéine d'intérêt, la protéine HoxA9, est fusionnée à la partie N-terminale de la Venus (VN). Comme nous l'avons introduit (**Introduction Chapitre IV.1.2**) la VN est capable de compléter avec la partie C-terminale de la Cerulean (CC) Ainsi, nous avons fusionné une banque de 12 000 ORFs humains différents avec le fragment CC, et intégré ces constructions dans différentes lignées cellulaires cancéreuses afin de tester puis d'identifier directement *in vivo* les nouveaux partenaires de la protéine HoxA9. Ce crible est encore en phase de test et de mise en place au sein du laboratoire et les résultats préliminaires semblent très encourageants. Il fait actuellement l'objet d'un dépôt de brevet.

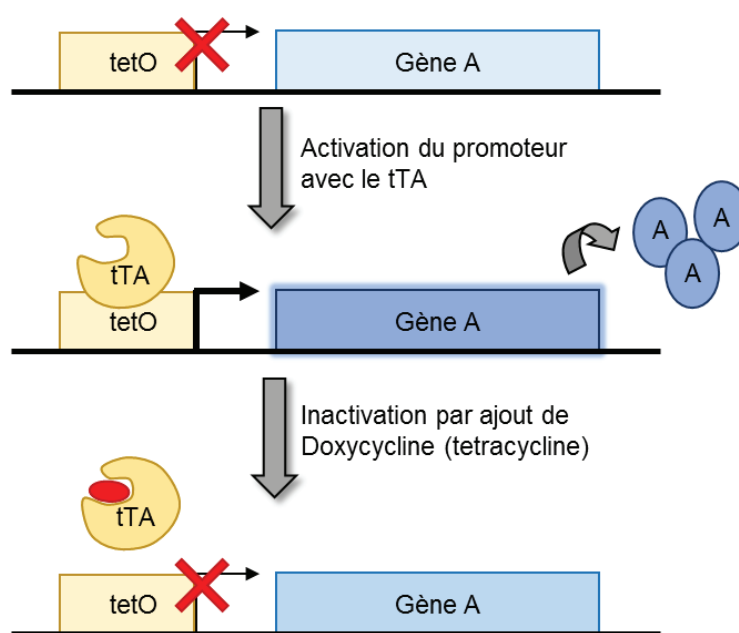


**Figure 4.2 : Etapes clés du cribre BiFC.** Dans une première étape les différentes lignées cellulaires sont infectées par des rétrovirus contenant la construction tetO :: CC-ORF IRES Puro. Cette construction est issue d'une banque de 12 000 ORFs humains. Nous sommes partis des lignées PC3 (cancer de la prostate) MDA-MB231 (cancer du sein), SKOV3 (cancer des ovaires) et HEK 293T (rein embryonnaire humain pour contrôle). Suite à l'infection, nous avons induit une pression de sélection avec un antibiotique, la puromycine, pour ne sélectionner que les cellules ayant intégré la construction d'intérêt. Suite à une amplification des différentes lignées afin de garder la représentativité des membres de la banque CC-ORF intégré, nous avons réalisé un séquençage, permettant d'établir la représentativité de la banque. La protéine d'intérêt du cribre VNHox est ensuite délivrée dans les cellules en même temps que l'activateur du promoteur tTA par trois différentes techniques : la transfection, l'électroporation ou l'apport direct des protéines par Virus Like Particules (VLPs). Les cellules fluorescentes Venus sont ensuite triées au FACS car ce sont les cellules dans lesquels une interaction potentielle s'est établie entre la protéine Hox et un des 12 000 ORFs. L'ADN génomique des cellules est extrait puis les intégrations sont amplifiées par PCR spécifiques pour être ensuite séquencées.

## 4.2.1. La banque CCORF

Afin d'identifier de nouveaux candidats pour l'interaction avec les protéines HOX par la BiFC, nous avons décidé d'opter pour une technique légèrement différente de celle utilisée pour le ciblage conduit contre la protéine Akt1 (Ding *et al.*, 2006) que nous avons présenté dans l'introduction de ce manuscrit. En effet, les insertions aléatoires de la cassette tetO-VC pourraient être soumises aux repliements de l'ADN et à sa régulation. Des gènes cibles de partenaires potentiels se situant dans des régions dérégulées et compactées pourraient alors nous échapper. De plus, cette technique entraîne la génération de mutants, de formes tronquées des protéines d'intérêt et donc encore une fois, un biais qui irait contre les intérêts du crible.

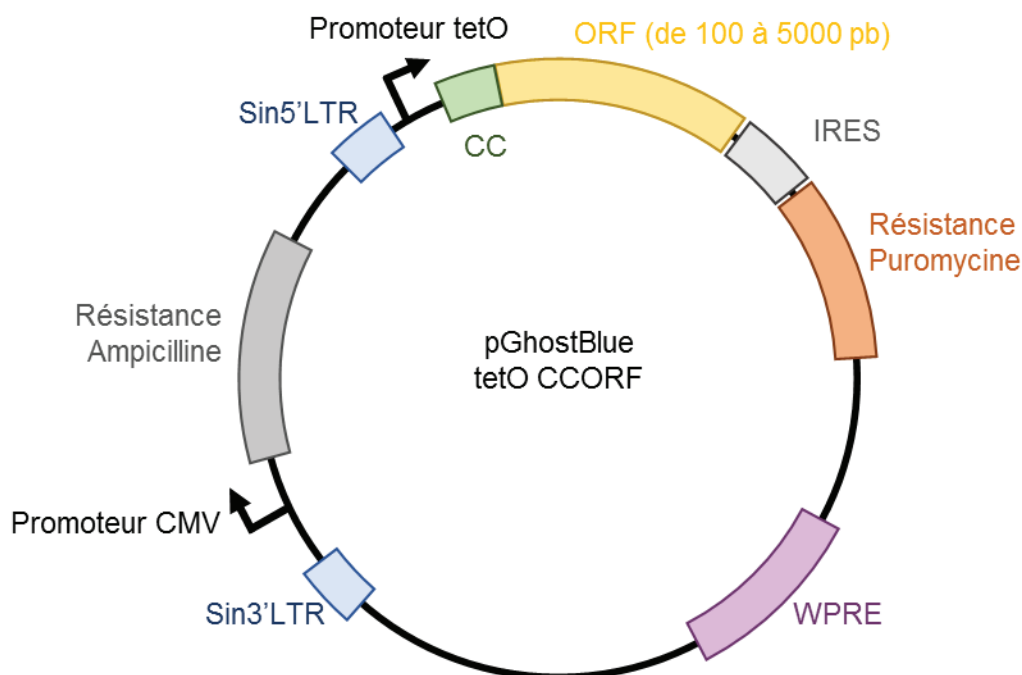
Nous avons donc opté pour l'utilisation d'une banque d'ORFs humains (ORFéome 3.1 ; Lamesch *et al.*, 2006). Cette banque contient environ 12 000 ORFs humains qui, après identification, correspondent à environ 8200 gènes humains différents. Ces cDNAs sont intégrés dans un vecteur lentiviral de type HIV sous un promoteur tet inductible (ou tet-Opérateur tetO) et en aval d'une cassette CC. Nous avons choisi ce promoteur répondant au système tet-ON/tet-OFF afin de nous affranchir d'une éventuelle toxicité d'une protéine artificiellement exprimée. En effet, nos partenaires potentiels ne seront exprimés qu'en présence d'un activateur de ce promoteur (Figure 4.2.1A) qui ne sera délivré qu'au dernier moment. Cette banque de plasmides a été générée en collaboration avec Philippe Mangeot (Figure 4.2.1B).



**Figure 4.2.1A :** Le système tet-OFF.

En collaboration avec le plateau de vectorologie de l'ENS de Lyon, nous avons ensuite généré une banque de rétrovirus permettant d'intégrer de façon stable les 8200 gènes dans différentes lignées d'intérêt. La quantité de particules virales a été calculée de manière à ce qu'une seule intégration se produise dans les lignées cellulaires dans la majorité des cas (30%), minimisant ainsi les doubles ou multiples intégrations qui nuiraient au crible. Nous avons sélectionné pour commencer trois lignées cancéreuses ainsi qu'une lignée test. Ces trois lignées sont les MDA-MB231 (cancer du sein), les PC3 (cancer de la prostate) ainsi que les SKOV3

(cancer des ovaires). Pour les tests et afin de sélectionner les candidats les plus spécifiques nous avons en parallèle mené le crible sur une lignée HEK 293T.



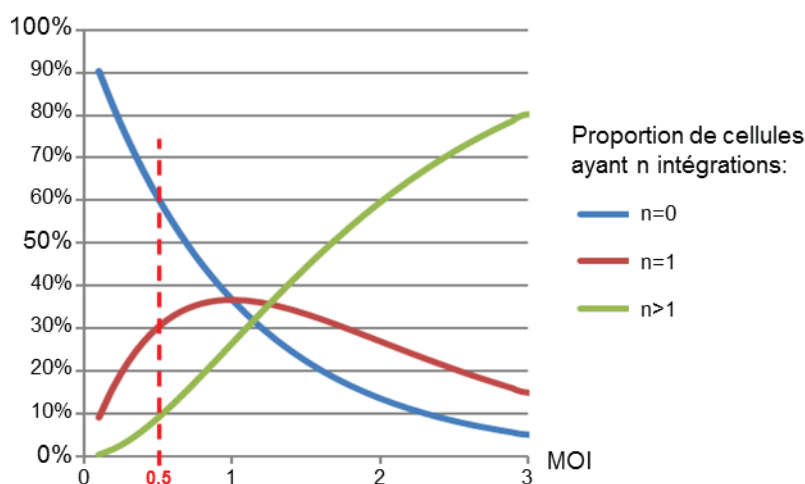
**Figure 4.2.1B : Plasmide pGhostBlue TetO-CCORF.** Avec CC pour la partie C-ter de la Cerulean. La banque de 12 000 ORFs de l'ORFeome v3.1 est intégrée entre deux sites de recombinaison non mentionnés ici à la suite de la cassette CC. L'IRES est le site d'entrée du ribosome. La séquence WPRE (Woodchuck Hepatitis Virus (WHP) Posttranscriptional Regulatory Element) est une séquence spécifique des retrovirus qui permet une stabilisation de l'ARN retroviral et augmente ainsi la qualité d'infection.

Les cellules ainsi transduites sont ensuite sélectionnées sous une pression antibiotique : la puromycine. Une fois la sélection terminée, une partie des cellules est lysée et leur ADN génomique extrait. Par une amplification spécifique nous isolons les insertions rétrovirales puis en collaboration avec le plateau de séquençage de l'IGFL nous vérifions l'intégration de la banque au sein des différentes lignées. (**Tableau 4.2.1C**)

	HEK 293T	PC3	MDA-MB231	SKOV3
Première Infection (Juin 2015)	<b>65%</b> (6 772/10 354)	<b>36%</b> (3 677/10 354)	<b>43%</b> (4 521/10 354)	Non infectées
Deuxième Infection (Juin 2016)	<b>22%</b> (2 243/10 354)	<b>55%</b> (5 652/10 354)	<b>14%</b> (1 417/10 354)	<b>44%</b> (4 604/10 354)

**Tableau 4.2.1C : Représentativité de la banque ORFeome v3.1 au sein des différentes lignées cellulaires.**

Les intégrations de notre banque d'ORFs n'étant pas de 100%, de nouvelles infections sont en cours afin d'augmenter la qualité de la transduction et la représentativité de la banque au sein des différentes lignées. Nous avons d'ores et déjà amélioré celle-ci au sein de la lignée PC3 (**Tableau 4.2.1C**). Ces différences peuvent s'expliquer par la nature même des cellules. En effet, les cellules cancéreuses ont des mécanismes de défense qui peuvent les rendre plus résistantes à des phénomènes infectieux, ou elles peuvent aussi posséder moins de récepteurs à leur surface (LDL récepteur) rendant les particules rétrovirales moins aptes à les infecter. Il est donc nécessaire de systématiquement s'assurer du titre infectieux des particules en fonction de chaque lignée cellulaire afin d'adapter au mieux la quantité de virus délivrée. De plus, par la réserve que nous mettons de n'avoir qu'une seule intégration par cellule, nous diminuons drastiquement le nombre de cellules infectées (**Figure 4.2.1D**).



**Figure 4.2.1D : Pourcentage de cellules infectées en fonction de la quantité de virus délivré par cellule (MOI).** Les courbes correspondent au pourcentage de cellules ayant aucune (bleue), une seule (rouge) ou plusieurs (verte) intégrations. Les intégrations répondent à une loi de Poisson. La valeur de MOI à 0,5 est le meilleur compromis. Elle correspond à deux fois plus de cellules que de particules.

#### 4.2.2. Expression de la protéine Hox dans les différentes lignées établies

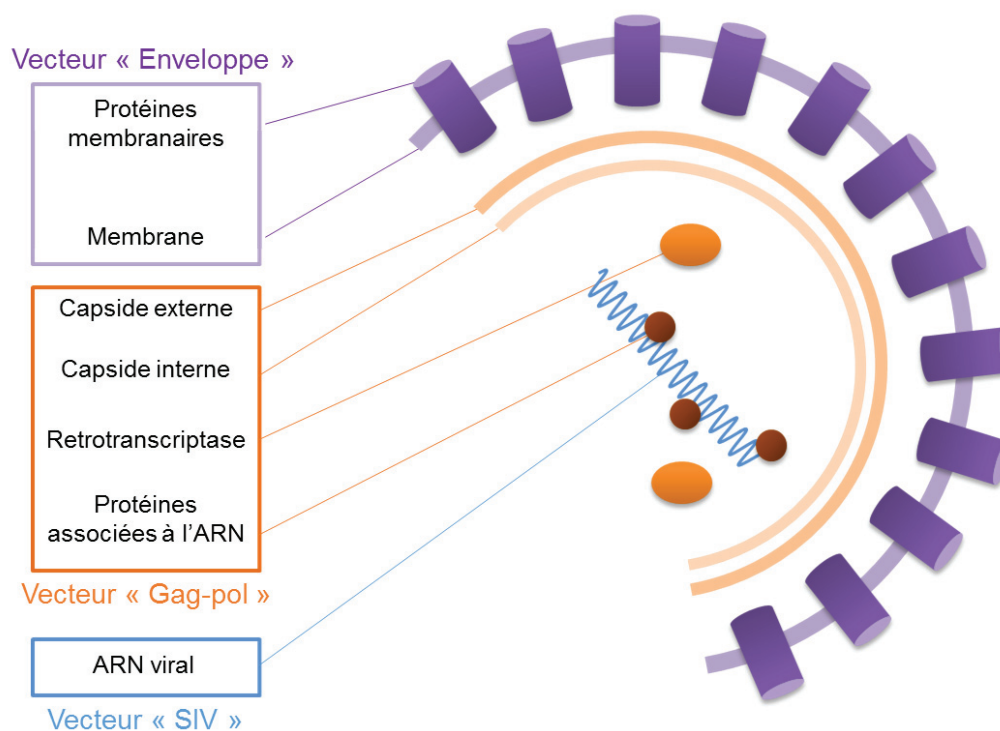
Afin d'introduire ensuite l'activateur de notre promoteur ainsi que la protéine HOX d'intérêt, différentes techniques s'offrent à nous.

La plus évidente est la transfection. Dans les cellules HEK 293T il est très facile de transférer un plasmide codant pour nos protéines d'intérêt. Dans cette lignée, la qualité de la transfection sera d'environ 70 à 90 % suivant le transfectant utilisé, avec une létalité elle aussi dépendante du produit. Cependant, ces produits ne sont pas aussi efficaces sur les différentes lignées cancéreuses. En effet, après des tests sur les MDA-MB231, les PC3 et les SKOV3, nous avons pu constater qu'aucun produit n'est capable de les transfecter de manière égale et au-delà de 40% d'efficacité.

Nous avons alors testé l'électroporation via le kit Neon® d'Invitrogen. Celui-ci nous a permis d'augmenter la qualité d'incorporation de nos vecteurs dans les cellules au détriment de la survie de celles-ci. En effet, seules moins de 50% des cellules survivent et sur ces dernières, 60 à 70% d'entre elles expriment nos protéines d'intérêt. Encore une fois, nous sommes bien en deçà de la qualité attendue pour ensuite pouvoir mener à bien un crible sur ces lignées.

Récemment, Eyckerman *et al.* ont mis au point une nouvelle technique basée sur les capacités d'encapsidation des rétrovirus (**Eyckerman *et al.*, 2016**) : les VLPs pour Virus Like Particule. Afin de pouvoir manipuler les rétrovirus en laboratoire leur matériel génétique a été séparé en trois parties : l'enveloppe (codée par un vecteur env et dans notre cas le VSVG), les

protéines de la capsid (gag-pol) et l'ARN qui est ensuite encapsidé. Ils ont alors détourné ce système en fusionnant une protéine d'intérêt (initialement la GFP) aux protéines de la capsid (Figure 4.2.2A). Cette fusion gag-protéine est alors encapsidée dans le virus, puis délivrée aux cellules. Ainsi, ce n'est plus un morceau d'ARN codant une protéine qui est délivré afin d'être intégré dans les cellules mais directement la protéine d'intérêt. L'expression de celle-ci est donc transitoire puisqu'aucun matériel génomique ne permet sa réexpression dans la cellule. De plus, contrairement aux constructions rétrovirales classiques, les VLPs ne contiennent aucun patrimoine génétique qui leur permettrait d'induire des mutations/insertions dans le génome. Ainsi il est aisé de les manipuler en confinement P2 sans nécessité de passer dans un confinement P3 beaucoup plus contraignant.

