



Étude de la régulation transcriptionnelle du gène *Hoxa5* chez la souris

Thèse

Félix-Antoine Bérubé-Simard

Doctorat en biologie cellulaire et moléculaire
Philosophiae Doctor (Ph.D.)

Québec, Canada

© Félix-Antoine Bérubé-Simard, 2014

RÉSUMÉ

Les gènes *Hox* codent pour des facteurs de transcription orchestrant l'identité antéro-postérieure du plan corporel des animaux à symétrie bilatérale. La souris *Hoxa5*^{-/-} a permis de démontrer que ce gène joue un rôle primordial dans la spécification des squelettes axial et appendiculaire, ainsi que dans l'ontogénie de plusieurs organes. À l'aide d'une approche de transgénèse et de délétions successives de la séquence intergénique *Hoxa4-Hoxa5*, j'ai identifié deux éléments régulateurs responsables de l'expression du gène *Hoxa5* dans les systèmes respiratoire et digestif: un fragment d'ADN *NcoI-SacI* de 163-pb possédant une activité de type activatrice et dirigeant l'expression au niveau du poumon, de l'estomac et de l'intestin, de même qu'un fragment *XbaI-BssHII* de 259-pb, nécessaire à une expression complète du gène *Hoxa5* au niveau du système digestif. Des expériences de retard sur gel (EMSA) et d'immunoprécipitation de la chromatine (ChIP) m'ont permis de démontrer la liaison du facteur de transcription YY1 à ces deux séquences d'ADN. En mutant ses sites de liaison dans un contexte de transgénèse, j'ai mis en évidence le rôle de YY1 comme activateur transcriptionnel du gène *Hoxa5* dans les organes. Il s'agit d'ailleurs d'un des rares exemples où la protéine YY1 ne réprime pas l'expression des gènes *Hox*. J'ai également appliqué la technique de ChIP pour confirmer que les facteurs de transcription à boîte homéo CDX4 et HOXB9 se lient tous les deux au fragment d'ADN *AvrII-Eco47III* de 164-pb situé à l'intérieur de l'élément MES. J'ai donc montré que la protéine HOXB9 participe à restreindre caudalement l'expression du gène *Hoxa5* au niveau de la prévertèbre 10, supportant ainsi le concept de prévalence postérieure.

De plus, j'ai généré deux lignées de souris transgéniques exprimant la recombinase *Cre* sous le contrôle de deux combinaisons de séquences régulatrices identifiées du gène *Hoxa5*. Ces lignées ont été caractérisées et fournissent de nouveaux outils utiles pour étudier la fonction de différents gènes dans certains tissus le long de l'axe antéro-postérieur.

Enfin, le locus *Hoxa5* produit 4 transcrits de 1.8, 5.0, 9.5 et 11-kb de longueur se chevauchant et pouvant produire une protéine *in vitro*. Cependant, j'ai démontré que seul le court transcrit de 1.8-kb, correspondant aux deux exons connus du gène *Hoxa5*, génère une protéine associée à la fonction du gène *in vivo*. Les différents résultats obtenus seront présentés et discutés.

ABSTRACT

Hox genes encode transcription factors, which orchestrate bilaterian anteroposterior patterning. Using *Hoxa5*^{-/-} mice as model, we have demonstrated that this gene plays a key role in axial and appendicular skeletal patterning as well as in the formation of several organs such as the respiratory and digestive tracts. Using a transgenesis approach and successive deletions in the *Hoxa4-Hoxa5* intergenic region, I have identified two distinct regulatory elements responsible for *Hoxa5* expression in respiratory and digestive tracts: a 163-bp *NcoI-SacI* DNA fragment having enhancer activity that drives expression in lung, stomach and intestine, and a 259-bp *XbaI-BssHII* fragment necessary for a complete *Hoxa5* digestive tract expression. Electrophoretic mobility shift (EMSA) and chromatin immunoprecipitation (ChIP) assays have demonstrated the capacity of the YY1 transcription factor to bind these two DNA sequences. By mutating its binding sites in a transgenesis context, I have highlighted the transcriptional activator role of the YY1 protein in *Hoxa5* organ expression, which is very interesting since few examples of *Hox* gene activation by YY1 are reported in the literature.

I have also generated two transgenic mice lines expressing the *Cre* recombinase under the control of two combinations of identified regulatory sequences. These lines have been characterized and provide useful genetic tools to study gene function in specific tissues along the anteroposterior axis.

I have also applied ChIP technology to demonstrate the *in vivo* binding of CDX4 and HOXB9 homeobox transcription factors to the 164-bp *AvrII-Eco47III* DNA fragment included in the MES regulatory element. Consequently, I have shown that the HOXB9 protein caudally participates to restrict the *Hoxa5* gene expression at the level of prevertebra 10, which supports the posterior prevalence concept.

Finally, the *Hoxa5* locus encompasses 4 overlapping transcripts of 1.8, 5.0, 9.5 and 11.0-kb that can produce a HOXA5 protein in an *in vitro* context. However, I have demonstrated that only the short transcript of 1.8-kb corresponding to the two known *Hoxa5* gene exons is transcribed into an *in vivo* HOXA5 protein associated to the gene function. Data will be presented and discussed.

Table des matières

RÉSUMÉ	iii
ABSTRACT	v
Table des matières	vii
Liste des tableaux	xi
Liste des figures	xiii
Liste des abréviations	xvii
REMERCIEMENTS	xix
CHAPITRE I : Introduction	1
1.1 La drosophile	3
1.1.1 L'émergence de la génétique développementale	3
1.1.2 Segmentation antéro-postérieure	6
1.1.2.1 Les gènes <i>gap</i>	7
1.1.2.2 Les gènes de parité segmentaire (<i>pair-rule</i>)	7
1.1.2.3 Les gènes de polarité segmentaire	7
1.1.3 Identité antéro-postérieure	8
1.1.3.1 Les gènes homéotiques sélecteurs	8
1.1.3.2 La structure des protéines codées par les gènes <i>Hox</i>	9
1.1.3.3 Le modèle de liaison sélective	12
1.1.3.4 Le modèle d'activité régulée	12
1.2 Les mammifères	13
1.2.1 Conservation des gènes homéotiques dans l'évolution	13
1.2.2 Organisation des gènes <i>Hox</i>	15
1.2.3 Principe de colinéarité	18
1.2.4 Spécificité et modifications post-traductionnelles	19
1.2.5 Spécificité et cofacteurs de liaison	19
1.2.6 Établissement du squelette axial	21
1.2.6.1 Les gènes <i>Hox</i> spécifiant la tête (paralogues 1 à 4)	22
1.2.6.2 Les gènes <i>Hox</i> spécifiant le tronc (paralogues 5 à 8)	23
1.2.6.3 Les gènes <i>Hox</i> spécifiant la partie caudale (paralogues 9 à 13)	26
1.2.7 Établissement du squelette paraxial	27
1.2.8 Rôles dans l'organogenèse	29
1.2.8.1 Système nerveux central (SNC) et cellules de la crête neurale (CCN)	29
1.2.8.2 Organes glandulaires	33
1.2.8.3 Système respiratoire	33
1.2.8.4 Système digestif	37
1.2.8.5 Formation des membres	40
1.2.8.6 Système urogénital	41
1.2.8.7 Glande mammaire	42
1.2.8.8 Barrière cutanée	43
1.2.8.9 Cellules hématopoïétiques	43
1.2.8.10 Angiogenèse et microenvironnement tumoral	44
1.2.8.11 Structures extra-embryonnaires	45
1.3 Régulation de l'expression des gènes <i>Hox</i>	45

1.3.1 L'autorégulation.....	45
1.3.2 La régulation croisée.....	47
1.3.3 La suppression phénotypique ou prévalence postérieure.....	49
1.3.4 Les facteurs agissant en <i>trans</i>	50
1.3.4.1 L'acide rétinoïque (RA).....	51
1.3.4.2 Les protéines CDX.....	56
1.3.4.3 Voies de signalisation FGF, Wnt et Notch.....	58
1.3.5 Partage et sélectivité d'éléments régulateurs.....	62
1.3.6 Régulation à distance.....	66
1.3.7 Organisation tridimensionnelle de la chromatine.....	71
1.3.8 Les protéines membres du groupe Polycomb (PcG).....	73
1.3.8.1 Complexe répressif Polycomb de type 1(PRC1).....	75
1.3.8.2 Complexe répressif Polycomb de type 2(PRC2).....	76
1.3.8.3 Facteur de transcription YY1 (complexe PRC0 ou PHORC).....	76
1.3.9 Les protéines membres du groupe Trithorax (TrxG).....	81
1.3.10 Transcrits multiples.....	82
1.3.11 Longs ARN non codants (lncRNAs).....	83
1.3.12 Régulation post-transcriptionnelle.....	84
1.3.12.1 MicroARNs (miRNAs).....	84
1.3.12.2 Protéines ribosomales.....	86
1.3.13 Interprétation des phénotypes et cassette de sélection.....	86
1.4 Le gène <i>Hoxa5</i>	87
1.4.1 Fonction du gène <i>Hoxa5</i>	89
1.4.1.1 Établissement du squelette axial.....	89
1.4.1.2 Établissement du squelette paxial et appendiculaire.....	92
1.4.1.3 Système nerveux central (SNC) et cellules de la crête neurale (CCN).....	93
1.4.1.4 Organes glandulaires.....	96
1.4.1.5 Système respiratoire.....	96
1.4.1.6 Système digestif.....	98
1.4.1.7 Système urogénital.....	99
1.4.1.8 Glande mammaire.....	100
1.4.1.8 Cellules hématopoïétiques.....	100
1.4.1.10 Angiogenèse et microenvironnement tumoral.....	101
1.4.2 Régulation de l'expression du gène <i>Hoxa5</i>	102
1.4.2.1 Élément Temporel.....	103
1.4.2.2 Élément BSC (Brachial Spinal Cord).....	103
1.4.2.3 Élément MES (Mesodermal Enhancer Sequence).....	103
1.4.2.4 Élément Poumon/Intestin.....	106
1.5 Projet de ma thèse de doctorat.....	107
CHAPITRE II: YY1 acts as a transcriptional activator of <i>Hoxa5</i> gene expression in mouse organogenesis.....	11
AVANT-PROPOS.....	113
ABSTRACT.....	115
INTRODUCTION.....	117
MATERIALS AND METHODS.....	119
RESULTS.....	126

DISCUSSION	149
ACKNOWLEDGMENTS	156
REFERENCES.....	156
CHAPITRE III: <i>Hoxa5/Cre</i> transgenic mice: novel tools for regional deletion along the anterior-posterior	161
AVANT-PROPOS	163
ABSTRACT	165
INTRODUCTION	167
RESULTS AND DISCUSSION	168
METHODS	173
ACKNOWLEDGMENTS	180
REFERENCES.....	180
CHAPITRE IV: <i>Hoxa5</i> regional expression along the developping anteroposterior axis requires several regulatory elements.....	183
AVANT-PROPOS	185
ABSTRACT	187
INTRODUCTION	189
MATERIALS AND METHODS	192
RESULTS	198
DISCUSSION	213
ACKNOWLEDGMENTS	220
REFERENCES.....	220
CHAPITRE V: Multiple Promoters and Alternative Splicing: <i>Hoxa5</i> Transcriptional Complexity in the Mouse Embryo.....	227
AVANT-PROPOS	229
ABSTRACT	231
INTRODUCTION	233
RESULTS	236
DISCUSSION	258
MATERIALS AND METHODS	266
ACKNOWLEDGMENTS	271
AUTHOR CONTRIBUTIONS.....	271
REFERENCES.....	272
CHAPITRE VI: Partial functional redundancy between <i>Hoxa5</i> and <i>Hoxb5</i> paralog genes during lung morphogenesis.....	281
AVANT-PROPOS	283
ABSTRACT	285
INTRODUCTION	287
MATERIALS AND METHODS	289
RESULTS	293
DISCUSSION	313
ACKNOWLEDGMENTS	316
GRANTS.....	317

DISCLOSURES	317
AUTHOR CONTRIBUTIONS	317
REFERENCES	317
Chapitre VII :	323
7.1 ADDENDUM AU CHAPITRE III (<i>Hoxa5/Cre</i> transgenic mice: novel tools for regional deletion along the anterior-posterior axis)	323
AVANT-PROPOS	323
7.2 Rôle de l'expression du gène <i>Yy1</i> au niveau de l'épithélium des voies respiratoires	335
AVANT-PROPOS	335
7.3 Rôle critique de la signalisation ERK/MAPK dans le développement du système respiratoire	343
AVANT-PROPOS	343
Chapitre VIII : DISCUSSION	347
8.1 L'élément poumon/intestin contient plusieurs séquences régulatrices importantes	347
8.1.1. Le facteur YY1 régule positivement l'expression du gène <i>Hoxa5</i> dans les poumons	350
8.2 Les lignées <i>Hoxa5/Cre</i> permettent d'étudier la fonction de gènes dans certains tissus le long de l'axe AP	367
8.3 L'élément MES contient plusieurs séquences régulatrices importantes	374
8.4. Caractérisation des lncRNAs du gène <i>Hoxa5</i>	380
8.5. Conclusion	381
RÉFÉRENCES	387
ANNEXE	435
AVANT-PROPOS	437
ABSTRACT	439
INTRODUCTION	441
MATERIALS AND METHODS	442
RESULTS	447
DISCUSSION	470
CONCLUSION	475
REFERENCES	478

Liste des tableaux

CHAPITRE II	
Tableau 2.1.....	121
CHAPITRE IV	
Tableau 4.1.....	195
CHAPITRE V	
Tableau 5.1.....	259
CHAPITRE VI	
Tableau 6.1.....	296
ANNEXE I	
Tableau 1.....	465
Tableau 2.....	473

Liste des figures

CHAPITRE I

<u>Figure 1.1</u> : Drosophile de type sauvage et mutante pour le gène <i>Antp</i>	5
<u>Figure 1.2</u> : Représentation schématique des complexes <i>Hox</i> et de leur expression chez la drosophile.....	11
<u>Figure 1.3</u> : Représentation schématique des complexes <i>Hox</i> chez la souris.....	17
<u>Figure 1.4</u> : Schéma d'une vue latérale d'un squelette de souris	25
<u>Figure 1.5</u> : Schéma représentatif de l'axe AP.....	31
<u>Figure 1.6</u> : Domaines d'expression des gènes <i>Hox</i> au niveau du cerveau primitif postérieur d'un embryon de souris au stade E11.5.....	35
<u>Figure 1.7</u> : Représentation schématique des 5 stades développementaux de la morphogenèse pulmonaire chez la souris et l'humain.....	39
<u>Figure 1.8</u> : Modèles représentant le phénomène de prévalence postérieure.....	53
<u>Figure 1.9</u> : Réseau d'interactions génétiques impliquant les gènes <i>Hox</i>	61
<u>Figure 1.10</u> : Modèle résumant le partage et la spécificité d'action des éléments régulateurs situé en 3' du complexe <i>Hoxb</i>	65
<u>Figure 1.11</u> : Régulation globale de l'expression du complexe <i>HoxD</i>	69
<u>Figure 1.12</u> : Expression des transcrits du gène <i>Hoxa5</i> au niveau de l'embryon de souris au stade développemental E12.5.....	91
<u>Figure 1.13</u> : Résumé des différents phénotypes observés chez les souris <i>Hoxa5</i> ^{-/-}	95
<u>Figure 1.14</u> : Résumé des régions régulatrices contrôlant l'expression du gène <i>Hoxa5</i> identifiées par transgénèse chez la souris.	105

CHAPITRE II

<u>Figure 2.1</u> : β -galactosidase expression profile of construct 1 during embryogenesis	129
<u>Figure 2.2</u> : Characterization of the <i>Hoxa4-Hoxa5</i> intergenic region in E13.5 F0 transgenic mouse embryos	131
<u>Figure 2.3</u> : Characterization of the 1.5-kb <i>XbaI-XbaI</i> DNA fragment in E13.5 F0 transgenic mouse embryos	133
<u>Figure 2.4</u> : Characterization of the 751-bp <i>BssHII-SacI</i> DNA fragment in E13.5 F0 transgenic mouse embryos	137
<u>Figure 2.5</u> : Characterization of the 259-bp <i>XbaI-BssHII</i> DNA fragment by EMSA	139
<u>Figure 2.6</u> : Characterization of the RARE and YY1 binding sites in E13.5 F0 transgenic mouse embryos	141
<u>Figure 2.7</u> : Characterization of the 163-bp <i>NcoI-SacI</i> DNA fragment by EMSA	145
<u>Figure 2.8</u> : <i>In vivo</i> detection of YY1 protein binding to Oligo C3 and Oligo G3 regions by ChIP analysis	147
<u>Figure 2.9</u> : Analysis of the lung phenotype in <i>Yyl</i> ^{flox/flox} <i>Dermo1</i> ^{+/-Cre} mutants.....	151

CHAPITRE III

<u>Figure 3.1</u> : Characterization of <i>Hoxa5/Cre</i> mouse lines during development and at adult age.....	171
<u>Figure 3.2</u> : Impact of the conditional <i>Hoxa5</i> deletion with the <i>Hoxa5/Cre</i> transgenes.....	175
<u>Figure 3.3</u> : Comparative analysis of the skeletal morphology of <i>Hoxa5</i> conditional mutants	177

CHAPITRE IV

<u>Figure 4.1</u> : Characterization of MES regulatory element intergenic region in E12.5 F0 transgenic mouse embryos	201
<u>Figure 4.2</u> : Characterization of the 331-bp <i>PstI-HindIII</i> DNA fragment by EMSA.....	203
<u>Figure 4.3</u> : Characterization of the 164-bp <i>AvrII-Eco47III</i> DNA fragment by EMSA	209
<u>Figure 4.4</u> : <i>In vivo</i> detection of HOXB9 protein binding to the AvEc 164-bp region.....	211
<u>Figure 4.5</u> : HOXB9 and CDX4 proteins do not physically interact	215

CHAPITRE V

<u>Figure 5.1</u> : Schematic representation of the different transcripts encompassing <i>Hoxa5</i> sequences in the e12.5 mouse embryo	239
<u>Figure 5.2</u> : Molecular characterization of <i>Hoxa5</i> transcripts by northern analysis	241
<u>Figure 5.3</u> : Evolutionary conservation of the <i>Hoxa5-Hoxa7</i> genomic region among animal species.	245
<u>Figure 5.4</u> : Transcriptional activity of the D1 and D2 putative promoters in e12.5 transgenic embryos.	247
<u>Figure 5.5</u> : Differential expression pattern of <i>Hoxa5</i> transcripts.....	251
<u>Figure 5.6</u> : Restricted spatial distribution of HOXA5 protein along the anteroposterior axis and in the respiratory and digestive tracts.	253
<u>Figure 5.7</u> : HOXA5 protein production from <i>Hoxa5</i> transcripts	257
<u>Figure 5.8</u> : Comparative expression patterns of <i>Hoxa5</i> transcripts in e12.5 wildtype, <i>Hoxa6</i> homozygous and <i>Hoxa5</i> ; <i>Hoxa6</i> transheterozygous mutants....	261
<u>Figure supplémentaire 5.1</u> : Molecular characterization of the <i>Hoxa5</i> alternate transcripts by 5'-RACE and RT-PCR.	277
<u>Figure supplémentaire 5.2</u> : <i>Hoxa5</i> expression in the developing hindgut.	279

CHAPITRE VI

<u>Figure 6.1</u> : Expression analyses of <i>Hox5</i> paralog genes in the developing lung.....	295
<u>Figure 6.2</u> : Comparative analyses of lungs from <i>Hoxa5</i> ; <i>Hoxb5</i> mutant embryos.....	299
<u>Figure 6.3</u> : Characterization of lung epithelium and microvasculature of E18.5 <i>Hoxa5</i> ; <i>Hoxb5</i> mutant embryos.	303

<u>Figure 6.4:</u> <i>Hoxa5</i> plays a predominant role in tracheal formation and it regulates <i>Fstl1</i> expression in the developing trachea.	305
<u>Figure 6.5:</u> Abnormal diaphragm innervation pattern in <i>Hoxa5;Hoxb5</i> embryos carrying <i>Hoxa5</i> mutant alleles.....	309
<u>Figure 6.6:</u> Comparative analyses of lungs from <i>Hoxa5;Hoxb5</i> adult mutant mice.....	311

CHAPITRE VII

<u>Figure 7.1.1:</u> Génotypage par Southern blot pour contrôler l'efficacité de la délétion conditionnelle du gène <i>Hoxa5</i> à l'aide des deux lignées <i>Hoxa5/Cre</i>	325
<u>Figure 7.1.2:</u> Génotypage par Southern blot pour contrôler l'efficacité de la délétion conditionnelle du gène <i>Hoxa5</i> à l'aide des deux lignées <i>Hoxa5/Cre</i>	327
<u>Figure 7.1.3:</u> Génotypage par Southern blot pour contrôler l'efficacité de la délétion conditionnelle du gène <i>Hoxa5</i> à l'aide des deux lignées <i>Hoxa5/Cre</i>	329
<u>Figure 7.1.4:</u> Génotypage par Southern blot pour contrôler l'efficacité de la délétion conditionnelle du gène <i>Hoxa5</i> à l'aide des deux lignées <i>Hoxa5/Cre</i>	331
<u>Figure 7.1.5:</u> Impact de la délétion conditionnelle du gène <i>Yy1</i> suite à l'action de la 14.5-kb <i>Hoxa5/Cre</i>	333
<u>Figure 7.2.1:</u> Analyse du phénotype pulmonaire des souris <i>Yy1^{flox/flox}; Shh^{+Cre}</i>	339
<u>Figure 7.2.2:</u> Détection de la liaison des protéines YY1 et BAF53A au niveau de la région promotrice du gène <i>Shh</i> chez la souris.....	341
<u>Figure 7.3.1:</u> Dymorphogenèse cranio-faciale et anomalies squelettiques observées chez les embryons <i>Mek1^{flox/flox};Mek2^{-/-};Dermo1^{+Cre}</i>	345

CHAPITRE VIII

<u>Figure 8.1:</u> Caractérisation du fragment d'ADN <i>BssHII-NcoI</i> de 588-pb chez des embryons transgéniques F0 au stade E13.5.....	353
<u>Figure 8.2:</u> Détection de la liaison de la protéine YY1 au niveau des régions C3 et G3 par ChIP chez la souris	355
<u>Figure 8.3:</u> Impact de différentes délétions alléliques des gènes <i>Yy1</i> et <i>Hoxa5</i> sur le poumon durant le développement.....	359
<u>Figure 8.4:</u> Regroupements de sites potentiels de liaison dans la séquence intergénique <i>Hoxa4-Hoxa5</i>	365
<u>Figure 8.5:</u> Impact de la délétion conditionnelle du gène <i>Hoxa5</i> sur le poumon suite à l'action de la 14.5-kb <i>Hoxa5/Cre</i>	373
<u>Figure 8.6:</u> Caractérisation du fragment d'ADN <i>PstI-HindIII</i> 331-pb de 588-pb par EMSA	377
<u>Figure 8.7:</u> Stratégie utilisée pour générer les variants alléliques pour les lncRNAs <i>Hoxa5</i>	383

ANNEXE I

<u>Figure 1:</u> Constructs used to generate transgenic animals.....	445
<u>Figure 2:</u> Conserved non-coding elements in the <i>Hoxb5</i> genes of vertebrates.....	449
<u>Figure 3:</u> Embryonic expression of mouse and zebrafish <i>Hoxb5</i> genes.....	453
<u>Figure 4:</u> Expression of <i>lacZ</i> reporter transgenes in mouse embryos.....	455

<u>Figure 5</u> : Analysis of reporter transgene expression on sagittal sections and in the internal organs of mouse embryos.	459
<u>Figure 6</u> : Expression of <i>EGFP</i> reporter transgenes in zebrafish embryos.....	461
<u>Figure 7</u> : Enhancer activity of individual <i>Hoxb5</i> non-coding elements in transgenic mouse embryos.	467
<u>Figure 8</u> : Enhancer activity of individual <i>Hoxb5</i> non-coding elements in transgenic zebrafish.	469
<u>Figure 9</u> : Phylogenetic analysis of <i>Hoxb5</i> loci from human, mouse, zebrafish and <i>Takifugu</i>	477

Liste des abréviations

AA	: acide aminé
ADN	: acide désoxyribonucléique
ADNc	: ADN complémentaire
AER	: crête apicale ectodermique
AML	: leucémie myéloïde aigue
ARN	: acide ribonucléique
AP	: antéro-postérieur
ARE	: élément d'autorégulation
ARNm	: ARN messenger
BX-C	: complexe de gènes <i>Bithorax</i>
BSC	: Brachial Spinal Cord
C	: vertèbre cervicale
Ca	: vertèbre caudale
CCN	: cellule de la crête neurale
CDRE	: élément de réponse au Cdx
ChIP	: immunoprécipitation de la chromatine
CT	: territoire chromosomique
Da	: Dalton
E	: stade embryonnaire
EE/ELCR	: région de contrôle précoce des membres
EMSA	: essai de retard sur gel
EMT	: transition épithéliale-mésenchymale
ER α	: Récepteur alpha aux estrogènes
ESC	: cellules souches embryonnaires
FGF	: facteurs de croissance fibroblastique
GCR	: région de contrôle global
HM	: homozygote
HOM-C	: complexe des gènes homéotiques
HSPC	: cellules souches hématopoïétiques et progénitrices
HT	: hétérozygote
ICD	: domaine interchromosomique
IL-16	: Interleukine 16
ION	: noyau olivaire inférieur
Kb	: kilobase
L	: vertèbre lombaire
lncRNA	: long ARN non codant
LMC	: colonne motrice latérale
LPM	: mésoderme de la plaque latérale
MEF	: fibroblastes embryonnaires de souris
MES	: Mesodermal Enhancer Sequence
miRNA	: microARN
MPOC	: maladies pulmonaires obstructives chroniques
mRNP	: particule ribonucléoprotéique messagère
MSC	: cellules souches mésenchymateuses

nt	: nucléotide
ORF	: cadre de lecture ouvert
pb	: paire de base
PcG	: groupe Polycomb
PD	: proximodistal
PGC	: colonne de la ceinture pectorale
PMC	: colonne phrénique
PRC	: complexe de répression Polycomb
PRE	: élément de réponse aux protéines Polycomb
PS	: ligne primitive
PSM	: mésoderme présomitique
pv	: prévertèbre
r	: rhombomère
RA	: acide rétinoïque
RACE	: amplification rapide des extrémités des ADN complémentaires
RAR	: récepteur de l'acide rétinoïque
RARE	: élément de réponse à l'acide rétinoïque
RXR	: récepteur des rétinoïdes X
s	: somite
S	: vertèbre sacrale
SNC	: système nerveux central
SP	: protéine de surfactant
T	: vertèbre thoracique
TALE	: Three Amino Acid Loop Extension
T-ALL	: leucémie lymphoblastique aigüe de type lymphocyte T
TCR	: récepteur des lymphocytes T
TRE	: élément de réponse aux protéines Trithorax
TrxG	: groupe Trithorax
UTR	: UnTranslated Region
ZPA	: zone d'activité polarisante

REMERCIEMENTS

Comme l'étendue des résultats présentés dans le cadre de ma thèse de doctorat a été acquise sous la direction du Dre Lucie Jeannotte, dans son laboratoire du Centre de Recherche en Cancérologie de l'Université Laval/Centre Hospitalier Universitaire de L'Hôtel-Dieu de Québec, c'est sans réelle surprise que j'entame cette section en lui rendant hommage. Tout d'abord, je la remercie sincèrement d'avoir bien voulu accueillir ce jeune blanc-bec à l'allure peu orthodoxe qui, tel «un chien dans un jeu de quilles», est apparu sans crier gare un beau jour de février, en se montrant désireux de compléter un diplôme de maîtrise. Je quitte son laboratoire avec des acquis inestimables: une démarche scientifique rigoureuse ainsi qu'une solide formation en biologie cellulaire et moléculaire du développement chez la souris. J'ai particulièrement apprécié sa capacité d'émerveillement sans borne face aux différents phénomènes développementaux. Combien de fois j'ai souri en entendant : "Regarde comme c'est beau on voit super bien les somites" ou bien " As-tu vu le gradient d'expression ? De toute beauté !" Enfin, j'ai su profiter de ses judicieux conseils autant pour la rédaction d'articles que de cette thèse.

Les huit années passées au Centre de Recherche de l'Hôtel-Dieu de Québec m'ont également permis de faire plusieurs rencontres plus intéressantes les unes que les autres. Pour ces beaux moments, j'aimerais remercier:

Le Dr Jean Charron, pour sa disponibilité, son expertise en clonage moléculaire ainsi que son humour particulier.

La rayonnante Marcelle Carter, pour son énergie qui rendrait jaloux même Grégory Charles ainsi que pour sa dévotion à la plateforme de transgénèse.

Le Dr Sébastien Tabariès, pour avoir pris de son temps de rédaction en fin de thèse afin de m'enseigner les techniques qu'il maîtrisait.

Les membres des laboratoires Charron-Jeannotte, pour leur soutien, leur aide ainsi que leur amitié.

Le Dr Jacques Côté ainsi que les membres de son laboratoire, pour leurs précieux conseils en ce qui à trait à la technique d'immunoprécipitation de la chromatine.

Enfin, comment passer sous silence la contribution des héros obscurs de mon quotidien:

Mathieu Boulanger, Joachim Bernier, Michel Lachance, Jean-Sébastien Gosselin et Éric Pouliot, pour leur amitié indestructible et leur don pour me faire décrocher lorsque le besoin s'en fit sentir (à utiliser avec modération car peu créer une dépendance)

Ma copine Geneviève Dion, pour avoir fait passer les onze dernières années à la vitesse de l'éclair, m'avoir épaulé dans les moments les plus sombres et accompagné dans ceux plus lumineux. Tu réussis encore aujourd'hui à me faire sentir comme quelqu'un d'extraordinaire.

Bien qu'il s'agisse d'une infime attention en considérant les trente et une dernières années au cours desquelles ils entreprirent de me fournir tous les outils à leur disposition pour que je puisse être heureux, c'est à la fois avec satisfaction et humilité que j'offre à mes parents cette thèse de doctorat. Sans eux, ni elle ni moi n'existerions aujourd'hui.

Dans un premier temps, je leur serai éternellement reconnaissant de ne pas être toxicomanes, violents, extrémistes intégristes religieux ou toutes autres formes de dérèglements conduisant à une hypoplasie des facultés de l'esprit humain. Il m'apparaît plutôt impossible, ou du moins prétentieux, d'envisager être autant accompli, à la fois personnellement et académiquement parlant, si mon intégrité physique et morale avaient été compromises à la maison.

Dans un deuxième temps, je leur suis également redevable de m'avoir élevé dans un environnement exempt de stress lié au culte du soi-disant succès. Je me suis toujours senti davantage accompagné gentiment, plutôt qu'escorté pieds et mains liés, dans mon parcours de vie. Jamais ils ne m'ont laissé supposer que je n'avais pas satisfait leurs attentes, ni n'ont tenté de me décourager lorsque j'avais un projet douteux, comme celui d'entreprendre une carrière de lutteur professionnel. La confiance aveugle qu'ils ont

toujours eue en moi s'est transmise par osmose, elle m'habite depuis plusieurs années maintenant et m'a permis de compléter mes études graduées.

J'ai toujours été conscient de la chance que j'ai eu de les avoir comme parents. Même si je ne leur apprendrai rien lorsqu'ils liront cette dernière phrase, il me fait plaisir de la leur répéter pour une X^{ième} fois: " Je vous aime et je ne vous échangerais même pas contre un p'tit truck qui dompe."

J'aimerais également dédier cette thèse à M. Lemieux, mon professeur de Français en troisième secondaire au séminaire Saint-François, décédé il y a quelques années des suites d'un cancer du cerveau. Il aura su transmettre le plaisir de la lecture à un jeune homme de 14 ans avec des titres tels que: "Thérèse et Pierrette à l'école des Saint-Anges, Les thanatonautes, La ferme des animaux, etc." Il restera à jamais le premier enseignant digne de ce nom à réussir à me transmettre sa passion.

CHAPITRE I: INTRODUCTION

La fascination de l'être humain pour la genèse de la vie ne date pas d'hier. Déjà dans la Grèce antique, les grands penseurs de l'époque étaient tourmentés par des questionnements fondamentalement philosophiques tels que : "Qui sommes-nous et d'où venons-nous?" D'ailleurs, c'est Aristote (approximativement entre l'an 332 et 300 av. J-C) qui fut le premier à avoir l'éclair de génie d'interrompre successivement la croissance d'un nouvel œuf de poule à chacun des vingt et un jours d'incubation pour y observer la formation des organes principaux. Père de l'embryologie en quelque sorte, il a été témoin de l'extraordinaire processus au cours duquel une cellule unique, soit l'œuf fécondé ou zygote, deviendra un organisme multicellulaire complexe suite à une série de divisions mitotiques successives. Il faudra toutefois attendre plusieurs siècles avant de comprendre que le développement d'un tel organisme nécessite un contrôle hiérarchique de l'information génétique menant à l'expression différentielle selon un profil spatio-temporel précis des gènes sélecteurs. En tant que bons descendants spirituels d'Aristote, plusieurs chercheurs de partout sur la planète se donnèrent donc comme mandat d'élucider les différents mécanismes moléculaires derrière la régulation pointue de la cascade d'événements génétiques responsables de la spécification anatomique, grâce à l'utilisation de divers modèles d'étude ayant chacun leurs forces et leurs faiblesses. Tout ça pour qu'ultimement notre intérêt concernant les mécanismes à l'origine de la vie nous permette de mieux comprendre différentes pathologies et ainsi repousser l'heure de notre mort.

La simplicité d'utilisation de la drosophile (mouche du vinaigre ou *Drosophila melanogaster*) comme modèle d'étude, ainsi que le haut niveau d'homologie entre son bagage génétique et celui des mammifères, a permis de faire des pas de géant dans l'identification de gènes essentiels au développement embryonnaire. Parmi tous ces gènes indispensables, les gènes *Hox* (appelés *HOM* chez la drosophile) font probablement partie des plus connus. Cette famille de gènes hautement conservés dans tout la clade des métazoaires, code pour des facteurs de transcription jouant un rôle clef dans l'établissement de l'identité structurale le long de l'axe antéro-postérieur (AP), notamment au niveau du squelette axial, du squelette appendiculaire, de même que dans l'ontogénie de nombreux organes. L'expression des gènes *Hox* au cours du développement embryonnaire est

contrôlée de manière spatio-temporelle par la combinaison d'éléments régulateurs qui agissent en *cis* et en *trans*, permettant ainsi l'établissement du profil dynamique d'expression de chaque gène *Hox*. Bien que l'on sache actuellement que la régulation raffinée de l'expression des gènes *Hox* à l'intérieur de leur complexe respectif est orchestrée via le partage, la compétition et la sélectivité de séquences d'ADN définies pour les promoteurs de gènes *Hox* spécifiques, plusieurs incertitudes demeurent en ce qui concerne les subtilités liées à leurs mécanismes d'action. Par ailleurs, l'état actuel de la recherche sur les gènes *Hox* consiste à établir précisément leur positionnement au sein des cascades signalétiques les impliquant, tant durant l'embryogenèse qu'après la naissance.

La Dre Lucie Jeannotte s'inscrit dans la veine des chercheurs qui se penchent sur la compréhension des gènes *Hox* sous toutes leurs facettes. Le laboratoire s'intéresse particulièrement au gène *Hoxa5* qui, de part sa position centrale au sein de son complexe et de son rôle capital dans l'organogenèse, en fait un bon modèle dans l'étude des gènes *Hox* en général. Les projets entrepris dans le laboratoire en utilisant la souris comme modèle peuvent être regroupés à l'intérieur de trois principaux axes de recherche. Tout d'abord, il s'agit de bien comprendre la fonction du gène *Hoxa5* durant l'embryogenèse. Ensuite, nous tentons de clarifier son rôle postnatal. Ces deux questions sont adressées principalement en utilisant la souris mutante *Hoxa5*^{-/-} (Jeannotte *et al.*, 1991, 1993) grâce à laquelle la fonction du gène *Hoxa5* dans la formation du squelette axial (Jeannotte *et al.*, 1993; Aubin *et al.*, 1998), du squelette appendiculaire (Aubin *et al.*, 2002b; McIntyre *et al.*, 2007) ainsi que dans l'ontogénie de divers organes (Aubin *et al.*, 1997, 1999, 2002a; Meunier *et al.*, 2003; Garin *et al.*, 2006; Mandeville *et al.*, 2006; Gendronneau *et al.*, 2010; Gendronneau *et al.*, 2012; Boucherat *et al.*, 2012) a été mise en évidence. Enfin, le troisième axe de recherche concerne l'étude des mécanismes de régulation qui gouvernent l'expression spatio-temporelle du gène *Hoxa5*. Une approche de transgénèse chez la souris a permis d'identifier quatre grandes régions régulatrices (Tuggle *et al.*, 1990; Larochelle *et al.*, 1999; Nowling *et al.*, 1999; Moreau et Jeannotte, 2002; Tabariès *et al.*, 2005). Les résultats obtenus et discutés dans ma thèse font majoritairement partie de cet axe.

Voici donc le canevas du chapitre d'introduction présenté dans le cadre de ma thèse. Je débute par un survol des gènes homéotiques chez la drosophile (section 1.1). Ensuite, je

résume les différentes fonctions des gènes *Hox* chez les mammifères (section 1.2). Le corps de l'introduction consiste en une revue exhaustive des mécanismes de régulation contrôlant l'expression des gènes *Hox* (section 1.3). Après quoi j'aborde les connaissances concernant les fonctions et la régulation de l'expression du gène *Hoxa5* (section 1.4). Enfin, je décris les différents objectifs inhérents à mon projet de recherche au doctorat (section 1.5).

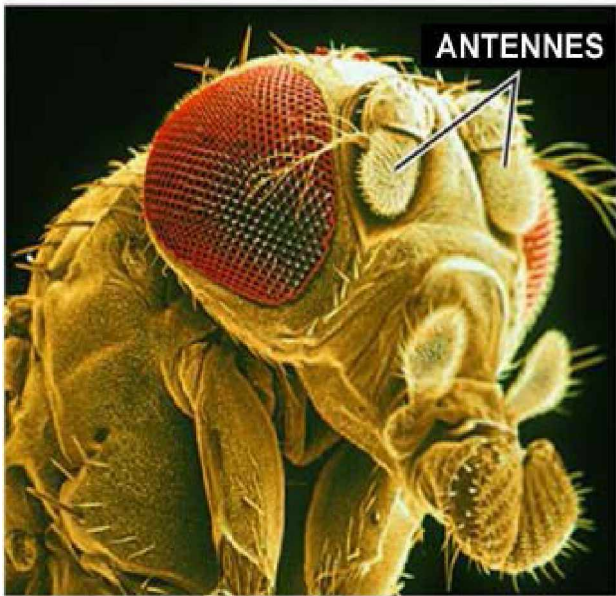
1.1 La drosophile

1.1.1 L'émergence de la génétique développementale

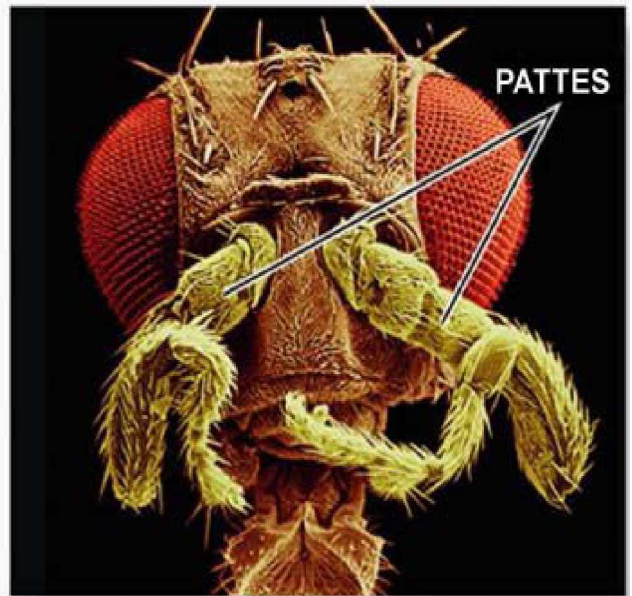
Cette histoire commence avec un zoologiste anglais nommé William Bateson qui défendait en 1884 une thèse à la fois rigoureuse et révolutionnaire sur les variations discontinues intitulée *Materials for the Study of Variation Treated with Especial Regards to Discontinuity in the Origin of Species*. Lorsqu'en 1900 il prend connaissance des travaux de Johann Gregor Mendel, il confirme les lois mendéliennes, présente ses résultats et les défend en Grande-Bretagne en 1902 dans *Mendel's Principles of Heredity*, où il observe leur application aux animaux ainsi qu'aux maladies humaines héréditaires. C'est en croisant des drosophiles qu'il en trouve une morte avec une paire de pattes en guise d'antennes (Figure 1.1). Il décide alors d'appeler "homéose" le phénomène de changement d'identité des segments qu'il décrit comme: "something is changed into the likeness of something else". Enfin, nous lui devons les termes : génétique, hétérozygote (HT), homozygote (HM) et épistasie.

Les années 1900 marquèrent le début de l'ascension en popularité des études menées sur la drosophile, où des généticiens comme Thomas Hunt Morgan qui cherchaient un organisme modèle peu dispendieux se reproduisant rapidement et s'entreposant facilement, s'en servirent pour écrire en 1919 *The Physical Basis of Heredity*. Aidé de Calvin Bridges, ils furent les premiers à décrire une mutation homéotique isolée qu'ils appelèrent *bithorax* (*Bx*) (Bridges et Morgan, 1923), où le troisième segment thoracique avait pris l'identité du deuxième. Enfin, il faut attendre les années 40 avant qu'Edward Bok Lewis commence à étudier le pseudo-allélisme en utilisant le complexe BX-C comme système modèle. Il découvrit plus d'une centaine de mutations homéotiques chez

Figure 1.1: (a) Drosophile de type sauvage où l'on observe une tête normale avec des antennes. (b) Drosophile mutante pour le gène *Antennapedia* (*Antp*) où la tête présente des pattes au lieu des antennes (tirée de Science VU/Dr. F. Rudolph Turner/Visuals Unlimited, Inc. Autorisation reçue de la part de(s) auteur(s).)



(a) Mouche normale



(b) Mouche mutante

drosophile (Lewis, 1978).

La drosophile comporte plusieurs segments céphaliques, trois segments thoraciques, huit segments abdominaux et dérive d'un embryon entièrement segmenté. L'embryogenèse comporte 17 stades répartis sur 22 heures (Hartenstein, 1993). De plus, chacun de ses segments possède son identité propre, c'est-à-dire que l'on retrouve des pattes sur le premier segment thoracique, alors que l'on observe à la fois des pattes et des ailes sur le second, et des pattes ainsi que des haltères sur le troisième.

1.1.2 Segmentation antéro-postérieure

L'établissement de la segmentation antéro-postérieure de l'embryon s'amorce au stade 14, durant lequel les noyaux du blastoderme deviennent complètement enfermés par la membrane plasmique. La synchronie mitotique est subséquentement perdue, la gastrulation débute et l'expression des gènes zygotiques (aussi appelés gènes de segmentation), contrôlés quant à eux par les gènes à effet maternel, est augmentée. Cette période critique du développement embryonnaire de la mouche s'appelle transition maternelle-zygotique (MZT) (Edgar et Datar, 1996). Grâce aux microtubules, l'ARNm maternel *Bicoid* (*Bd*) peut être transporté spécifiquement vers l'extrémité antérieure (Berleth *et al.*, 1988) de l'embryon, tandis que l'ARNm maternel *Oskar* (*Osk*) se retrouve à l'extrémité postérieure (Ephrussi *et al.*, 1991; Kim-Ha *et al.*, 1991). Les produits de ces deux gènes généreront à leur tour trois gradients d'expression de facteurs de transcription: *Hunchback* (Nusslein-Volhard et Wieschaus, 1980) qui est responsable des structures antérieures, ainsi que *Caudal* et *Nanos* qui s'occupent des structures postérieures (Macdonald et Struhl, 1986; Mlodzik et Gehring, 1987 ; Wang et Lehmann, 1991; Dahanukar et Wharton, 1996). Enfin, l'activation des gènes de segmentation par ces gènes à effet maternel induira le fractionnement de l'embryon en 14 parasegments constitués de la partie postérieure et antérieure de segments adjacents (Martinez-Arias et Lawrence, 1985).

1.1.2.1 Les gènes *gap*

Chaque membre de cette famille est exprimé selon un profil correspondant à une ou deux larges bandes chevauchant plusieurs segments futurs (Knipple *et al.*, 1985; Dalton *et al.*, 1989; Mohler *et al.*, 1989; Pignoni *et al.*, 1990; Capovilla *et al.*, 1992), et l'absence de segments contigus est observée chez les individus mutants pour ces gènes (Nusslein-Volhard et Wieschaus, 1980). Les gènes *Gap* antérieurs nécessaires au développement de la tête et du thorax sont: *Hunchback (Hb)*, *Ocelliless (Oc)*, *Empty spiracles (Ems)* et *Buttonhead (Btd)*. Le gène *Kruppel (Kr)* façonne les segments thoraciques, tandis que les gènes *Knirps (Kni)* et *Giant (Gt)* sont exprimés au niveau des futurs segments abdominaux. L'expression des gènes *Gap* est stabilisée et maintenue par répression mutuelle (Jäckle *et al.*, 1986; Harding et Levine, 1988). Enfin, la majorité des facteurs de transcription encodés par ces gènes sont des protéines à doigts de zinc.

1.1.2.2 Les gènes de parité segmentaire (*pair-rule*)

Situés en aval des gènes *Gap* dans la cascade développementale établissant l'axe AP de l'embryon de drosophile, ils sont le cœur du phénomène de segmentation et établissent la polarité à l'intérieur de chaque segment. Cette famille comprend sept membres soit *Hairy (H)*, *Even-skipped (Eve)*, *Runt*, *Fushi Tarazu (Ftz)*, *Odd-paired (Opa)*, *Odd-skipped (Odd)* ainsi que *Sloppy-paired (Slp)*, qui divisent l'embryon en quatorze parasegments où les cellules positives pour chacun des gènes de parité segmentaire sont séparées de leurs voisines par une rangée de cellules négatives. Au final, on obtient un embryon zébré le long de l'axe AP (Hafen *et al.*, 1984; Lawrence, 1992).

1.1.2.3 Les gènes de polarité segmentaire

Il existe huit membres connus de la famille des gènes de polarité segmentaire. Il s'agit des gènes *engrailed*, *hedgehog*, *wingless*, *patched*, *armadillo*, *cubitus interruptus*, *fused*, *gooseberry* et *pangolin*. Toujours au sein de la cascade développementale mettant en place l'axe AP de la drosophile, ils relayent le signal entre la fin du processus de segmentation AP, régi par les gènes de parité segmentaire, et le début du processus d'établissement de l'identité AP propre à chaque segment qui est contrôlé par les gènes

homéotiques sélecteurs. La mutation de ces gènes cause généralement une délétion de portions de chacun des segments et génère par la même occasion l'image miroir des segments résiduels qui coïncident, ou alors des duplications en tandem de ces régions (Nusslein-Volhard et Wieschaus, 1980; Perrimon et Mahowald, 1987; Martinez Arias *et al.*, 1988; Perrimon et Smouse, 1989).

1.1.3 Identité antéro-postérieure

1.1.3.1 Les gènes homéotiques sélecteurs

Contrairement aux gènes de segmentation décrits à la section précédente, les gènes homéotiques ne sont pas essentiels à la formation des différents segments corporels, mais le sont plutôt pour assigner à ceux-ci une identité unique. (Lewis, 1978). Le fait qu'une mutation dans un gène homéotique n'affecte pas le nombre de segments, mais bien son identité, en est la preuve. Chez la drosophile, les disques imaginaux, structures embryonnaires se développant de manière semblable aux somites chez les mammifères, donneront naissance aux antennes, aux yeux, aux haltères, aux ailes ainsi qu'aux pattes. La notion de «transformation homéotique» réfère donc au changement de la destinée normale des cellules d'un disque imaginal qui survient chez l'embryon. Les gènes homéotiques de la mouche, répondant également à l'appellation gènes *HOM*, sont regroupés à l'intérieur de deux complexes localisés tous les deux sur le bras droit du troisième chromosome (Lewis, 1978; Kaufman *et al.*, 1980). Il s'agit des complexes *Bithorax* (*BX-C*) et *Antennapedia* (*ANT-C*). Le complexe *Bithorax* englobe trois gènes homéotiques soit *Ultrabithorax* (*Ubx*), *Abdominal-A* (*Abd-A*) et *Abdominal-B* (*Abd-B*) (Lewis, 1978; Bender *et al.*, 1983; Karch *et al.*, 1985; Sanchez-Herrero *et al.*, 1985; Tiong *et al.*, 1985). Ces trois gènes contrôlent l'identité des parasegments 5-13, soit respectivement le troisième segment thoracique dans le cas du gène *Ubx*, ainsi que les segments abdominaux en ce qui concerne les gènes *Abd-A* et *Abd-B*. Les loci homéotiques regroupés dans le complexe *Antennapedia* correspondent aux gènes *Labial* (*Lab*), *Proboscipedia* (*Pb*), *Deformed* (*Dfd*), *Sex combs reduced* (*Scr*) and *Antennapedia* (*Antp*) (Lewis *et al.*, 1980a, 1980b; Wakimoto et Kaufman, 1981; Merrill *et al.*, 1987). Quant à ces cinq gènes, ils modulent l'identité régionale des structures antérieures au cinquième parasegment, c'est-à-dire que les gènes *Lab*, *Pb* et *Dfd* s'occupent

des segments céphaliques, tandis que les gènes *Scr* et *Antp* se chargent des deux premiers segments thoraciques.

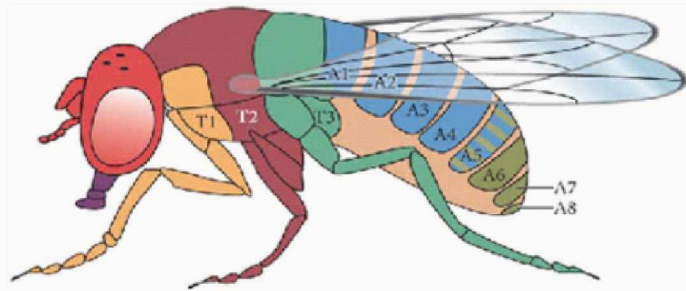
De plus, la technique d'hybridation *in situ* de l'ARNm a permis de détecter les transcrits propres à chacun des différents gènes *HOM* à l'intérieur de leur site d'expression embryonnaire spécifique, tels que représentés à la Figure 1.2 (Harding *et al.*, 1985; Akam, 1987).

Enfin, les travaux d'Edward Bok Lewis auront permis de mettre en évidence deux fascinants concepts qui seront abordés davantage ultérieurement. Tout d'abord, il révéla la corrélation directe entre l'ordre organisationnel des gènes *HOM* au sein du chromosome 3 et leur profil d'expression le long de l'axe AP, communément appelé principe de colinéarité (section 1.2.3). Sa deuxième observation fut que la perte du complexe *BX-C* transforme en deuxième segment thoracique tous les segments postérieurs à ce dernier. De plus, comme les mutations affectant la fonction spécifique d'un segment causaient des transformations homéotiques au niveau du segment non affecté le plus antérieur, il conclut que le devenir de la structure dépendait du signal postérieur. Ces observations initiales ouvrirent la porte au concept de prévalence postérieure (section 1.3.3). Fait intéressant, la mutation du gène *Scr* engendre aussi bien une antériorisation du deuxième segment thoracique en premier segment thoracique, qu'une postériorisation des structures labiales en structures maxillaires (Wakimoto et Kaufman, 1981; Struhl, 1983; Pattatucci *et al.*, 1991; Pederson *et al.*, 1996).