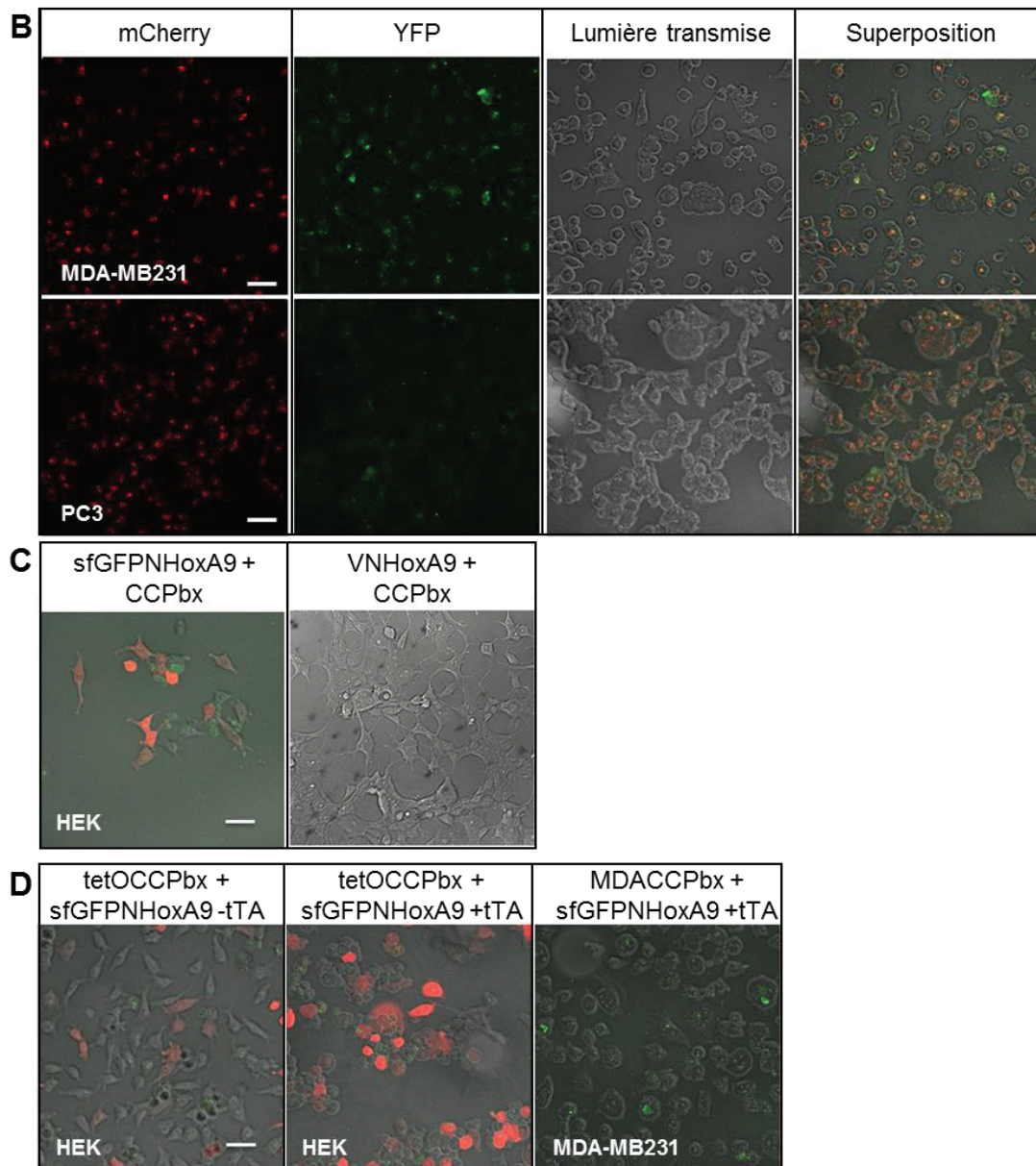


**Figure 4.2.2A : Composition d'une particule virale.** Le vecteur d'enveloppe va permettre de déterminer la cible d'infection. Nous utilisons le vecteur codant pour la protéine d'enveloppe VSVG qui est capable d'infecter les cellules humaines en utilisant le LDL récepteur à la surface. Le vecteur *Gag-pol* permet la formation de particules virale. C'est aux protéines Gag permettant la formation de la Capsid que sont fusionnés les protéines d'intérêt utilisées dans les VLPs. Elles sont ainsi encapsidées avec le reste du matériel du virus. Le vecteur « SIV » ou « HIV » en fonction des particules permet l'intégration d'un ADNc dans les cellules. Dans le cas des VLPs il n'est pas introduit dans les particules.

Initialement, la technique a été mise au point sur des cellules HEK 293T (Eyckerman *et al.*, 2016). Nous avons alors voulu tester s'il était ainsi possible d'apporter transitoirement dans nos lignées cancéreuses une protéine fluorescente pour commencer. Pour cela nous avons construit deux vecteurs, l'un codant pour une protéine gag-YFP et l'autre pour gag-mCherry. Nous avons successivement transduit les différentes lignées cellulaires avec les particules vertes puis rouges, ou rouges puis vertes, ou les deux simultanément. Nous avons pu observer d'une part au microscope confocal que toutes les cellules étaient bien transduites (Figure 4.2.2B). Le signal reste cependant ponctiforme dans les cellules, ne se répartissant pas de façon uniforme. D'autre part nous avons quantifié au FACS le pourcentage de cellules fluorescentes et nous arrivons à environ 100% de cellules fluorescentes (entre 98 et >99% pour chaque lignée). Les



cellules ne subissent pas le même stress qu'avec les autres techniques et la létalité est fortement diminuée.



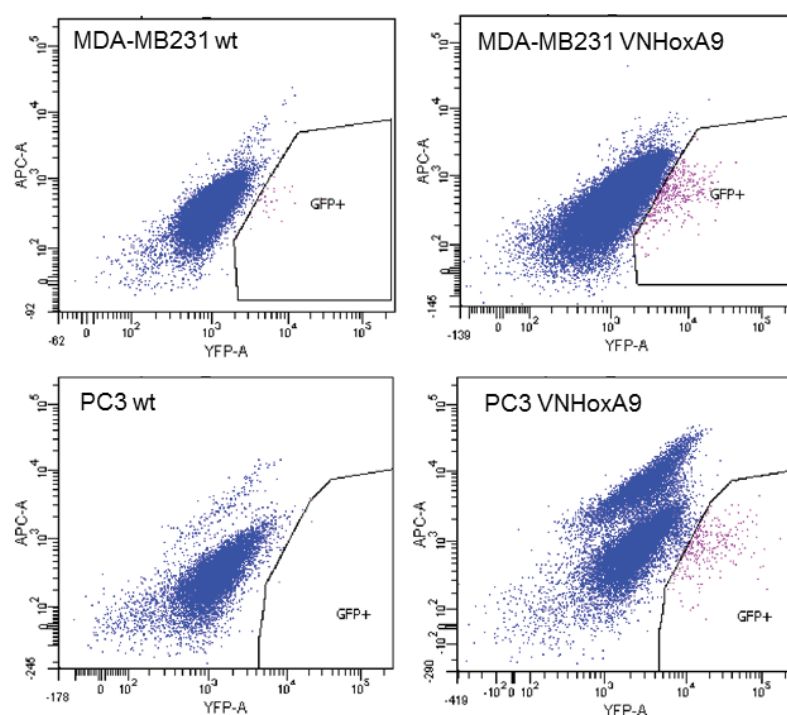
**Figure 4.2.2 :** Les VLPs permettent de délivrer des protéines fonctionnelles dans les différentes lignées cellulaires. Dans les cellules HEK la protéine mCherry est utilisée comme témoin de la qualité de transfection (*C et D*). (**B**) Des VLPs contenant la protéine mCherry ainsi que des VLPs contenant la protéine YFP sont délivrées à des cellules MDA-MB231 et PC3 sauvages puis sont observées au microscope confocal 24h après l'infection. Le signal des deux protéines fluorescentes est observable dans 100% des cellules. (**C**) Des cellules HEK ont été préalablement transfectées avec des constructions exprimant la protéine CCPbx1 ainsi que la protéine mCherry. Des VLPs sont ajoutées 18h après la transfection sur les cellules pour délivrer les protéines sfGFPNHoxA9 ou VNHoxA9. La BiFC est observée 6h après et l'on peut détecter un signal de BiFC dans les noyaux des cellules, démontrant que la protéine est correctement délivrée et fonctionnelle. (**D**) Des cellules HEK ont été transfectées par une construction contenant CCPbx1 en aval du promoteur tet-opérateur (tetO) ainsi que par une construction permettant l'expression de la mCherry. Les protéines sfGFPNHoxA9 ainsi que l'activateur sont apportés par les VLPs après 18h de transfection. On peut observer un signal BiFC dans les cellules, confirmant ainsi la robustesse du système. Pour valider celui-ci dans le contexte du crible, nous avons apporté les

protéines sfGFPN<sub>HoxA9</sub> et tTA par VLPs dans la lignée cellulaire MDA-MB231 contenant la banque CCORF et là aussi nous pouvons observer un signal de BiFC dans certaines cellules, révélant des cofacteurs potentiels de la protéine HoxA9.

Nous avons alors testé une construction VN<sub>HoxA9</sub> ainsi que sfGFPN<sub>HoxA9</sub> sur des cellules préalablement transfectées par CC-Pbx (**Figure 4.2.2C**). Nous avons pu constater qu'il était possible de voir un signal de BiFC localisé dans les noyaux des cellules HEK 293T. Des tests ont ensuite été réalisés dans les conditions du crible, c'est-à-dire directement sur les lignées PC3 et MDA-MB231 possédant l'intégration de la banque CC-ORF, par ajout successif de VLPs contenant l'activateur tTA puis de la fusion sfGFPN<sub>HoxA9</sub> (**Figure 4.2.2D**). Dans ce cas aussi nous pouvons observer un signal de BiFC dans les noyaux des cellules. Il est donc possible de visualiser des IPPs par la combinaison de la BiFC et des VLPs. Il est encore nécessaire de mettre au point cette technique, à savoir si la sfGFP est plus adaptée ou non que la Venus, à quel moment est-il plus judicieux d'ajouter les VLPs, si l'activateur tTA est vraiment nécessaire car d'autres expériences nous ont montré que même sans activateur le promoteur tet-opérateur a une activité minimale assez forte.

#### 4.2.3. Identification des candidats par séquençage

Suite à la transfection ou électroporation des constructions VN<sub>HoxA9</sub> et VN<sub>HoxA9</sub> muté dans le motif HX sur les lignées HEK 293T, PC3 et MDA-MB231 CCORF, nous avons pu visualiser un signal BiFC au microscope confocal. Nous avons alors réalisé un tri cellulaire FACS sur la fluorescence Venus (YFP) de nos cellules, isolant ainsi les cellules contenant potentiellement une intégration codant pour un partenaire de la protéine HoxA9 mutée ou non dans son motif HX. Une très faible proportion des cellules initiales est triée comme positive (<0,01%) ce qui implique un tri long et fastidieux d'une population très rare (**Figure 4.2.3A**).



**Figure 4.2.3A :** Une faible proportion des cellules électroporées est triée au FACS grâce au signal BiFC. Suite à l'électroporation de VN<sub>HoxA9</sub> et VN<sub>HoxA9</sub>HX dans les cellules MDA-MB231 et PC3, les cellules fluorescentes sont isolées. Les fenêtres de tri ont été déterminées à partir d'une population de cellules non fluorescentes. Ne sont ici présentés que les résultats avec VN<sub>HoxA9</sub>, le tri avec la forme mutée dans le motif HX ayant strictement le même profil.

L'ADN génomique de ces cellules positives a été extrait puis les intégrations amplifiées par PCR avant d'être séquencées. Nous avons alors généré des listes de candidats potentiels pour chaque contexte cellulaire, et en fonction de la mutation du motif HX. Suivant la

représentativité de chaque candidat dans le séquençage (le nombre de fois où il a été identifié sur l'ensemble des lectures de la puce de séquençage par rapport au nombre de fois où il est présent dans des cellules non triées) nous pouvons extrapoler sur l'enrichissement de ce candidat grâce au tri. Cet enrichissement traduit une interaction potentielle avec notre protéine HoxA9 dans le contexte cellulaire donné.

Ce crible pilote nous a donné une liste de candidats potentiels de la protéine humaine HoxA9 plus ou moins enrichis en fonction des contextes de cancer de la prostate ou du sein (**Tableau 4.2.3B**).

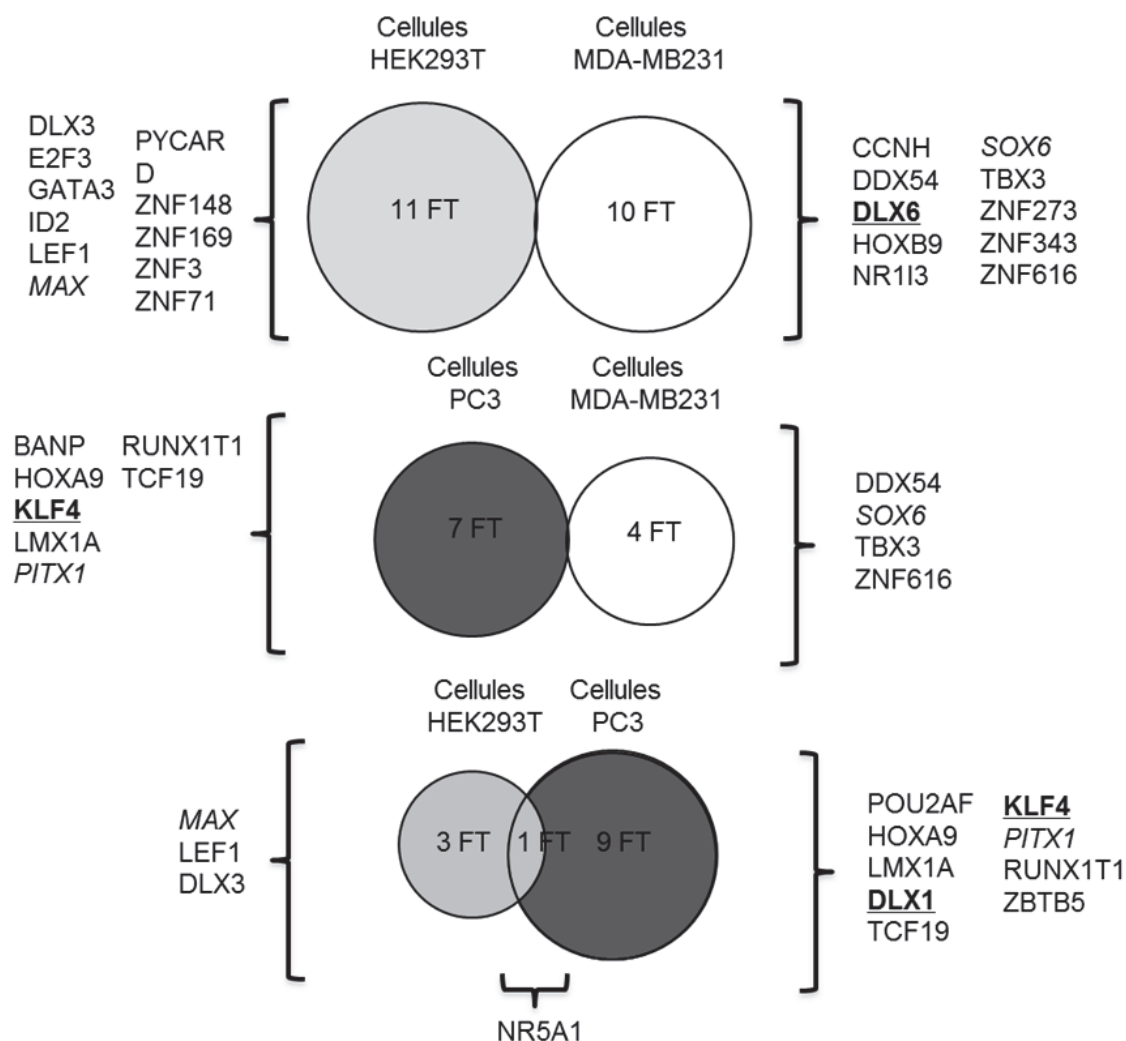
	HEK	MDA	PC3
<b>ATF3</b> ( <i>activating transcription factor 3</i> )	x23	présent	absent
<b>BANF1</b> ( <i>barrier to autointegration factor 1</i> )	x0,2	x74	absent
<b>BANP</b> ( <i>BTG3 associated nuclear protein</i> )	présent	présent	x13
<b>BZW2</b> ( <i>basic leucine zipper and W2 domains 2</i> )	x10	absent	absent
<b>C14orf166</b> ( <i>chromosome 14 open reading frame 166</i> )	présent	x26	présent
<b>CBX8</b> ( <i>chromobox homolog 8 (Pc class homolog, Drosophila)</i> )	x1,5	absent	x21
<b>DLX1</b> ( <i>distal-less homeo box 1</i> )	x0,2	absent	x14
<b>DLX6</b> ( <i>distal-less homeo box 6</i> )	x5	x118	absent
<b>E2F3</b> ( <i>E2F transcription factor 3</i> )	x8	présent	présent
<b>ENC2</b> ( <i>BTB/POZ KELCH domain protein</i> )	présent	présent	x11,5
<b>GATA3</b> ( <i>GATA binding protein 3</i> )	x7	présent	présent
<b>HMGNA4</b> ( <i>high mobility group nucleosomal binding domain 4</i> )	présent	x14	x7
<b>HOP</b> ( <i>homeodomain-only protein</i> )	présent	x0,1	x0,1
<b>HOXA9</b> ( <i>homeo box A9</i> )	x0,1	présent	x9,5
<b>ID2</b> ( <i>inhibitor of DNA binding 2, dominant negative helix-loop-helix protein</i> )	x9,5	présent	présent
<b>ING3</b> ( <i>inhibitor of growth family, member 3</i> )	x0,8	x1,3	x47
<b>KLF4</b> ( <i>Kruppel-like factor 4</i> )	x1,8	présent	x15
<b>LMX1A</b> ( <i>LIM homeobox transcription factor 1, alpha</i> )	présent	x1,1	x12
<b>NFYB</b> ( <i>nuclear transcription factor Y, beta</i> )	x15	présent	absent
<b>NR5A1</b> ( <i>nuclear receptor subfamily 5, group A, member 1</i> )	présent	absent	x12
<b>PHF23</b> ( <i>PHD finger protein 23</i> )	présent	présent	x17
<b>PITX1</b> ( <i>paired-like homeodomain transcription factor 1</i> )	x1,4	x0,5	x17
<b>POLR2H</b> ( <i>polymerase (RNA) II (DNA directed) polypeptide H</i> )	x11	absent	absent
<b>POU2AF1</b> ( <i>POU domain, class 2, associating factor 1</i> )	présent	absent	x11
<b>RNF138</b> ( <i>ring finger protein 138</i> )	x10	absent	absent
<b>RNF32</b> ( <i>ring finger protein 32</i> )	présent	absent	x11,5
<b>RUNXT1</b> ( <i>runt-related transcription factor 1; translocated to, 1 (cyclin D-related)</i> )	absent	absent	x21
<b>SOX6</b> ( <i>SRY (sex determining region Y)-box 6</i> )	présent	x44	présent
<b>TCF19</b> ( <i>transcription factor 19 (SC1)</i> )	présent	présent	x15
<b>THAP11</b> ( <i>THAP domain containing 11</i> )	présent	x0,8	x24,5
<b>ZBTB5</b> ( <i>zinc finger and BTB domain containing 5</i> )	présent	absent	x85
<b>ZCCH4</b> ( <i>zinc finger, CCHC domain containing 4</i> )	x20	absent	absent
<b>ZNF148</b> ( <i>zinc finger protein 148 (pHZ-52)</i> )	x16	présent	présent
<b>ZNF273</b> ( <i>zinc finger protein 273</i> )	présent	x7	x9
<b>ZNF697</b> ( <i>zinc finger protein 697</i> )	x10	présent	absent

**Tableau 4.2.3B :** Les candidats isolés par le crible contre la protéine VNHoxA9 sont plus ou moins enrichis en fonction des contextes cellulaires. Une partie seulement des candidats séquencés est ici présentées. Le facteur d'enrichissement est calculé par rapport à la représentation de chaque membre dans les cellules non triées. « Absent » signifie que le gène n'a pas été intégré ni séquencé dans les

cellules non triées. « Présent » signifie que le gène a bien été intégré et séquencé dans les cellules non triées mais qu'il n'est pas enrichi par le tri.