1.1.3.2 La structure des protéines codées par les gènes *Hom*

Tous les gènes *Hom* possèdent une séquence de 180 pb hautement conservée à 90 % d'homologie entre tous les animaux bilatéraux, appelée boîte homéo, codant pour un domaine protéique de 60 acides aminés (aa), appelé domaine homéo, dont la conformation de gènes cibles. En effet, c'est grâce à l'organisation des trois hélices α du domaine homéo en motif de type hélice-boucle-hélice (HLH) (Desplan *et al.*, 1985; Hoey et Levine, 1988), que les protéines *HOM* sont en mesure de reconnaître spécifiquement la séquence nucléotidique minimale constituée des quatre nucléotides ATTA, et ainsi jouer son rôle de facteur de transcription en activant ou réprimant les gènes cibles appelés gènes effecteurs

Figure 1.2: (Centre) Représentation schématique des deux complexes de gènes *Hom* (*Hom-C*), soit les complexes *Antennapedia* (*ANT-C*) et *Bithorax* (*BX-C*). (Bas) Les différents profils d'expression des gènes *Hom* au niveau du blastoderme embryonnaire. Les métamères sont numérotés de 1 à 14 et ceux qui sont plus foncés coïncident avec les sites où l'expression est la plus forte. (Haut) Les différents profils d'expression des gènes *Hom* selon l'axe antéro-postérieur (AP) chez la drosophile adulte. (tirée de Carstens, 2004. Autorisation reçue de la part de(s) auteur(s).))



Complexe Antennapedia

Complexe Bithorax

Lab Pb Dfd Scr Antp

Ubx AbdA AbdB



(Affolter *et al.*, 1990). La boîte homéo a été mise en évidence en comparant la séquence des gènes *Antp*, *Ubx* et *Ftz* (McGinnis *et al.*, 1984; Scott et Weiner, 1984). De plus, le pourcentage élevé d'homologie entre les domaines homéo des différentes protéines HOM se traduit par une redondance fonctionnelle *in vitro* de liaison à l'ADN (Desplan *et al.*, 1988; Hoey et Levine, 1988; Kalionis et O'Farrell, 1993; Ekker *et al.*, 1994). Toutefois, ces même protéines dictent *in vivo* des identités régionales uniques aux futures structures, ce qui soulève la question: " Comment les protéines HOM reconnaissent et régulent-elles de façon spécifique différents gènes cibles (Hayashi et Scott, 1990; Biggin et McGinnis, 1997) ? Deux modèles non mutuellement exclusifs se partagent l'explication, soit le modèle de liaison sélective et le modèle d'activité régulée.

1.1.3.3 Le modèle de liaison sélective

Ce modèle stipule que la spécificité de liaison des protéines à domaine homéo pour leurs gènes cibles dépend de la nature de ses cofacteurs. L'une des expériences clés qui a servi à défendre ce modèle fut réalisée sur le gène de parité segmentaire *Ftz*. Ce gène, en plus d'être localisé dans le complexe *ANT-C*, possède une boîte homéo hautement conservée très similaire à celle des gènes *Hom*. De façon intéressante, la délétion du domaine homéo de la protéine FTZ n'affecte aucunement son activité (Fitzpatrick *et al.*, 1992; Copeland *et al.*, 1996; Hyduk et Percival-Smith, 1996). Par ailleurs, la seule interaction de son cofacteur α FTZ-F1, capable de reconnaître un motif sur l'extrémité N-terminale de la protéine FTZ, est suffisante pour son recrutement aux promoteurs de ses gènes cibles (Guichet *et al.*, 1997). Un autre exemple implique le cofacteur Extradenticle (EXD) qui est en mesure de modifier la structure de l'homéoprotéine SCR afin que 2 chaînes latérales d'aa basiques (impliquant Arg3 et His12) normalement inutilisées puissent participer à la liaison à l'ADN (Joshi *et al.*, 2010). Enfin, il s'agit du modèle le plus accepté et davantage d'exemples le renforçant seront décrits chez les mammifères (section 1.2.5).

1.1.3.4 Le modèle d'activité régulée

Ce modèle se base principalement sur les observations relatives aux expériences de pontages covalents ADN-protéines suite à une exposition aux rayons UV (Biggin et

McGinnis, 1997; Liang et Biggin, 1998; Carr et Biggin, 1999). Effectivement, l'immunoprécipitation des complexes ADN-protéines à domaine homéo aurait révélé une large distribution de ces protéines sur les différents fragments testés. Donc, contrairement au modèle de liaison sélective où le complexe protéique se forme avant que la protéine à boîte homéo ne se lie à la séquence de son gène cible, le modèle d'activité régulée soutient que la spécificité de liaison des homéoprotéines est dépendante des protéines qui, suite au contact entre l'homéoprotéine et son gène cible, viennent solidifier l'interaction via les sites de liaison avoisinant. Par exemple, la protéine HOM Deformed (DFD) est en mesure de reconnaître *in vivo* le site de liaison à l'ADN de ses gènes cibles, et ce même en l'absence de son cofacteur Extradenticle (EXD), une protéine à domaine homéo de la famille TALE. Toutefois, l'état activateur du facteur de transcription DFD ne serait enclenché que si la protéine EXD se lie à proximité (Li *et al.*, 1999). Enfin, les détracteurs de ce modèle le qualifient de réductionniste, puisqu'ils avancent que le nombre de molécules accessibles à l'intérieur d'une cellule pour conférer une telle spécificité d'action est beaucoup trop faible comparativement au nombre de sites de liaison pour les protéines HOM (Krause et al., 1988; Nasiadka et al., 2000).

1.2 Les mammifères

1.2.1 Conservation des gènes homéotiques dans l'évolution

Les différentes analyses phylogéniques comparatives des séquences d'ADN des gènes homéotiques, appelés gènes *Hox* chez les vertébrés, ont démontré un très haut niveau de conservation durant l'évolution (Carrasco *et al.*, 1984; Duboule, 2007; DeRobertis, 2008; Gehring *et al.*, 2009). En effet, un seul complexe ancestral de 6 gènes (*Hom-C*), organisé à partir de rondes successives de duplication du gène homéotique originel et retrouvé chez les insectes primitifs, serait à l'origine de tous les gènes *Hox* (Valentine *et al.*, 1996). Puis, suite à la complexification des organismes, notamment au niveau axial avec l'apparition des chordés et des vertébrés, il y aurait eu deux rondes successives de duplication des complexes en entier (Ruddle et al., 1994; Nam et Nei, 2005). Afin d'acquérir les diversités morphologiques presque infinies retrouvées chez certaines espèces

de poissons, une dernière vague de duplication a fait passer le nombre de complexes *Hox* de quatre à huit lors de l'apparition des téléostéens ou poissons à nageoires rayonnées (Crow *et al.*, 2006). Enfin, comme toutes ces duplications ont généré des gènes ayant la même fonction et qu'ultimement les cellules tendent à ne pas dépenser de l'énergie inutilement (répliquer, transcrire et traduire des gènes à redondance complète dans ce cas-ci), il y a eu la perte de certains gènes *Hox* au sein de leur complexe (Hoegg et Meyer, 2005). Il existe aussi un autre modèle prônant que les espèces plus jeunes d'un point de vue évolutif, comme les mammifères par exemple, auraient également eu un ancêtre avec huit complexes *Hox*, potentiellement entre l'ancêtre des téléostes et les polyptères, et qu'il y aurait eu une ou plusieurs rondes de consolidation faisant passer à quatre le nombre de complexes (Duboule, 2007). Dans ce modèle, la complexité des organismes ne découle pas de la duplication des complexes *Hox*, mais plutôt de l'apparition d'éléments de régulation globaux à effets longue distance.

Grâce aux analyses d'homologies de séquence, plusieurs classes différentes de gènes *Hox* ont été identifiées. Parmi celles-ci, deux classes principales se distinguent : les gènes *Hox* de type *Paired (Prd)* et ceux de type *Antennapedia (Antp)*. Cette dernière classe peut également être subdivisée en trois, soit les gènes *Hox*, le complexe ParaHox et le complexe NK (Garcia-Fernandez, 2005). Quant aux classes moins classiques, il s'agit des gènes *Hox* de type TALE, POU, LIM et SINE (Holland et Takahashi, 2005). Comme son nom l'indique, la classe TALE pour " Three Amino acid Loop Extension" se distingue par son domaine homéo qui compte 63 aa comparativement aux 60 retrouvés normalement.

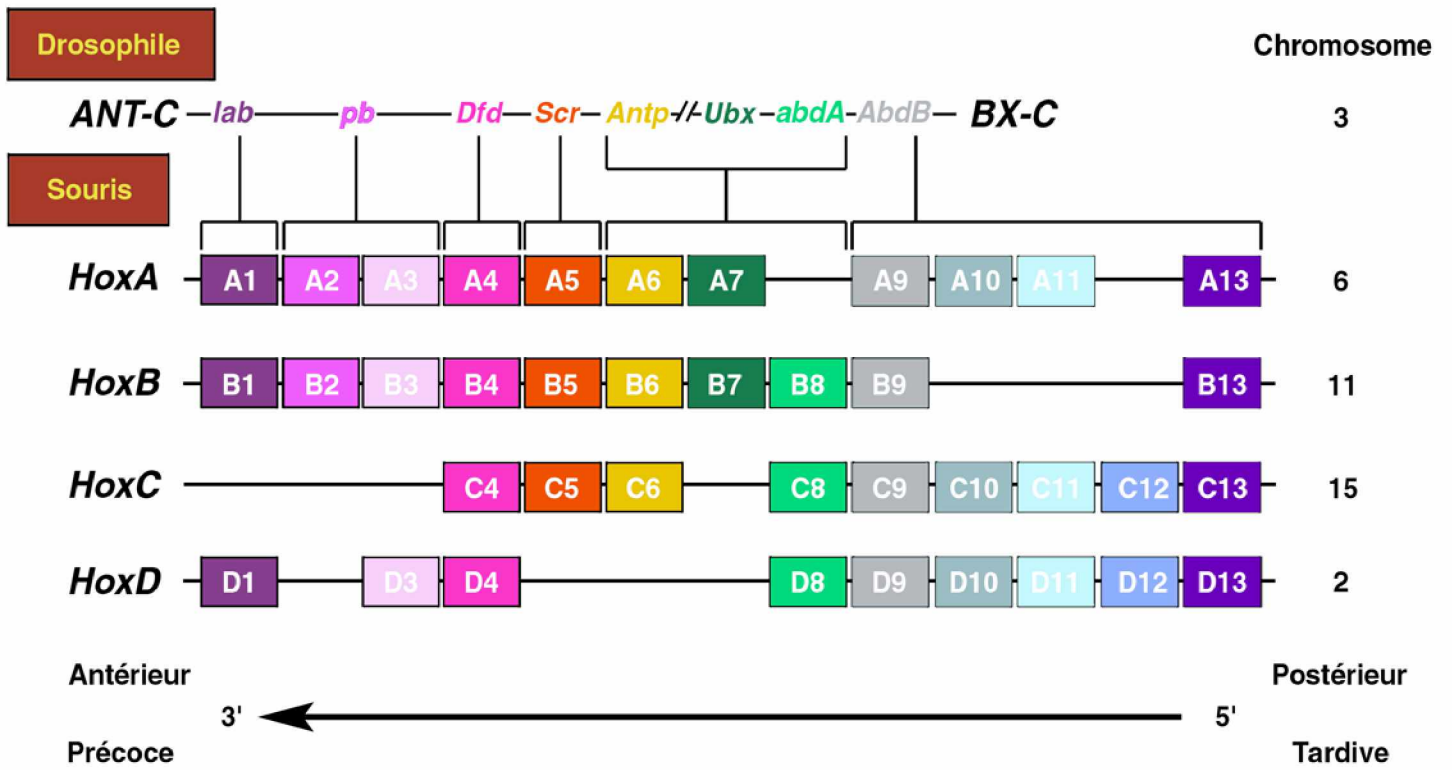
L'importante conservation des gènes homéotiques entre les espèces ne s'arrête pas uniquement à leur structure protéique ainsi qu'à leur organisation génomique. En effet, la découverte de la séquence correspondant à la boîte homéo chez la drosophile a ouvert la porte à l'étude des gènes homéotiques chez d'autres organismes modèles, puisque l'utilisation de cette séquence comme sonde a dévoilé l'existence de gènes *Hox* chez tous les métazoaires étudiés jusqu'à aujourd'hui. Tout cela, additionné aux progrès technologiques en biologie cellulaire et moléculaire, ont permis de souligner également la très grande conservation au niveau de leur fonction (sections 1.2.6-1.2.8) et de la régulation de leur expression (section 1.3).

1.2.2 Organisation des gènes *Hox*

Trente-neuf gènes *Hox* de type *Antp* sont retrouvés chez les mammifères. Ils sont organisés en quatre complexes (identifiés *HoxA* à *D*), lesquels sont situés sur quatre chromosomes différents. Il s'agit des chromosomes 6, 11, 15 et 2 chez la souris, comparativement aux chromosomes 7, 17, 12 et 2 chez l'humain (Figure 1.3). Chacun des gènes *Hox* peut également être classé dans un des treize groupes paralogues. Ces groupes de gènes paralogues ont été définis en fonction du degré d'homologie suite à l'analyse comparative de leur séquence protéique, de même qu'en fonction des propriétés potentiellement équivalentes de leur protéine (Greer *et al.*, 2000). Il a aussi été proposé qu'un complexe *Hox* agisse comme une unité fonctionnelle, appelée métagène, où aucun des gènes *Hox* ne peut exercer la fonction du complexe à lui seul (Duboule, 2007; Mainguy *et al.*, 2007). On compte entre neuf et onze gènes *Hox* par complexe (McGinnis et Krumlauf, 1992) où chacun d'entre eux possède au moins un paralogue dans un des autres complexes ainsi qu'un équivalent, appelé orthologue, dans le complexe HOM-C de drosophile. Au jour embryonnaire (E) 7.2, l'initiation de l'expression des gènes *Hox1* et *Hox2* se limite à la partie postérieure de la ligne primitive (PS), une région destinée à devenir le mésoderme extra-embryonnaire. Cette expression précoce dans la partie postérieure de l'épiblaste ne concerne pas directement les structures embryonnaires axiales et paraxiales, puisque celles-ci dérivent à la fois des cellules plus antérieures de la PS, de même que certaines cellules avoisinantes (Lawson *et al.*, 1991). Toutefois, l'initiation séquentielle de la transcription des gènes *Hox* déterminera le moment précis où les domaines d'expression atteindront successivement la partie antérieure de la PS, communément appelée la région du nœud, à partir de laquelle il y aura extension des axes embryonnaires (Forlani *et al.*, 2003; Wacker *et al.*, 2004).

Leur domaine s'étend aussi au tube neural ainsi qu'au mésoderme paraxial jusqu'à ce qu'une frontière rostrale prédéterminée par un gradient d'acide rétinoïque (voir section

Figure 1.3: (Haut) Représentation schématique des deux complexes *ANT-C* et *BX-C* de gènes *HOM* chez la drosophile, de même que le chromosome sur lequel ils sont situés. (Bas) Représentation schématique des quatre complexes de gènes *Hox* chez la souris (*HoxA* à *HoxD*), localisés respectivement sur les chromosomes 6, 11, 15 et 2. Les différents gènes *Hox* sont alignés verticalement en fonction de leur orthologue chez la drosophile, ainsi que leur groupe paralogue. La flèche noire symbolise le sens de la transcription de même que la relation de colinéarité entre l'organisation des gènes *Hox* au sein de leur complexe et leur expression spatio-temporelle. (tirée de Aubin et Jeannotte, 2001. Autorisation reçue de la part de(s) auteur(s).))



1.3.4.1) soit atteinte (Deschamps et Wijgerde, 1993; Roelen *et al.*, 2002; Oosterveen *et al.*, 2003a). Par conséquent, la précision temporelle de l'activation des gènes *Hox* lors des étapes initiales de l'établissement de leur domaine d'expression est cruciale pour dicter l'identité régionale du profil corporel.

1.2.3 Principe de colinéarité

Tel qu'observée initialement pour les gènes *HOM* chez la drosophile (Lewis, 1978), la conservation de l'organisation des gènes *Hox* à l'intérieur de leur complexe reflète la conservation de leur expression spatio-temporelle durant l'embryogenèse (Krumlauf, 1994). En d'autres termes, les gènes *Hox* situés plus en 3' des complexes ont une expression plus précoce et dans des domaines d'expression plus antérieurs que ceux situés en 5'. C'est ce phénomène qu'on appelle la colinéarité et qui génère des domaines d'expression restreints se chevauchant. Tout ça pour établir ultimement une adresse génétique, déterminée par le ratio et la nature des protéines Hox présentes dans une région donnée, qui modulera les caractéristiques spécifiques à cette même région (Kessel et Gruss, 1991). Cette adresse, aussi connue sous l'appellation "code *Hox*" serait potentiellement responsable de la spécification des différentes structures embryonnaires le long des axes (Duboule, 1998). Cependant, le programme précoce d'expression des gènes *Hox*, transmis à l'intérieur de la région du nœud, ne correspond pas au code *Hox* définitif transmis aux cellules du mésoderme présomitique et du mésoderme de la plaque latérale (Eloy-Trinquet et Nicolas, 2002; Forlani *et al.*, 2003). Selon le principe de colinéarité, il existe un lien intime entre l'arrangement chromosomique des gènes *Hox* et leur régulation spatio-temporelle. Comme les cellules du neurectoderme possèdent le même génome que celles du mésoderme présomitique et que toutes doivent interpréter le même signal temporel en coordonnées spatiales différentes, leur devenir structurel respectif doit être influencé par une régulation différentielle plus tardivement. Ces différents modes de régulation seront abordés à la section 1.3.

1.2.4 Spécificité et modifications post-traductionnelles

L'une des stratégies qui semblent être utilisées par la cellule pour moduler la spécificité d'action *in vivo* des protéines HOX, pouvant s'appliquer aussi bien au modèle de liaison sélective (section 1.1.3.3) qu'au modèle d'activité régulée (section 1.1.3.4), consiste à modifier la structure de ces protéines. La phosphorylation représente la modification post-traductionnelle touchant les protéines Hox la mieux documentée. En effet, des études biochimiques ont démontré que les protéines HOXA5 et HOXB5 pouvaient être phosphorylées (Odenwald *et al.*, 1989 ; Wall *et al.*, 1992). De façon intéressante, une étude *in vivo* a également montré le rôle essentiel de deux motifs phosphorylables, conservés chez les protéines HOX du groupe paralogue 10, pour remplir leur rôle répresseur de la formation des côtes au niveau des vertèbres lombaires (Guerreiro *et al.*, 2012). La modulation de la spécificité de liaison des protéines HOX contrôlée par leur phosphorylation est un mécanisme conservé durant l'évolution, puisqu'on le retrouve également chez la mouche, notamment pour la protéine UBX (Taghli-Lamalle *et al.*, 2008). De plus, la phosphorylation des résidus sérine/thréonine en position C-ter de la protéine UBX serait effectuée par la Caséine Kinase 2 (CK2) afin d'empêcher la répression du gène *Distal-less*.

Quant aux autres modifications post-traductionnelles, très peu de choses sont connues jusqu'à maintenant. Dans un modèle de cellules cancéreuses du sein, l'inhibition de l'acétylation engendre une augmentation de la transactivation contrôlée par la protéine HOXB7. Toutefois, l'étude ne permet pas de discriminer entre un contrôle direct de sa fonction protéique, indirect via l'acétylation d'un cofacteur protéique ou au niveau de la chromatine même modifiant ainsi l'accessibilité au site de liaison (Chariot *et al.*, 1999).

1.2.5 Spécificité et cofacteurs de liaison

Les facteurs de transcription à domaine homéo de la classe TALE, mentionnés à la section 1.2.1, sont probablement les cofacteurs de liaison associés aux protéines HOX les plus connus. Ils sont en grande partie à l'origine du modèle de liaison sélective, abordé à la section 1.1.3.3, et peuvent être classés en deux familles différentes : PBC et MEINOX. Cette dernière famille peut également être subdivisée en deux sous-familles soit les

protéines MEIS et PREP (Fognani *et al.*, 2002; Mukherjee et Burglin, 2007). Chez les mammifères, la famille PBC compte quatre gènes *Pbx*, alors que les sous-familles *Meis* et *Prep* regroupent respectivement trois et deux gènes.

PBX est un acronyme pour "Pre-B cell leukemia homeobox" étant donné que la protéine PBX1, isolée en protéine fusionnée au domaine d'activation transcriptionnelle de la protéine E2A, a été identifiée dans un bon nombre de leucémies lymphocytiques Pré-B aigües (Kamps *et al.*, 1990; Nourse *et al.*, 1990). De plus, le pouvoir oncogénique associé à la protéine de fusion E2A/PBX1 dépendrait de ses interactions avec différentes protéines HOX (Slupsky *et al.*, 2001). Le rôle des protéines PBX en tant que cofacteurs liés à la spécificité d'action des protéines HOX a également été souligné grâce aux souris *Pbx1^{-/-}*, mourant à E15.5 d'hypoplasie/aplasie de plusieurs organes, où l'on observe une transformation homéotique du deuxième arc branchial en premier arc branchial (Selleri *et al.*, 2001). D'ailleurs, c'est l'interaction entre le domaine homéo des protéines PBX et le motif YPWM (penta/hexapeptide), conservé au sein des gènes *Hox* des groupes paralogues 1-8 (Di Rocco *et al.*, 2001; LaRonde-LeBlanc et Wolberger, 2003; Moens et Selleri, 2006; Merabet *et al.*, 2009), qui fait office de modulateur de la spécificité de liaison des protéines HOX à l'ADN (Lufkin, 1996; Shen *et al.*, 1996). De la même façon, les protéines MEIS peuvent interagir spécifiquement avec les protéines HOX des groupes paralogues 9-13 (Shen *et al.*, 1997a).

MEIS est un acronyme pour "Myeloid Ecotropic viral Integration Site" en raison de sa découverte par mutagenèse d'insertion (Moskow *et al.*, 1995; Steelman *et al.*, 1997). Les souris déficientes pour le gène *Meis1* ont un phénotype de létalité embryonnaire à E14.5 ainsi qu'une déficience en cellules souches hématopoïétiques dites "définitives" (Azcoitia *et al.*, 2005; Carramolino *et al.*, 2010). De la même façon, la surexpression du gène *Meis1* joue un rôle dans la transformation leucémique chez l'humain et la souris, ainsi que dans la genèse tumorale en général (Wong *et al.*, 2007; Crijins *et al.*, 2007; Grubach *et al.*, 2008). Enfin, PREP est un acronyme pour "PBX Regulatory Protein" (Berthelsen *et al.*, 1998a). Le phénotype des embryons *Prep1^{-/-}* est très sévère et s'accompagne d'une mort utérine avant E7.5 (Fernandez *et al.*, 2010), alors qu'un taux de mortalité embryonnaire élevé à E17.5, associé à des cellules souches hématopoïétiques défectueuses, est observé dans les

mutants hypomorphes *Prep1ⁱⁱ* (Ferreti *et al.*, 2006; Di Rosa *et al.*, 2007). Quant aux souris *Prep1ⁱⁱ* qui échappent à la mortalité embryonnaire, elles développent des tumeurs avec une forte incidence (Longobardi *et al.*, 2010). Toutefois, bien que les mutations *Meis1^{-/-}* et *Prep1ⁱⁱ* affectent les mêmes organes, il ne semble pas y avoir d'interaction génétique entre les deux puisque les souris doubles hétérozygotes sont d'apparence normale (Penkov *et al.*, 2013). Les protéines MEIS et PREP sont en mesure d'interagir avec les protéines PBX (Wagner *et al.*, 2001; Diaz *et al.*, 2007; Oriente *et al.*, 2008). Ainsi, un complexe tertiaire HOX/PBX/MEIS ou HOX/PBX/PREP permet au domaine homéo de la protéine HOX de reconnaître spécifiquement la séquence TNAT(G/T)(G/A) du gène cible (Berthelsen *et al.*, 1998; Shen *et al.*, 1999).

Un autre exemple d'interaction protéines HOX/CDX a également été répertorié (Lafontaine *et al.*, 2012). Bien que cet hétérodimère protéique soit en mesure de lier des éléments de réponse aux gènes *Hox*, les effets fonctionnels dans la modulation de la spécificité ainsi que dans l'établissement du profil antéro-postérieur demeurent inconnus.

1.2.6 Établissement du squelette axial

Le squelette axial (colonne vertébrale) dérive des somites, structures embryonnaires mésodermiques qui se forment entre E7.75-E8.0 et qui sont organisées symétriquement des deux côtés le long de l'axe antéro-postérieur (Dubrulle et Pourquié, 2004). Bien que chacun des somites soit similaire d'un point de vue morphologique, ils sont à l'origine d'éléments squelettiques (*vertebrae*) aux caractéristiques morphologiques uniques, selon leur position le long de l'axe AP (Brent et Tabin, 2002). Les somites se différencient en trois parties aux destinées différentes: le dermomyotome, le sclérotome et syndétome (Christ *et al.*, 2007). Ces parties donneront respectivement naissance au derme ainsi qu'aux muscles, au squelette axial et aux tendons interconnectant les muscles aux tissus squelettiques. Sous l'action de plusieurs gènes dont celle des gènes *Hox*, le sclérotome des quatre et demie premières somites généreront le crâne occipital (Goodrich, 1930), tandis que celui des somites restantes subira une nouvelle vague de différenciation et de segmentation, durant laquelle la moitié postérieure d'une somite fusionne à la moitié antérieure de sa voisine la

plus caudale, pour donner chez la souris sept vertèbres cervicales (C), treize thoraciques (T), six lombaires (L), quatre sacrales (S) et une trentaine de caudales (Ca).

Plusieurs études, réalisées principalement chez la souris, ont révélé la complexité de l'activité des gènes *Hox* tout au long de l'axe AP. Leur action combinée mène, entre autres, à l'établissement du squelette axial organisé, où chacune des vertèbres se voit octroyer des propriétés anatomiques propres (Krumlauf, 1994; Wellik, 2007; Malo *et al.*, 2010). En effet, les transformations homéotiques observées chez les souris mutantes avec perte de fonction de gènes *Hox* consistent principalement en des transformations d'antériorisation, alors que celles générées lors d'un gain de fonction de gènes *Hox* sont plutôt de type postériorisation (Kessel *et al.*, 1990; Jegalian *et al.*, 1992; Charité *et al.*, 1994; Pollock *et al.*, 1995). Suite à ces études, il est devenu évident que plusieurs gènes *Hox* adjacents agissent de concert lors de la spécification des vertèbres (Malo *et al.*, 2009; Wellik, 2009). De plus, la génération des mutants composés *Hox*, incluant plusieurs membres du même groupe paralogue, résulte presque toujours en un phénotype aggravé (Horan *et al.*, 1995; Wellik et Capecchi, 2003; McIntyre *et al.*, 2007), ce qui confirme par la même occasion l'idée suggérée dans le passé qu'il y aurait redondance fonctionnelle partielle entre les membres d'un même groupe paralogue (Gaunt *et al.*, 1989; Kessel et Gruss, 1991; McGinnis et Krumlauf, 1992; Burke *et al.*, 1995). Enfin, grâce aux différentes mutations générées chez la souris, les gènes *Hox* ont pu être classés en trois groupes selon leurs phénotypes au niveau du squelette axial. Il s'agit des gènes *Hox* spécifiant la tête (paralogues 1 à 4), le tronc (paralogues 5 à 8) et la partie caudale (paralogues 9 à 13). Un schéma récapitulant les domaines d'expression des principaux gènes *Hox* impliqués dans l'élaboration du squelette axial est présenté à la Figure 1.4.

1.2.6.1 Les gènes *Hox* spécifiant la tête (paralogues 1 à 4)

Tout d'abord, on observe des défauts directement au niveau du crâne chez les souris mutantes pour chacun des gènes *Hox* membres des groupes paralogues 1 et 2 (Lufkin *et al.*, 1991; Gendron-Maguire *et al.*, 1993; Rijli *et al.*, 1993; Davenne *et al.*, 1999; Trainor et Krumlauf, 2000). Ces altérations de la boîte crânienne sont également décrites pour les mutants *Hoxa3* (Chisaka *et al.*, 1992). Quant aux vertèbres cervicales, il y a transformation

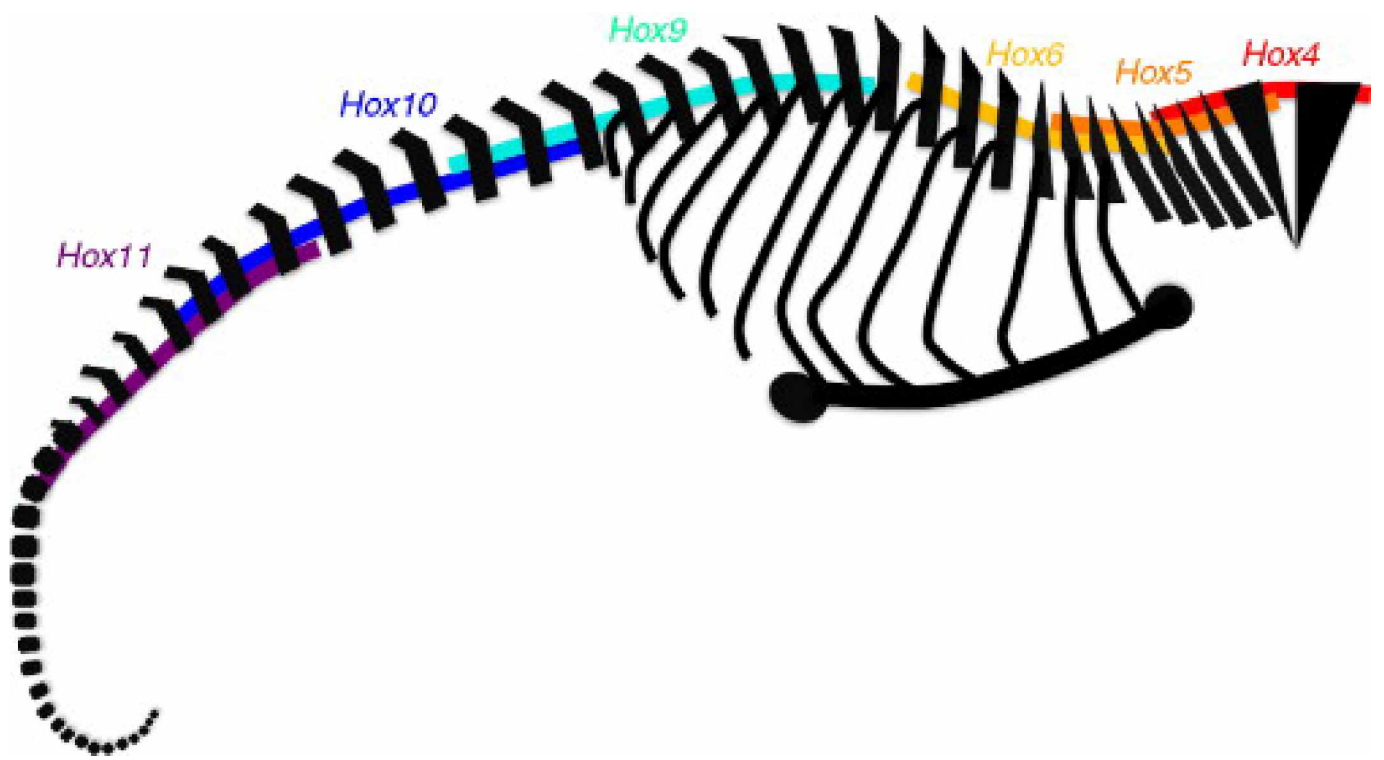
des C1 et C2 chez les souris mutantes pour les gènes *Hoxb3* de même que *Hoxd3* (Chisaka et Capecchi, 1991; Condie et Capecchi, 1993; Manley et Capecchi, 1997). Il faut souligner également que les souris mutantes pour le gène *Hoxd3* meurent de dislocation cervicale. D'ailleurs, les gènes *Hoxa3* et *Hoxd3* sont responsables de la formation de l'Atlas (C1). De la même façon, chacune des souris mutantes pour un membre du groupe paralogue 4 a subi une transformation d'antériorisation de C3 en C2 (Horan *et al.*, 1994; Kostic et Capecchi, 1994). Le rôle capital des gènes membres du paralogue 4 en ce qui a trait à l'établissement de la région cervicale a également été démontré chez la souris triple mutante *Hoxa4/Hoxb4/Hoxd4*, où les C2 à C5 adoptent une morphologie propre à l'Atlas (C1), alors que C6 et C7 ressemblent davantage à l'Axe (C2) (Horan *et al.*, 1995).

1.2.6.2 Les gènes *Hox* spécifiant le tronc (paralogues 5-8)

Comme le portrait du phénotype squelettique axial engendré par la perte de fonction du gène *Hoxa5* sera dépeint en profondeur ultérieurement (section 1.4.1), il ne sera pas abordé dans cette section. Toutefois, une petite parenthèse historique pour souligner que le mutant *Hoxa5* est le premier mutant *Hox* où l'on a observé simultanément des transformations d'antériorisation et de postériorisation (Jeannotte *et al.*, 1993). En ce qui concerne les souris mutantes *Hoxb5*, les vertèbres C6 à T1 acquièrent une identité plus antérieure, alors que le point d'attachement de la ceinture scapulaire est déplacé antérieurement. Puisque l'interférence transcriptionnelle de la mutation *Hoxc4* sur *Hoxc5* rend l'interprétation du phénotype squelettique de ce dernier ambiguë, les résultats ne seront pas décrits dans cette section (Boulet et Capecchi, 1996). D'ailleurs, les phénotypes potentiellement inhérents à la stratégie utilisée pour l'obtention de la perte de fonction d'un gène *Hox* seront volontairement omis et feront l'objet d'un paragraphe au chapitre régulation (section 1.3.14) Tout comme pour les gènes *Hox* du groupe paralogue 4, ceux du groupe 5 exercent également des fonctions essentielles dans le modelage de la région cervicale. Effectivement, les vertèbres C3 à T1 des souris triples mutantes *Hoxa5/Hoxb5/Hoxc5* ont l'apparence de l'Axe (C2) (McIntyre *et al.*, 2007).

Les souris triples mutantes *Hoxa6/Hoxb6/Hoxc6* ont également été générées et elles présentent des transformations homéotiques au niveau des C7 à T6, parmi lesquelles les

Figure 1.4: Schéma d'une vue latérale d'un squelette de souris où l'extrémité antérieure est à droite. Les vertèbres du squelette axial, de même que les autres éléments squelettiques, sont représentés en noir, alors que les lignes de couleur représentent les profils d'expression colinéaires des gènes *Hox* obtenus suite à des expériences de perte de fonction des différents paralogues (tirée de Wellik *et al.*, 2009. Autorisation reçue de la part de(s) auteur(s)).



trois premières vertèbres thoraciques ont acquis une identité cervicale (McIntyre *et al.*, 2007). De plus, des analyses comparatives s'intéressant à la transition cervicothoracique suggèrent que les gènes *Hox* du groupe paralogue 6 fixent le début de la région thoracique sur le squelette axial (Gaunt, 1994; Burke *et al.*, 1995). La souris transgénique exprimant ectopiquement *Hoxb6* vient supporter cette hypothèse puisqu'elle induit la formation de côtes ectopiques (Vinagre *et al.*, 2010).

Enfin, la mutation des différents gènes *Hox* au sein du groupe paralogue 8 illustre bien à la fois la redondance partielle et la spécificité fonctionnelle. Dans un premier temps, les phénotypes squelettiques observés chez les mutants *Hoxc8* et *Hoxd8* sont assez similaires: des transformations homéotiques d'antériorisation de T8 en T7 ainsi que de L1 en T13. De son côté, la mutation du gène *Hoxd8* entraîne également un changement de l'identité de la vertèbre T7 en T6 (Le Mouellic *et al.*, 1992; Turet *et al.*, 1993; van den Akker *et al.*, 2001). En ce qui concerne le gène *Hoxb8*, son rôle semble davantage spécifique à la formation du squelette paraxial (section 1.2.7).

1.2.6.3 Les gènes *Hox* spécifiant la partie caudale (paralogues 9-13)

Les souris où l'on retrouve individuellement une perte de fonction pour les gènes *Hoxa9*, *Hoxc9* et *Hoxd9* subissent respectivement des transformations homéotiques d'antériorisation aux niveaux des vertèbres L1 à L5, T10 à L1, et L3 à L5 ainsi que S2 à Ca1 (Suemori, 1995; Fromental-Ramain *et al.*, 1996a). De plus, la génération de quadruple mutants pour les gènes *Hoxa9/Hoxb9/Hoxc9/Hoxd9* a pour effet de modifier l'identité des vertèbres T8 à L4 en T7 (McIntyre *et al.*, 2007).

La perte de fonction du gène *Hoxa10* affecte la morphologie des vertèbres L1, L2, S1 et S4. (Satokata et Uchiyama, 1995; Wahba *et al.*, 2001). De son côté, la perte de fonction du gène *Hoxd10* mène à l'antériorisation de S1 à Ca1 (Carpenter *et al.*, 1997). De façon intéressante, les souris triple mutantes pour les gènes *Hoxa10/Hoxc10/Hoxd10* sont caractérisées par l'absence totale de vertèbres lombaires et sacrales, suite à l'apparition de côtes sur celles-ci. À l'inverse, on observe l'inhibition de la formation des côtes chez la souris transgénique qui exprime le gène *Hoxa10* de façon ectopique au niveau du mésoderme paraxial (Carapuco *et al.*, 2005). Par conséquent, il semble que les gènes *Hox*

appartenant au dixième groupe paralogue remplissent la fonction d'établir la transition de la région thoracique vers la région lombo-sacrale.

La perte de fonction du gène *Hoxa11* provoque à la fois des transformations homéotiques d'antériorisation et de postériorisation, qui mènent respectivement à un changement de S1 en L6 et de T13 en L1 (Hogan *et al.*, 1988). Les souris mutantes pour le gène *Hoxd11* acquièrent quant à elles une identité plus antérieure au niveau des vertèbres S2 à Ca1 (Izpisúa-Belmonte *et al.*, 1991; Davis *et al.*, 1995; Favier *et al.*, 1996). La perte de fonction triple pour les gènes *Hoxa11*, *Hoxc11* et *Hoxd11* résulte en la perte complète des vertèbres sacrales qui se transforment en lombaires, de même que des anomalies au niveau du pelvis (Wellik et Capecchi, 2003). De plus, l'habileté des membres du groupe paralogue 11 à générer des caractéristiques sacrales fut également démontrée grâce à des expériences de surexpression, où des embryons transgéniques en développement exprimant *Hoxa11* dans le PSM auraient des vertèbres adjacentes fusionnées (ou sacralisation) (Carapuco *et al.*, 2005).

En ce qui concerne les gènes *Hox* membres du groupe paralogue 13, les mutations individuelles de perte de fonction des gènes *Hoxc13* et *Hoxd13* présentent respectivement des changements au niveau de Ca5 à Ca10 et S4 à Ca3 (Dollé *et al.*, 1993a; Godwin et Capecchi, 1998).

Une petite parenthèse pour mentionner que l'initiation de l'expression des gènes *Hox* des groupes paralogues 10 à 13 est fortement retardée chez les embryons de serpents et de *caeciliidae*, ce qui correspond à la transition torso-caudale de même qu'à la fin de l'élongation de l'axe (Woltering *et al.*, 2009; Di-Poï *et al.*, 2010). Ces résultats indiquent que les régions dorsales, remarquablement longues et contenant des côtes propres à ce type d'animal, auraient pu émerger suite à un changement temporel dans l'activation des gènes *Hox* postérieurs. De plus, il a été suggéré que le nombre plus élevé de somites dans ces animaux résulterait d'une vitesse de segmentation plus rapide comparativement à celle de la spécification du squelette axial (Gomez *et al.*, 2008).

Enfin, pour vérifier si une fonction squelettique pouvait être attribuable à un complexe *Hox* spécifique en tant que métagène, la délétion de chacun des complexes fut

entreprise (Zákány et Duboule, 1999b; Medina-Martinez *et al.*, 2000; Suemori et Noguchi, 2000; Spitz *et al.*, 2001; Zákány *et al.*, 2001; Kmita *et al.*, 2005). L'absence de phénotypes squelettiques exacerbés comparativement aux mutations individuelles suggère que les complexes *Hox* sont fortement redondants, ce qui pourrait s'expliquer par le fait qu'ils descendent du même complexe *HOM* ancestral. Cependant, le niveau relativement faible d'homologie entre les séquences intergéniques des gènes *Hox* à l'intérieur des différents groupes paralogues, suggérant une régulation de leur expression propre à chacun des complexes (organisation tridimensionnelle de la chromatine, ARN non-codants, éléments de régulation agissant à distance, *etc.*), rend cette hypothèse peu probable. Ces différents mécanismes de régulation seront d'ailleurs dépeints à la section 1.3.

1.2.7 Établissement du squelette paraxial

Tel que mentionné précédemment à la section 1.2.3, le squelette paraxial, comprenant le sternum qui dérive du mésoderme de la plaque latérale (LPM), de même que les côtes s'y rattachant qui dérivent du mésoderme présomitique (PSM), ne répond pas entièrement au principe de colinéarité (Nowicki et Burke, 2000), en opposition au squelette axial qui dérive uniquement du PSM. Bien que les gènes *Hox* des groupes paralogues 5, 6 et 9 soient les plus importants dans la spécification du squelette paraxial, tel que représenté au schéma en Figure 1.5, des membres appartenant à d'autres groupes paralogues ont également un certain rôle à jouer. Par exemple, la perte de fonction du gène *Hoxb4* entraîne une mortalité périnatale associée à une fission sternale (Ramírez-Solis *et al.*, 1993). Ensuite, les souris mutantes pour le gène *Hoxb7* et *Hoxb8* présentent également des malformations du sternum lorsqu'ils sont mutés de façon individuelle (Burke *et al.*, 1995; Chen *et al.*, 1998; van den Akker *et al.*, 2001). En ce qui concerne les souris triple mutantes *Hoxa5/Hoxb5/Hoxc5*, elles possèdent un sternum plus court avec absence de la première sternèbre. Il y a également absence de fusion au niveau du xyphoïde (McIntyre *et al.*, 2007). On observe les même phénotypes, mais légèrement aggravés, chez les triple mutants *Hoxa6/Hoxb6/Hoxc6* à savoir un sternum anormal et très court, une absence complète du xyphoïde et la présence de plusieurs côtes fusionnées (McIntyre *et al.*, 2007). Enfin, bien que des anomalies ont été détectées au niveau du sternum chez certaines souris simple

mutantes (Fromental-Ramain *et al.*, 1996a; Chen et Capecchi, 1997), le rôle principal des gènes *Hox* appartenant au groupe paralogue 9 dans l'établissement du squelette paraxial semble être la production des côtes flottantes, puisqu'on voit apparaître une cage thoracique contenant entre treize et quatorze sternèbres chez les individus quadruple mutants *Hoxa9/Hoxb9/Hoxc9/Hoxd9* au lieu de sept (McIntyre *et al.*, 2007).

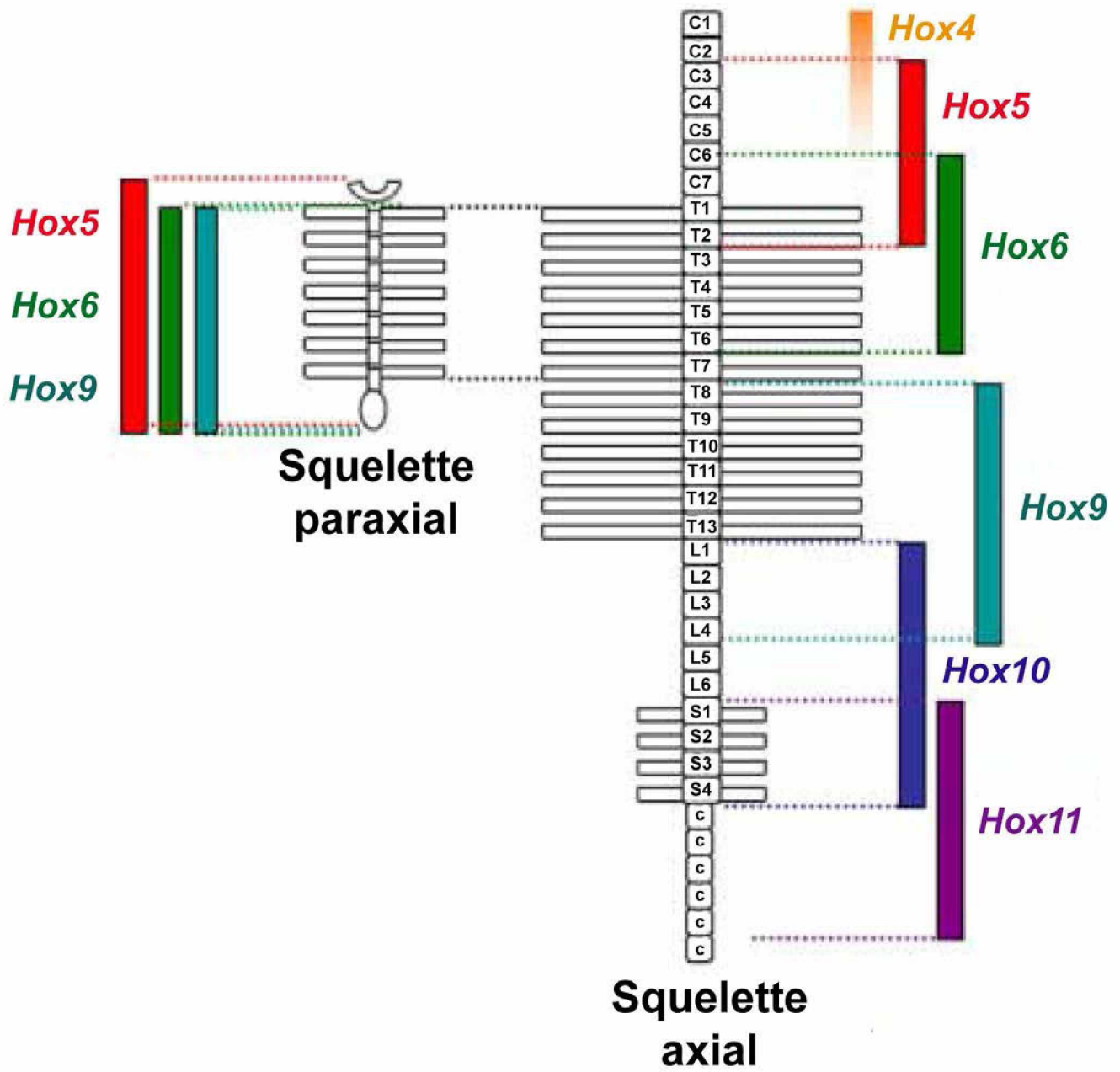
1.2.8 Rôles dans l'organogenèse

À l'origine, l'une des principales raisons ayant poussé certains membres de la communauté scientifique à faire le saut chez la souris afin d'étudier les gènes *Hox* fut de vérifier si leur expression répondait également au principe de colinéarité dans le modelage de structures complexes le long de l'axe AP. Ils constatèrent rapidement que la colinéarité s'appliquait également au niveau de nombreux processus d'organogenèse et que la perte de fonction de certains gènes *Hox* pouvait également compromettre la survie de l'animal (Krumlauf, 1994; Aubin et Jeannotte, 2001).

1.2.8.1 Système nerveux central (SNC) et cellules de la crête neurale (CCN)

Lors de la neurulation, le tube neural donnera naissance au SNC (cerveau et moelle épinière) de même qu'aux CCN (neurones moteurs, cartilages, cellules pigmentaires, *etc.*) (Santagati et Rijli, 2003). En ce qui a trait au cerveau, les gènes *Hox* ne sont pas exprimés au niveau du cerveau antérieur (prosencephale), ni au niveau du cerveau moyen (mésencéphale). Par conséquent, la limite d'expression antérieure des gènes *Hox* se situe au niveau du cerveau postérieur (rhombencéphale), ou plus particulièrement au niveau de son deuxième segment, appelé rhombomère 2 (r2) (Moens et Prince, 2002; Glover *et al.*, 2006). D'ailleurs, l'expression colinéaire des gènes *Hox* contrôle la segmentation du rhombencéphale permettant l'organiser des propriétés régionales essentielles à l'orchestration de ses fonctions hautement conservées (Maconochie *et al.*, 1996; Kiecker et Lumsden, 2005; Nolte et Krumlauf, 2006; Parrish *et al.*, 2009). Parmi celles-ci, on retrouve la coordination de l'activité motrice, de la posture, de l'équilibre et des cycles du sommeil. Le rhombencéphale est également responsable de fonctions inconscientes essentielles

Figure 1.5: Schéma représentatif de l'axe AP où chacune des vertèbres est représentée et identifiée en noir, alors que les rectangles de couleur symbolisent les régions du squelette affectées chez les souris mutantes pour les différents membres des groupes paralogues *Hox*. Les gènes *Hox* en 3' des complexes sont exprimés dans les domaines les plus antérieurs, ce qui s'accorde avec le principe de colinéarité. (Gauche) Squelette paraxial/abaxial. (Droite) Squelette axial (tirée de Wellick, 2007. Autorisation reçue de la part de(s) auteur(s).)



comme le rythme de la respiration et la circulation sanguine. Les gènes *Hox* appartenant aux groupes paralogues 1 à 4 sont les principaux joueurs impliqués dans l'établissement du SNC, alors que plusieurs gènes *Hox* des groupes paralogues 1 à 10 sont responsables de la migration et de la spécification des CCN en neurones moteurs (Lufkin, 1996; Lumsden et Krumlauf, 1996; Oury *et al.*, 2006; Matis *et al.*, 2007; Geisen *et al.*, 2008). Un schéma récapitulatif montrant l'expression colinéaire des gènes *Hox* au niveau du SNC est présenté à la Figure 1.6.

À l'intérieur de la colonne vertébrale, les corps cellulaires des neurones moteurs, qui projettent les axones le long d'un sentier périphérique défini sont regroupés en différents niveaux organisationnels (colonnes motrices, divisions, niches, *etc.*). Ces sous-types neuronaux occupent tous une position fixe dans la colonne vertébrale dictée par le code *Hox* (Carpenter, 2002). Les gènes *Hox* des trois premiers groupes paralogues s'occupent essentiellement des neurones du visage (Carpenter *et al.*, 1993; Dollé *et al.*, 1993b; Goddard *et al.*, 1996; Gavalas *et al.*, 1998; Studer *et al.*, 1998; Bell *et al.*, 1999; Jungbluth *et al.*, 1999; Rossel et Capecchi, 1999; Gaufo *et al.*, 2000, 2003; McNulty *et al.*, 2005; Santagati *et al.*, 2005). D'ailleurs, les souris nouveau-nés mutantes pour le gène *Hoxa1* meurent d'anoxie résultant de l'absence du noyau moteur des nerfs faciaux vague et glosso-pharyngien. Quant aux gènes *Hox* membres des groupes paralogues 4-6, 7-9 et 10, ils modulent respectivement les neurones moteurs de la colonne latérale (LMC) des membres supérieurs, de la ceinture pectorale (PGC) thoracique et des neurones LMC lombaires associés aux membres inférieurs (Carpenter *et al.*, 1997; Tiret *et al.*, 1998; Lance-Jones *et al.*, 2001; Liu *et al.*, 2001; Dasen *et al.*, 2003, 2005; Choe *et al.*, 2006; Tarchini et Duboule, 2006; Jung *et al.*, 2010). De plus, les gènes *Hoxa5* et *Hoxc5* sont exprimés au niveau des neurones moteurs du nerfs phrénique qui innerve le diaphragme (Philippidou *et al.*, 2012) (voir aussi section 1.4.3.1) Enfin, la génération de souris mutantes *HoxA^{flox/flox} HoxB^{del/del} Wnt1Cre* a permis de découvrir que la délétion combinée de ces deux complexes dans les CCN entraîne des malformations au niveau de la mâchoire (dents, cartilages et muscles) (Vieux-Rochas *et al.*, 2013).