#### 4.2.4. Confirmation des candidats en BiFC

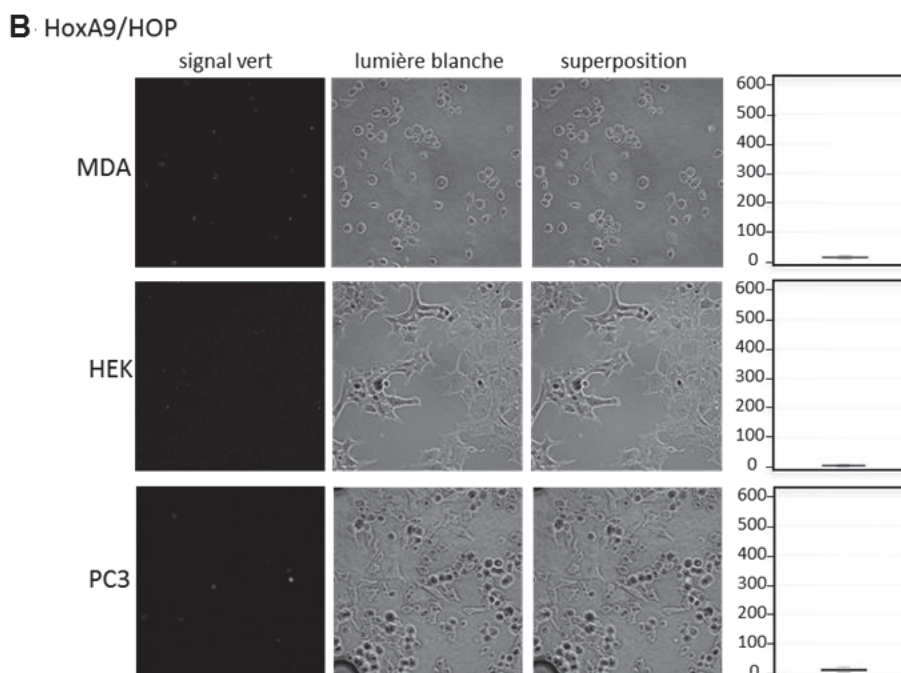
Afin de prouver la robustesse du crible et avant la mise au point de tous les paramètres, nous avons testé différents candidats isolés à partir du crible pilote. Pour cela nous avons choisi de comparer des FTs isolés uniquement dans la lignée PC3, ou MDA-MB231, ou HEK 293T (**Figure 4.2.4A**).



**Figure 4.2.4A :** Certaines interactions entre HoxA9 et des FTs sont spécifiques à une lignée cellulaire. Les FTs qui étaient présents et séquencés dans les lignées cellulaires non triées ont été analysés et comparés suite au séquençage des cellules issues du crible contre la protéine HoxA9. Ainsi 11 FTs interagissent avec la protéine HoxA9 dans les cellules HEK293T mais pas dans les cellules MDA-MB231, et 10 dans les MDA-MB231 mais pas dans les cellules HEK 293T.

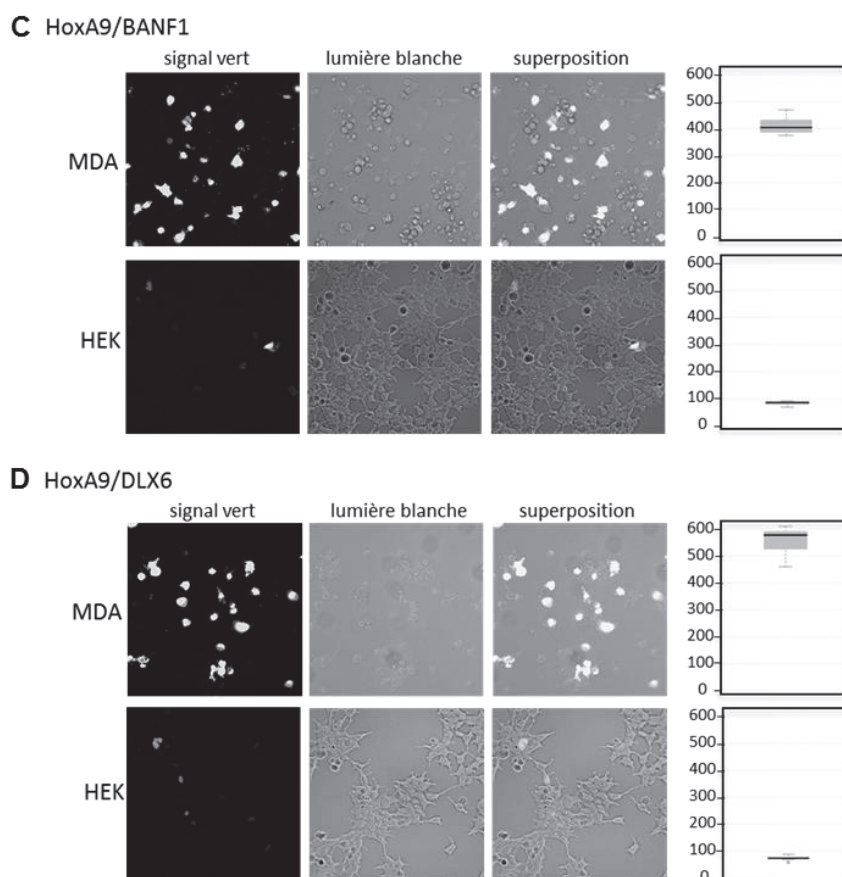
Nous avons alors isolé les candidats à partir de la banque CC-ORF et cloné ces derniers dans un vecteur d'expression en cellules mammifères (pCDNA3) afin de vérifier les interactions par des expressions transitoires du CC-Cofacteur en présence de VN-HoxA9. La protéine HOPX a été utilisée en guise de contrôle car elle est bien intégrée dans les trois lignées mais jamais enrichie par le crible. On peut constater qu'elle n'interagit pas non plus avec

HoxA9 lors de son expression transitoire dans les différentes lignées (**Figure 4.2.4B**). Le gène candidat *Dlx6* est particulièrement enrichi dans les cellules MDA-MB231. Lors de l'expression transitoire de cette protéine candidate dans les trois lignées, on ne constate une forte interaction avec HoxA9 que dans le contexte des cellules MDA-MB231 (**Figure 4.2.4D**). De la même façon, le gène *ID2* est enrichi dans les cellules contrôles HEK 293T et on confirme l'interaction entre les protéines ID2 et HoxA9 uniquement dans le contexte cellulaire HEK 293T en BiFC (non montré).



**Figure 4.2.4B :** La protéine HOPX n'interagit pas avec la protéine HoxA9 dans toutes les lignées cellulaires. La co-expression des protéines VNHoxA9 et CCHOPX dans les différentes lignées cellulaires n'entraîne pas la formation d'un signal BiFC. Le gène *HOPX* n'étant pas enrichi par le crible, cela constitue donc un bon contrôle négatif de notre tri cellulaire.

Par opposition, les protéines candidates KLF4 et BANF1 dont l'interaction avec la protéine HoxA9 semblait spécifique des contextes cellulaires MDA-MB231 et PC3 respectivement dans le crible pilote, semblent capables d'interagir avec la protéine HoxA9 dans les deux contextes cellulaires de cancer indifféremment (**Figure 4.2.4C**). La confirmation est donc nécessaire afin de découvrir des partenaires bien spécifiques d'un contexte particulier.



**Figure 4.2.4 C et D :** Les protéines BANF1 et DLX6 sont capables d’interagir avec la protéine HoxA9 dans un contexte de cancer du sein. Les co-expressions des protéines VNHoxA9 et CCBANF1 (C) ou CCDLX6 (D) dans les cellules MDA-MB231 entraînent la formation d’un complexe BiFC. Le gène *BANF1* n’a été trouvé enrichi que dans les cellules PC3 lors du crible. Cela nous démontre la nécessité de contrôler systématiquement les interactions isolées par le crible.

#### 4.2.5. L’échec des précédents cribles

Il est important de mentionner que cette technique n’est pas celle initialement prévue lors de l’établissement de mon projet de thèse. En effet, celle-ci a mûri et beaucoup évolué au cours des années passées. Il est largement décrit que la surexpression de HoxA9 est suffisante pour entraîner l’hyperprolifération des cellules souches hématopoïétiques de souris (**Sitwala et al., 2008**) et que l’ajout de Meis ou Pbx va permettre d’accentuer ce phénomène (**Thorsteindottir et al., 2001 ; Li et al., 2012**). De plus, afin de comprendre les événements reliés au très mauvais pronostic des leucémies myéloïdes aiguës induites par HoxA9, de gènes cibles et des sites de liaison génomiques ont été identifiés dans des cellules souches hématopoïétiques d’Homme (**Dorsam et al., 2004**) ainsi que de souris (**Huang et al., 2012**) surexprimant une protéine HoxA9.

Ces résultats nous ouvraient des portes afin de mieux comprendre les mécanismes des protéines Hox dans les cancers et nous voulions donc trouver des partenaires du complexe Hox/TALE dans un contexte leucémique. Pour cela nous comptons immortaliser des cellules souches hématopoïétiques humaines à l’aide de plasmides permettant d’intégrer les protéines VNHoxA9 ainsi que CNMeis1a, ou CNPbx3. Puis à partir des lignées établies, nous comptons

appliquer la même technique de crible que Ding et al., c'est-à-dire apporter de façon aléatoire la cassette avec le promoteur Tet-opérateur ainsi que la CC et cribler des partenaires au hasard.

Or nous n'avons jamais réussi à obtenir une lignée à partir de ces intégrations initiales dans les cellules souches hématopoïétiques humaines (obtenues à partir de sang de cordon ou de moelle osseuse adulte). Nous avons aussi tenté de générer ces lignées dans les cellules souches hématopoïétiques de souris, et dans ce cas nous avons bien obtenu une hyperprolifération des cellules, jusqu'à un palier, puis celles-ci cessaient de se multiplier. Or pour le crible nous avons besoin d'une quantité de matériel bien supérieure à ce que nous avons pu obtenir de la sorte. La génération des outils, la mise en place et la mise au point des techniques ont demandé beaucoup d'efforts et méritent d'être mentionné car la littérature ne renseigne pas sur l'immortalisation de cellules souches hématopoïétiques humaines de la sorte. Finalement nous avons adapté le modèle en nous orientant sur des lignées cellulaires comme nous l'avons présenté plus haut.

# DISCUSSION

---

## 1. CHAPITRE 1 : VERS DE NOUVELLES INTERFACES D'INTERACTION

### 1.1. POURQUOI LE MOTIF HX EST LONGTEMPS RESTE L'UNIQUE MOTIF D'INTERACTION DES COFACTEURS TALE AU SEIN DES PROTEINES HOX ?

L'étude de l'interaction des protéines Hox avec leurs partenaires TALE a toujours été concentrée sur l'interaction directe entre la protéine Hox et son partenaire PBC. En effet, de nombreux tests ont été réalisés *in vitro*, tels que des gels retards, mais seulement avec les protéines Hox et PBC ou des parties tronquées de ces dernières, en l'absence de Meis et avec des sites artificiels.

Comme nous l'avons aussi introduit, par cristallographie, des portions tronquées des protéines Hox et PBC ont ainsi été analysées au contact de l'ADN, révélant d'autant plus le rôle du motif HX dans cette interaction. Ces structures ne révèlent donc pas le d'autres régions protéiques pouvant aussi intervenir dans cette interaction. De plus le rôle du motif HX n'a pas été creusé d'avantage au sein d'un complexe trimérique contenant la protéine Meis.

Enfin, il existe peu de sites physiologiques connus pour être régulés par des complexes Hox/Pbx/Meis qui permettraient une dissection moléculaire des interactions au sein de ce complexe trimérique *in vitro*. L'interaction du complexe Scr/Exd sur un site utilisé pour la régulation du gène cible *forkhead (fkh)* montre ainsi le rôle de résidus paralogues-spécifiques additionnels pour la formation du complexe sur l'ADN (Joshi et al., 2007), ce qui souligne l'importance d'utiliser des sites physiologiques pour révéler de nouvelles interactions en cristallographie.

### 1.2. LES PROTEINES HOX UTILISENT DES INTERFACES PARALOGUES SPECIFIQUES ADDITIONNELLES POUR INTERAGIR AVEC LES COFACTEURS TALE

Pourtant certains motifs spécifiques responsables de l'interaction des protéines Hox avec leurs partenaires TALEs ont pu être identifiés. C'est le cas par exemple du motif UbDA dans les protéines de drosophile Ubx et AbdA (Balavoine *et al.*, 2002 ; Merabet *et al.*, 2007 ; Saadaoui *et al.*, 2011 ; Hudry *et al.*, 2012 ; Foos *et al.*, 2015), ou encore du motif TDWM dans la protéine AbdA d'insecte (Merabet *et al.*, 2011 ; Katherine *et al.*, 2011 ; Hudry *et al.*, 2012). Ces motifs s'ajoutent donc au motif HX pour créer de nouvelles interfaces spécifiques avec les cofacteurs TALE.

Dans le cas des protéines de souris, il a été montré *in vitro* et *in vivo* que des protéines Hox des groupes 6 à 10 pouvaient former des complexes avec leurs partenaires Pbx et Meis lorsque leur motif HX était muté (Hudry *et al.*, 2012). En effet, pour ces protéines la mutation du motif HX entraîne la perte d'interaction directe entre la protéine Hox et le partenaire Pbx, mais cette interaction est sauvée par l'ajout de la protéine Meis au sein du complexe supposant



ainsi la présence d'autres motifs responsables de cette redondance. Pourtant aucun motif commun à toutes ces protéines n'a pu être mis en évidence, soulignant le fait que cette redondance pourrait reposer sur des motifs spécifiques. Ces résultats restaient à démontrer de manière plus large avec des protéines Hox humaines et d'autres groupes de paralogie.

Nous avons alors analysé de façon large tous les groupes de paralogie en utilisant de préférence des protéines Hox humaines, et quand c'était possible, plusieurs représentants de chaque groupe de paralogie. Ce que nous avons pu constater c'est que, à l'instar des protéines Hox de souris, il existe bien une redondance fonctionnelle du motif HX pour l'ensemble des groupes de paralogie centraux et postérieurs, redondance qui dépend systématiquement de la présence additionnelle de Meis. L'interaction au sein des complexes Hox/Pbx/Meis pour les protéines Hox des groupes antérieurs 1 et 2 semble reposer exclusivement sur le motif HX *in vitro*.

Suite à cette analyse, nous avons cherché à déterminer si, au sein des protéines Hox centrales et postérieures, il était possible de trouver les motifs susceptibles d'être des interfaces d'interaction avec les cofacteurs TALE. Les prédictions de ces SLiMs (Short Linear Motif, **Introduction Chapitre III.2**) ont permis de mettre en évidence que leur nombre varie de trois motifs maximum pour les protéines Hox des groupes 7 à 13, à plus de cinq motifs pour les protéines Hox des groupes 3 à 6. Nous avons pu alors montrer que l'interaction entre certaines protéines Hox et leur partenaire TALE utilisait l'un de ces petits motifs. C'est le cas de la protéine HoxA7 ainsi que de la protéine HoxC8. Dans le cas de la protéine HoxA9, nous avons prouvé une redondance fonctionnelle du motif HX avec certains résidus spécifiques de l'HD pour l'interaction avec le partenaire TALE *in vitro* et *in vivo*.

Ces résultats tendent à montrer que, pour certains groupes de paralogie, il existe des motifs spécifiques utilisés pour réaliser ces interactions complexes. Le motif que nous avons mis en évidence dans la protéine HoxA7 est retrouvé chez toutes les protéines Hox mammifères du groupe 7. Quant au motif trouvé dans la protéine HoxC8, lui n'est retrouvé que dans les protéines HoxC de mammifère du groupe 8.

De plus, ces modes d'interaction alternatifs ne sont retrouvés que dans les complexes trimériques incluant la protéine Meis qui permet de sauver l'interaction entre les protéines Hox et Pbx lors de la mutation du motif HX. On peut supposer que des interactions directes entre la protéine Hox et la protéine Meis ont lieu plutôt qu'avec la protéine Pbx pour constituer les complexes trimériques. Des structures cristallographiques des complexes entre les trois protéines sont nécessaires afin de confirmer ou non ces hypothèses.

Enfin, il est important de noter que, du fait de leur implication dans de nombreux cancers (**Bhatlekar et al., 2014 ; Sitwala et al., 2008**), les protéines Hox sont les cibles de plusieurs études visant à induire une apoptose dans les cancers. Les protéines Hox sont impliquées dans les mécanismes oncogéniques avec la participation de leurs partenaires TALEs d'où le nombre important d'études visant à abolir cette interaction. Pour cela, un peptide imitant le motif HX a été créé et valorisé comme un traitement anti-cancéreux prometteur. Or ce peptide a aussi la propriété de cibler les motifs HX de toutes les protéines Hox dans tous les contextes cellulaires, d'autant plus qu'au sein d'un même cancer certaines protéines Hox peuvent avoir un rôle oncogénique alors que d'autres ont un rôle suppresseur de tumeur. Afin de rendre ce type de traitements plus spécifiques et efficaces, l'étude des nouveaux petits motifs d'interaction prend tout son sens. En effet, en ciblant spécifiquement une de ces nouvelles interfaces d'interaction,

il serait possible de cibler un complexe Hox/TALE en particulier, ou un seul groupe de paralogie plutôt que les protéines Hox dans leur ensemble.

### 1.3. VERS UNE VALIDATION FONCTIONNELLE DE LA DISSECTION MOLECULAIRE

Afin de pouvoir nous diriger vers de telles ambitions thérapeutiques, il faudrait avant tout valider le rôle des motifs identifiés *in vitro* et dans des cultures cellulaires, dans des modèles *in vivo*. Il est primordial de déterminer l'intérêt fonctionnel de ces interfaces d'interactions spécifiques à certains groupes de paralogie.