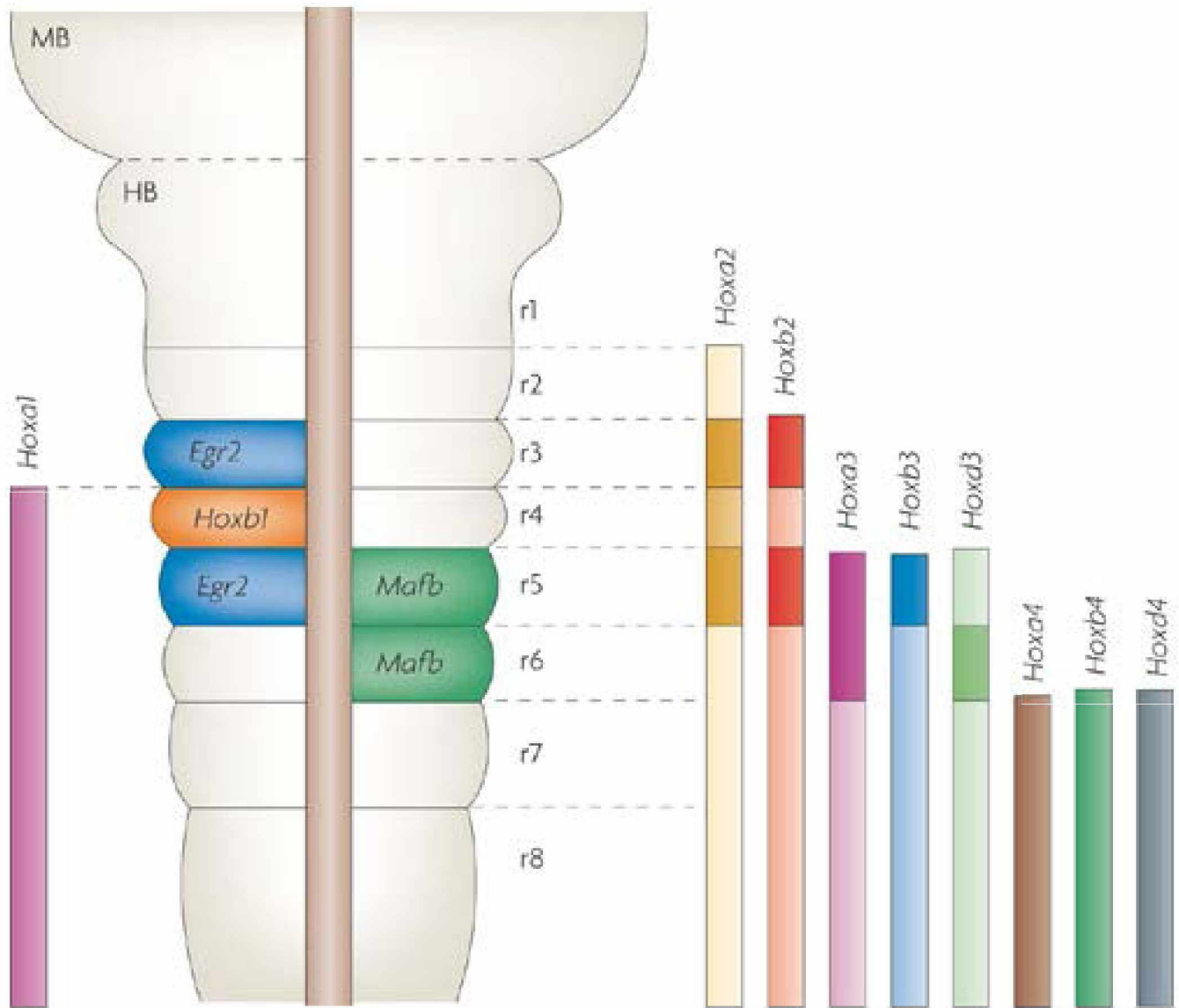
1.2.8.2 Organes glandulaires

L'expression des gènes *Hox* au niveau des CCN, qui migrent et participent à l'ontogenèse des arches branchiaux, joue par le fait même un rôle important dans le développement des organes glandulaires (Stein *et al.*, 1996). Ainsi, l'étude des souris mutantes pour les gènes *Hox* appartenant aux trois premiers groupes paralogues a démontré leur importance au niveau du thymus, de même que des glandes thyroïde et parathyroïde. Par exemple, l'inactivation simultanée des gènes *Hoxa1* et *Hoxb1* résulte en l'agenèse du thymus ainsi que de la glande parathyroïde qui dérivent tous deux du troisième arche branchial (Rossel et Capecchi, 1999). L'absence complète de thymus et de la glande parathyroïde constitue également une des composantes du phénotype des souris déficientes pour le gène *Hoxa3* (Chisaka et Capecchi, 1991; Manley et Capecchi, 1995, 1997, 1998; Su *et al.*, 2001). On compte aussi un ganglion spinal (ganglion de Frierip) de plus chez les souris exprimant le gène *Hoxb8* sous le contrôle du promoteur du gène *RARβ2* (Charité *et al.*, 1994).

1.2.8.3 Système respiratoire

Durant la gestation, le poumon se développe afin de fournir au nouveau-né un organe capable de maintenir la respiration ainsi que les échanges gazeux. Le développement du poumon est un phénomène continu durant l'embryogenèse, qui se poursuit après la naissance, et il est divisé en cinq stades basés sur des caractéristiques anatomiques et histologiques. Ces cinq stades sont schématisés à la Figure 1.7. Premièrement, lors du stade I (embryonnaire), on voit apparaître le bourgeon pulmonaire rudimentaire, jaillissant de l'endoderme ventral de l'intestin antérieur primitif et poussant à l'intérieur du mésoderme splanchnique adjacent vers le jour E9.5 (Ten Have-Opbroek, 1991; Minoo et King, 1994). On assiste ensuite à l'élongation rapide de ce bourgeon qui se scinde en deux bourgeons pulmonaires primaires asymétriques à E10.5, où le bourgeon de droite est plus gros que celui de gauche. Simultanément, la gouttière laryngotrachéale se forme à partir de l'épithélium de l'intestin antérieur primitif, appelé aussi intestin céphalique primitif. Cette gouttière s'allongera et se divisera en 2 selon l'axe dorsoventral pour former l'œsophage primitif et la trachée rudimentaire. La septation entre la trachée et l'oesophage s'achève au jour E11.5. Deuxièmement, lors du stade II (pseudoglandulaire), on

Figure 1.6: Les différents domaines d'expression des gènes *Hox* au niveau du cerveau primitif postérieur d'un embryon de souris au stade E11.5. Les différents niveaux d'expression des gènes *Hox* au sein de leur domaine respectif, qui s'étend de la partie caudale du cerveau postérieur primitif jusqu'à la frontière délimitée par un rhombomère spécifique, sont représentés en fonction de l'intensité de la couleur. HB, cerveau postérieur primitif; MB, cerveau médian primitif; r, rhombomère (tirée de Guthrie *et al.*, 2007. Autorisation reçue de la part de(s) auteur(s).)



assiste chez la souris à une ramification active donnant naissance au lobe pulmonaire gauche ainsi qu'aux quatre lobes droits, de même qu'à l'ébauche de l'arbre respiratoire au jour E12.5 (Warburton *et al.*, 2000). Durant cette période, il y a apparition des toutes premières cellules musculaires lisses autour des espaces aériens. Quant au stade III (canaliculaire), il est caractérisé par le développement du réseau de capillaires et des acini pulmonaires au jour E16.5. C'est durant cette période que débute la diminution graduelle du mésenchyme pulmonaire, de même que l'apparition des cellules caractéristiques des voies aériennes proximales (cellules ciliées, cellules de Clara et cellules de goblet ou productrices de mucus). On assiste également à la spécification des premiers pneumocytes de types I et II, responsables des échanges gazeux et de la sécrétion des surfactants (Burri et Moschopulos, 1992). Le stade IV (sacculaire) coïncide avec une septation rudimentaire des espaces aériens périphériques qui prennent alors l'allure de sacs au jour E17.5. Il y a également une différenciation intensive des pneumocytes de types I et II pour que ces cellules caractérisant les voies aériennes distales deviennent ultimement les cellules épithéliales les plus abondantes du poumon à la naissance (J0) (Wert *et al.*, 1993). Finalement, lors du stade V (alvéolaire) qui se poursuivra de J5 jusqu'à J30 chez la souris, on observe la septation des saccules menant à la formation du réseau alvéolaire (Soutiere *et al.*, 2004). L'augmentation intensive du nombre d'alvéoles sert à accroître la surface d'échanges gazeux. De plus, la formation des alvéoles s'accompagne du dépôt de l'élastine pulmonaire afin de permettre l'expansion et la compression alvéolaire au fur et à mesure que la surface d'échange gazeux augmente.

Plusieurs gènes *Hox* membres des groupes paralogues 1 à 8 sont exprimés niveau du mésenchyme des voies respiratoires selon un profil d'expression spatio-temporel spécifique. Cette observation suggère que ces gènes *Hox* pourraient donc participer au développement proximo-distal du tractus respiratoire selon leur positionnement au sein de leur complexe (Cardoso, 1995; Bogue *et al.*, 1996; Herriges *et al.*, 2012). Par contre, à l'exception de la mutation du gène *Hoxa5* qui sera présenté en détails à la section 1.4.3.3, les souris simple mutantes pour les autres gènes *Hox* ne présentent aucun phénotype pulmonaire apparent, suggérant qu'il existe une redondance fonctionnelle entre ces gènes. D'ailleurs, le fait que l'on retrouve des nouveau-nés décédés d'insuffisance respiratoire chez

les souris délétées pour les complexes *HoxB* (*Hoxb1* à *Hoxb9*) et *HoxC* en entier (Medina-Martinez *et al.*, 2000; Suemori et Noguchi, 2000) supporte l'hypothèse de la redondance fonctionnelle. Ensuite, bien que l'histologie n'a révélé aucun défaut apparent, les poumons des souris double mutantes pour les gènes *Hoxa1* et *Hoxb1* sont d'allure plus petits et possèdent moins de lobes que ceux simple mutants pour ces gènes (Rossel et Capecchi, 1999). Enfin, même s'il y a une redondance partielle entre le rôle pulmonaire joué par *Hoxb5* et celui joué par *Hoxa5*, la mutation du gène *Hoxa5* est la seule identifiée parmi les gènes *Hox* jusqu'à ce jour conduisant à une mort périnatale associée à la détresse respiratoire (Aubin *et al.*, 1997). La mutation du gène *Hoxb5* a quant à elle démontré son rôle dans le processus de ramification, la spécification des cellules de goblet de même que dans la structure postnatale des espaces aériens (Volpe *et al.*, 1997; Volpe *et al.*, 2000a; Volpe *et al.*, 2000b; Volpe *et al.*, 2003; Boucherat *et al.*, 2013).

1.2.8.4 Système digestif

Sous la gouttière laryngo-trachéale, la partie caudale de l'intestin proximal primitif, appelé aussi intestin abdominal primitif, donnera naissance au système digestif. Le segment antérieur formera l'œsophage, l'estomac ainsi que la moitié supérieure du duodénum. Pour sa part, l'intestin abdominal primitif médian se spécifiera en la moitié distale du duodénum, en jéjunum, en iléon, en caecum et ses appendices, en colon ascendant ainsi que les deux premiers tiers du colon transverse. Quant à l'intestin abdominal primitif distal, il génère le dernier tiers du colon transverse, le colon descendant, le colon sigmoïde, le rectum ainsi que la partie supérieure du canal anal.

De la même façon que pour les organes décrits précédemment, les gènes *Hox* jouent également un rôle dans l'établissement de l'identité régionale le long du tube digestif. D'ailleurs, il a été stipulé initialement que le rôle premier des gènes *Hox* chez les invertébrés était de définir les domaines fonctionnels du système digestif et que cette fonction aurait été récupérée par les vertébrés durant l'évolution pour la spécification du squelette (Burke *et al.*, 1995). On observe une mortalité périnatale associée à un défaut de la couche musculaire de l'œsophage chez les souris mutantes pour le gène *Hoxc4*, ce qui les empêche de se nourrir (Boulet et Capecchi, 1996). De son côté, la souris surexprimant

Figure 1.7: Représentation schématique du système respiratoire lors des 5 stades développementaux de la morphogenèse pulmonaire chez la souris et l'humain. (tirée de <http://www.cincinnatichildrens.org/research/divisions/p/pulmonary-bio/lungmorphogenesis/>; Laboratoire du Dr Whitsett. Autorisation reçue de la part de(s) auteur(s).)

Stades du développement pulmonaire: souris vs humain

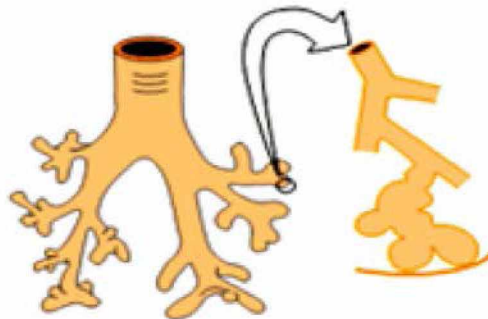


Bourgeon pulmonaire

I. Embryonnaire
Souris: E9-12
Humain: Sem 3-7



II Pseudoglandulaire
Souris: E12-15
Humain: Sem 5-17



III. Canaliculaire
Souris: E15-17
Humain: Sem 16-26



IV. Sacculaire
Souris: E17-J0
Humain: Sem 26-36



V. Alvéolaire
Souris: J0-J25
Humain: Sem 36-3 ans

le gène *Hoxa4* par l'entremise du promoteur SV40 meurt quelque temps avant l'âge adulte d'une accumulation de matières fécales au niveau du colon (Wolgemuth et al., 1989). Le gène *Hoxc8* semble également jouer un rôle important dans la spécification du tube digestif, puisque des malformations ont été soulignées suite à la perte de sa fonction (Le Mouellic et al., 1992; Tiret et al., 1998), de même que suite à son expression ectopique dans les domaines d'expression du gène *Hoxa4* (Pollock et al., 1992), comme la désorganisation de l'épithélium gastrique ainsi que la présence de lésions hamartomateuse au niveau de la couche musculaire. Une délétion du complexe *HoxD* en entier, ou même de la séquence génomique comprenant les gènes *Hoxd4* à *Hoxd13*, entraîne des anomalies au niveau des sphincters iléocaecal et anal (Záchány et Duboule, 1999a; Záchány et al., 2001). De plus, l'inactivation ciblée des gènes *Hoxd12* ou *Hoxd13* affecte uniquement le sphincter anal (Kondo et al., 1996). Par conséquent, plusieurs gènes du complexe *HoxD* sont exprimés dans une région de l'intestin médian primitif correspondant au futur bourgeon du caecum (Levin et al., 1997; Sekimoto et al., 1998; Pitera et al., 1999; Kawazoe et al., 2002), alors que les gènes *Hoxd12* et *Hoxd13* y sont exclus (Dolle et al., 1991; Kmita et al., 2000; Spitz et al., 2005).

1.2.8.5 Formation des membres

La génération de multiples mutations au niveau des complexes *Hox* a confirmé leur rôle indispensable dans le développement normal des membres, durant lequel les gènes *Hoxa9-Hoxa13* et *Hoxd9-Hoxd13* participent à la fois à la croissance des membres et à l'organisation des structures, comme la taille et le nombre de doigts, par exemple (Rijli et Chambon, 1997; Kondo et al., 1998a; Zákány et Duboule, 2007). Durant la phase précoce, les principes généraux qui dictent l'expression des gènes *Hox* afin de spécifier les membres selon l'axe AP sont comparables entre les complexes *HoxA* et *HoxD*, faisant intervenir une régulation spatio-temporelle colinéaire semblable à celle du tronc. Par contre, durant la phase tardive où l'on assiste à la spécification des doigts selon l'axe proximo-distales (PD), les profils d'expression des gènes *HoxA* et *HoxD* sont de moins en moins similaires, suggérant qu'ils font intervenir des mécanismes de régulation transcriptionnelle acquis séparément suite à la duplication des complexes.

Trois régions composent les membres chez l'adulte. Celles-ci sont le stylopode, zone antérieure composée de l'humérus (membre supérieur) ou du fémur (membre inférieur), le zeugopode, zone postérieure composée du radius et de l'ulna (supérieur) ou du tibia et de la fibula (inférieur), ainsi que de l'autopode, zone distale composée des carpes, métacarpes et phalanges (supérieur) ou des tarses, métatarses et orteils (inférieur). Alors que les expériences de gain de fonction ont mis en évidence l'importance des gènes *Hox* des groupes paralogues 9 à 13 dans l'établissement d'un profil morphologique selon l'axe AP (Knezevic *et al.*, 1997; Herault *et al.*, 1998; van den Akker *et al.*, 2001; Kmita *et al.*, 2002; Utsch *et al.*, 2002; Caronia *et al.*, 2003; Zákány *et al.*, 2004; Williams *et al.*, 2006), les pertes de fonction combinées de ces gènes entraînent plutôt des troncations sévères le long de l'axe PD (Davis *et al.*, 1995; Fromental-Ramain *et al.*, 1996a; McIntyre *et al.*, 2007). Par ailleurs, les gènes des paralogues *Hox10* et *Hox11* façonnent respectivement le stylopode et le zeugopode (Wellik et Capecchi, 2003), alors que ceux appartenant aux groupes paralogues 11 à 13 spécifient l'autopode de manière dose-dépendante (Zákány *et al.*, 1997a; Zákány et Duboule, 1999b). Enfin, l'importance capitale des gènes *Hox* lors du développement des membres, a été démontrée suite à la génération de la souris double mutante *HoxA*(conditionnelle)/*HoxD*, où l'on observe une troncation sévère de ses membres. Des études récentes ont démontré que ces troncations résulteraient de l'effet combiné de l'absence d'expression du gène *Shh* dans la zone d'activité polarisante (ZPA) (Kmita *et al.*, 2005; Galli *et al.*, 2010; Xu et Wellik, 2011) et d'un mécanisme indépendant au gène *Shh*, au niveau de la crête apicale ectodermique (AER) (Sheth *et al.*, 2012).

1.2.8.6 Système urogénital

Comme le mésoderme néphrogénique se forme également le long de l'axe AP, il n'est pas surprenant que différents groupes paralogues *Hox* contribuent au développement du rein (Wellik, 2011). Plus de 25 des 39 gènes *Hox* sont exprimés de façon colinéaire dans différentes structures incluant le canal de Wolff, le bourgeon de l'uretère, le mésenchyme non différencié ainsi que les structures différenciées dérivant du mésenchyme condensé (Patterson et Potter, 2004). Par exemple, la perte de fonction combinée des gènes *Hox* du groupe paralogue 11 perturbe l'organogenèse du rein en occasionnant son agénésie complète (Patterson *et al.*, 2001; Wellik *et al.*, 2002). Les souris mutantes pour les gènes

Hox du groupe paralogue 13 ont un défaut de connexion entre l'uretère et la vessie (Warot et al., 1997; Di-Poï et al., 2007). Un dédoublement du rein a été également observé lors d'expériences de surexpression du gène *Hoxb7* (Argao et al., 1995).

Une expression colinéaire de plusieurs gènes situés en 5' des complexes *HoxA* et *HoxD* est également retrouvée au niveau du tractus génital. Par exemple chez la femelle, le gène *Hoxa7* est exprimé au niveau des follicules ovariens (Ota et al., 2006), *Hoxa9* au niveau des oviductes, *Hoxa10* ainsi que *Hoxa11* au niveau du col de l'utérus, et *Hoxa13* au niveau de la vulve (Taylor et al., 1997; Du et Taylor, 2004). Quant aux mâles, les gènes *Hox* plus postérieurs sont responsables du développement des lobes séparés de la prostate, un organe glandulaire se développant à partir du sinus urogénital (Hostikka et Capecchi, 1997; Economides et Capecchi, 2003; Huang et al., 2007; Javed et Langley, 2013). De plus, la perte de fonction individuelle ou combinée des gènes des groupes paralogues 10 et 11, appartenant aux complexes *HoxA* et *HoxD*, engendre la stérilité chez les souris mâles et femelles, de même qu'une transformation homéotique d'antériorisation respectivement au niveau du canal déférent et de l'utérus (Hsieh-Li et al., 1995; Raines et al., 2013). La stérilité observée chez les souris femelles mutantes pour les gènes *Hoxa10* et *Hoxa11* viendrait d'un défaut lors de l'implantation des embryons dans l'utérus puisque ces gènes moduleraient la réponse à la progestérone et à l'œstrogène (Benson et al., 1996; Taylor et al., 1998; Eun Kwon et Taylor, 2004). Enfin, la mutation des gènes *Hoxa13* et *Hoxd13* cause une malformation du tubercule génital (Kondo et al., 1997; Warot et al., 1997; Scott et al., 2005). Ces mutations sont aussi retrouvées chez l'humain et son associée aux syndromes main-pied-génital et synpolydactyly.

1.2.8.7 Glande mammaire

L'analyse des souris mutantes pour un ou plusieurs gènes *Hox* a également révélé l'importance de ceux-ci dans le développement de la glande mammaire. Tout d'abord, le gène *Hoxc6* est exprimé dans la glande des souris pubères et adultes, alors que son expression diminue durant la gestation (Friedmann et al., 1994; Garcia-Gasca et Spyropoulos, 2000). La perte de fonction du gène *Hoxc6* entraîne une diminution de la ramification des canaux. Les gènes *Hox* du groupe paralogue 9 contrôlent le

développement de la glande mammaire en réponse à la gestation, puisque la perte combinée des gènes *Hoxa9*, *Hoxb9* et *Hoxd9* se traduit par une hypoplasie de la glande mammaire, un développement lobuloalvéolaire aberrant et l'incapacité de produire du lait (Chen et Capecchi, 1999). On observe également un problème de lactation causé par des défauts d'expansion alvéolaire chez les souris déficientes pour le gène *Hoxd10* (Lewis, 2000). Enfin, le gène *Hoxb13* semble jouer un rôle au niveau du remodelage de la glande mammaire en réponse aux hormones, puisqu'il régule directement de façon négative la transcription du récepteur α aux estrogènes (ER α) et positivement celle de l'interleukine 16 (IL-16) (Shah *et al.*, 2013).

1.2.8.8 Barrière cutanée

Certains gènes *Hox* sont exprimés au niveau de l'épiderme où ils remplissent des fonctions de maintien de l'intégrité de la peau, de cicatrisation et de formation des poils. Par exemple, les souris mutantes pour le gène *Hoxd3* cicatrisent plus difficilement (Uyeno *et al.*, 2001). Le gène *Hoxc4* serait quant à lui impliqué dans la différenciation des kératynocytes (Rieger *et al.*, 1994). En absence du gène *Hoxb13*, la différenciation épithéliale serait compromise empêchant la régénération optimale suite aux blessures (Mack *et al.*, 2003, 2005), alors que la perte de fonction du gène *Hoxc13* inhibe la formation de tous les poils (Godwin et Capecchi, 1998).

1.2.8.9 Cellules hématopoïétiques

Les gènes *Hox* sont exprimés dans les cellules souches hématopoïétiques ainsi que les progéniteurs (HSPC) selon un profil spécifique à la lignée et au stade de différenciation, de façon similaire à ce qui est observé au cours du développement précoce de l'embryon. Par exemple, les gènes *Hoxb3*, *Hoxb4* et *Hoxa9* sont fortement exprimés dans les cellules hématopoïétiques pluripotentes, alors que les gènes *Hoxb8* et *Hoxa10* sont exprimés dans les cellules engagées dans la voie myéloïde. De plus, les différentes lignées hématopoïétiques possèdent leur propre profil d'expression de gènes *Hox*. Par exemple, les gènes du complexe *HoxA* sont exprimés majoritairement dans les cellules myéloïdes, les gènes du complexe *HoxB* dans les cellules érythroïdes, les gènes du complexe *HoxC* dans

les cellules lymphoïdes, alors qu'on ne retrouve aucun gène exprimé appartenant au complexe *HoxD* (Giampaolo *et al.*, 1994; Kawagoe *et al.*, 1999; Pineault *et al.*, 2002). Par ailleurs, des expériences de gain de fonction chez la souris, au cours desquelles plusieurs gènes *Hox* différents ont été surexprimés, induisent l'expansion des populations de cellules souches et de cellules progénitrices, tout en bloquant la différenciation (Sauvageau *et al.*, 1995, 1997; Daga *et al.*, 2000; Buske *et al.*, 2001; Thorsteinsdottir *et al.*, 2002; Amsellem *et al.*, 2003; Fischbach *et al.*, 2005). La régulation aberrante de l'expression de nombreux gènes *Hox* est associée ou directement reliée à plusieurs leucémies. D'ailleurs, on retrouve fréquemment la surexpression de gènes du complexe *HoxA* suite à la transformation leucémique (Afonja *et al.*, 2000; Drabkin *et al.*, 2002; Bach *et al.*, 2010; Sio *et al.*, 2013), comme les gènes *Hoxa9* et *Meis1* identifiés chez environ la moitié des patients souffrant de leucémie myéloïde aigue (AML) (Lawrence *et al.*, 2005; Palmqvist *et al.*, 2007). De plus, la substitution des domaines homéo des protéine HOXA1 et HOXA9 est suffisant pour convertir le phénotype leucémique faiblement pénétrant initié par la protéine HOXA1 en une maladie agressive spécifique à la protéine HOXA9 et *vice versa* (Breitinger *et al.*, 2012). Le complexe de gènes *HoxA* est également impliqué dans plusieurs leucémies lymphoblastiques aiguës de type lymphocytes T (T-ALL) où une translocation chromosomique amène des parties du complexe *HoxA* sous le contrôle d'un activateur fort, dirigeant normalement l'expression du récepteur des lymphocytes T (TCR) (Soulier *et al.*, 2005; Speleman *et al.*, 2005). De plus, les gènes *Hox* participent à des translocations génomiques créant des protéines de fusion HOX hautement leucémogéniques avec les protéines du complexe de pore nucléaire NUP96/98 (Iwasaki *et al.*, 2005; Lin *et al.*, 2005; Nakamura, 2005; Palmqvist *et al.*, 2007; Novak *et al.*, 2012).

1.2.8.10 Angiogenèse et microenvironnement tumoral

Les gènes *Hox* sont en mesure d'induire la formation tumorale par la promotion de facteurs pro-angiogéniques (VEGF, bFGF, IL-8, ANGPTL-2, les intégrines $\alpha 5\beta 1$ et $\alpha 5\beta 3$, MMP-14, uPAR), de facteurs pro-motilité (amphiréguline, épériréguline, neuréguline) ainsi que de la transition épithéliale-mésenchymale (EMT) associée à l'activation du sentier TGF β (Boudreau et Varner, 2004; Mace *et al.*, 2005; Hayashida *et al.*, 2010). De plus, certains gènes *Hox*, comme le gène *Hoxd10*, empêchent la migration cellulaire endothéliale

et bloquent l'angiogenèse induite par le bFGF et le VEGF *in vivo* au niveau de la membrane chorioalantoïque (Myers *et al.*, 2002) .

1.2.8.11 Structures extra-embryonnaires

Une fois de plus, l'analyse des souris mutantes pour les différents gènes *Hox* a mis en évidence que les gènes en 5' du complexe *HoxA*, soit les gènes *Hoxa10* à *Hoxa13*, jouent un rôle primordial dans le développement du placenta (Scotti et Kmita, 2012). En effet, les embryons mutants pour le gène *Hoxa13* meurent aux alentours de E12.5 en raison d'un défaut de l'expansion vasculaire fœtale au niveau du labyrinthe placentaire, nécessaire à l'échange entre le sang maternel et fœtal (Shaut *et al.*, 2008).

1.3 Régulation de l'expression des gènes *Hox*

La complexité grandissante des structures retrouvées chez les différents organismes vivants, au fur et à mesure que l'on s'éloigne de la racine de l'arbre phylogénétique de la vie, corrèle avec la panoplie de mécanismes régulateurs utilisés pour contrôler précisément l'expression spatio-temporelle des gènes *Hox*. En effet, même si leur expression est initiée très tôt durant l'embryogenèse par certains activateurs, d'autres stratégies de régulation sont nécessaires pour maintenir leur expression afin que les gènes homéotiques puissent remplir leurs différentes fonctions développementales (Struhl, 1982). C'est pourquoi l'étude de la régulation transcriptionnelle des gènes *Hox*, notamment en circonscrivant les éléments régulateurs localisés en *cis* et en déterminant les facteurs se liant à ces séquences d'ADN en *trans*, est une étape clef pour comprendre l'une des premières étapes de cette hiérarchie régulationnelle.

1.3.1 L'autorégulation

La régulation de l'expression des gènes *Hox* par leur propre produit protéique, mécanisme que l'on appelle autorégulation, a été décrite initialement chez la drosophile. Ce mécanisme est impliqué principalement dans le maintien des domaines d'expression et ce, que ce soit en prenant part à une boucle d'autorégulation positive comme dans le cas

des gènes *Deformed* et *Ultrabithorax* (Casanova *et al.*, 1985, 1987; Peifer *et al.*, 1987; Kuziora et McGinnis, 1988), ou d'autorégulation négative comme dans le cas du gène *labial* (Chouinard et Kaufman, 1991).

Chez les mammifères, suite à l'initiation de l'expression des gènes *Hox* par le RA (voir la section 1.3.4.1), et par les facteurs de transcription KROX20 et KREISLER, l'expression des gènes *Hox* doit être maintenue pour que le programme développemental puisse s'effectuer normalement. Par exemple, au niveau des rhombomères, cette tâche est prise en charge par les mécanismes d'autorégulation et de régulation croisée, discutés à la section suivante, notamment en ce qui concerne les gènes *Hox* des groupes paralogues 1 à 4 appartenant aux complexes *HoxA* et *HoxB*. Par exemple, la comparaison des régions régulatrices du gène *Hoxb1*, faite chez plusieurs vertébrés, a permis d'identifier trois séquences aux motifs conservés qui sont indispensables au développement normal du quatrième rhombomère (r4) du cerveau postérieur primitif. De plus, l'analyse fonctionnelle de souris transgéniques a démontré l'importance d'un élément d'autorégulation (ARE), sur lequel se lie des hétérodimères HOXB1-PBX1 de même que PBX1-PREP1, pour restreindre l'expression du gène *Hoxb1* au niveau de r4 (Studer *et al.*, 1994; Pöpperl *et al.*, 1995; Gavalas *et al.*, 1998; Ferretti *et al.*, 2005). En ce qui concerne le gène *Hoxa2*, en plus d'être exprimé au niveau des troisième et quatrième rhombomères, ainsi que dans le tube neural jusqu'à la frontière séparant r1 et r2, il est le seul gène *Hox* présent au niveau du deuxième rhombomère, tel qu'illustré à la Figure 1.6. Cette expression spécifique à l'intérieur de r2 est maintenue grâce à l'action de la protéine HOXA2 ainsi que de ses cofacteurs PBX1 et MEIS1 (Lampe *et al.*, 2004). De la même façon, les membres du groupe paralogue 3, *Hoxa3* et *Hoxb3*, maintiennent également leur expression grâce à un mécanisme d'autorégulation. En effet, la comparaison des éléments régulateurs localisés *in cis* des gènes *Hoxa3* du poulet et de la souris a montré qu'un élément conservé de type activateur/amplificateur contrôle la phase tardive de son expression grâce à une boucle d'autorégulation impliquant l'hétérodimère protéique HOXA3/PBX1A (Manzanares *et al.*, 2001). Quant au gène *Hoxb3*, son expression dynamique et complexe au niveau du cerveau postérieur primitif en développement, ainsi que dans les cellules de la crête neurale qui sont associées, est contrôlée par trois éléments régulateurs situés *in cis*. Par une approche jumelant la mutagenèse à la transgénèse afin d'identifier les facteurs agissant *in trans*

responsables du maintien de son expression chez la souris, la protéine HOXB3 a été identifiée comme étant essentielle à sa propre expression (Yau *et al.*, 2002). Enfin, lorsqu'il est introduit dans un environnement génétique *Hoxa4*^{-/-}, l'expression du transgène *Hoxa4/lacZ* est soit complètement abolie, soit significativement réduite dans les embryons de souris âgées entre E10.5 et E12 (Packer *et al.*, 1998).

1.3.2 La régulation croisée

Tel que mentionné à la section précédente, le mécanisme impliquant la régulation croisée des différents produits des gènes *Hox*, découvert initialement chez la drosophile, permet également à ceux-ci de maintenir l'intégrité de leurs domaines d'expression. Par exemple, les mutants pour le gène *Ubx* présentent une absence de la frontière postérieure d'expression du gène *Antp*, suggérant une régulation négative de l'expression du gène *Antp* par la protéine UBX (Hafen *et al.*, 1984). De la même façon, le produit protéique des gènes homéotiques postérieurs *AbdA* et *AbdB* réprimerait l'expression du gène *Ubx* (Struhl et White, 1985; White et Wilcox, 1985). Fait intéressant, le même mécanisme de régulation croisée peut avoir des répercussions différentes selon le tissu à l'intérieur duquel il opère, comme par exemple le gène *Scr* qui est positivement régulé par la protéine ANTP au niveau du mésoderme viscéral, et négativement au niveau de l'épiderme (Reuter et Scott, 1990). Tel que mentionné préalablement à la section 1.1.3.1, les études menées chez la drosophile concernant l'expression ectopique des gènes homéotiques ont révélé des transformations homéotiques circonscrites aux parties antérieures de leur domaine d'expression endogène (Mann et Hogness, 1990). L'hypothèse découlant de ce constat veut que l'absence de phénotypes dans les parties postérieures au domaine d'expression normal des gènes homéotiques exprimés ectopiquement résulte d'une régulation croisée négative, orchestrée par les protéines HOM présentes dans le tissu. Autrement dit, l'expression ectopique est possible dans les régions exemptes des gènes suppresseurs. Ce phénomène de suppression phénotypique est appelé la prévalence postérieure (Gonzalez-Reyes et Morata, 1990; Mann et Hogness, 1990) et sera décrit à la section suivante.

Plusieurs exemples de régulation croisée ont également été rapportés chez la souris. Par exemple, les souris *Hoxb1*^{-/-} montrent un développement anormal du cerveau postérieur

primitif entre les deuxième et cinquième rhombomères inclusivement, qui s'explique par l'absence d'expression du gène *Hoxb1* au niveau de la frontière entre r3 et r4 (Studer *et al.*, 1996, 1998; Barrow *et al.*, 2000; Di Rocco *et al.*, 2001), de même que par l'expression anormale des autres gènes *Hox* impliqués dans cette cascade développementale, comme l'expression du gène *Hoxb2* au niveau du quatrième rhombomère ainsi que des CNN qui en dérivent (Maconochie *et al.*, 1997). De plus, la génération de souris transgéniques exprimant de façon ectopique le gène *Hoxb3*, spécifiquement au niveau du quatrième rhombomère grâce à l'élément de type activateur/amplificateur appartenant au gène *Hoxb2*, entraîne l'abolition de l'expression du gène *Hoxb1* au niveau de cette structure, de même que des défauts neuronaux similaires à ceux observés chez les souris mutantes *Hoxb1*^{-/-} (Wong *et al.*, 2011). Des analyses moléculaires ont confirmé l'existence d'une boucle de régulation croisée négative exercée par la protéine HOXB3 sur l'expression du gène *Hoxb1*. Concernant le gène *Hoxb3*, les deux sites de liaison potentiels aux protéines HOX situés au sein d'un élément régulateur ne reconnaissent pas que la protéine HOXB3 de façon spécifique, mais également la protéine HOXB4, ce qui suggère la présence d'un mécanisme de régulation croisée pour que l'expression du gène *Hoxb3* soit correctement dirigée au niveau du cerveau postérieur primitif et du tube neural (Yau *et al.*, 2002). Quant à la séquence impliquée dans l'expression tardive du gène *Hoxa3*, elle est reconnue par les hétérodimères comprenant la protéine PBX1A associée au produit de n'importe lequel des gènes *Hox* membres du groupe paralogue 3 (Manzanares *et al.*, 2001). Enfin, des études de délétion des gènes situés en 5' des complexes *HoxA* et *HoxD* ont révélé une dérégulation globale des gènes *Hox* appartenant à ces complexes, suggérant la présence d'un mécanisme par lequel les protéines HOX fonctionnellement dominantes établissent et/ou maintiennent les paramètres spatiaux d'expression des autres gènes de *Hox* au sein d'un tissu donné (Sheth *et al.*, 2013).

Bien que les études portant sur les mécanismes d'autorégulation et de régulation croisée n'aient été réalisées que dans le contexte du cerveau postérieur primitif, il n'en demeure pas moins que la régulation de l'expression des gènes situés en 3' des complexes *Hox* a été plus approfondie que celle des gènes en 5'. Des études plus récentes ont toutefois démontré que ces mécanismes sont également présents chez ces gènes *Hox* et qu'ils

participeraient à la prévalence postérieure chez les vertébrés (Yekta *et al.*, 2004, 2008; Woltering et Durston, 2008; Plaza *et al.*, 2009; Durston *et al.*, 2011).

1.3.3 La suppression phénotypique ou prévalence postérieure

Tel que mentionné à la section précédente, le modèle de prévalence postérieure découle des travaux chez la drosophile, où l'expression des gènes *HOM* plus postérieurs masque celle des plus antérieurs (Bachiller *et al.*, 1994; Duboule et Morata, 1994; Peichel *et al.*, 1997). Ces travaux ont également permis d'identifier deux mécanismes impliqués : la régulation transcriptionnelle croisée, qui fait l'objet de cette section, et la régulation post-transcriptionnelle impliquant les microARN (miRNA), qui sera décrite à la section 1.3.13.1.

Ainsi, l'utilisation du promoteur du gène *Hsp70* pour diriger ectopiquement l'expression du gène *Ubx* génère des transformations homéotiques de postériorisation dans les segments antérieurs à son domaine d'expression endogène, alors que les segments postérieurs, où la protéine AbdB est exprimée, ne présentent aucun phénotype (Gonzalez-Reyes et Morata, 1990; Mann et Hogness, 1990). De la même façon, les protéines UBX et AbdA régulent respectivement l'expression du gène *Dpp* dans le mésoderme viscéral, de façon positive au niveau du septième parasegment et de façon négative au niveau des huitième et douzième parasegments, en liant toutes deux la même séquence de type activatrice/amplificatrice (Panganiban *et al.*, 1990; Reuter *et al.*, 1990), mais sur des sites différents excluant l'hypothèse de la compétition directe (Li et McGinnis, 1999). Par conséquent, l'hypothèse voulant qu'un encombrement stérique produit par la protéine HOM dite "suppresseuse" empêche la liaison d'une autre protéine à l'ARE et module ainsi la transcription semble envisageable (Capovilla et Botas, 1998). Enfin, la mutation du gène *Extra sex combs (Esc)*, membre de la famille *Polycomb (PcG)* décrite à la section 1.3.8, génère une surexpression globale des gènes *HOM*, alors que la transformation homéotique observée est dictée par la surexpression du gène *AbdB* (Duboule et Morata, 1994). Des données rapportées relativement récemment ébranlent quelque peu le concept de prévalence postérieure, puisque les quatre gènes homéotiques les plus en 3' du complexe *Ant-C (Lab, Pbx, Def et Scr)* seraient dominants sur le gène *Antp*. Suite à des expériences

de perte et de gain de fonction, il semblerait que le gène *Antp* serait le premier gène *Hox* ancestral et que ceux situés de part et d'autre auraient évolué par duplication en "images miroirs" (Gehring *et al.*, 2009). Une schématisation des différents concepts liés à la prévalence postérieure est présentée à la Figure 1.8.

L'hypothèse voulant que le phénomène de prévalence postérieure soit conservé chez les vertébrés a été formulée suite à l'incapacité d'un traitement avec le RA à rétablir l'expression des gènes *Hox* situés les plus en 3' des complexes chez le poulet (section 1.3.4.1.), alors que les gènes situés plus en 5' des complexes occasionnent une duplication des structures lorsqu'ils sont exprimés de façon ectopique au niveau du site de traitement (Duboule et Morata, 1994). De plus, comme les transformations homéotiques associées à l'expression ectopique d'un gène *Hox* surviennent toujours dans les régions antérieures au domaine d'expression endogène à ce dernier (Cribbs *et al.*, 1995), plusieurs hypothèses ont été émises pour tenter d'expliquer le mode d'action de la prévalence postérieure. Par exemple, l'efficacité supérieure des produits des gènes *Hox* postérieurs à reconnaître le site de liaison à l'ADN, ou bien à interagir avec les cofacteurs comparativement aux gènes plus antérieurs (LaRonde-LeBlanc et Wolberger, 2003), pourrait expliquer ce phénomène. De plus, cette observation suggère que le treizième résidu du domaine homéo ainsi que le bras N-terminal, qui diffèrent entre les protéines HOX antérieures et postérieures, joueraient un rôle central dans l'efficacité de liaison à l'ADN. D'ailleurs des expériences où il y a échange de domaine homéo entre deux protéines abondent dans ce sens (Zhao et Potter, 2002). La génération de souris transgéniques exprimant ectopiquement certains gènes *Hox* postérieurs ont également confirmé la légitimité du concept de prévalence postérieure chez les mammifères (van der Hoeven *et al.*, 1996; Peichel *et al.*, 1997). Même si la prévalence postérieure demeure un concept relativement accepté dans le domaine, aucune preuve moléculaire directe n'a jamais vraiment confirmé sa validité hors de tout doute. Donc, afin d'avoir une vue d'ensemble représentative de la régulation de l'expression des gènes *Hox* en général, il est important que la lumière soit faite sur ce phénomène.

1.3.4 Les facteurs agissant en *trans*

Mis à part les mécanismes d'autorégulation et de régulation croisée, qui ont été

respectivement décrits dans les sections 1.3.1 ainsi que 1.3.2, et qui font intervenir les produits des gènes *Hox* dans la modulation de leur propre expression, plusieurs autres facteurs sont en mesure de réguler l'activité transcriptionnelle des gènes *Hox* en se fixant directement à leur séquences d'ADN. Les premières pistes quant aux acteurs potentiels agissant en *trans* au niveau des éléments régulateurs de l'expression des gènes *Hox* vinrent des données d'hybridation *in situ* concernant leur expression spatio-temporelle. En effet, puisqu'ils sont initiés durant la gastrulation et qu'ils sont essentiels à la spécification de l'axe AP, l'acide rétinoïque (RA), les facteurs de transcription KROX20, KREISLER ainsi que CDX, de même que certains membres des voies de signalisation Wnt, FGF et Notch (Pearson *et al.*, 2005; Lengerke *et al.*, 2008) constituent des acteurs potentiels dans le contrôle de l'expression des gènes *Hox*. Par la suite, les approches de transgénèse faisant intervenir des souris *Hox/lacZ* ont permis, dans un premier temps, de confirmer les résultats d'hybridation *in situ* de plusieurs gènes *Hox* localisés en 3' des complexes, et, dans un deuxième temps, de circonscrire les régions régulatrices d'intérêt par délétions successives (Schughart *et al.*, 1991; Vogels *et al.*, 1993; van der Hoeven *et al.*, 1996; Gould *et al.*, 1997; Charité *et al.*, 1998; Sharpe *et al.*, 1998; Manzanares *et al.*, 2001). Quant aux gènes *Hox* situés plus en 5', ils semblent répondre à d'autres mécanismes régulateurs. D'ailleurs, la régulation à distance de l'expression des gènes situés en 5' du complexe *HoxD* au niveau des membres fera l'objet de la section 1.3.5.

1.3.4.1 L'acide rétinoïque

L'acide rétinoïque, un métabolite actif de la vitamine A, est synthétisé essentiellement par la rétinaldéhyde déshydrogénase de type II (RALDH2) à l'intérieur du mésoderme proximal dès la gastrulation (Durstun *et al.*, 1989), alors qu'il est dégradé par le cytochrome P450 (CYP26). Suite à sa synthèse, le RA est pris en charge par la protéine de transport CRABP1/2 qui s'occupe de son importation nucléaire et de le présenter à ses récepteurs cibles (Dollé *et al.*, 1989). Il existe deux classes de récepteurs nucléaires appartenant à la superfamille des hormones stéroïdiennes qui sont respectivement les récepteurs de l'acide rétinoïque (RAR), capables de lier à la fois toutes les formes *trans* ainsi que la forme 9-*cis* RA, et les récepteurs aux rétinoïdes X (RXR), qui reconnaissent seulement la forme 9-*cis*. Ces deux classes se divisent respectivement en trois membres:

Figure 1.8: (A) La prévalence postérieure au sein des gènes *Hox* chez la drosophile. L'effet dominant des gènes plus postérieurs sur les plus antérieurs est représenté. (B) La prévalence postérieure contrôlée par l'action des microARNs (miRNAs). *miR-10* et *miR-196* répriment tous les deux l'expression de certains gènes *Hox*. Un seul complexe est schématisé et l'action sur le groupe paralogue est représentée par une flèche rouge. (C) Postulat du modèle de Gehring pour un état basal ancestral représenté par la fonction du gène *Antennapedia (Hox6)*, au niveau du second segment thoracique, qui est dominée à la fois par les gènes *Hox* antérieurs et postérieurs (tirée de Durston *et al.*, 2011. Autorisation reçue de la part de(s) auteur(s).)

A

Prévalence postérieure chez la drosophile

ANTP < BXC < AbdA < AbdB
HOX6 < HOX7 < HOX8 < HOX9

B

Prévalence postérieure médiée par les miRNA



C

Modèle de Gehring

HOX1 > HOX2 > HOX3 > HOX4 > HOX5 > HOX6 < HOX7 < HOX8 < HOX9

NR1B1 (RAR α), NR1B2 (RAR β), NR1B3 (RAR γ), NR2B1 (RXR α), NR2B2 (RXR β) et NR1B3 (RXR γ), chacun possédant de multiples isoformes (Chambon, 1996). Une fois activés, ces récepteurs s'hétérodimérisent au niveau d'un site que l'on nomme "élément de réponse à l'acide rétinoïque" (RARE) (Kastner *et al.*, 1997). Il s'agit d'un motif consensus en tandem séparé généralement par 2 ou 5 pb (DR2 ou DR5). Le nombre important d'exemples recensés concernant les altérations du profil d'expression de gènes *Hox* induites par le RA, et ce dans les embryons de plusieurs espèces différentes de vertébrés (Conlon et Rossant, 1992; Maden et Holder, 1992; Marshall *et al.*, 1992; Morriss-Kay, 1993), fait des RAR/RXR des candidats logiques pour le contrôle de l'expression des gènes *Hox*, d'autant plus que les souris mutantes pour le récepteur RAR γ présentent des transformations homéotiques (Lohnes *et al.*, 1993) dont certaines rappellent celles observées chez les souris *Hoxb4*^{-/-} (Ramirez-Solis *et al.*, 1993). Plusieurs RARE fonctionnels ont été identifiés dans les séquences d'ADN flanquant certains gènes *Hox* comme *Hoxa1*, *Hoxb1*, *Hoxa4*, *Hoxb4* et *Hoxd4* (Langston et Gudas, 1992; Pöpperl et Featherstone, 1993; Marshall *et al.*, 1994; Dupé *et al.*, 1997; Gould *et al.*, 1998; Packer *et al.*, 1998; Zhang *et al.*, 2000; Huang *et al.*, 2002). En ce qui concerne le RARE situé dans la séquence intergénique *Hoxa4-Hoxa5*, bien que le traitement des souris transgéniques *Hoxa4/lacZ* au RA induise une expression ectopique du gène rapporteur au niveau du tube neural ainsi que du cerveau postérieur primitif, la boucle d'autorégulation est nécessaire à l'induction (Packer *et al.*, 1998). De plus, le RARE localisé dans la séquence intergénique *Hoxb4-Hoxb5* est également responsable de l'expression des gènes *Hoxb5*, *Hoxb6* et *Hoxb8* au niveau du cerveau postérieur primitif (Oosterveen *et al.*, 2003a). De plus, il a été observé dans la lignée cellulaire de tératocarcinome ovarien humain PA-1, transformée par la protéine N-RAS, que l'activation de l'expression des gènes du complexe *HoxA* par RA est retardée (Doerksen *et al.*, 1996). Ceci s'explique, entre autres, par une compétition au niveau d'un élément régulateur, situé en 5' du promoteur du gène *Hoxa4* chez l'humain, entre un site RARE, agissant conjointement avec un site de liaison pour la protéine AP2 comme activateurs transcriptionnels, et une boucle de rétrocontrôle négatif enclenchée par la surexpression du gène *Ap2* dépendante de la protéine N-RAS.

En réponse au RA, le rôle de facteur de transcription rempli par les différents récepteurs nucléaires activés permet également d'initier des modifications au niveau de la

structure de la chromatine et ce en recrutant aussi bien des coactivateurs que des corépresseurs (Weston *et al.*, 2003b; Chamberyron et Bickmore, 2004). En effet, en absence de RA, les hétérodimères RAR/RXR induisent un état chromatinien transcriptionnellement inaccessible en mobilisant des corépresseurs tels que les récepteurs nucléaires corépresseurs (NCoR) et les déacétylases des histones (HDAC) (Weston *et al.*, 2003a). Bien que l'effet répresseur sur l'expression des gènes *Hox* qu'exercent ces récepteurs en l'absence de leur ligand soit peu documenté, au moins un RARE régulant négativement leur expression a été décrit dans le cas du gène *Hoxb1* (Studer *et al.*, 1994). En présence de RA, on observe un changement allostérique du domaine d'interaction protéique qui permet cette fois-ci le recrutement de coactivateurs possédant une activité histone acétyltransférase (HAT), telle que la protéine de liaison à CREB (CBP/p300). De plus, à l'aide de lignées de cellules souches embryonnaires (ESC) mutantes pour RAR γ , il a été démontré que ce récepteur est essentiel à l'activation des complexes *HoxA* et *HoxB*. On observe également une cinétique de réorganisation épigénomique, où ni l'abolition complète de la marque répressive de trimétylation de la lysine en position 27 de la queue de l'histone H3 (H3K27me3), exécutée par les protéines Polycomb, ni la présence de RAR γ ne sont requises à l'établissement de l'état bivalent de transcription des complexes *Hox*. Toutefois, le signal RA/RAR γ est nécessaire à l'effacement de la marque H3K27me3 et à l'activation des gènes *Hox* lors de la différenciation des ESC (Kashyap *et al.*, 2011). Une autre étude a permis de découvrir que la production de RA par le mésoderme paraxial active des RARs liés constitutivement à certains sites génomiques qui à leur tour initient le recrutement additionnel de RAR au niveau de domaines chromatiniens à proximité des gènes *Hox* appartenant aux groupes paralogues 1 à 5 (Mahony *et al.*, 2011). De plus, le système de différenciation *in vitro* des ESC de souris en neurones moteurs spinaux (Wichterle *et al.*, 2002) a révélé le recrutement de RARs en 3' des complexes, de même que la perte de la marque H3K27me3, menant à l'expression des facteurs de transcription HOXA4, HOXA5, HOXC4 et HOXC5 au niveau des neurones moteurs cervicaux (Peljto *et al.*, 2010; Mazzoni *et al.*, 2013). Les sections 1.3.7 et 1.3.8 traiteront respectivement de l'organisation tridimensionnelle de la chromatine et des protéines du groupe Polycomb (PcG) de façon plus approfondie.

Enfin, bien qu'une quantité importante de connaissances sur l'activation des gènes *Hox* par RA a été acquise en étudiant la spécification du cerveau postérieur primitif (Wilkinson *et al.*, 1989; Morrison *et al.*, 1996; Folberg *et al.*, 1997; Gavalas et Krumlauf, 2000; Niederreither *et al.*, 2000), la régulation de l'expression des gènes *Hox* par le RA participe également à l'ontogenèse de plusieurs autres organes (Wilson *et al.*, 1953). En effet, l'altération du niveau d'expression de plusieurs acteurs du sentier signalitique de RA (*Raldh2*, *Rara* β/γ , *Rxra* β/γ , *Cyp26*) entraîne des malformations sévères au niveau du cœur (Mic *et al.*, 2002), des membres (Riddle *et al.*, 1993), de même que plusieurs organes émergeant de l'intestin antérieur primitif tels que le pharynx et le pancréas (Bayha *et al.*, 2009), ainsi qu'un rôle prépondérant dans le développement pulmonaire (Rossant *et al.*, 1991; Malpel *et al.*, 2000; Mollard *et al.*, 2000; Wongtrakool *et al.*, 2003; Desai *et al.*, 2006; Chen *et al.*, 2007). De plus, le site RARE situé dans la séquence intergénique *Hoxa4-Hoxa5* est essentiel et suffisant à l'expression du transgène *Hoxa4/lacZ* au niveau du poumon et de l'intestin proximal (Packer *et al.*, 1998, 2000), alors que l'induction de l'expression des gènes *Hoxb5* et *Hoxb6* est détectée après traitement de fibroblastes de poumons embryonnaires de souris (MEF) au RA (Bogue *et al.*, 1994; Volpe *et al.*, 1997; Chinoy *et al.*, 1998).