Comme nous l'avons déjà introduit, il a été largement décrit que la protéine Hoxa9 de souris était capable à elle seule d'immortaliser des cellules souches hématopoïétiques humaines. De plus, lors de la surexpression conjointe de la protéine Hoxa9 et de son partenaire Meis dans ce même contexte, l'effet pro-prolifératif et anti-différenciation est d'autant plus grand, supposant une action conjointe de ces deux protéines. Un premier moyen de valider notre approche serait de tester les différentes formes de notre protéine HoxA9 humaine dans des cellules souches hématopoïétiques. Nous pourrions nous attendre à obtenir une diminution de l'effet oncogénique de la protéine HoxA9 en coopération avec Meis lorsque le motif HX et les résidus clés de l'HD sont simultanément mutés. Dans le cas contraire, nous pourrions nous demander si le contexte des cellules hématopoïétiques est bien le bon pour l'utilisation de cette interface et nous pourrions trouver un autre contexte où ce motif pourrait avoir un rôle important.

Un autre exemple différent pour étudier ces nouvelles interfaces serait de s'orienter cette fois vers une lignée cellulaire cancéreuse établie, telle que les cellules MCF7 de cancer du sein, dans lesquelles on sait que certaines protéines Hox jouent un rôle clé dans les capacités invasives et prolifératrices des cellules (**Bhatlekar et al., 2014**). C'est le cas par exemple de la protéine HoxA7 dans le cancer du sein (**Cantile et al., 2003**). Elle est surexprimée dans les cancers du sein, laissant supposer son rôle oncogénique dans ce contexte. Nous pourrions donc exprimer d'avantage cette protéine dans la lignée MCF7 avec les différentes mutations et déterminer si le rôle de la protéine HoxA7 sur les capacités invasives et prolifératives des cellules MCF7 dépend du motif GAGA ou non. Si tel est le cas, ce rôle pourra ensuite être confirmé par l'addition de peptides mimant la séquence de ce motif comme décrit pour le motif HX (**Morgan et al., 2012**).

### 1.4. INTERET DES DIFFERENTS MODES D'INTERACTION HOX-TALE

Ces différents modes d'interaction sont importants pour deux aspects principaux : la **diversité fonctionnelle**, ainsi que la **spécificité fonctionnelle** des protéines Hox (**Merabet et al., 2011**).

Comme nous l'avons déjà présenté, les protéines Hox ont des rôles très différents au sein d'un même contexte cellulaire et surtout une même protéine Hox peut aussi avoir des rôles différents d'un contexte cellulaire à l'autre. Cette diversité fonctionnelle repose certainement sur la présence d'un contexte protéique radicalement différent entre les différents contextes cellulaires. Des co-facteurs tissu spécifiques peuvent expliquer que dans certains cas une protéine Hox agira par exemple comme un oncogène très puissant, c'est le cas de la protéine

Hoxa9 dans la leucémie, ou comme un anticancéreux, dans le cancer du sein. La présence de plusieurs interfaces d'interaction entre les protéines Hox et les TALE permet de libérer d'autres interfaces ouvrant ainsi le recrutement de nouveaux partenaires en fonction des contextes cellulaires.

Au sein d'un même contexte cellulaire, les protéines Hox ont des rôles qui peuvent être très spécifiques et différents. Avec des interfaces d'interaction spécifiques des différents groupes de paralogie, chaque protéine Hox va pouvoir interagir avec les partenaires TALE de façon spécifique et potentiellement ouvrir à la reconnaissance de sites spécifiques sur l'ADN. On observe ainsi une spécificité fonctionnelle apportée par l'utilisation de modes alternatifs de formation des complexes.

## 2. CHAPITRE 2 : RECHERCHE DE NOUVEAUX PARTENAIRES DES PROTEINES HOX HUMAINES EN CONTEXTE ONCOGENIQUE

Nos résultats préliminaires nous laissent entrevoir la possibilité d'étendre notre méthode de crible afin de répondre à plusieurs questions biologiques. En effet, l'intégration de 80% des gènes de l'ORFéome v8.1 contenus dans la banque CC-ORF au sein de la lignée des cellules HEK nous permet d'espérer une telle représentativité dans les autres lignées cellulaires. De plus, la validation des candidats isolés par le crible apporte la preuve nécessaire à la validation de ce concept. Il est cependant important d'améliorer certaines des étapes du cribles qui sont, à l'heure actuelle, sources de biais non négligeables.

### 2.1. AMELIORER LA REPRESENTATIVITE DE LA BANQUE

Afin d'assurer la meilleure représentativité possible de la banque dans les lignées cellulaires, deux étapes doivent être améliorées. La première correspond à la génération initiale de la banque CC-ORF. En effet, comme nous l'avons présenté, cette banque a été générée par un clonage de type GATEWAY dans un vecteur contenant le promoteur tet-operateur ainsi que la cassette CC. Le premier biais de cette technique dans le cadre du clonage d'une banque est celui de l'intégration favorisée des petits ADNc. Nous avons créé une première banque de plasmides que nous avons séquencée avant l'intégration dans les cellules. Nous avons ainsi mis en évidence que l'essentiel des gènes manquant par rapport à la banque ORFeome initiale étaient les membres de grande taille. Afin de palier à ce premier biais, nous pourrions trier les candidats par leur taille avant la recombinaison en faisant migrer la banque plasmidique linéarisée. Puis nous pourrions réaliser les recombinaisons en plusieurs expériences différentes avant de rassembler tous les clones pour la production de particules virales cette fois enrichie des gènes de grande taille.

Un autre biais repose sur le fait que, d'une lignée cellulaire à l'autre, les particules virales apportant les gènes n'ont pas les mêmes capacités d'infection. Nous avons en premier généré des lignées PC3 et MDA-MB231 CCORF en tenant compte du titre infectieux établi sur les cellules HEK. Face à la faible représentativité de la banque, nous avons donc établi le titre infectieux pour chaque lignée cellulaire indépendamment les unes des autres et nous nous sommes rendu compte que la qualité varie ici d'un facteur 10 suivant la lignée que nous utilisons. Cette approche a permis d'améliorer la représentativité de la banque dans les cellules PC3 (passée de 36% à 55%). Ainsi établir des paramètres d'infection optimum devient une priorité. Les paramètres sur lesquels il est primordial de jouer sont les productions des particules, afin de les concentrer le plus possible, la spinoculation, l'ajout de Polybrène qui favorise l'entrée des particules dans les cellules. Nous ne pouvons pas augmenter la quantité de particules par cellule, ce qui permettrait de gagner en nombre d'intégrations mais nous ferait diminuer les simples intégrations à la faveur de multiples. Nous introduirions ainsi un biais dans la sélection des candidats.

### 2.2. ETENDRE LE CRIBLE A DIFFERENTES LIGNEES

Une fois les paramètres d'infection correctement établis, nous allons nous diriger vers un choix plus large de lignées cellulaires pour les tests d'interaction. En effet les lignées que nous avons présentées ne sont pas les seules dans lesquelles nous pourrions chercher de

nouveaux partenaires des protéines Hox. Les cellules MCF7 de cancer du sein, HeLa du cervix, HL60 de leucémie, *etc.* pourraient aussi être d'excellents candidats.

Nous avons discuté dans la revue **Merabet et Dard (2013)** ainsi que dans l'introduction de l'importance des interactions Hox-cofacteurs pouvant être spécifiques d'un contexte. Nos résultats montrent que la stratégie utilisée permet en effet de révéler des interactions pouvant être spécifiques d'un type cellulaire. Par exemple l'interaction avec Dlx6 est spécifique de la lignée MDA-MB231, ce qui pourrait être le cas d'un certain nombre d'interactions (**Tableau Résultats**). Il est important de noter que le crible réalisé avec HoxA9 dans les lignées MDA-MB231 et PC3 correspond à des contextes cellulaires dans lesquels la protéine HoxA9 a une activité anti- ou pro-oncogénique respectivement. Cela permet ainsi de potentiellement révéler des partenaires responsables de ces deux activités opposées.

Ces partenaires candidats isolés à partir de cellules de cancer du sein MDA-MB231 pourraient être comparés à d'autres partenaires candidats issus de nouveaux cribles sur d'autres lignées de type cancer du sein. Il en existe de nombreuses qui répondent ou non à l'inhibition des complexes Hox/Pbx (**Morgan et al., 2012**), telles que les lignées SKBR3, ZR75 ou MCF7. Nous pourrions alors établir une liste de candidats qui auraient un profil \*cancer du sein\* ou simplement \*cellule du sein\* pour les partenaires communs.

Enfin, dans un but de valorisation de ce projet, proposer différentes lignées permettrait à d'autres laboratoires de s'approprier cette technique de crible avec leurs propres candidats à tester dans le contexte qui correspond le mieux à leur question biologique.

### 2.3. REALISER LE CRIBLE AVEC DIFFERENTES PROTEINES HOX SAUVAGES OU MUTEES

Notre système permet aussi de poser la question de la spécificité des interactions par rapport à une protéine Hox, et d'en connaître les détails moléculaires. Sachant que plusieurs protéines peuvent être exprimées dans un même tissu, il serait intéressant de savoir si celles-ci sont capables d'établir des interactions identiques. Cette question revêt une importance d'autant plus grande qu'il est démontré que des protéines Hox différentes peuvent avoir des fonctions opposées dans certains cancers comme celui du sein (**Morgan et al., 2012**). Est-ce que les activités opposées de HoxA5 et HoxA9 dans le cancer du sein reposent sur des interactions avec différents cofacteurs par exemple ?

Dans la même logique, des protéines Hox avec des fonctions similaires ou opposées dans les cancers pourraient utiliser des interfaces d'interaction différentes, ou un même motif, mais d'autres cofacteurs. Le motif le plus intéressant à étudier dans ce contexte serait bien sûr le motif HX, qui est décrit pour être nécessaire à l'activité oncogénique de plusieurs protéines Hox. Il serait ainsi intéressant de savoir si la protéine HoxA9 mutée dans son motif HX établit les mêmes interactions que la protéine HoxA9 sauvage dans les cellules MDA-MB231 et PC3. Ceci permettrait de révéler des interactions pouvant potentiellement être responsables de l'importance du motif HX pour l'activité pro-oncogénique de HoxA9 dans les cellules PC3. Ce type d'analyse pourrait aussi être comparée avec d'autres protéines Hox mutées dans le motif HX. Cela permettrait de mettre en avant la diversité fonctionnelle et moléculaire pour un même motif dans différentes protéines Hox : différentes interactions pourraient très certainement être révélées, illustrant l'importance du contexte protéique pour l'utilisation d'un motif, comme déjà discuté (**Merabet et al., 2009**)

#### 2.4. APPORTER LE CANDIDAT PAR LES VLPs

Une autre étape critique de notre technique de crible correspond à l'introduction de l'activateur tetA ainsi que de la protéine d'intérêt. Cette étape repose actuellement sur une transfection ou une électroporation suivant le type cellulaire d'intérêt. Or d'une expérimentation à l'autre le taux de transfection/électroporation est extrêmement variable et surtout en deçà des valeurs espérées (moins de 40% pour les MDA-MB231 par exemple). Ce taux faible entraîne la nécessité d'un tri plus long et plus fastidieux au FACS pour isoler les cellules fluorescentes positives. D'où l'intérêt de l'utilisation des VLPs qui sont en cours de mise au point. Les résultats préliminaires montrent une délivrance de protéines fluorescentes dans presque 100% des cellules et une viabilité cellulaire bien meilleure qu'avec les autres techniques. En effet, l'électroporation est la technique la plus délétère pour les cellules puisque presque 50% des cellules ne survivent pas au choc électrique.

Un autre avantage des VLPs est qu'ils ne semblent pas montrer de différence évidente pour l'apport des protéines entre les différents types cellulaires. Ils pourraient alors être utilisés indifféremment sur toutes les lignées cellulaires, impliquant une diminution du coût et du temps de mise en place de la technique pour l'apport de la protéine d'intérêt, quelle qu'elle soit. Ainsi une seule et même production de ces particules pourrait permettre d'introduire la protéine dans toutes les lignées et entraîner un gain de temps conséquent, protéger des contaminations et de la létalité induites par l'électroporation, et avoir une meilleure répétabilité d'une expérience à l'autre.

#### 2.5. LA BiFC AVEC LA sfGFP : RAPIDITE DE MATURATION

Il a été montré que la protéine super folder GFP (sfGFP) est capable de compléter beaucoup plus rapidement que la protéine Venus lorsqu'elle est scindée en deux moitiés (**Zhou et al., 2011**). Grâce à cette propriété nous espérons nous affranchir de la saturation de la production de protéines dans les cellules et détecter au plus tôt les interactions stabilisées par la BiFC.

Nous avons montré que la partie N-terminale de la sfGFP (sfGFPN) est capable de compléter avec la partie C-terminale de la Cerulean en produisant un signal GFP. Nous espérons alors pouvoir utiliser la rapidité de maturation de la sfGFP à notre avantage dans le crible. En effet, il serait envisageable de délivrer une protéine de fusion entre la sfGFPN et une protéine d'intérêt directement dans les cellules grâce aux VLPs précédemment décrits. Par la combinaison de la BiFC avec la sfGFP et les VLPs nous arriverions à diminuer drastiquement le temps nécessaire pour voir les interactions entre la protéine d'intérêt et les candidats potentiel : il n'est plus nécessaire d'attendre que la protéine soit produite par la cellule cible, et le temps de maturation de la BiFC est diminué. Cela nous permettrait alors de piéger des interactions beaucoup plus rapidement.

#### 2.6. ETABLIR DES PARAMETRES DE SEQUENÇAGE OPTIMAUX

Après l'isolation des cellules fluorescentes, l'identification des candidats est une étape clé et délicate. En effet, partant d'une simple copie intégrée sur l'ADN génomique humain ; soit quelques centaines de paires de bases dans 3,4 milliards de paires de bases pour environ 24 000 gènes ; il est difficile d'amplifier spécifiquement l'intégration d'intérêt. Un nombre

important de cycles de PCRs est nécessaire afin de créer une banque et cela introduit un biais dans la représentativité des candidats. En effet, les petits fragments étant amplifiés de manière préférentielle par la polymérase, ils seront sur-représentés par rapport aux plus grands. La création d'amorces plus spécifiques est alors soulevée. L'actuelle technique de séquençage repose sur des amorces situées dans les sites de recombinaison gateway généré lors de la création de la banque. Afin de s'affranchir de ce biais, nous proposons de découper l'ADN génomique extrait des cellules en petits fragments et, par reconnaissance des sites gateway avec une amorce spécifique fusionnée à la biotine, récupérer uniquement les fragments biotinylés. Cette technique isolerait les ADNc préalablement intégrés et permettrait une amplification plus spécifique d'un matériel déjà trié, diminuant le nombre de cycles de PCR nécessaires pour avoir un matériel suffisant et éliminer l'ADN génomique qui ne nous intéresse pas. Cette stratégie est en cours d'essai.

## 2.7. VERS UNE BANQUE ORDONNÉE DE CANDIDATS

Pour rendre plus accessibles les candidats isolés par le crible il serait nécessaire de posséder une banque ordonnée pour la validation individuelle des cofacteurs candidats. Jusqu'à présent cette validation nécessite de re-cloner le cofacteur candidat en fusion avec CC à partir d'un ADNc dans un vecteur d'expression cellulaire. Cette étape étant longue et fastidieuse, elle est difficilement envisageable pour tous les cofacteurs candidats identifiés à partir du crible. Nous souhaiterions donc pouvoir cloner de manière ordonnée, par gateway, l'ensemble des gènes codant pour des facteurs de transcription de l'ORFéome v9.1 en fusion avec CC. Cette nouvelle banque ordonnée d'environ 1000 gènes pourrait ainsi être directement disponible pour des tests BiFC avec n'importe quelle protéine candidate. Elle serait donc également très utile pour l'ensemble de la communauté, au-delà de la validation post-crible en lignées cellulaires.

## 2.8. LE CHOIX DU PROMOTEUR

Lors du design du crible nous avons choisi d'exprimer nos cofacteurs candidats sous l'influence d'un promoteur tet-opérateur. Cette idée était venue du premier crible qui a été réalisé en utilisant la BiFC pour la protéine Akt-1 (**Ding et al., 2006**) dans les cellules HeLa. En effet, la cassette contenant le promoteur tet-opérateur en amont de la partie C-terminale de la Venus était intégrée de façon aléatoire dans le génome comme nous l'avons décrit dans l'introduction. L'idée principale était de s'affranchir d'une construction qui pourrait avoir un effet toxique sur les cellules. En effet, le promoteur tet-opérateur n'est normalement actif qu'en présence d'un activateur et donc éteint dans les cellules non stimulées (voir **Résultats**). Dans le cas où la protéine ainsi exprimée a un rôle suppresseur de tumeur ou un rôle oncogénique celui-ci ne pourrait pas introduire de biais lors de la génération de la banque de cellules en éliminant un candidat ou en favorisant l'un par rapport aux autres.

Or nous avons observé que le promoteur tet-opérateur n'est en fait pas complètement éteint et entraîne un taux basal d'expression. De plus nous avons sélectionné les lignées ayant intégré la banque par une pression de sélection par antibiotiques (puromycine) sans ajouter d'activateur du promoteur tet-opérateur, reposant cette sélection sur la seule activité minimale de ce promoteur. Les paramètres du crible ayant changé depuis son design jusqu'à sa mise en place et confrontés à des réalités biologiques qui ne pouvaient être prévues, l'utilité de ce promoteur en tant que promoteur activable est largement discutable. Il est pourtant justifiable

par le fait que son niveau d'expression étant très faible, il permet de limiter les faux positifs qui seraient dû à une trop forte expression des co-facteurs candidats dans les cellules. La BiFC stabilisant les interactions transitoires et faibles, l'expression réduite de nos candidats dans ce contexte n'est pas un frein mais plutôt un avantage.





## MATERIELS ET METHODES

Nous présentons ici les constructions ayant été utilisées pour le **Chapitre III des Résultats**. L'ensemble de ces constructions ont été clonées dans un vecteur **pCDNA3** en aval d'un promoteur CMV permettant l'expression des protéines dans les cellules mammifères ainsi que par l'utilisation de kit de production de protéines en lysat de réticulocytes de lapin.

	Clone	Source	Mutations	Clonage
1	HoxB1	addgene # 8520		EcoRI XbaI
2	HoxB1 HX	(from addgene # 8520)	FDWM → FAAA	EcoRI XbaI
3	VNHoxB1	(from addgene # 8520)		BamHI EcoRI XbaI
4	VNHoxB1 HX	(from addgene # 8520)	FDWM → FAAA	BamHI EcoRI XbaI
5	HoxA1	Merabet		EcoRI XbaI
6	HoxA1 HX	Merabet	FDWM → FAAA	EcoRI XbaI
7	VNHoxA1	Merabet		EcoRI XbaI
8	VNHoxA1 HX	Merabet	FDWM → FAAA	EcoRI XbaI
9	HoxB2	addgene # 8522		XhoI XbaI
10	HoxB2 HX	(from addgene # 8522)	FPWM → FAAA	XhoI XbaI
11	VNHoxB2	(from addgene # 8522)		BamHI XhoI XbaI
12	VNHoxB2 HX	(from addgene # 8522)	FPWM → FAAA	BamHI XhoI XbaI
13	HoxB3	addgene # 8523		XhoI XbaI
14	HoxB3 HX	(from addgene # 8523)	FPWM → FAAA	XhoI XbaI
15	VNHoxB3	(from addgene # 8523)		BamHI XhoI XbaI
16	VNHoxB3 HX	(from addgene # 8523)	FPWM → FAAA	BamHI XhoI XbaI
17	Hoxd4	addgene #8549		XhoI XbaI
18	Hoxd4 HX	addgene # 21002	YPWM → YPVM	XhoI XbaI
19	VNHoxd4	(from addgene # 8549)		BamHI XhoI XbaI
20	VNHoxd4 HX	(from addgene # 21002)	YPWM → YPVM	BamHI XhoI XbaI
21	HoxA5	Merabet		EcoRI XhoI
22	HoxA5 HX	Merabet	FPWM → FAAA	EcoRI XhoI
23	VNHoxA5	Merabet		EcoRI XhoI
24	VNHoxA5 HX	Merabet	FPWM → FAAA	EcoRI XhoI
25	VNHoxA5 DN90	Merabet		EcoRI XhoI
26	VNHoxA5 DN90 HX	Merabet	FPWM → FAAA	EcoRI XhoI
27	VNHoxA5 DN160	Merabet		EcoRI XhoI
28	VNHoxA5 HX DN160	Merabet	FPWM → FAAA	EcoRI XhoI
29	HoxB6	ORFeome v3.1		XhoI XbaI
30	HoxB6 HX	(from ORFeome v3.1)	YPWM → YAAA	XhoI XbaI
31	VNHoxB6	(from ORFeome v3.1)		BamHI XhoI XbaI
32	VNHoxB6 HX	(from ORFeome v3.1)	YPWM → YAAA	BamHI XhoI XbaI
33	HoxA7	Merabet		EcoRI XhoI
34	HoxA7 HX	Merabet	YPWM → YAAA	EcoRI XhoI
35	VNHoxA7	Merabet		EcoRI XhoI
36	VNHoxA7 HX	Merabet	YPWM → YAAA	EcoRI XhoI
37	VNHoxA7 DN 32	Merabet		EcoRI XhoI
38	VNHoxA7 DN 32 HX	Merabet	YPWM → YAAA	EcoRI XhoI
39	VNHoxA7 DN 47	Merabet		EcoRI XhoI
40	VNHoxA7 DN 47 HX	Merabet	YPWM → YAAA	EcoRI XhoI
41	VNHoxA7 ΔGA	Merabet	RSGYGAGAGAFASTV → RSGYASTV	EcoRI XhoI
42	VNHoxA7 HX ΔGA	Merabet	RSGYGAGAGAFASTV → RSGYASTV YPWM → YAAA	EcoRI XhoI
43	HoxC8	addgene # 21001		XhoI XbaI
44	HoxC8 HX	(from addgene # 21001)	FPWM → FAAA	XhoI XbaI
45	VNHoxC8	(from addgene # 21001)		BamHI EcoRI XbaI
46	VNHoxC8 HX	(from addgene # 21001)	FPWM → FAAA	BamHI EcoRI XbaI
47	VNHoxC8 DN109	(from addgene # 21001)		BamHI EcoRI XbaI
48	VNHoxC8 DN109 HX	(from addgene # 21001)	FPWM → FAAA	BamHI EcoRI XbaI
49	VNHoxC8 ΔE	(from addgene # 21001)		BamHI EcoRI XbaI
50	VNHoxC8 HX ΔE	(from addgene # 21001)	FPWM → FAAA	BamHI EcoRI XbaI

51	HoxB9	addgene # 27022		XhoI XbaI
52	HoxB9 HX	(from addgene # 27022)	W → A	XhoI XbaI
53	VNHoxB9	(from addgene # 27022)		BamHI EcoRI XbaI
54	VNHoxB9 HX	(from addgene # 27022)	W → A	BamHI EcoRI XbaI
55	HoxD10	addgene # 21007		XhoI XbaI
56	HoxD10 HX	(from addgene # 21007)	W → A	XhoI XbaI
57	VNHoxD10	(from addgene # 21007)		BamHI EcoRI XbaI
58	VNHoxD10 HX	(from addgene # 21007)	W → A	BamHI EcoRI XbaI
59	Hoxa11	addgene #8517		XhoI XbaI
60	VNHoxa11	(from addgene # 8517)		BamHI XhoI XbaI
61	Hoxd12	addgene #8553		XhoI XbaI
62	VNHoxd12	(from addgene # 8553)		BamHI XhoI XbaI
63	HoxB13	addgene # 8684		XhoI XbaI
64	VNHoxB13	(from addgene # 8684)		BamHI XhoI XbaI
65	HoxA9	Merabet		EcoRI XhoI
66	HoxA9HX	Merabet	W → A	EcoRI XhoI
67	VNHoxA9	Merabet		EcoRI XhoI
68	VNHoxA9 HX	Merabet	W → A	EcoRI XhoI
69	VNHoxA9 DN187	Merabet		EcoRI XhoI
70	VNHoxA9 DN187 HX	Merabet	W → A	EcoRI XhoI
71	VNHoxA9 HD	Merabet		EcoRI XhoI
72	VNHoxA9 HDA1	Merabet		EcoRI XhoI
73	VNHoxA9 HDA1 HXm	Merabet	W → A	EcoRI XhoI
74	VNHoxA1 HDA9	Merabet		EcoRI XhoI
75	VNHoxA1 HDA9 HXm	Merabet	FDWM → FAAA	EcoRI XhoI
76	VNHoxA9 K4A	Merabet	HD TRKKRCPYT → TRKARCPYT	EcoRI XhoI
77	VNHoxA9 HX K4A	Merabet	W → A HD TRKKRCPYT → TRKARCPYT	EcoRI XhoI
78	VNHoxA9 C6Q	Merabet	HD TRKKRCPYT → TRKKRQPYT	EcoRI XhoI
79	VNHoxA9 HX C6Q	Merabet	W → A HD TRKKRCPYT → TRKKRQPYT	EcoRI XhoI
80	VNHoxA9 P7G	Merabet	HD TRKKRCPYT → TRKKRCGYT	EcoRI XhoI
81	VNHoxA9 HX P7G	Merabet	W → A HD TRKKRCPYT → TRKKRCGYT	EcoRI XhoI
82	VNHoxA9 C6QP7G	Merabet	HD TRKKRCPYT → TRKKRQGYT	EcoRI XhoI
83	VNHoxA9 HX C6QP7G	Merabet	W → A HD TRKKRCPYT → TRKKRQGYT	EcoRI XhoI
84	VNHoxA9 M24R	Merabet	HD FNMVLRDRR → FNRVLRDRR	EcoRI XhoI
85	VNHoxA9 HX M24R	Merabet	W → A HD FNMVLRDRR → FNRVLRDRR	EcoRI XhoI
86	VNHoxA9 D29A	Merabet	HD FNMVLRDRR → FNMVLRARR	EcoRI XhoI
87	VNHoxA9 HX D29A	Merabet	W → A HD FNMVLRDRR → FNMVLRARR	EcoRI XhoI
88	VNHoxA9 M56W	Merabet	HD RMKMKKIN → RMKWKKIN	EcoRI XhoI
89	VNHoxA9 HX M56W	Merabet	W → A HD RMKMKKIN → RMKWKKIN	EcoRI XhoI
90	VNHoxA9 M24RD29A	Merabet	HD FNMVLRDRR → FNRVLRDRR	EcoRI XhoI
91	VNHoxA9 HX M24RD29A	Merabet	W → A HD FNMVLRDRR → FNRVLRDRR	EcoRI XhoI
92	VNHoxA9 D29AM56W	Merabet	HD LTRDRR → LTRARR HD RMKMKKIN → RMKWKKIN	EcoRI XhoI
93	VNHoxA9 HX D29AM56W	Merabet	W → A HD LTRDRR → LTRARR HD RMKMKKIN → RMKWKKIN	EcoRI XhoI
94	VNHoxA9 M24RM56W	Merabet	HD FNMVLRDRR → FNRVLRDRR HD RMKMKKIN → RMKWKKIN	EcoRI XhoI
95	VNHoxA9 HXM24RM56W	Merabet	W → A HD FNMVLRDRR → FNRVLRDRR HD RMKMKKIN → RMKWKKIN	EcoRI XhoI
96	VNHoxA9 3m	Merabet	HD FNMVLRDRR → FNRVLRARR HD RMKMKKIN → RMKWKKIN	EcoRI XhoI
97	VNHoxA9 HX 3m	Merabet	W → A HD FNMVLRDRR → FNRVLRARR HD RMKMKKIN → RMKWKKIN	EcoRI XhoI
98	CC Pbx1	Merabet		EcoRI XhoI
99	Pbx1	Merabet		EcoRI XhoI
100	Meis1a	Merabet		EcoRI XhoI
101	mCherry	Merabet		EcoRI XhoI

Nous présentons ici les constructions ayant été utilisées pour le **Chapitre IV des Résultats**. Le vecteur de clonage est précisé.

	Clone	Source	Vecteur	Mutation	Clonage
1	Dlx6	ORFeome v3.1	pCDNA3		BamHI EcorI XbaI
2	CCDlx6	(from ORFeome v3.1)	pCDNA3		BamHI EcorI XbaI
3	KLF4	ORFeome v3.1	pCDNA3		BamHI EcorI XbaI
4	CCKLF4	(from ORFeome v3.1)	pCDNA3		BamHI EcorI XbaI
5	ID2	ORFeome v3.1	pCDNA3		BamHI EcorI XbaI
6	CCID2	(from ORFeome v3.1)	pCDNA3		BamHI EcorI XbaI
7	BANF1	ORFeome v3.1	pCDNA3		BamHI EcorI XbaI
8	CCBANF1	(from ORFeome v3.1)	pCDNA3		BamHI EcorI XbaI
9	HOPX	ORFeome v3.1	pCDNA3		BamHI EcorI XbaI
10	CCHOPX	(from ORFeome v3.1)	pCDNA3		BamHI EcorI XbaI
11	VNHoxA9 gag	Merabet	GagPol		
12	sfGFPNHoxA9 gag	Merabet	GagPol		
13	mCherry	P. Mangeot	GagPol		
14	YFP	P. Mangeot	GagPol		
15	tetOCCPbx1	Merabet	HIV		
16	tetOGFP	P. Mangeot	HIV		
17	VSVG	P. Mangeot	VSVG		
18	GagPol	P. Mangeot	GagPol		
19	tTA gag	P. Mangeot	GagPol		
	Clone	Source	Vecteur	Mutation	Clonage
1	VNHoxA9iCNMeis1a	Merabet	SIV		
2	VNHoxA9iCNPbx3	Merabet	SIV		
3	VNHoxA9imCherry	Merabet	SIV		
4	VNHoxA9	Merabet	SIV		
5	CNMeis1aimCherry	Merabet	SIV		
6	CNMeis1a	Merabet	SIV		
7	CNPbx3imCherry	Merabet	SIV		
8	CNPbx3	Merabet	SIV		



**Agrawal** P, Habib F, Yelagandula R, Shashidhara LS. Genome-level identification of targets of Hox protein Ultrabithorax in *Drosophila*: novel mechanisms for target selection. *Sci Rep.* **2011**;1:205. doi: 10.1038/srep00205. Epub 2011 Dec 22. PubMed PMID: 22355720; PubMed Central PMCID: PMC3244697.

**Alharbi** RA, Pettengell R, Pandha HS, Morgan R. The role of HOX genes in normal hematopoiesis and acute leukemia. *Leukemia.* 2013 Apr;27(5):1000-8. doi: 10.1038/leu.2012.356. Epub **2012** Dec 5. Review. PubMed PMID: 23212154.

**Amin** S, Donaldson IJ, Zannino DA, Hensman J, Rattray M, Losa M, Spitz F, Ladam F, Sagerström C, Bobola N. *Hoxa2* selectively enhances Meis binding to change a branchial arch ground state. *Dev Cell.* **2015** Feb 9;32(3):265-77. doi: 10.1016/j.devcel.2014.12.024. Epub 2015 Jan 29. PubMed PMID: 25640223; PubMed Central PMCID: PMC4333904.

**Baëza** M, **Viala** S, **Heim** M, Dard A, Hudry B, Duffraisse M, Rogulja-Ortmann A, Brun C, Merabet S. Inhibitory activities of short linear motifs underlie Hox interactome specificity in vivo. *Elife.* **2015** Apr 14;4. doi: 10.7554/eLife.06034. PubMed PMID: 25869471; PubMed Central PMCID: PMC4392834.

**Balavoine** G, de Rosa R, Adoutte A. Hox clusters and bilaterian phylogeny. *Mol Phylogenet Evol.* **2002** Sep;24(3):366-73. Review. PubMed PMID: 12220978.

**Bender** W, Akam M, Karch F, Beachy PA, Peifer M, Spierer P, Lewis EB, Hogness DS. Molecular Genetics of the Bithorax Complex in *Drosophila melanogaster*. *Science.* **1983** Jul 1;221(4605):23-9. PubMed PMID: 17737996.

**Bergiers** I, Bridoux L, Nguyen N, Twizere JC, Rezsöhazy R. The homeodomain transcription factor *Hoxa2* interacts with and promotes the proteasomal degradation of the E3 ubiquitin protein ligase RCHY1. *PLoS One.* **2013** Nov 7;8(11):e80387. doi: 10.1371/journal.pone.0080387. eCollection 2013. PubMed PMID:24244684; PubMed Central PMCID: PMC3820564.

**Bertolino** E, Reimund B, Wildt-Perinic D, Clerc RG. A novel homeobox protein which recognizes a TGT core and functionally interferes with a retinoid-responsive motif. *J Biol Chem.* **1995** Dec 29;270(52):31178-88. PubMed PMID: 8537382.

**Bharathan** G, Janssen BJ, Kellogg EA, Sinha N. Did homeodomain proteins duplicate before the origin of angiosperms, fungi, and metazoa? *Proc Natl Acad Sci U S A.* **1997** Dec 9;94(25):13749-53. PubMed PMID: 9391098; PubMed Central PMCID: PMC28378.

**Bhatlekar** S, Fields JZ, Boman BM. HOX genes and their role in the development of human cancers. *J Mol Med (Berl).* **2014** Aug;92(8):811-23. doi:10.1007/s00109-014-1181-y. Epub 2014 Jul 5. Review. PubMed PMID: 24996520.

**Björnsson** JM, Larsson N, Brun AC, Magnusson M, Andersson E, Lundström P, Larsson J, Repetowska E, Ehinger M, Humphries RK, Karlsson S. Reduced proliferative capacity of

hematopoietic stem cells deficient in Hoxb3 and Hoxb4. *Mol Cell Biol.* **2003** Jun;23(11):3872-83. PubMed PMID: 12748289; PubMed Central PMCID: PMC155209.

**Bondos** SE, Tan XX, Matthews KS. Physical and genetic interactions link hox function with diverse transcription factors and cell signaling proteins. *Mol Cell Proteomics.* **2006** May;5(5):824-34. Epub Feb 2. PubMed PMID: 16455680.

**Bridoux** L, Deneyer N, Bergiers I, Rezsöházy R. Molecular Analysis of the HOXA2-Dependent Degradation of RCHY1. *PLoS One.* 2015 Oct 23;10(10):e0141347. doi: 10.1371/journal.pone.0141347. eCollection **2015**. PubMed PMID: 26496426; PubMed Central PMCID: PMC4619689.

**Buljan** M, Chalancon G, Dunker AK, Bateman A, Balaji S, Fuxreiter M, Babu MM. Alternative splicing of intrinsically disordered regions and rewiring of protein interactions. *Curr Opin Struct Biol.* **2013** Jun;23(3):443-50. doi:10.1016/j.sbi.2013.03.006. Epub 2013 May 22. Review. PubMed PMID: 23706950.

**Bürglin** TR, Affolter M. Homeodomain proteins: an update. *Chromosoma.* **2016** Jun;125(3):497-521. doi: 10.1007/s00412-015-0543-8. Epub 2015 Oct 13. Review. PubMed PMID: 26464018; PubMed Central PMCID: PMC4901127.

**Bürglin** TR, Ruvkun G. New motif in PBX genes. *Nat Genet.* **1992** Aug;1(5):319-20. PubMed PMID: 1363814.

**Bürglin** TR. Analysis of TALE superclass homeobox genes (MEIS, PBC, KNOX, Iroquois, TGIF) reveals a novel domain conserved between plants and animals. *Nucleic Acids Res.* **1997** Nov 1;25(21):4173-80. PubMed PMID: 9336443; PubMed Central PMCID: PMC147054.

**Bürglin** TR. Homeodomain subtypes and functional diversity. *Subcell Biochem.* **2011**;52:95-122. doi: 10.1007/978-90-481-9069-0\_5. PubMed PMID: 21557080.

**Cantile** M, Pettinato G, Procino A, Feliciello I, Cindolo L, Cillo C. In vivo expression of the whole HOX gene network in human breast cancer. *Eur J Cancer.* **2003** Jan;39(2):257-64. PubMed PMID: 12509959.

**Caré** A, Silvani A, Meccia E, Mattia G, Peschle C, Colombo MP. Transduction of the SkBr3 breast carcinoma cell line with the HOXB7 gene induces bFGF expression, increases cell proliferation and reduces growth factor dependence. *Oncogene.* **1998** Jun 25;16(25):3285-9. PubMed PMID: 9681827.

**Chan** HM, La Thangue NB. p300/CBP proteins: HATs for transcriptional bridges and scaffolds. *J Cell Sci.* **2001** Jul;114(Pt 13):2363-73. Review. PubMed PMID: 11559745.

**Chan** SK, Pöpperl H, Krumlauf R, Mann RS. An extradenticle-induced conformational change in a HOX protein overcomes an inhibitory function of the conserved hexapeptide

motif. *EMBO J.* **1996** May 15;15(10):2476-87. PubMed PMID:8665855; PubMed Central PMCID: PMC450180.

**Chariot** A, van Lint C, Chapelier M, Gielen J, Merville MP, Bours V. CBP and histone deacetylase inhibition enhance the transactivation potential of the HOXB7 homeodomain-containing protein. *Oncogene.* **1999** Jul 8;18(27):4007-14. PubMed PMID: 10435624.