1.3.4.2 Les protéines CDX

Les gènes *Cdx* codent pour des facteurs de transcription à boîte homéo qui sont homologues à la protéine Caudal, laquelle est exprimée initialement dans la région postérieure de l'embryon de drosophile au niveau de l'abdomen et de l'intestin postérieur primitif (Macdonald et Struhl, 1986; Mlodzik et Gehring, 1987). Chez la souris, on retrouve trois protéines CDX homologues (CDX1, CDX2 et CDX4) qui se lient à la séquence d'ADN (A/C)TTTATA(A/G) afin de réguler directement les gènes *Hox* au niveau du mésoderme ainsi que du neurectoderme, et ce de manière dose dépendante (Margalit *et al.*, 1993; Subramanian *et al.*, 1995; Pownall *et al.*, 1996; Charité *et al.*, 1998; Isaacs *et al.*, 1998; Gaunt *et al.*, 2004). De plus, les phénotypes observés chez les différentes souris mutantes pour les gènes *Cdx* concernant l'établissement de l'axe AP suggèrent qu'ils affectent l'expression des gènes *Hox* de la région cervicale à la région caudale du squelette axial (van den Akker *et al.*, 2002; Houle *et al.*, 2003). Cette régulation spatio-temporelle

s'effectue grâce aux éléments de réponse aux CDX (CDRE) présents dans les séquences d'ADN adjacentes aux gènes *Hox* ciblés. Ainsi, des CDRE fonctionnels orchestrent l'expression des gènes *Hoxa7*, *Hoxb8*, *Hoxc8* (Knittel *et al.*, 1995; Shashikant *et al.*, 1995; Subramanian *et al.*, 1995; Haerry et Gehring, 1996; Taylor *et al.*, 1997; Charité *et al.*, 1998; Gaunt, 2000; Gaunt *et al.*, 2004).

Le gène *Cdx1* est exprimé initialement à la fois dans l'ectoderme et le mésoderme de la PS au jour E7.5 (Meyer et Gruss, 1993). Un peu plus tard, on le retrouve au niveau de la corde spinale présomptive, de certaines populations de CCN, du mésoderme paraxial ainsi qu'au niveau des somites, où son expression deviendra restreinte au dermomyotome à E10.5. Au même stade embryonnaire, on observe également une expression mésenchymale du gène *Cdx1* au niveau des bourgeons des membres supérieurs et, de plus faible intensité, au niveau des membres inférieurs. De plus, les souris *Cdx1*^{-/-} possèdent un spectre de transformations homéotiques d'antériorisation, localisé entre C1 et T8 inclusivement, et associé à une postériorisation du domaine d'expression des gènes *Hoxc6* et *Hoxa7* (Subramanian *et al.*, 1995).

Parmi les trois gènes *Cdx*, c'est le gène *Cdx2* qui a l'expression la plus précoce en débutant à E3.5 au niveau des cellules trophoblastiques (Beck *et al.*, 1995). Il est aussi retrouvé sensiblement dans les mêmes tissus que le gène *Cdx1*, en plus de la base de l'allantois et de l'intestin postérieur primitif. De plus, les embryons *Cdx2*^{-/-} sont incapables de s'implanter adéquatement et meurent à E3.5. C'est donc en étudiant les embryons *Cdx2*^{+/-} que des transformations homéotiques comparables à celles observées chez les mutants *Cdx1*^{-/-} ont pu être répertoriées. De façon intéressante, malgré le fait que la protéine CDX2 a une frontière rostrale d'expression située au niveau du PSM, la perte de sa fonction altère tout de même l'expression des gènes *Hox* au niveau cervical, suggérant une interaction précoce avec ses gènes *Hox* cibles au niveau du PSM (van den Akker *et al.*, 2002). D'ailleurs, l'expression des gènes *Hoxb1*, *Hoxd4* et *Hoxb8* est diminuée chez les embryons mutants *Cdx2* HM générés par agrégation d'embryons tétraploïdes. Enfin, les souris HT pour le gène *Cdx2* sont plus petites à la naissance et possèdent une queue tronquée, suggérant pour la première fois l'implication des protéines CDX dans l'élongation de l'axe AP (Chawengsaksophak *et al.*, 1997; 2004). À noter également le rôle du gène *Cdx2* dans

l'ontogenèse du système digestif étant exprimé dans les cellules prolifératives de la crypte (Freund *et al.*, 1998).

Tout comme le gène *Cdx1*, l'expression du gène *Cdx4* est initiée dès E7.5, mais de façon plus faible et plus postérieure au niveau de la PS. Comparativement aux gènes *Cdx2* et *Cdx1* qui s'éteignent respectivement dans la partie postérieure à E12.5 et E12.0, l'expression du gène *Cdx4* disparaît après E10.5 (Gamer et Wright, 1993). Les souris *Cdx4*^{-/-} sont viables, fertiles, et ne présentent aucun phénotype apparent. Toutefois, l'implication du gène *Cdx4* dans l'établissement du squelette axial est révélée chez les mutants doubles HT avec l'un ou l'autre de ses homologues (van Nes et al., 2006). D'ailleurs, le phénotype de troncation de l'axe AP observé chez les individus mutants *Cdx2*^{+/-}*Cdx4*^{-/-} est sauvé en forçant l'expression des gènes *Hoxa5* et *Hoxb8*, suggérant la régulation de leur expression par les protéines CDX lors de ce processus développemental (Young *et al.*, 2009). La régulation de l'expression du gène *Hoxd4a* par *Cdx4* lors de la spécification des hémangioblastes a également été observée chez le poisson zèbre (Amali *et al.*, 2013).

1.3.4.3 Voies de signalisation FGF, Wnt et Notch

Le réseau complexe d'interactions génétiques contrôlant l'expression des gènes *Hox* lors des processus d'élongation et de spécification de l'axe AP, de même que la segmentation mésodermique, font intervenir différents signaux tels que ceux contrôlés par le RA (section 1.3.4.1), les facteurs de croissance fibroblastique (FGF), de même que par les protéines WNT et NOTCH (Deschamps et van Nes, 2005). En plus de l'action directe du RA sur la transcription des gènes *Hox*, et l'action potentiellement directe du FGF et de la protéine WNT, ces voies contrôlent également indirectement l'expression des gènes *Hox* par l'intermédiaire des facteurs de transcription CDX. Un résumé sous forme schématisée de l'intégration de ces différents sentiers signalétiques dans la régulation de l'expression des gènes *Hox* est présenté à la Figure 1.9.

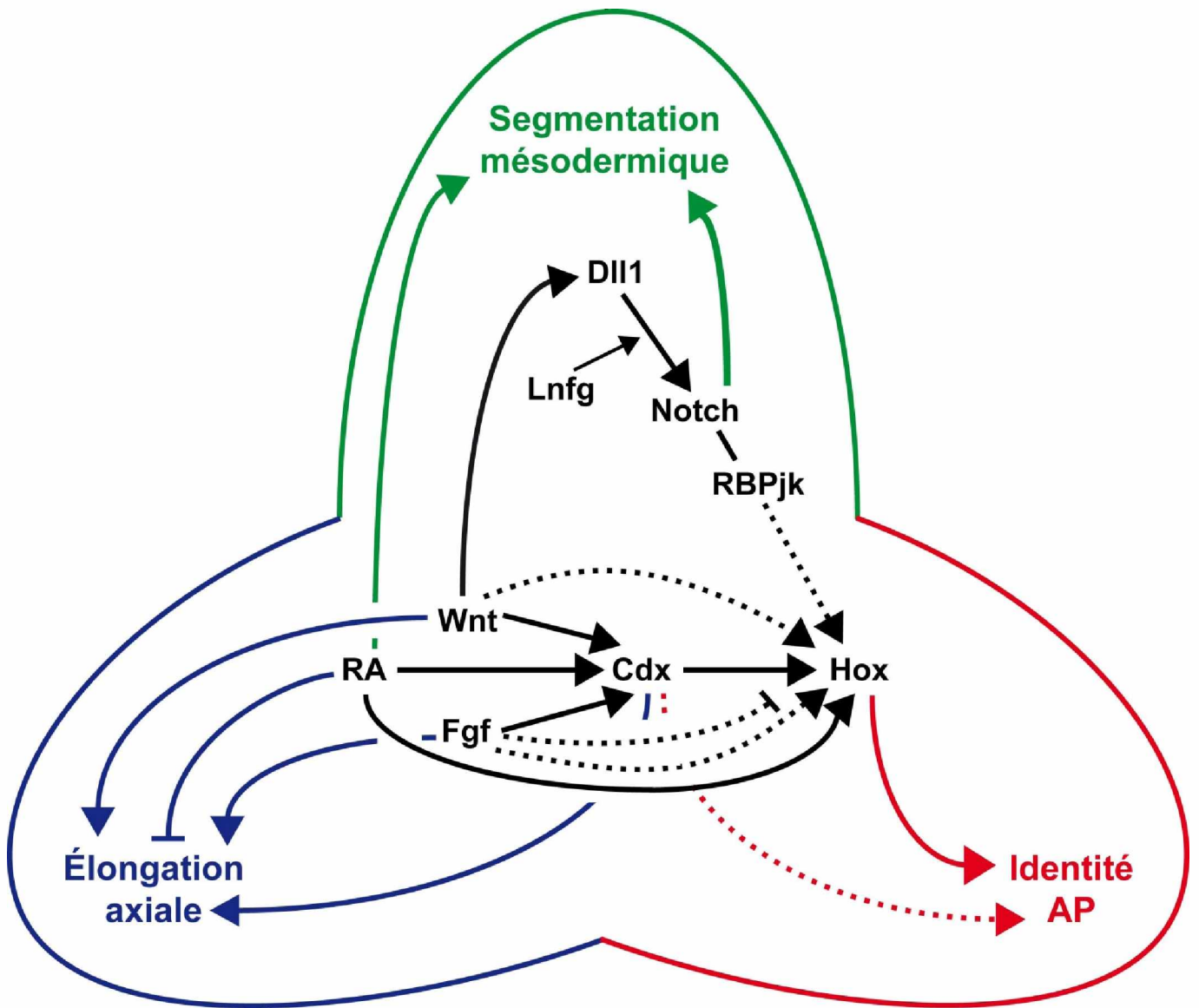
Bien qu'aucune évidence moléculaire n'ait pu confirmer si le rôle qu'exercent les FGFs sur l'expression des gènes *Hox* est direct ou indirect, plusieurs exemples dans la littérature montrent que l'altération de la voie FGF altère l'expression de certains gènes *Hox* (Ciruna et Rossant, 2001). Par exemple, la perte partielle de la transmission du signal FGF,

causée par une mutation hypomorphe au niveau du gène codant pour le récepteur FGFR1, entraîne une postérieurisation du profil d'expression des gènes *Hoxd4* ainsi que *Hoxb9* au niveau du mésoderme latéral (Partanen *et al.*, 1998). L'induction de l'expression du gène *Hoxb8* dans la partie caudale d'embryons de souris, suite à l'administration de FGF, est également observée (Bel-Vialar *et al.*, 2002; Delfino-Machín *et al.*, 2005). De plus, des expériences de greffes hétérotopiques de l'isthme, source de FGF à la frontière des cerveaux moyen et postérieur, abolissent l'expression du gène *Hoxa2* au niveau des CCN dérivant du second arc branchial (Trainor *et al.*, 2002). Ces résultats corrént avec la transformation homéotique observée au niveau du second arc branchial des souris mutantes *Hoxa2*^{-/-}, de même qu'avec la perte de son expression lors d'une expression ectopique du gène *Fgf8* dans un système de greffe caille-poulet (Irving et Mason, 2000).

De la même façon, malgré les indices suggérant que le sentier signalétique WNT contrôlerait l'expression des gènes *Hox* durant la progression antérieure de la PS (Forlani *et al.*, 2003), il est impossible de confirmer qu'il s'agit d'un rôle direct ou indirect faute de données moléculaires. Par contre, l'analyse de différents mutants *Wnt3a* montre une postériorisation du domaine d'expression des gènes *Hoxd3*, *Hoxd13*, ainsi que *Hoxb4* (Ikeya et Takada, 2001).

Plusieurs évidences moléculaires concernant l'action indirecte des voies de signalisation RA, FGF et Wnt sur l'expression des gènes *Hox* via l'intermédiaire des protéines CDX sont recensées dans la littérature (Isaacs *et al.*, 1998; Ikeya et Takada, 2001; Sakai *et al.*, 2001; Walshe *et al.*, 2002; Abu-Abed *et al.*, 2003; Chawengsaksophak *et al.*, 2004; Pilon *et al.*, 2006; Shimizu *et al.*, 2006; Skromne *et al.*, 2007). Les protéines Cdx permettent, entre autres, l'intégration du signal complexe provenant des gradients opposés de RA et de FGF, ainsi que du sigal WNT, afin de transmettre l'information aux gènes *Hox* pour qu'ils puissent remplir leurs fonctions. Par exemple, le signal RA active l'expression des gènes *Hox* en 3', alors que le signal FGF active l'expression des gènes plus caudaux membres des paralogues *Hox6* à *Hox9* (Bel-Vialar *et al.*, 2002; Nordström *et al.*, 2006), qui contrôlent l'identité épinière brachiale et thoracique lors de la spécification régionale du CNS (Liu *et al.*, 2001). De plus, les gradients rostro-caudaux opposés de RA et de FGF, de concert avec la siganlisation WNT, contrôlent également l'expression des gènes *Hox*

Figure 1.9: Le réseau d'interactions génétiques impliquant les gènes *Hox* dans le contrôle de l'élongation axiale (orange), de la segmentation du mésoderme (vert) ainsi que dans l'établissement de l'axe AP (violet). La relation entre les processus morphogénétiques initiés par les voies de signalisation impliquant FGF, WNT, RA ainsi que NOTCH, passant ou non par les protéines CDX, et l'expression des gènes *Hox* est représentée. Les lignes pleines correspondent à des interactions confirmées, alors que les lignes pointillées consistent en des interactions qui n'ont pas été confirmées au niveau moléculaire (tirée de Deschamps et Van Nes, 2005. Autorisation reçue de la part de(s) auteur(s).)



lors de la spécification des somites (Dubrulle *et al.*, 2001; Aulehla *et al.*, 2003; Dubrulle et Pourquié, 2004a). Il ne s'agit cependant pas d'un modèle universel, puisque les sentiers RA et FGF peuvent également agir dans la même direction lors de l'établissement de l'axe AP (Pownall *et al.*, 1998; Diez del Corral et Storey, 2004; Bayha *et al.*, 2009) et que l'action des sentiers WNT et FGF sur l'élongation axiale ainsi que sur la formation des somites, via l'intermédiaire des protéines CDX, ne font pas forcément intervenir les gènes *Hox* (Savory *et al.*, 2009).

En plus de l'influence des morphogènes, l'expression des gènes *Hox* dans le PSM antérieur est affectée par la mutation de gènes impliqués dans le processus de segmentation, comme la perte ou le gain de fonction de membres du sentier de signalisation NOTCH (inactivation de l'effecteur *RBPjk*, activation d'une forme dominante négative du ligand DLL1) (Zákány *et al.*, 2001; Cordes *et al.*, 2004). De plus, l'effet sur l'expression des gènes *Hox* découlant de l'altération de la voie WNT au niveau du PSM régulerait directement la voie NOTCH (Hofmann *et al.*, 2004; Galceran *et al.*, 2004).

1.3.5 Partage et sélectivité d'éléments régulateurs

Plusieurs éléments de réponse aux différents facteurs agissant en *trans*, décrits à la section 1.3.4, font partie de régions régulatrices situées à l'intérieur des séquences intergéniques d'ADN qui séparent les loci des gènes *Hox* au sein du même complexe. Par conséquent, des mécanismes complexes de régulation contrôlant l'expression spatio-temporelle des gènes *Hox* ont été acquis au cours de l'évolution. Le partage et la sélectivité d'un élément régulateur pour différents promoteurs de gènes *Hox* en sont des exemples.

Le partage de séquences régulatrices agissant en *cis* permet de cibler l'expression de gènes *Hox* voisins dans les mêmes tissus ou cellules. Par exemple, la séquence intergénique *Hoxb3-Hoxb4* renferme un élément nécessaire à leur expression conjointe au niveau du cerveau postérieur (Gutman *et al.*, 1994; Gould *et al.*, 1997). Ensuite, un élément situé à l'intérieur de la séquence intronique du gène *Hoxb4* agit positivement sur l'établissement des limites d'expression des gènes *Hoxb4* et *Hoxb5* au niveau du tube neural (Whiting *et al.*, 1991). De plus, la séquence intergénique *Hoxb4-Hoxb5* contient une autre

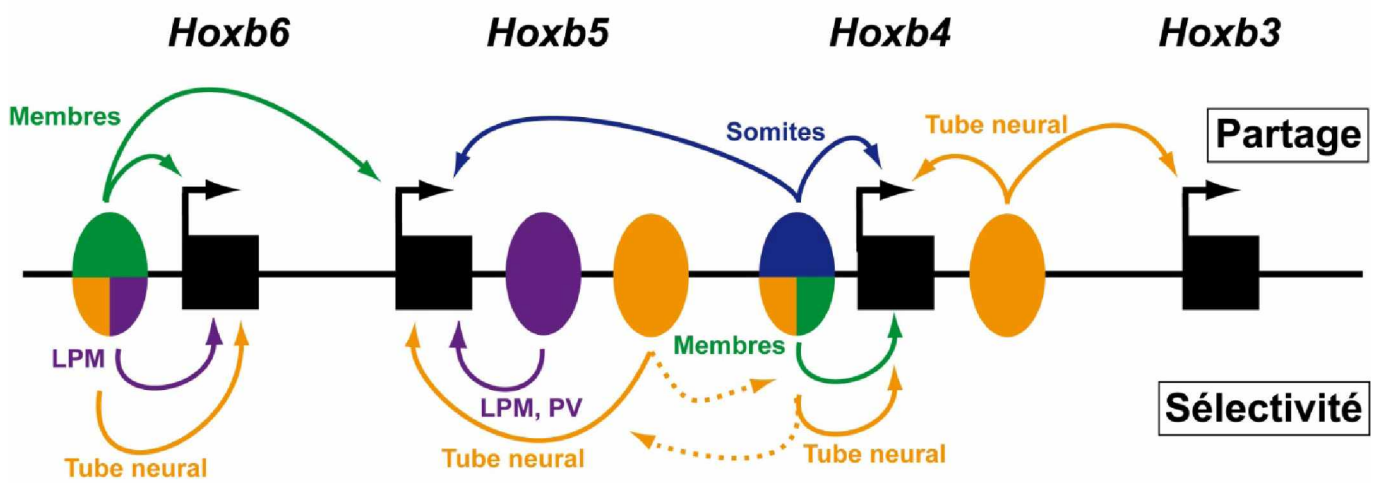
séquence d'ADN agissant en *cis*, ayant pour fonction de promouvoir l'expression de ces deux gènes au niveau des somites. Enfin, un élément localisé en 5' du gène *Hoxb6* dirige son expression, de même que celle du gène *Hoxb5* dans les membres (Sharpe *et al.*, 1998). De la même façon, les gènes *Hoxd10* et *Hoxd11* partagent deux éléments régulateurs différents (Zákány *et al.*, 1997b). Il s'agit d'une séquence régulatrice agissant positivement sur leur expression dans le mésoderme paraxial, alors que l'autre empêche leur transcription afin de fixer leur frontière antérieure au niveau du squelette axial.

Toutefois, puisque les domaines d'expression des gènes *Hox* voisins ne se chevauchent que partiellement, des stratégies conférant une spécificité à certaines séquences régulatrices doivent également s'ajouter aux mécanismes partagés afin que le protocole de différenciation initié par le code *Hox* soit respecté. La grande région d'ADN englobant les locus des gènes *Hoxb3* à *Hoxb6* est également riche en éléments régulateurs spécifiques à certains promoteurs. Un schéma récapitulatif des séquences régulatrices qui sont respectivement partagées ou spécifiques aux promoteurs de ces gènes *Hox* est présenté à la Figure 1.10 (Sharpe *et al.*, 1998). On remarque, dans un premier temps, que l'expression au niveau des membres, dirigée par un élément situé entre les gènes *Hoxb4* et *Hoxb5*, n'agit que sur *Hoxb4*. Ensuite, leur expression temporelle au niveau du tube neural est aussi régulée différemment puisque celle du gène *Hoxb4* est plus précoce que celle du gène *Hoxb5*. De plus, il existerait une compétition pour la séquence d'ADN agissant en *cis* sur l'expression neurale de ces gènes, où le promoteur du gène *Hoxb4* serait préférentiellement activé. Enfin, l'élément qui dirige conjointement l'expression des gènes *Hoxb5* et *Hoxb6* au niveau des membres est responsable de l'expression spécifique du gène *Hoxb6* à l'intérieur du SNC ainsi que du LPM.

L'étude du complexe *HoxD* a également permis de faire la lumière sur un système complexe de partage et de sélectivité d'éléments régulateurs contrôlant l'expression de ces gènes au niveau des membres. Une analyse par "tilling array" ainsi que des délétions ciblées dans un contexte de transgène ont permis de mettre en évidence que, bien que la relation entre l'ordre des gènes au sein du complexe *HoxD* et leur frontière d'expression rostrale respective respecte la colinéarité spatiale, leur distribution n'est pas complètement progressive le long de l'axe. En effet, il existe deux frontières antéro-postérieures, associées

Figure 1.10: Modèle résumant les différents cas de partage, de même que les actions spécifiques des éléments régulateurs situés dans les séquences non codantes des gènes *Hoxb* antérieurs. Les différentes séquences régulatrices sont représentées par des ovales et dirigent l'expression des gènes *Hoxb* au niveau des membres (vert), des somites (bleu), LPM (violet) et dans le tube neural (jaune). Les flèches pointillées représentent des interactions qui varient en fonction du promoteur (Tirée de Sharpe *et al.*, 1998. Autorisation reçue de la part de(s) auteur(s).)

Complexe *HoxB*



aux subdivisions fonctionnelles de la moëlle épinière, qui sont contrôlées par deux blocs transcriptionnels différents (Tschopp *et al.*, 2012). Ce type de régulation faisant intervenir des éléments qui contrôlent globalement l'expression de plusieurs gènes *Hox*, et ce à distance, fait l'objet de la section suivante.

1.3.6 Régulation à distance

L'expression spatio-temporelle finement contrôlée des gènes *Hox* lors du développement des différentes structures complexes des animaux bilatériens nécessite la mise en place d'un réseau régulateur sophistiqué intégrant de nombreuses stratégies complémentaires. Ainsi, des séquences régulatrices séparées par plusieurs kilobases (kb) de leurs promoteurs cibles s'ajoutent aux éléments plus classiques, situés à proximité des gènes qu'ils régulent, pour orchestrer adéquatement la transcription coordonnée des gènes *Hox* dans leurs domaines respectifs.

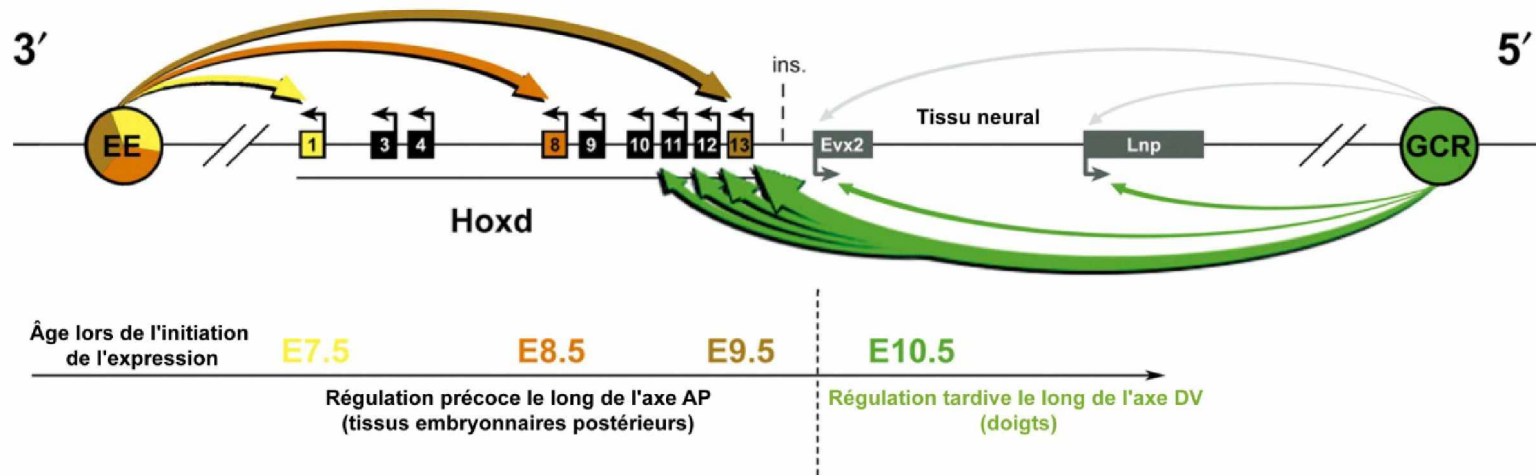
La régulation à distance est un mécanisme conservé dans plusieurs espèces, notamment chez la drosophile (Dorsett, 1999), où plusieurs séquences activatrices/amplificatrices (IAB), situées dans la région intergénique *AbdA-AbdB*, exercent un contrôle transcriptionnel sur ces deux gènes, de même que sur *Ubx* (Bender *et al.*, 1983; Estrada *et al.*, 2002; Sipos et Gyurkovig, 2005). Par exemple, l'élément activateur/amplificateur IAB5 dirige spécifiquement l'expression du gène *AbdB*, et ce malgré 55 kb de séquence d'ADN ainsi qu'au moins trois éléments isolateurs qui le séparent du promoteur. De plus, un élément (PTE) localisé à 59 kb en 5' du promoteur du gène *AbdB*, donc à 114 kb de l'élément IAB5, est nécessaire pour que ce dernier remplisse ses fonctions de séquences activatrices/amplificatrices (Ho *et al.*, 2011).