**Chatr-Aryamontri** A, Breitkreutz BJ, Oughtred R, Boucher L, Heinicke S, Chen D, Stark C, Breitkreutz A, Kolas N, O'Donnell L, Reguluy T, Nixon J, Ramage L, Winter A, Sellam A, Chang C, Hirschman J, Theesfeld C, Rust J, Livstone MS, Dolinski K, Tyers M. The BioGRID interaction database: 2015 update. *Nucleic Acids Res.* **2015** Jan;43(Database issue):D470-8. doi: 10.1093/nar/gku1204. Epub 2014 Nov 26. PubMed PMID: 25428363; PubMed Central PMCID: PMC4383984.

**Chen** J, Zhu S, Jiang N, Shang Z, Quan C, Niu Y. HoxB3 promotes prostate cancer cell progression by transactivating CDCA3. *Cancer Lett.* **2013** Apr 28;330(2):217-24. doi: 10.1016/j.canlet.2012.11.051. Epub 2012 Dec 5. PubMed PMID: 23219899.

**Cheng** W, Liu J, Yoshida H, Rosen D, Naora H. Lineage infidelity of epithelial ovarian cancers is controlled by HOX genes that specify regional identity in the reproductive tract. *Nat Med.* **2005** May;11(5):531-7. Epub 2005 Apr 10. PubMed PMID:15821746.

**Choo** SW, White R, Russell S. Genome-wide analysis of the binding of the Hox protein Ultrabithorax and the Hox cofactor Homothorax in *Drosophila*. *PLoS One.* 2011 Apr 5;6(4):e14778. doi: 10.1371/journal.pone.0014778. PubMed PMID: 21483667; PubMed Central PMCID: PMC3071696.

**Comelli** L, Marchetti L, Arosio D, Riva S, Abdurashidova G, Beltram F, Falaschi A. The homeotic protein HOXC13 is a member of human DNA replication complexes. *Cell Cycle.* 2009 Feb 1;8(3):454-9. Epub **2009** Feb 19. PubMed PMID: 19182517.

**Crocker** J, Abe N, Rinaldi L, McGregor AP, Frankel N, Wang S, Alsawadi A, Valenti P, Plaza S, Payre F, Mann RS, Stern DL. Low affinity binding site clusters confer hox specificity and regulatory robustness. *Cell.* **2015** Jan 15;160(1-2):191-203. doi: 10.1016/j.cell.2014.11.041. Epub 2014 Dec 31. PubMed PMID: 25557079; PubMed Central PMCID: PMC4449256.

**Crooks** GM, Fuller J, Petersen D, Izadi P, Malik P, Pattengale PK, Kohn DB, Gasson JC. Constitutive HOXA5 expression inhibits erythropoiesis and increases myelopoiesis from human hematopoietic progenitors. *Blood.* **1999** Jul 15;94(2):519-28. PubMed PMID: 10397719.

**Daga** A, Podesta M, Capra MC, Piaggio G, Frassoni F, Corte G. The retroviral transduction of HOXC4 into human CD34(+) cells induces an in vitro expansion of clonogenic and early progenitors. *Exp Hematol.* **2000** May;28(5):569-74. PubMed PMID: 10812247.



**Derelle** R, Lopez P, Le Guyader H, Manuel M. Homeodomain proteins belong to the ancestral molecular toolkit of eukaryotes. *Evol Dev.* **2007** May-Jun;9(3):212-9. PubMed PMID: 17501745.

**Diella** F, Gould CM, Chica C, Via A, Gibson TJ. Phospho.ELM: a database of phosphorylation sites--update **2008**. *Nucleic Acids Res.* 2008 Jan;36(Database issue):D240-4. Epub 2007 Oct 25. PubMed PMID: 17962309; PubMed Central PMCID:PMC2238828.

**Ding** Z, Liang J, Lu Y, Yu Q, Songyang Z, Lin SY, Mills GB. A retrovirus-based protein complementation assay screen reveals functional AKT1-binding partners. *Proc Natl Acad Sci U S A.* **2006** Oct 10;103(41):15014-9. Epub 2006 Oct 3. PubMed PMID: 17018644; PubMed Central PMCID: PMC1636760.

**Dinkel** H, Van Roey K, Michael S, Davey NE, Weatheritt RJ, Born D, Speck T, Krüger D, Grebnev G, Kuban M, Strumillo M, Uyar B, Budd A, Altenberg B, Seiler M, Chemes LB, Glavina J, Sánchez IE, Diella F, Gibson TJ. The eukaryotic linear motif resource ELM: 10 years and counting. *Nucleic Acids Res.* **2014** Jan;42(Database issue):D259-66. doi: 10.1093/nar/gkt1047. Epub 2013 Nov 7. PubMed PMID: 24214962; PubMed Central PMCID: PMC3964949.

**Doidge** R, Mittal S, Aslam A, Winkler GS. The anti-proliferative activity of BTG/TOB proteins is mediated via the Caf1a (CNOT7) and Caf1b (CNOT8) deadenylase subunits of the Ccr4-not complex. *PLoS One.* **2012**;7(12):e51331. doi:10.1371/journal.pone.0051331. Epub 2012 Dec 7. PubMed PMID: 23236473; PubMed Central PMCID: PMC3517456.

**Donaldson** IJ, Amin S, Hensman JJ, Kutejova E, Rattray M, Lawrence N, Hayes A, Ward CM, Bobola N. Genome-wide occupancy links Hoxa2 to Wnt- $\beta$ -catenin signaling in mouse embryonic development. *Nucleic Acids Res.* **2012** May;40(9):3990-4001. doi: 10.1093/nar/gkr1240. Epub 2012 Jan 5. PubMed PMID: 22223247; PubMed Central PMCID: PMC3351182.

**Dorsam** ST, Ferrell CM, Dorsam GP, Derynck MK, Vijapurkar U, Khodabakhsh D, Pau B, Bernstein H, Haqq CM, Largman C, Lawrence HJ. The transcriptome of the leukemogenic homeoprotein HOXA9 in human hematopoietic cells. *Blood.* **2004** Mar 1;103(5):1676-84. Epub 2003 Nov 6. PubMed PMID: 14604967.

**Dostie** J, Ferraiuolo M, Pause A, Adam SA, Sonenberg N. A novel shuttling protein, 4E-T, mediates the nuclear import of the mRNA 5' cap-binding protein, eIF4E. *EMBO J.* **2000** Jun 15;19(12):3142-56. PubMed PMID: 10856257; PubMed Central PMCID: PMC203362.

**Dostie** J, Lejbkowitz F, Sonenberg N. Nuclear eukaryotic initiation factor 4E (eIF4E) colocalizes with splicing factors in speckles. *J Cell Biol.* **2000** Jan 24;148(2):239-47. PubMed PMID: 10648556; PubMed Central PMCID: PMC2174286.

**Dunker** AK, Lawson JD, Brown CJ, Williams RM, Romero P, Oh JS, Oldfield CJ, Campen AM, Ratliff CM, Higgs KW, Ausio J, Nissen MS, Reeves R, Kang C, Kissinger CR, Bailey RW, Griswold MD, Chiu W, Garner EC, Obradovic Z. Intrinsically disordered protein. *J Mol Graph Model.* **2001**;19(1):26-59. Review. PubMed PMID: 11381529.

**Dyson** HJ, Wright PE. Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol.* **2005** Mar;6(3):197-208. Review. PubMed PMID: 15738986.

**Eyckerman** S, Titeca K, Van Quickelberghe E, Cloots E, Verhee A, Samyn N, De Ceuninck L, Timmerman E, De Sutter D, Lievens S, Van Calenbergh S, Gevaert K, Tavernier J. Trapping mammalian protein complexes in viral particles. *Nat Commun.* **2016** Apr 28;7:11416. doi: 10.1038/ncomms11416. PubMed PMID: 27122307; PubMed Central PMCID: PMC4853472.

**Faber** J, Krivtsov AV, Stubbs MC, Wright R, Davis TN, van den Heuvel-Eibrink M, Zwaan CM, Kung AL, Armstrong SA. HOXA9 is required for survival in human MLL-rearranged acute leukemias. *Blood.* 2009 Mar 12;113(11):2375-85. doi: 10.1182/blood-2007-09-113597. Epub **2008** Dec 3. PubMed PMID: 19056693; PubMed Central PMCID: PMC2656267.

**Fan** R, Bonde S, Gao P, Sotomayor B, Chen C, Mouw T, Zavazava N, Tan K. Dynamic HoxB4-regulatory network during embryonic stem cell differentiation to hematopoietic cells. *Blood.* **2012** May 10;119(19):e139-47. doi: 10.1182/blood-2011-12-396754. Epub 2012 Mar 21. PubMed PMID: 22438249; PubMed Central PMCID: PMC3362371.

**Fischbach** NA, Rozenfeld S, Shen W, Fong S, Chrobak D, Ginzinger D, Kogan SC, Radhakrishnan A, Le Beau MM, Largman C, Lawrence HJ. HOXB6 overexpression in murine bone marrow immortalizes a myelomonocytic precursor in vitro and causes hematopoietic stem cell expansion and acute myeloid leukemia in vivo. *Blood.* **2005** Feb 15;105(4):1456-66. Epub 2004 Nov 2. PubMed PMID: 15522959.

**Florence** B, McGinnis W. A genetic screen of the Drosophila X chromosome for mutations that modify Deformed function. *Genetics.* **1998** Dec;150(4):1497-511. PubMed PMID: 9832527; PubMed Central PMCID: PMC1460420.

**Foos** N, Maurel-Zaffran C, Maté MJ, Vincentelli R, Hainaut M, Berenger H, Pradel J, Saurin AJ, Ortiz-Lombardía M, Graba Y. A flexible extension of the Drosophila ultrabithorax homeodomain defines a novel Hox/PBC interaction mode. *Structure.* **2015** Feb 3;23(2):270-9. doi: 10.1016/j.str.2014.12.011. PubMed PMID: 25651060.

**Galant** R, Walsh CM, Carroll SB. Hox repression of a target gene: extradenticle-independent, additive action through multiple monomer binding sites. *Development.* **2002** Jul;129(13):3115-26. PubMed PMID: 12070087.

**Gebelein** B, McKay DJ, Mann RS. Direct integration of Hox and segmentation gene inputs during Drosophila development. *Nature.* **2004** Oct 7;431(7009):653-9. PubMed PMID: 15470419.

**Gehring** WJ, Qian YQ, Billeter M, Furukubo-Tokunaga K, Schier AF, Resendez-Perez D, Affolter M, Otting G, Wüthrich K. Homeodomain-DNA recognition. *Cell.* **1994** Jul 29;78(2):211-23. Review. PubMed PMID: 8044836.

**Hamid** SM, Cicek S, Karamil S, Ozturk MB, Debelec-Butuner B, Erbaykent-Tepedelen B, Varisli L, Gonen-Korkmaz C, Yorukoglu K, Korkmaz KS. HOXB13 contributes to G1/S and G2/M checkpoint controls in prostate. *Mol Cell Endocrinol.* **2014** Mar 5;383(1-2):38-47. doi: 10.1016/j.mce.2013.12.003. Epub 2013 Dec 8. PubMed PMID: 24325868.

**Harding** K, Wedeen C, McGinnis W, Levine M. Spatially regulated expression of homeotic genes in *Drosophila*. *Science.* **1985** Sep 20;229(4719):1236-42. Review. PubMed PMID: 3898362.

**Harding** KW, Gellon G, McGinnis N, McGinnis W. A screen for modifiers of Deformed function in *Drosophila*. *Genetics.* **1995** Aug;140(4):1339-52. PubMed PMID: 7498774; PubMed Central PMCID: PMC1206698.

**Hass** MR, Liow HH, Chen X, Sharma A, Inoue YU, Inoue T, Reeb A, Martens A, Fulbright M, Raju S, Stevens M, Boyle S, Park JS, Weirauch MT, Brent MR, Kopan R. SpDamID: Marking DNA Bound by Protein Complexes Identifies Notch-Dimer Responsive Enhancers. *Mol Cell.* **2015** Aug 20;59(4):685-97. doi: 10.1016/j.molcel.2015.07.008. Epub 2015 Aug 6. PubMed PMID: 26257285; PubMed Central PMCID: PMC4553207.

**Hu** YL, Fong S, Ferrell C, Largman C, Shen WF. HOXA9 modulates its oncogenic partner Meis1 to influence normal hematopoiesis. *Mol Cell Biol.* **2009** Sep;29(18):5181-92. doi: 10.1128/MCB.00545-09. Epub 2009 Jul 20. PubMed PMID: 19620287; PubMed Central PMCID: PMC2738281.

**Huang** Y, Sitwala K, Bronstein J, Sanders D, Dandekar M, Collins C, Robertson G, MacDonald J, Cezard T, Bilenky M, Thiessen N, Zhao Y, Zeng T, Hirst M, Hero A, Jones S, Hess JL. Identification and characterization of Hoxa9 binding sites in hematopoietic cells. *Blood.* **2012** Jan 12;119(2):388-98. doi: 10.1182/blood-2011-03-341081. Epub 2011 Nov 9. PubMed PMID: 22072553; PubMed Central PMCID: PMC3257007.

**Huang** Y, Sitwala K, Bronstein J, Sanders D, Dandekar M, Collins C, Robertson G, MacDonald J, Cezard T, Bilenky M, Thiessen N, Zhao Y, Zeng T, Hirst M, Hero A, Jones S, Hess JL. Identification and characterization of Hoxa9 binding sites in hematopoietic cells. *Blood.* **2012** Jan 12;119(2):388-98. doi: 10.1182/blood-2011-03-341081. Epub 2011 Nov 9. PubMed PMID: 22072553; PubMed Central PMCID: PMC3257007.

**Hudry** B, Remacle S, Delfini MC, Rezsöházy R, Graba Y, Merabet S. Hox proteins display a common and ancestral ability to diversify their interaction mode with the PBC class cofactors. *PLoS Biol.* **2012**;10(6):e1001351. doi: 10.1371/journal.pbio.1001351. Epub 2012 Jun 26. PubMed PMID: 22745600; PubMed Central PMCID: PMC3383740.

**Hudry** B, Thomas-Chollier M, Volovik Y, Duffraisse M, Dard A, Frank D, Technau U, Merabet S. Molecular insights into the origin of the Hox-TALE patterning system. *Elife.* **2014** Mar 18;3:e01939. doi: 10.7554/eLife.01939. PubMed PMID: 24642410; PubMed Central PMCID: PMC3957477.

**Hudry** B, Viala S, Graba Y, Merabet S. Visualization of protein interactions in living *Drosophila* embryos by the bimolecular fluorescence complementation assay. *BMC Biol.*

**2011** Jan 28;9:5. doi: 10.1186/1741-7007-9-5. PubMed PMID: 21276241;PubMed Central PMCID: PMC3041725.

**Hueber** SD, Bezdán D, Henz SR, Blank M, Wu H, Lohmann I. Comparative analysis of Hox downstream genes in *Drosophila*. *Development*. **2007** Jan;134(2):381-92. Epub 2006 Dec 13. PubMed PMID: 17166915.

**Hueber** SD, Lohmann I. Shaping segments: Hox gene function in the genomic age. *Bioessays*. **2008** Oct;30(10):965-79. doi: 10.1002/bies.20823. Review. PubMed PMID: 18798525.

**Hyman** CA, Bartholin L, Newfeld SJ, Wotton D. *Drosophila* TGIF proteins are transcriptional activators. *Mol Cell Biol*. **2003** Dec;23(24):9262-74. PubMed PMID: 14645536; PubMed Central PMCID: PMC309625.

**In der Rieden** PM, Mainguy G, Woltering JM, Durston AJ. Homeodomain to hexapeptide or PBC-interaction-domain distance: size apparently matters. *Trends Genet*. **2004** Feb;20(2):76-9. Review. PubMed PMID: 14746988.

**Javed** S, Langley SE. Importance of HOX genes in normal prostate gland formation, prostate cancer development and its early detection. *BJU Int*. **2014** Apr;113(4):535-40. doi: 10.1111/bju.12269. Epub 2013 Aug 13. Review. PubMed PMID:23937390.

**Joshi** R, Passner JM, Rohs R, Jain R, Sosinsky A, Crickmore MA, Jacob V, Aggarwal AK, Honig B, Mann RS. Functional specificity of a Hox protein mediated by the recognition of minor groove structure. *Cell*. **2007** Nov 2;131(3):530-43. PubMed PMID: 17981120; PubMed Central PMCID: PMC2709780.

**Jung** C, Kim RS, Lee SJ, Wang C, Jeng MH. HOXB13 homeodomain protein suppresses the growth of prostate cancer cells by the negative regulation of T-cell factor 4. *Cancer Res*. **2004** May 1;64(9):3046-51. PubMed PMID: 15126340.

**Jung** C, Kim RS, Zhang HJ, Lee SJ, Jeng MH. HOXB13 induces growth suppression of prostate cancer cells as a repressor of hormone-activated androgen receptor signaling. *Cancer Res*. **2004** Dec 15;64(24):9185-92. PubMed PMID: 15604291.

**Kennison** JA, Tamkun JW. Dosage-dependent modifiers of polycomb and antennapedia mutations in *Drosophila*. *Proc Natl Acad Sci U S A*. **1988** Nov;85(21):8136-40. PubMed PMID: 3141923; PubMed Central PMCID: PMC282375.

**Kerppola** TK. Simultaneous visualization of multiple protein interactions using multicolor bimolecular fluorescence complementation (BiFC) analysis. *Cold Spring Harb Protoc*. **2013** Sep 1;2013(9):892-5. doi: 10.1101/pdb.prot077172. PubMed PMID: 24003203.

**Kerppola** TK. Visualization of molecular interactions using bimolecular fluorescence complementation analysis: characteristics of protein fragment complementation. *Chem Soc*

Rev. **2009** Oct;38(10):2876-86. doi: 10.1039/b909638h. Epub 2009 Sep 4. Review. PubMed PMID: 19771334; PubMed Central PMCID: PMC2980501.

**Kmita** M, Tarchini B, Zàkány J, Logan M, Tabin CJ, Duboule D. Early developmental arrest of mammalian limbs lacking HoxA/HoxD gene function. *Nature*. **2005** Jun 23;435(7045):1113-6. PubMed PMID: 15973411.

**Knoepfler** PS, Kamps MP. The pentapeptide motif of Hox proteins is required for cooperative DNA binding with Pbx1, physically contacts Pbx1, and enhances DNA binding by Pbx1. *Mol Cell Biol*. **1995** Oct;15(10):5811-9. PubMed PMID: 7565734; PubMed Central PMCID: PMC230833.

**Kodama** Y, Hu CD. Bimolecular fluorescence complementation (BiFC): a 5-year update and future perspectives. *Biotechniques*. **2012** Nov;53(5):285-98. doi: 10.2144/000113943. Review. PubMed PMID: 23148879.

**Kroon** E, Kros J, Thorsteinsdottir U, Baban S, Buchberg AM, Sauvageau G. Hoxa9 transforms primary bone marrow cells through specific collaboration with Meis1a but not Pbx1b. *EMBO J*. **1998** Jul 1;17(13):3714-25. PubMed PMID: 9649441; PubMed Central PMCID: PMC1170707.

**Krumlauf** R. Hox genes in vertebrate development. *Cell*. **1994** Jul 29;78(2):191-201. Review. PubMed PMID: 7913880.

**Lambert** B, Vandeputte J, Remacle S, Bergiers I, Simonis N, Twizere JC, Vidal M, Rezsöházy R. Protein interactions of the transcription factor Hoxa1. *BMC Dev Biol*. **2012** Oct 22;12:29. doi: 10.1186/1471-213X-12-29. PubMed PMID: 23088713; PubMed Central PMCID: PMC3514159.

**Langdale** JA. Plant morphogenesis. More knots untied. *Curr Biol*. **1994** Jun 1;4(6):529-31. Review. PubMed PMID: 7922374.

**LaRonde-LeBlanc** NA, Wolberger C. Structure of HoxA9 and Pbx1 bound to DNA: Hox hexapeptide and DNA recognition anterior to posterior. *Genes Dev*. **2003** Aug 15;17(16):2060-72. PubMed PMID: 12923056; PubMed Central PMCID: PMC196259.

**Lee** JH, Lin H, Joo S, Goodenough U. Early sexual origins of homeoprotein heterodimerization and evolution of the plant KNOX/BELL family. *Cell*. **2008** May 30;133(5):829-40. doi: 10.1016/j.cell.2008.04.028. PubMed PMID: 18510927.

**Lewis** EB. A gene complex controlling segmentation in *Drosophila*. *Nature*. **1978** Dec 7;276(5688):565-70. PubMed PMID: 103000.

**Li** Z, Zhang Z, Li Y, Arnovitz S, Chen P, Huang H, Jiang X, Hong GM, Kunjamma RB, Ren H, He C, Wang CZ, Elkahloun AG, Valk PJ, Döhner K, Neilly MB, Bullinger L, Delwel R, Löwenberg B, Liu PP, Morgan R, Rowley JD, Yuan CS, Chen J. PBX3 is an important cofactor of HOXA9 in leukemogenesis. *Blood*. **2013** Feb 21;121(8):1422-31. doi:

10.1182/blood-2012-07-442004. Epub 2012 Dec 20. PubMed PMID: 23264595; PubMed Central PMCID: PMC3578957.

**Luo** L, Yang X, Takihara Y, Knoetgen H, Kessel M. The cell-cycle regulator geminin inhibits Hox function through direct and polycomb-mediated interactions. *Nature*. **2004** Feb 19;427(6976):749-53. PubMed PMID: 14973489.

**Lüscher** B, Larsson LG. The basic region/helix-loop-helix/leucine zipper domain of Myc proto-oncoproteins: function and regulation. *Oncogene*. **1999** May 13;18(19):2955-66. Review. PubMed PMID: 10378692.

**Maconochie** M, Nonchev S, Morrison A, Krumlauf R. Paralogous Hox genes: function and regulation. *Annu Rev Genet*. **1996**;30:529-56. Review. PubMed PMID: 8982464.

**Mahaffey** JW. Assisting Hox proteins in controlling body form: are there new lessons from flies (and mammals)? *Curr Opin Genet Dev*. **2005** Aug;15(4):422-9. Review. PubMed PMID: 15979870.

**Makiyama** K, Hamada J, Takada M, Murakawa K, Takahashi Y, Tada M, Tamoto E, Shindo G, Matsunaga A, Teramoto K, Komuro K, Kondo S, Katoh H, Koike T, Moriuchi T. Aberrant expression of HOX genes in human invasive breast carcinoma. *Oncol Rep*. **2005** Apr;13(4):673-9. PubMed PMID: 15756441.

**Mann** RS, Lelli KM, Joshi R. Hox specificity unique roles for cofactors and collaborators. *Curr Top Dev Biol*. **2009**;88:63-101. doi: 10.1016/S0070-2153(09)88003-4. Review. PubMed PMID: 19651302; PubMed Central PMCID: PMC2810641.

**Marchetti** L, Comelli L, D'Innocenzo B, Puzzi L, Luin S, Arosio D, Calvello M, Mendoza-Maldonado R, Peverali F, Trovato F, Riva S, Biamonti G, Abdurashidova G, Beltram F, Falaschi A. Homeotic proteins participate in the function of human-DNA replication origins. *Nucleic Acids Res*. **2010** Dec;38(22):8105-19. doi:10.1093/nar/gkq688. Epub 2010 Aug 6. PubMed PMID: 20693533; PubMed Central PMCID:PMC3001063.

**Mark** M, Rijli FM, Chambon P. Homeobox genes in embryogenesis and pathogenesis. *Pediatr Res*. **1997** Oct;42(4):421-9. Review. PubMed PMID: 9380431.

**McGinnis** W, Hart CP, Gehring WJ, Ruddle FH. Molecular cloning and chromosome mapping of a mouse DNA sequence homologous to homeotic genes of *Drosophila*. *Cell*. **1984** Oct;38(3):675-80. PubMed PMID: 6091896.

**McGinnis** W, Krumlauf R. Homeobox genes and axial patterning. *Cell*. **1992** Jan 24;68(2):283-302. Review. PubMed PMID: 1346368.

**Merabet** S, Catala F, Pradel J, Graba Y. A green fluorescent protein reporter genetic screen that identifies modifiers of Hox gene function in the *Drosophila* embryo. *Genetics*. **2002** Sep;162(1):189-202. PubMed PMID: 12242233; PubMed Central PMCID: PMC1462243.

**Merabet** S, Dard A. Tracking context-specific transcription factors regulating hox activity. *Dev Dyn*. **2014** Jan;243(1):16-23. doi: 10.1002/dvdy.24002. Epub 2013 Jul 26. Review. PubMed PMID: 23794379.

**Merabet** S, Hudry B, Saadaoui M, Graba Y. Classification of sequence signatures: a guide to Hox protein function. *Bioessays*. **2009** May;31(5):500-11. doi: 10.1002/bies.200800229. Review. PubMed PMID: 19334006.

**Merabet** S, Hudry B. On the border of the homeotic function: re-evaluating the controversial role of cofactor-recruiting motifs: the role of cofactor-recruiting motifs in conferring Hox evolutionary flexibility may critically depend on the protein environment. *Bioessays*. **2011** Jul;33(7):499-507. doi:10.1002/bies.201100019. Epub 2011 May 5. Review. PubMed PMID: 21544844.

**Merabet** S, Kambris Z, Capovilla M, Bérenger H, Pradel J, Graba Y. The hexapeptide and linker regions of the AbdA Hox protein regulate its activating and repressive functions. *Dev Cell*. **2003** May;4(5):761-8. PubMed PMID: 12737810.

**Merabet** S, Litim-Mecheri I, Karlsson D, Dixit R, Saadaoui M, Monier B, Brun C, Thor S, Vijayraghavan K, Perrin L, Pradel J, Graba Y. Insights into Hox protein function from a large scale combinatorial analysis of protein domains. *PLoS Genet*. **2011** Oct;7(10):e1002302. doi: 10.1371/journal.pgen.1002302. Epub 2011 Oct 27. PubMed PMID: 22046139; PubMed Central PMCID: PMC3203194.

**Merabet** S, Lohmann I. Toward a new twist in Hox and TALE DNA-binding specificity. *Dev Cell*. **2015** Feb 9;32(3):259-61. doi: 10.1016/j.devcel.2015.01.030. PubMed PMID: 25669880.

**Merabet** S, Saadaoui M, Sambrani N, Hudry B, Pradel J, Affolter M, Graba Y. A unique Extradenticle recruitment mode in the *Drosophila* Hox protein Ultrabithorax. *Proc Natl Acad Sci U S A*. **2007** Oct 23;104(43):16946-51. Epub 2007 Oct 17. PubMed PMID: 17942685; PubMed Central PMCID: PMC2040397.

**Miller** GJ, Miller HL, van Bokhoven A, Lambert JR, Werahera PN, Schirripa O, Lucia MS, Nordeen SK. Aberrant HOXC expression accompanies the malignant phenotype in human prostate. *Cancer Res*. **2003** Sep 15;63(18):5879-88. PubMed PMID:14522913.

**Morgan** R, Boxall A, Harrington KJ, Simpson GR, Gillett C, Michael A, Pandha HS. Targeting the HOX/PBX dimer in breast cancer. *Breast Cancer Res Treat*. **2012** Nov;136(2):389-98. doi: 10.1007/s10549-012-2259-2. Epub 2012 Sep 30. PubMed PMID:23053648.

**Morgan** R, Boxall A, Harrington KJ, Simpson GR, Michael A, Pandha HS. Targeting HOX transcription factors in prostate cancer. *BMC Urol*. **2014** Feb 5;14:17. doi: 10.1186/1471-2490-14-17. PubMed PMID: 24499138; PubMed Central PMCID: PMC3942264.

**Morgan** R, In der Rieden P, Hooiveld MH, Durston AJ. Identifying HOX paralog groups by the PBX-binding region. *Trends Genet.* **2000** Feb;16(2):66-7. PubMed PMID: 10652532.

**Mukherjee** K, Brocchieri L, Bürglin TR. A comprehensive classification and evolutionary analysis of plant homeobox genes. *Mol Biol Evol.* **2009** Dec;26(12):2775-94. doi: 10.1093/molbev/msp201. Epub 2009 Sep 4. PubMed PMID: 19734295; PubMed Central PMCID: PMC2775110.

**Mukherjee** K, Bürglin TR. Comprehensive analysis of animal TALE homeobox genes: new conserved motifs and cases of accelerated evolution. *J Mol Evol.* **2007** Aug;65(2):137-53. Epub 2007 Jul 30. PubMed PMID: 17665086.

**Murali** T, Pacifico S, Yu J, Guest S, Roberts GG 3rd, Finley RL Jr. DroID 201a comprehensive, integrated resource for protein, transcription factor, RNA and gene interactions for Drosophila. *Nucleic Acids Res.* **2011** Jan;39(Database issue):D736-43. doi: 10.1093/nar/gkq1092. Epub 2010 Oct 29. PubMed PMID: 21036869; PubMed Central PMCID: PMC3013689.

**Nakamura** T, Largaespada DA, Lee MP, Johnson LA, Ohyashiki K, Toyama K, Chen SJ, Willman CL, Chen IM, Feinberg AP, Jenkins NA, Copeland NG, Shaughnessy JD Jr. Fusion of the nucleoporin gene NUP98 to HOXA9 by the chromosome translocation t(7;11)(p15;p15) in human myeloid leukaemia. *Nat Genet.* **1996** Feb;12(2):154-8. PubMed PMID: 8563753.

**Naora** H, Montz FJ, Chai CY, Roden RB. Aberrant expression of homeobox gene HOXA7 is associated with müllerian-like differentiation of epithelial ovarian tumors and the generation of a specific autologous antibody response. *Proc Natl Acad Sci U S A.* **2001** Dec 18;98(26):15209-14. Epub 2001 Dec 11. PubMed PMID:11742062; PubMed Central PMCID: PMC65008.

**Neduva** V, Linding R, Su-Angrand I, Stark A, de Masi F, Gibson TJ, Lewis J, Serrano L, Russell RB. Systematic discovery of new recognition peptides mediating protein interaction networks. *PLoS Biol.* **2005** Dec;3(12):e405. Epub 2005 Nov 15. PubMed PMID: 16279839; PubMed Central PMCID: PMC1283537.

**Neuteboom** ST, Peltenburg LT, van Dijk MA, Murre C. The hexapeptide LFPWMR in Hoxb-8 is required for cooperative DNA binding with Pbx1 and Pbx2 proteins. *Proc Natl Acad Sci U S A.* **1995** Sep 26;92(20):9166-70. PubMed PMID: 7568094; PubMed Central PMCID: PMC40945.

**Noyes** MB, Christensen RG, Wakabayashi A, Stormo GD, Brodsky MH, Wolfe SA. Analysis of homeodomain specificities allows the family-wide prediction of preferred recognition sites. *Cell.* **2008** Jun 27;133(7):1277-89. doi: 10.1016/j.cell.2008.05.023. PubMed PMID: 18585360; PubMed Central PMCID: PMC2478728.

**Ohno** Y, Yasunaga S, Janmohamed S, Ohtsubo M, Saeki K, Kurogi T, Mihara K, Iscove NN, Takihara Y. Hoxa9 transduction induces hematopoietic stem and progenitor cell activity through direct down-regulation of geminin protein. *PLoS One.* **2013**;8(1):e53161. doi:



10.1371/journal.pone.0053161. Epub 2013 Jan 11. PubMed PMID: 23326393; PubMed Central PMCID: PMC3543444.

**Ohno** Y, Yasunaga S, Ohtsubo M, Mori S, Tsumura M, Okada S, Ohta T, Ohtani K, Kobayashi M, Takihara Y. Hoxb4 transduction down-regulates Geminin protein, providing hematopoietic stem and progenitor cells with proliferation potential. *Proc Natl Acad Sci U S A*. **2010** Dec 14;107(50):21529-34. doi: 10.1073/pnas.1011054107. Epub 2010 Nov 22. PubMed PMID: 21098278; PubMed Central PMCID: PMC3003101.

**Orchard** S. Molecular interaction databases. *Proteomics*. **2012** May;12(10):1656-62. doi: 10.1002/pmic.201100484. Review. PubMed PMID: 22611057.

**Oshima** M, Endoh M, Endo TA, Toyoda T, Nakajima-Takagi Y, Sugiyama F, Koseki H, Kyba M, Iwama A, Osawa M. Genome-wide analysis of target genes regulated by HoxB4 in hematopoietic stem and progenitor cells developing from embryonic stem cells. *Blood*. **2011** Apr 14;117(15):e142-50. doi: 10.1182/blood-2010-12-323212. Epub 2011 Feb 22. PubMed PMID: 21343615; PubMed Central PMCID: PMC3296623.

**Papagiannouli** F, Schardt L, Grajcarek J, Ha N, Lohmann I. The Hox gene Abd-B controls stem cell niche function in the *Drosophila* testis. *Dev Cell*. **2014** Jan 27;28(2):189-202. doi: 10.1016/j.devcel.2013.12.016. Erratum in: *Dev Cell*. 2014 Feb 24;28(4):474. PubMed PMID: 24480643.

**Passner** JM, Ryoo HD, Shen L, Mann RS, Aggarwal AK. Structure of a DNA-bound Ultrabithorax-Extradenticle homeodomain complex. *Nature*. **1999** Feb 25;397(6721):714-9. PubMed PMID: 10067897.

**Pearson** JC, Lemons D, McGinnis W. Modulating Hox gene functions during animal body patterning. *Nat Rev Genet*. **2005** Dec;6(12):893-904. Review. PubMed PMID: 16341070.

**Penkov** D, Mateos San Martín D, Fernandez-Díaz LC, Rosselló CA, Torroja C, Sánchez-Cabo F, Warnatz HJ, Sultan M, Yaspo ML, Gabrieli A, Tkachuk V, Brendolan A, Blasi F, Torres M. Analysis of the DNA-binding profile and function of TALE homeoproteins reveals their specialization and specific interactions with Hox genes/proteins. *Cell Rep*. **2013** Apr 25;3(4):1321-33. doi: 10.1016/j.celrep.2013.03.029. Epub 2013 Apr 18. PubMed PMID: 23602564.

**Perrin** L, Monier B, Ponzelli R, Astier M, Semeriva M. *Drosophila* cardiac tube organogenesis requires multiple phases of Hox activity. *Dev Biol*. **2004** Aug 15;272(2):419-31. PubMed PMID: 15282158.

**Pineault** N, Helgason CD, Lawrence HJ, Humphries RK. Differential expression of Hox, Meis1, and Pbx1 genes in primitive cells throughout murine hematopoietic ontogeny. *Exp Hematol*. **2002** Jan;30(1):49-57. PubMed PMID: 11823037.

**Piper** DE, Batchelor AH, Chang CP, Cleary ML, Wolberger C. Structure of a HoxB1-Pbx1 heterodimer bound to DNA: role of the hexapeptide and a fourth homeodomain helix in complex formation. *Cell*. **1999** Feb 19;96(4):587-97. PubMed PMID: 10052460.

**Prévôt** D, Voeltzel T, Birot AM, Morel AP, Rostan MC, Magaud JP, Corbo L. The leukemia-associated protein Btg1 and the p53-regulated protein Btg2 interact with the homeoprotein Hoxb9 and enhance its transcriptional activation. *J Biol Chem*. **2000** Jan 7;275(1):147-53. PubMed PMID: 10617598.

**Ramachandran** S, Liu P, Young AN, Yin-Goen Q, Lim SD, Laycock N, Amin MB, Carney JK, Marshall FF, Petros JA, Moreno CS. Loss of HOXC6 expression induces apoptosis in prostate cancer cells. *Oncogene*. **2005** Jan 6;24(1):188-98. PubMed PMID: 15637592.

**Raman** V, Martensen SA, Reisman D, Evron E, Odenwald WF, Jaffee E, Marks J, Sukumar S. Compromised HOXA5 function can limit p53 expression in human breast tumours. *Nature*. **2000** Jun 22;405(6789):974-8. PubMed PMID: 10879542.

**Rezsohazy** R, Saurin AJ, Maurel-Zaffran C, Graba Y. Cellular and molecular insights into Hox protein action. *Development*. **2015** Apr 1;142(7):1212-27. doi:10.1242/dev.109785. Review. PubMed PMID: 25804734.

**Rezsohazy** R. Non-transcriptional interactions of Hox proteins: inventory, facts, and future directions. *Dev Dyn*. **2014** Jan;243(1):117-31. doi: 10.1002/dvdy.24060. Epub 2013 Nov 7. Review. PubMed PMID: 24115586.

**Rieckhof** GE, Casares F, Ryoo HD, Abu-Shaar M, Mann RS. Nuclear translocation of extradenticle requires homothorax, which encodes an extradenticle-related homeodomain protein. *Cell*. **1997** Oct 17;91(2):171-83. PubMed PMID: 9346235.

**Robertson** LK, Bowling DB, Mahaffey JP, Imiolczyk B, Mahaffey JW. An interactive network of zinc-finger proteins contributes to regionalization of the Drosophila embryo and establishes the domains of HOM-C protein function. *Development*. **2004** Jun;131(12):2781-9. Epub.2004 May 13. PubMed PMID: 15142974.

**Rohs** R, West SM, Sosinsky A, Liu P, Mann RS, Honig B. The role of DNA shape in protein-DNA recognition. *Nature*. **2009** Oct 29;461(7268):1248-53. doi: 10.1038/nature08473. PubMed PMID: 19865164; PubMed Central PMCID: PMC2793086.

**Rubin** E, Wu X, Zhu T, Cheung JC, Chen H, Lorincz A, Pandita RK, Sharma GG, Ha HC, Gasson J, Hanakahi LA, Pandita TK, Sukumar S. A role for the HOXB7 homeodomain protein in DNA repair. *Cancer Res*. **2007** Feb 15;67(4):1527-35. PubMed PMID: 17308091.

**Saadaoui** M, Merabet S, Litim-Mecheri I, Arbeille E, Sambrani N, Damen W, Brena C, Pradel J, Graba Y. Selection of distinct Hox-Extradenticle interaction modes fine-tunes Hox protein activity. *Proc Natl Acad Sci U S A*. **2011** Feb 8;108(6):2276-81. doi: 10.1073/pnas.1006964108. Epub 2011 Jan 24. PubMed PMID:21262810; PubMed Central PMCID: PMC3038764.

**Salsi** V, Ferrari S, Ferraresi R, Cossarizza A, Grande A, Zappavigna V. HOXD13 binds DNA replication origins to promote origin licensing and is inhibited by geminin. *Mol Cell Biol.* **2009** Nov;29(21):5775-88. doi: 10.1128/MCB.00509-09. Epub 2009 Aug 24. PubMed PMID: 19703996; PubMed Central PMCID: PMC2772751.

**Sauvageau** G, Lansdorp PM, Eaves CJ, Hogge DE, Dragowska WH, Reid DS, Largman C, Lawrence HJ, Humphries RK. Differential expression of homeobox genes in functionally distinct CD34+ subpopulations of human bone marrow cells. *Proc Natl Acad Sci U S A.* **1994** Dec 6;91(25):12223-7. PubMed PMID: 7527557; PubMed Central PMCID: PMC45409.

**Sauvageau** G, Thorsteinsdottir U, Eaves CJ, Lawrence HJ, Largman C, Lansdorp PM, Humphries RK. Overexpression of HOXB4 in hematopoietic cells causes the selective expansion of more primitive populations in vitro and in vivo. *Genes Dev.* **1995** Jul 15;9(14):1753-65. PubMed PMID: 7622039.

**Sauvageau** G, Thorsteinsdottir U, Hough MR, Hugo P, Lawrence HJ, Largman C, Humphries RK. Overexpression of HOXB3 in hematopoietic cells causes defective lymphoid development and progressive myeloproliferation. *Immunity.* **1997** Jan;6(1):13-22. PubMed PMID: 9052833.

**Saxena** S, Yuan P, Dhar SK, Senga T, Takeda D, Robinson H, Kornbluth S, Swaminathan K, Dutta A. A dimerized coiled-coil domain and an adjoining part of geminin interact with two sites on Cdt1 for replication inhibition. *Mol Cell.* **2004** Jul 23;15(2):245-58. PubMed PMID: 15260975.

**Shah** CA, Wang H, Bei L, Plataniias LC, Eklund EA. HoxA10 regulates transcription of the gene encoding transforming growth factor beta2 (TGFbeta2) in myeloid cells. *J Biol Chem.* **2011** Jan 28;286(4):3161-76. doi: 10.1074/jbc.M110.183251. Epub 2010 Nov 18. PubMed PMID: 21087928; PubMed Central PMCID: PMC3024808.

**Shepherd** JC, McGinnis W, Carrasco AE, De Robertis EM, Gehring WJ. Fly and frog homoeo domains show homologies with yeast mating type regulatory proteins. *Nature.* **1984** Jul 5-11;310(5972):70-1. PubMed PMID: 6429549.

**Sitwala** KV, Dandekar MN, Hess JL. HOX proteins and leukemia. *Int J Clin Exp Pathol.* **2008** Mar 30;1(6):461-74. PubMed PMID: 18787682; PubMed Central PMCID: PMC2480589.

**Sivanantharajah** L, Percival-Smith A. Differential pleiotropy and HOX functional organization. *Dev Biol.* **2015** Feb 1;398(1):1-10. doi: 10.1016/j.ydbio.2014.11.001. Epub 2014 Nov 11. Review. PubMed PMID: 25448696.

**Slattery** M, Ma L, Négre N, White KP, Mann RS. Genome-wide tissue-specific occupancy of the Hox protein Ultrabithorax and Hox cofactor Homothorax in *Drosophila*. *PLoS One.* **2011** Apr 5;6(4):e14686. doi: 10.1371/journal.pone.0014686. PubMed PMID: 21483663; PubMed Central PMCID: PMC3071676.

**Slattery** M, Riley T, Liu P, Abe N, Gomez-Alcala P, Dror I, Zhou T, Rohs R, Honig B, Bussemaker HJ, Mann RS. Cofactor binding evokes latent differences in DNA binding specificity between Hox proteins. *Cell*. **2011** Dec 9;147(6):1270-82.doi: 10.1016/j.cell.2011.10.053. PubMed PMID: 22153072; PubMed Central PMCID: PMC3319069.

**Sonenberg** N, Gingras AC. The mRNA 5' cap-binding protein eIF4E and control of cell growth. *Curr Opin Cell Biol*. **1998** Apr;10(2):268-75. Review. PubMed PMID:9561852.

**Sorge** S, Ha N, Polychronidou M, Friedrich J, Bezdán D, Kaspar P, Schaefer MH, Ossowski S, Henz SR, Mundorf J, Rätzer J, Papagiannouli F, Lohmann I. The cis-regulatory code of Hox function in *Drosophila*. *EMBO J*. **2015** Nov 12;34(22):2862. doi: 10.15252/embj.201593321. PubMed PMID: 26564906; PubMedCentral PMCID: PMC4682644.

**Stevens** KE, Mann RS. A balance between two nuclear localization sequences and a nuclear export sequence governs extradenticle subcellular localization. *Genetics*. **2007** Apr;175(4):1625-36. Epub 2007 Feb 4. PubMed PMID: 17277370; PubMed Central PMCID: PMC1855138.

**Suzuki** M, Ueno N, Kuroiwa A. Hox proteins functionally cooperate with the GCbox-binding protein system through distinct domains. *J Biol Chem*. **2003** Aug 8;278(32):30148-56. Epub 2003 May 27. PubMed PMID: 12771139.

**Takatori** N, Saiga H. Evolution of CUT class homeobox genes: insights from the genome of the amphioxus, *Branchiostoma floridae*. *Int J Dev Biol*. **2008**;52(7):969-77. doi: 10.1387/ijdb.072541nt. PubMed PMID: 18956327.

**Talaty** P, Emery A, Holthusen K, Everly DN Jr. Identification of transmembrane protein 134 as a novel LMP1-binding protein by using bimolecular fluorescence complementation and an enhanced retroviral mutagen. *J Virol*. **2012** Oct;86(20):11345-55. Epub 2012 Aug 1. PubMed PMID: 22855487; PubMed Central PMCID: PMC3457160.

**Taminiau** A, Draime A, Tys J, Lambert B, Vandeputte J, Nguyen N, Renard P, Geerts D, Rezsöhazy R. HOXA1 binds RBCK1/HOIL-1 and TRAF2 and modulates the TNF/NF- $\kappa$ B pathway in a transcription-independent manner. *Nucleic Acids Res*. **2016** Jul 5. pii: gkw606. [Epub ahead of print] PubMed PMID: 27382069.

**Thorsteinsdottir** U, Kroon E, Jerome L, Blasi F, Sauvageau G. Defining roles for HOX and MEIS1 genes in induction of acute myeloid leukemia. *Mol Cell Biol*. **2001** Jan;21(1):224-34. PubMed PMID: 11113197; PubMed Central PMCID: PMC88796.

**Tompa** P, Davey NE, Gibson TJ, Babu MM. A million peptide motifs for the molecular biologist. *Mol Cell*. **2014** Jul 17;55(2):161-9. doi: 10.1016/j.molcel.2014.05.032. Review. PubMed PMID: 25038412.

**Tompa P.** Intrinsically disordered proteins: a 10-year recap. *Trends Biochem Sci.* **2012** Dec;37(12):509-16. doi: 10.1016/j.tibs.2012.08.004. Epub 2012 Sep 16. Review. PubMed PMID: 22989858.

**Topisirovic I, Borden KL.** Homeodomain proteins and eukaryotic translation initiation factor 4E (eIF4E): an unexpected relationship. *Histol Histopathol.* **2005** Oct;20(4):1275-84. Review. PubMed PMID: 16136508.

**Topisirovic I, Guzman ML, McConnell MJ, Licht JD, Culjkovic B, Neering SJ, Jordan CT, Borden KL.** Aberrant eukaryotic translation initiation factor4E-dependent mRNA transport impedes hematopoietic differentiation and contributes to leukemogenesis. *Mol Cell Biol.* **2003** Dec;23(24):8992-9002. PubMed PMID: 14645512; PubMed Central PMCID: PMC309660.

**Topisirovic I, Kentsis A, Perez JM, Guzman ML, Jordan CT, Borden KL.** Eukaryotic translation initiation factor 4E activity is modulated by HOXA9 at multiple levels. *Mol Cell Biol.* **2005** Feb;25(3):1100-12. PubMed PMID: 15657436;PubMed Central PMCID: PMC544005.

**Tour E, Hittinger CT, McGinnis W.** Evolutionarily conserved domains required for activation and repression functions of the Drosophila Hox protein Ultrabithorax. *Development.* **2005** Dec;132(23):5271-81. PubMed PMID: 16284118.

**Van Roey K, Uyar B, Weatheritt RJ, Dinkel H, Seiler M, Budd A, Gibson TJ, Davey NE.** Short linear motifs: ubiquitous and functionally diverse protein interaction modules directing cell regulation. *Chem Rev.* **2014** Jul 9;114(13):6733-78. doi: 10.1021/cr400585q. Epub 2014 Jun 13. Review. PubMed PMID:24926813.

**Vinagre T, Moncaut N, Carapuço M, Nóvoa A, Bom J, Mallo M.** Evidence for amyotomal Hox/Myf cascade governing nonautonomous control of rib specification within global vertebral domains. *Dev Cell.* **2010** Apr 20;18(4):655-61. doi: 10.1016/j.devcel.2010.02.011. PubMed PMID: 20412779.

**Walsh CM, Carroll SB.** Collaboration between Smads and a Hox protein in target gene repression. *Development.* **2007** Oct;134(20):3585-92. Epub 2007 Sep 12. PubMed PMID: 17855427.

**Waltregny D, Alami Y, Clause N, de Leval J, Castronovo V.** Overexpression of the homeobox gene HOXC8 in human prostate cancer correlates with loss of tumor differentiation. *Prostate.* **2002** Feb 15;50(3):162-9. PubMed PMID: 11813208.

**Witt LM, Gutzwiller LM, Gresser AL, Li-Kroeger D, Cook TA, Gebelein B.** Atonal, Senseless, and Abdominal-A regulate rhomboid enhancer activity in abdominal sensory organ precursors. *Dev Biol.* **2010** Aug 15;344(2):1060-70. doi: 10.1016/j.ydbio.2010.05.011. Epub 2010 May 15. PubMed PMID: 20478292; PubMed Central PMCID: PMC2914175.

**Wright** PE, Dyson HJ. Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J Mol Biol.* **1999** Oct 22;293(2):321-31. Review. PubMed PMID: 10550212.

**Wu** X, Chen H, Parker B, Rubin E, Zhu T, Lee JS, Argani P, Sukumar S. HOXB7, a homeodomain protein, is overexpressed in breast cancer and confers epithelial-mesenchymal transition. *Cancer Res.* **2006** Oct 1;66(19):9527-34. Erratum in: *Cancer Res.* 2015 Jun 1;75(11):2401. PubMed PMID: 17018609.

**Yamashita** T, Tazawa S, Yawei Z, Katayama H, Kato Y, Nishiwaki K, Yokohama Y, Ishikawa M. Suppression of invasive characteristics by antisense introduction of overexpressed HOX genes in ovarian cancer cells. *Int J Oncol.* **2006** Apr;28(4):931-8. PubMed PMID: 16525643.

**Zhao** JJ, Lazzarini RA, Pick L. Functional dissection of the mouse Hox-a5 gene. *EMBO J.* **1996** Mar 15;15(6):1313-22. PubMed PMID: 8635464; PubMed Central PMCID: PMC450034.

**Zhong** YF, Holland PW. The dynamics of vertebrate homeobox gene evolution: gain and loss of genes in mouse and human lineages. *BMC Evol Biol.* **2011** Jun 16;11:169.doi: 10.1186/1471-2148-11-169. Erratum in: *BMC Evol Biol.* 2011;11:204. PubMed PMID: 21679462; PubMed Central PMCID: PMC3141429.

**Zhou** J, Lin J, Zhou C, Deng X, Xia B. An improved bimolecular fluorescence complementation tool based on superfolder green fluorescent protein. *Acta Biochim Biophys Sin (Shanghai).* **2011** Mar;43(3):239-44. doi: 10.1093/abbs/gmq128. Epub 2011 Jan 27. PubMed PMID: 21273204.



## ANNEXES

---





## De la PROPRIETE INTELLECTUELLE à l'INTÉRESSEMENT

### Liste des INVENTEURS et REPARTITION des PARTS INVENTIVES entre les inventeurs

**Lister de façon exhaustive tous les inventeurs(\*) mais seulement les inventeurs, quels que soient leurs statuts et leurs employeurs, inventeurs qui apparaîtront sur le brevet.**  
**Cette liste doit être complète et refléter la réalité de la contribution inventive de chaque inventeur.**

**Les signatures certifient l'accord entre les inventeurs sur leurs parts inventives respectives.**

Noms et prénoms des inventeurs (*)	Lieu d'exercice lors de l'obtention des résultats (n° d'unité, entreprise...)	Employeur lors de l'obtention des résultats	Contribution à l'invention (% de part inventive)	Date	Signature
MERABET Samir	IGFL, UMR5242	CNRS	31	01/04 /2016	
MANGEOT Philippe	CIRI, UMR5308	INSERM	24	01/04 /2016	
DARD Amélie	IGFL, UMR5242	CNRS	15	01/04 /2016	
BLEICHER Françoise	IGFL, UMR5242	UCBL	20	01/04 /2016	
HUGUES Sandrine	IGFL, UMR5242	CNRS	5	01/04 /2016	
GILLET Benjamin	IGFL, UMR5242	CNRS	5	01/04 /2016	
<b>TOTAL</b>			100		

(\*) Est inventeur toute personne, chercheur, ingénieur, technicien, ayant contribué directement à la conception et/ou à la réalisation de tout ou partie d'une invention. Il doit y avoir un apport inventif personnel à la concrétisation de l'invention. Cette désignation se fait de façon objective, indépendamment de toute autre considération, qu'elle soit honorifique, de préséance, hiérarchique ou financière.  
 Ne pas confondre co-inventeur d'une invention et co-auteur d'une publication scientifique.

## Fiche individuelle d'INVENTEUR

Veillez **remplir UNE FICHE PAR INVENTEUR FONCTIONNAIRE OU AGENT PUBLIC** et joindre une **copie d'un bulletin de salaire délivré au moment de l'invention**. Si vous le souhaitez, vous pouvez masquer les parties chiffrées de votre rémunération.



Pour les **personnels contractuels (non permanents)**, joindre une copie du contrat de travail ou de la décision d'engagement. Joindre une copie de **l'accord de cession de droit** s'il en existe un.

INVENTEUR	Situation lors de l'obtention des résultats		Situation actuelle (si différente)	
NOM : MERABET  Nom de jeune fille :  Prénom : Samir  Date de naissance : 01/06/1974  Nationalité : Française  Adresse personnelle : 41 chemin de Figuerolles 13700 Marignane	Employeur : CNRS  Statut : DR2  Dates début/fin du contrat de travail : 10/2005	Adresse professionnelle : IGFL, UMR 5242, ENS-Lyon, 46 Allée d'Italie, 69007 Lyon tel : 0426731317 0622821188  mail : samir.merabet@ens-lyon.fr	Employeur :  Statut :  Dates début/fin du contrat de travail : /	Adresse professionnelle : (adresse complète, tél., méi.)
Contribution de l'inventeur à l'invention (succinctement) : Idée originale du projet. Mise en place du projet. Suivi des différentes étapes du projet.				
S'agit-il d'une invention : <input checked="" type="checkbox"/> De mission <input type="checkbox"/> Hors mission attribuable <input type="checkbox"/> Hors mission non attribuable <input type="checkbox"/> Commenter si besoin :				

Si l'inventeur est contractuel, stagiaire, étudiant, un engagement de confidentialité a-t-il été signé ?  oui  non

**L'inventeur certifie ne pas avoir divulgué d'informations sur son invention.**

**Date et signature : 04/03/2016**

## VISAS

### Responsable Projet :

Nom : MERABET

Prénom : Samir

Date : 04/03/2016

Signature :



### Directeur du laboratoire :

Nom : RUGGIERO

Prénom : Florence

Avis motivé :

Date : 04/03/2016

Signature :





## RESUME

---

Les gènes Hox sont présents dans la majorité des espèces du règne animal et sont nécessaires à la différenciation coordonnée des cellules le long de différents axes longitudinaux au cours du développement embryonnaire. Ils sont impliqués dans le maintien de l'homéostasie de nombreux tissus à l'âge adulte. Des mutations affectant leur expression et/ou leur fonction sont ainsi retrouvées dans de nombreux cancers chez l'Homme.

Les gènes Hox codent pour des facteurs de transcription reconnaissant des séquences nucléotidiques très similaires. L'interaction avec une classe évolutivement conservée de cofacteurs, les protéines Pbx et Meis, permet aux protéines Hox de reconnaître des sites de liaison plus spécifiques. Cette interaction a d'abord été décrite pour dépendre d'un petit motif commun aux protéines Hox, l'hexapeptide (HX). Cependant, des analyses récentes ont montré que ce motif pouvait en fait être dispensable *in vivo*, soulignant une capacité étonnante des protéines Hox à pouvoir potentiellement utiliser différents motifs pour interagir avec les mêmes cofacteurs.

Mon travail de thèse s'inscrit dans la problématique du rôle des petits motifs dans les interactions Hox-cofacteur. Un premier projet a consisté à réaliser une analyse systématique du mode d'interaction de chaque représentant des groupes de paralogie des protéines Hox humaines avec leurs cofacteurs Pbx/Meis. Ce travail a révélé de nouveaux modes d'interaction pour plusieurs protéines Hox. Un deuxième projet a consisté à mettre en place un nouveau système de crible moléculaire pour identifier des partenaires de la protéine humaine HoxA9 sauvage ou mutée dans son motif HX dans différentes lignées cellulaires.

L'ensemble de mon travail de thèse ouvre ainsi de nouvelles perspectives sur notre compréhension du mode moléculaire d'action des protéines Hox et de leurs cofacteurs, que cela soit en contexte développemental normal ou pathologique.