Tel que mentionné à la section 1.3.5, la majorité des informations recueillies jusqu'à maintenant en ce qui concerne la régulation à distance chez la souris provient des études faites sur le complexe *HoxD*. Toutefois, des exemples de régulation à distance impliquant d'autres complexes *Hox* ont aussi été répertoriés. C'est le cas d'un élément régulateur, localisé à environ 15 kb en 5' du promoteur du gène *Hoxc8*, qui est responsable d'établir sa limite antérieure d'expression au niveau des somites et du tube neural (Bradshaw *et al.*,

1996). Ceci étant dit, cette section est consacrée essentiellement aux différents modules régulant à distance l'expression des gènes *Hoxd*. D'ailleurs, une représentation schématique des séquences régulatrices agissant à distance sur l'expression des gènes appartenant au complexe *HoxD* est présentée en Figure 1.11.

En premier lieu, un élément régulateur localisé en 3' à l'extérieur du complexe a été proposé comme remplissant la fonction de diriger l'expression colinéaire précoce des gènes *Hoxd* au niveau du LPM du territoire présomptif du membre le long de l'axe AP. Cet élément, appelé la région de contrôle précoce des membres (ELCR ou EE), a été découvert grâce à des approches d'inversions/délétions de régions d'ADN du complexe *HoxD* chez la souris (Zákány et al., 2004). Par conséquent, l'expression des gènes du complexe *HoxD* au niveau du mésenchyme du bourgeon de membre en développement suit les mêmes règles de colinéarité que l'expression précoce dans les tissus axiaux et paraxiaux. À l'autre extrémité, on retrouve une région de contrôle global (GCR), située à 240 kb en 5' du complexe *Hoxd*, qui dirige l'expression coordonnée de gènes répartis sur un grand domaine chromosomique (Spitz et al., 2003), dont les gènes *Hoxd10* à *Hoxd13*, de même que *Evx2* et *Lnp*. Cette région contrôle non seulement l'expression de ces 6 gènes au niveau du mésenchyme distal des membres lors de la formation des doigts, mais abrite également des éléments dirigeant l'expression des gènes *Evx2* et *Lnp* au niveau neural. Cependant, une région isolatrice empêche l'action des séquences régulatrices neurales sur les gènes *Hoxd* postérieurs (Kmita et al., 2002b). La découverte du GCR vient également appuyer l'hypothèse de l'acquisition tardive dans l'évolution de la fonction des gènes *Hoxd* dans le développement des membres, comparativement à leur fonction ancestrale le long de l'axe AP. (Spitz et al., 2001; Kmita et al., 2002a). Ainsi, la GCR présente chez les poissons cartilagineux ne contient que les éléments neuraux (Spitz et al., 2003). De plus, ces observations ont fait ressurgir l'hypothèse voulant que la duplication de complexe *Hox* favoriserait l'apparition de différentes régions régulatrices globales. Des expériences utilisant une approche de translocations pangénomiques appelée PANTHERE (PAN-genomic Translocation for Heterologous Enhancer RE- shuffling), visant l'échange hétérologue de séquences activatrices/amplificatrices, ont permis de placer les gènes appartenant au complexe *HoxC* sous le contrôle de la totalité du paysage régulateur du

Figure 1.11: Régulation globale de l'expression du complexe *HoxD*. (Gauche) L'activation séquentielle 3'-5' des gènes du complexe *HoxD* est représentée selon l'hypothèse de l'activateur/amplificateur global précoce (EE) responsable de l'activation temporelle colinéaire des gènes *Hoxd* le long de l'axe AP. Seulement trois flèches sont représentées pour alléger la figure: *Hoxd1* (jaune), *Hoxd8* (orange) et *Hoxd13* (brun). L'élément EE du schéma correspond à la région de contrôle précoce des membres (ELCR) proposée par Zákány et collègues (Zákány *et al.*,2004) (Droite). L'influence de la région de contrôle global (GCR) sur les gènes *Hoxd* postérieurs, de même que les gènes avoisinants (*Evx2* et *Lmp*), au niveau des doigts (vert) et du tube neural (gris). L'épaisseur des flèches vertes représentent la force d'action inversement proportionnelle à la distance de l'élément GCR sur les gènes *Hoxd* postérieurs (tirée de Deschamps et Van Nes, 2005. Autorisation reçue de la part de(s) auteur(s).)



complexe *HoxD* (Tschopp *et al.*, 2011). De façon intéressante, le locus chimérique obtenu est en mesure d'interpréter l'information dictée par ces éléments régulateurs exogènes et de remplir le rôle normalement pris en charge par le complexe *HoxD*. Ces résultats abondent donc dans le sens d'une évolution optant pour l'apparition de nouveaux modules régulateurs au lieu de nouvelles cibles transcriptionnelles lors des processus de néofonctionalisation.

En plus des bourgeons de membres, deux autres régions régulatrices, potentiellement situées à l'extrémité télomérique du chromosome 2 chez la souris, seraient responsables de la dichotomie d'expression des gènes *Hoxd* dans le rein en développement (Di-Pöi *et al.*, 2007). En effet, ces deux éléments régulateurs dirigerait à distance respectivement l'expression des gènes *Hoxd9* à *Hoxd12* au niveau du mésenchyme et *Hoxd1* à *Hoxd9* au niveau du bourgeon urétral.

Enfin, bien que la représentation schématique de la régulation à distance contrôlant l'expression spatio-temporelle des gènes appartenant au complexe *HoxD* est généralement linéaire (en deux dimensions), il faut garder à l'esprit que le contact entre un élément régulateur et un promoteur séparés par plusieurs centaines de kb n'est envisageable que dans un modèle en trois dimensions, lequel sera abordé davantage dans la prochaine section (1.3.7). Par exemple, tel que mentionné précédemment, le développement des membres nécessite deux vagues de transcription des gènes *Hoxd*: une vague précoce lors de la formation du stylo-pode ainsi que du zeugopode, et une vague tardive lors de la formation de l'autopode (Tschopp et Duboule, 2010). Suite à la vague précoce de transcription, le réarrangement de la structure de la chromatine des loci *Hoxd* permet de dissocier progressivement les régions promotrices des gènes *Hoxd1*, *Hoxd3* et *Hoxd4* de la région régulatrice située en 3' du complexe (EE/ELCR). Parallèlement, les changements conformationnels de la chromatine au niveau des loci *Hoxd* promeuvent le contact entre l'élément situé en 5' du complexe (GCR) et les régions promotrices des gènes *Hoxd9*-*Hoxd13* afin d'initier la vague tardive spécifiant l'autopode (Andrey *et al.*, 2013). Ces changements sont également associés à la présence/absence de la marque répressive H3K27me3.

1.3.7 Organisation tridimensionnelle de la chromatine

Notre compréhension de la régulation transcriptionnelle en général a évolué d'une vision réductionniste, centrée sur des facteurs de transcription reconnaissant spécifiquement un élément agissant en *cis* sur l'expression de son gène cible, à une vision globale où les modifications de la chromatine ainsi que son remodelage font partie intégrante de la régulation de l'expression des gènes.

Tout d'abord, des études récentes suggèrent que la localisation spatiale des gènes à l'intérieur du noyau n'est pas aléatoire, mais plutôt organisée de façon à faciliter spécifiquement la régulation de l'expression des gènes remplissant une fonction caractéristique au type cellulaire impliqué (Misteli, 2004). Cette hypothèse prend tout son sens en ce qui concerne les gènes organisés en complexe, puisque la régulation de leur expression doit être coordonnée lors de la différenciation cellulaire et du développement embryonnaire. Les structures formées par les chromosomes au niveau du noyau interphasique, appelées territoires chromosomiques, constituent l'un des compartiments subnucléaires à l'intérieur duquel les gènes ne sont pas accessibles par la machinerie transcriptionnelle (Kosak et Groudine, 2004). Grâce à de nouvelles techniques telles que la capture conformationnelle chromosomique (3C), la relation entre la colinéarité et la structure tridimensionnelle des complexes *Hox* a pu être décrite. Ainsi, lors de la stimulation de la transcription du complexe *HoxB*, induite par un court traitement des ESC de souris au RA, on assiste à une décondensation généralisée de la chromatine au niveau des loci *Hoxb* au moment où seulement le gène *Hoxb1* est activé, lequel forme une boucle chromatinienne relaxée à l'extérieur de son territoire chromosomique (CT) vers le centre du noyau, où il devient accessible à la machinerie transcriptionnelle (Chambeyron et Bickmore, 2004; Wurtele et Cartrand, 2006; Noordermeer *et al.*, 2011). De la même façon, l'initiation de l'expression du gène *Hoxb1* lors de la gastrulation coïncide avec son expulsion du CT à l'intérieur des noyaux des cellules appartenant à la partie postérieure de la ligne primitive (Chambeyron *et al.*, 2005). Quant au gène *Hoxb9*, il faut attendre au jour E9.5 pour qu'il soit exclu du CT dans la partie postérieure du tube neural. Cependant, l'activation des gènes *Hoxd* postérieur lors de la vague tardive d'expression spécifiant l'autopode le long de l'axe PD se produit sans qu'ils soient relocalisés à l'extérieur de leur

CT (Morey *et al.*, 2007). Donc, l'expulsion progressive des loci *Hox* à l'extérieur des CT serait nécessaire à la mise en place de la colinéarité spatio-temporelle le long de l'axe AP et non à l'expression des gènes *Hox* à proprement dit.

La conversion de la chromatine compacte transcriptionnellement inactive vers sa forme relaxée active nécessite l'intervention de plusieurs facteurs dont ceux remodelant la chromatine de même que les enzymes modifiant les histones. Parmi les protéines responsables de la compaction de la chromatine, on retrouve la protéine isolatrice à doigts de zinc CTCF (Bushey *et al.*, 2009; Nègre *et al.*, 2010), laquelle est en mesure de promouvoir les contacts chromatiniens à distance, que ce soit en liant directement la chromatine ou en s'associant à des complexes répressifs impliquant les protéines de la famille Polycomb (section 1.3.8) (Bantignies *et al.*, 2010; Hodgson et Brock, 2011). D'ailleurs, la modélisation tridimensionnelle des complexes *Hox* de différentes espèces suggère que l'action des protéines CTCF aurait été conservée durant l'évolution (Ferraiuolo *et al.*, 2010). De plus, des régions enrichies pour l'homologue murin de la protéine isolatrice GAGA ont également été localisées au niveau des séquences d'ADN intergéniques de plusieurs gènes *Hox* (Srivastava *et al.*, 2013).

Le processus de décondensation de la chromatine est généralement associé à des modifications post-traductionnelles au niveau de la queue des histones. En effet, des expériences d'immunoprécipitation de la chromatine à grande échelle dans la lignée cellulaire de carcinome embryonnaire F9 ont permis de mettre en évidence la liaison des récepteurs RXR α ainsi que RAR γ au niveau du RARE du gène *Hoxa1*, et ce avant même son induction par le RA (Gillespie et Gudas, 2007a; 2007b). Suite au traitement, un enrichissement de la protéine coactivatrice p300 et de l'ARN polymérase II s'amorce au niveau du RARE, alors que la protéine membre de la famille Polycomb SUZ12 est diminuée. Parallèlement, la structure bivalente de la chromatine est observée à la fois par la présence de marques de l'état inactif tel que H3K27me3, ainsi que des marques de l'état actif comme H3ac et H3k4me2. De la même façon, le traitement de cellules progénitrices neuronales avec le RA entraîne la perte des marques H3k27me3 sur la chromatine des gènes membres des groupes paralogues *Hox1* à *Hox5* et est associé à l'acquisition de l'identité spinale cervicale, alors que le signal WNT ou FGF fait de même pour les gènes

des groupes paralogues *Hox1* à *Hox9* (Mazzoni *et al.*, 2013). Quant au traitement des cellules P19 avec le RA, il provoque la perte du complexe protéique Polycomb ASXL1/BAP1 et l'ubiquitination de l'histone H2B, menant à l'activation de gènes *Hox* (Lee *et al.*, 2013). Enfin, au moment de la phase tardive d'expression des gènes postérieurs appartenant au complexe *HoxD* lors de la formation des doigts, la diminution de la protéine Polycomb Ring1B et la perte de la marque H3K27me3, associées à la décompaction de la chromatine, se produit spécifiquement dans les cellules distales du membre postérieur. De plus, l'élément GCR colocalise spatialement avec les gènes en 5' du complexe *HoxD* (Montavon *et al.*, 2011; Williamson *et al.*, 2012).

1.3.8 Les protéines du groupe Polycomb (PcG)

Les gènes *Polycomb* (*Pc*) ont été découverts par le Dr P. Lewis en 1947 alors qu'il observait une mutation dominante produisant un peigne sexuel ectopique sur les deuxième et troisième pattes de la mouche mâle adulte, résultat de la surexpression du gène *Scr*. Il démontrait quelques années plus tard que les mutants homozygotes pour la perte de fonction d'un gène *Pc* présentaient une postériorisation globale des segments embryonnaires, confirmant leur rôle de régulateurs négatifs agissant en *trans* sur les gènes *Hom* (Struhl, 1983; Paro, 1990; Beuchle *et al.*, 2001; Schuettengruber *et al.*, 2007). L'analyse informatique comparative du génome de la drosophile a prédit approximativement 40 gènes appartenant au groupe *Polycomb* (*PcG*) (Jürgens, 1985), lequel peut être subdivisé en trois différentes sous-catégories: le complexe répressif Polycomb de type 1 (PRC1), de type 2 (PRC2), ainsi que le complexe répressif associé à la protéine PHO (PhoRC), également connu sous le nom de complexe répressif de type 0 (PRC0) (Bantignies et Cavalli, 2011; Schwartz et Pirrotta, 2007, 2008). Ces trois sous-catégories seront abordées davantage dans les prochaines sections (1.3.8.1-1.3.8.3).

Afin de maintenir respectivement les états transcriptionnels silencieux ou actifs des gènes *Hox*, les protéines du groupe Polycomb (PcG) et celles du groupe Trithorax (TrX) (section 1.3.9) doivent pouvoir jouer leur rôle antagoniste au niveau de leurs cibles communes (Francis et Kingston, 2001; Ringrose et Paro, 2004; Grimaud *et al.*, 2006). Ainsi, des éléments régulateurs du gène *Ubx* situés en 5' de son promoteur, au niveau

desquels des sites de liaison pour les protéines FTZ, EVE et HB sont retrouvés, contiennent également des éléments de réponse aux protéines Polycomb (PRE) ou Trithorax (TRE) (Muller et Bienz, 1991, 1992; Chan *et al.*, 1994a; Chiang *et al.*, 1995). Il ne s'agit toutefois pas d'une règle absolue, puisque l'expression spécifique de protéines PcG au niveau de différents sites exprimant ou non le gène *Eng* a permis de constater la présence de protéines appartenant à trois complexes PcG et ce peu importe l'état transcriptionnel du gène (Langlais *et al.*, 2012). Ces résultats suggèrent donc que la liaison des protéines PcG n'a pas d'incidence majeure sur l'expression du gène *Eng*.

Des études récentes identifiant différentes séquences d'ADN, comme les îlots CpG par exemple, qui semblent contrôler plusieurs gènes importants lors du développement (*Kreisler, Zfp2, Hoxd11, Hoxd12, etc.*) suggèrent que des éléments similaires aux PRE existeraient également chez les mammifères (Sing *et al.*, 2009; Mendenhall *et al.*, 2010; Woo *et al.*, 2010). Ces observations corrént avec les données recueillies antérieurement suggérant la conservation du rôle des protéines PcG dans le maintien de l'état silencieux des gènes *Hox*. Tout d'abord, la mutation de certains gènes *PcG* chez la souris entraîne l'expression ectopique de gènes *Hox* au niveau antérieur de l'embryon, comme par exemple les gènes *Hoxb3, Hoxd4, Hoxb6* et *Hoxc6* (Akasaka *et al.*, 2001; Rastegar *et al.*, 2004; Kobrossy *et al.*, 2006). De plus, la mutation du gène *M33* provoque l'expression précoce du gène *Hoxd11* (Bel-Vialar *et al.*, 2000). Ensuite, tout comme chez la drosophile, des protéines se liant à l'ADN de façon spécifique seraient nécessaires au recrutement des protéines PRC2, notamment Yin Yang 1 (YY1), l'homologue de la protéine PHO chez la drosophile, qui fait l'objet de la section 1.3.8.3, JARID2 (Peng *et al.*, 2009; Shen *et al.*, 2009; Landeira *et al.*, 2010; Li *et al.*, 2010; Pasini *et al.*, 2010), IKK α (Mulero *et al.*, 2013), *etc.* De façon similaire au mode d'action observé chez la mouche, l'étude des cellules de mammifères a démontré que les protéines PcG sont requises pour la compaction de la chromatine (Eskeland *et al.*, 2010), de même que la formation de boucles d'ADN (Tiwari *et al.*, 2008). Ces deux phénomènes seraient potentiellement associés à la mise en place des marques de mono-ubiquitinylation (Wang *et al.*, 2004; de Napoles *et al.*, 2004 Cao *et al.*, 2005) et de triméthylation (Cao *et al.*, 2002; Ringrose *et al.*, 2004; Li *et al.*, 2010) des différentes lysines présentes sur la queue des histones. Par exemple, le domaine protéique méthyltransférase SET est essentiel à l'expression normale du gène *Hoxc8* dans les

fibroblastes humains (Milne *et al.*, 2002). Enfin, l'implication des longs ARN non codants (lncRNA) dans le recrutement des protéines PcG a également été mis en évidence dans les cellules de mammifères (Bracken et Helin, 2009; Grewal, 2010; Guenther et Young, 2010). Le rôle des lncRNA dans la régulation transcriptionnelle des gènes *Hox* est d'ailleurs abordé à la section 1.3.11.

1.3.8.1 Complexe répressif Polycomb de type 1 (PRC1)

Les protéines membres du complexe PRC1 ont été isolées d'embryons de drosophile (Franke *et al.*, 1992; Shao *et al.*, 1999) sous forme d'un complexe d'environ 2MDa, contenant entre autres PC (Polycomb), PH (Polyhomeotic), PSC (Posterior Sex Comb) et SCM (Sex Comb on Midleg), de même qu'une trentaine d'autres peptides incluant plusieurs facteurs associés à la TBP (TATA Binding Protein) (Saurin *et al.*, 2001). Elles sont généralement recrutées à la chromatine grâce à l'interaction de leur chromodomaine avec la marque déposée par les membres du complexe PRC2 (Wang *et al.*, 1994) (section 1.3.8.2). Toutefois, des mécanismes alternatifs de recrutement à la chromatine sont aussi employés, comme l'interaction directe avec des protéines se liant à l'ADN telles que Zeste, PSQ (Pipsqueak) et PHO (Pleiohomeotic) (Saurin *et al.*, 2001; Mahmoudi *et al.*, 2003; Huang et Chang, 2004). Les protéines PRC1 remplissent leurs rôles de condensation de la chromatine, ainsi que de répression de l'expression des loci associés, grâce à leur fonction E3 ubiquitine ligase. Cette fonction agit respectivement sur l'histone H2A, menant à la répression transcriptionnelle (Endoh *et al.*, 2012), et sur la Géminine qui inhibe la réplication de l'ADN et la différenciation cellulaire en modulant sa stabilité. D'ailleurs, la perte de fonction du gène *Scmh1*, codant pour une protéine membre du complexe PRC1 interagissant avec la Géminine, entraîne l'augmentation de l'expression des gènes *Hoxb4* et *Hoxa9* dans les cellules hématopoïétiques (Yasunaga *et al.*, 2012).

Enfin, bien que l'inactivation individuelle de différents membres du complexe PRC1 chez la souris, tels que BMI1 (van der Lugt *et al.*, 1994), M33 (Coré *et al.*, 1997), MEL18 (Akasaka *et al.*, 1996) et RING1A (del Mar Lorente *et al.*, 2000), entraîne des transformations homéotiques de postériorisation du squelette axial associées à l'antériorisation des domaines d'expression de certains gènes *Hox* dans le mésoderme

paraxial, ces phénotypes sont relativement légers. Ces données suggèrent qu'il existe une redondance fonctionnelle entre les différents membres, ou encore que l'assemblage des complexes multimériques PRC1 est régulé finement de façon spatio-temporelle plutôt que globale, régionalisant par la même occasion leur spectre d'activité.

1.3.8.2 Complexe répressif Polycomb de type 2 (PRC2)

Suite à des études génétiques et moléculaires démontrant l'interaction directe des protéines ESC, E(Z) et SUZ12 (Jones *et al.*, 1998; Tie *et al.*, 1998) avec différents membres du complexe PRC1, ces protéines ont également été isolées au sein du même complexe chez la drosophile, soit le complexe répressif Polycomb de type 2 ou PRC2 (Ng *et al.*, 2000). Ces protéines sont caractérisées par leur activité histone méthyltransférase, responsable de la mise en place de la marque H3K27me3, permettant par le fait même l'initiation de la compaction de la chromatine localement ainsi que le recrutement des protéines membres du complexe PRC1. De plus, contrairement aux gènes codant pour les membres du complexe PRC1 chez la souris, les différentes mutations individuelles pour les gènes *Eed*, *Ezh2* et *Suz12* mènent à une létalité embryonnaire lors de la gastrulation additionnée de plusieurs défauts sévères au niveau de l'axe AP (Shumacher *et al.*, 1996; O'Carroll *et al.*, 2001; Pasini *et al.*, 2004). Ces résultats reflètent le rôle global de ces protéines sur la régulation génique. Toutefois, la souris mutante hypomorphe pour le gène *Eed* montre des transformations homéotiques comparables à celles retrouvées chez les mutants pour les gènes *PRC1*. Une analyse par approche de délétion conditionnelle devrait être envisagée pour valider le rôle de ces gènes sur l'expression des gènes *Hox*.

1.3.8.3 Facteur de transcription YY1 (complexe PRC0 ou PHORC)

Le facteur de transcription Yin Yang 1 (YY1), protéine à doigts de zinc de la famille Gli-Krüppel et homologue chez les mammifères de la protéine PcG drosophilienne PHO, a été isolé grâce à plusieurs systèmes expérimentaux différents (Berns et Bohenzky., 1987; Chang *et al.*, 1989; Hariharan *et al.*, 1991; Park et Atchison, 1991; Shi *et al.*, 1991; Flanagan *et al.*, 1992). Le gène codant pour cette protéine contient 5 exons, dont le premier code pour 55% de la protéine YY1. Cette protéine multifonctionnelle doit son nom à la

dualité de ses rôles de répresseur et d'activateur transcriptionnels. De plus, YY1 peut agir également à titre de protéine se liant à l'élément initiateur (Inr) dirigeant et initiant ainsi la transcription *in vitro* (Smale et Baltimore, 1989; Shi *et al.*, 1991). D'ailleurs, elle peut interagir physiquement avec l'ARN polymérase II à la fois *in vivo* et *in vitro* (Usheva et Shenk, 1996). Plusieurs domaines différents composent le facteur de transcription YY1, comme les 4 doigts de zinc de type C₂H₂ responsables de la liaison spécifique à l'ADN (Hariharan *et al.*, 1991) ainsi que de sa fonction répressive (Lee *et al.*, 1994, 1995; Lewis *et al.*, 1995), une séquence de 11 acides aminés chargés positivement (acide) responsable de sa fonction activatrice et une séquence de 11 histidines, un acide aminé chargé négativement (basique), qui neutralise potentiellement le domaine activateur. Toutefois, la mutation de ce domaine ne semble pas affecter la fonction répressive de la protéine YY1 (Becker *et al.*, 1994). Enfin, un domaine composé à 55% de glycines ainsi que d'alanines semble également posséder une activité de répression transcriptionnelle (Lee *et al.*, 1994; Yang *et al.*, 1996). Les différents doigts de zinc reconnaissent une séquence consensus 5'-(C/g/a)(G/t)(C/t/a)CATN(T/a)(T/g/c)-3' (Hyde-DeRuyscher *et al.*, 1995; Yant *et al.*, 1995).

Il existe chez les mammifères deux facteurs de transcription possédant un degré d'homologie avec YY1. Le premier est le facteur de transcription REX1 (Reduced EXpression 1), codé par le gène *Zfp42* et dont l'expression est réduite après traitement à l'acide rétinoïque de cellules F9 (Holser *et al.*, 1989). L'homologie entre les deux protéines se situe uniquement au niveau des doigts de zinc qui sont à 75 % identiques. Le facteur de transcription YY2 a aussi été découvert (Nguyen *et al.*, 2004) et 56% de sa séquence protéique est identique à celle de YY1, dont une très forte homologie au niveau des doigts de zinc. Bien que son rôle biologique ne soit pas tout à fait connu présentement, des études effectuées sur le gène *IFNβ* suggèrent que le facteur de transcription YY2 antagoniserait l'action négative de YY1 (Klar et Bode, 2005). De plus, les changements d'expression du gène *Yy2* dans différentes régions du cerveau en développement suggèrent son implication dans la régulation de gènes transcrits à ce niveau (Drews *et al.*, 2009).

La production de la protéine YY1 est initiée dans l'oocyte non-fécondé (Donohoe *et al.*, 1999) et est localisée au niveau du cytoplasme. Après la fécondation, il faut attendre le stade deux cellules avant qu'elle soit exprimée au noyau pour jouer ses rôles de protéine

membre du complexe répressif Polycomb de type 0 (PRC0), appelé aussi complexe répressif associé à la protéine PHO (PHORC), et de facteur de transcription. Par la suite, son expression ubiquitaire sera maintenue dans toutes les cellules de l'embryon de même que chez l'adulte. Toutefois, le niveau de l'ARNm de YY1 est significativement plus élevée au niveau des somites, des bourgeons de membres ainsi que de l'extrémité caudale. De plus, les souris mutantes pour le gène *Yy1* meurent approximativement au jour E6.0 (Donohoe *et al.*, 1999). Ces embryons mutants semblent s'implanter et induire la décidualisation utérine de façon normale, mais se résorbent rapidement après. De plus, bien que la majorité des embryons HT pour le gène *Yy1* soient phénotypiquement normaux, 17% de ceux-ci montrent un retard de croissance ainsi que des défauts de neurulation pouvant mener à l'exencéphalie. Ces différentes observations suggèrent que le facteur de transcription YY1 pourrait jouer un rôle dans le développement plus tardif de l'embryon comme lors de l'organogenèse, par exemple. Cette hypothèse a été confirmée de plusieurs façons, dont en générant des souris mutantes hypomorphes pour le gène *Yy1*, produisant seulement 25 % de la quantité normale de protéines (Affar *et al.*, 2006). Au lieu de se résorber peu de temps avant la gastrulation, ces souris mutantes sont menées à terme, mais elles ne sont pas en mesure de respirer normalement et meurent peu de temps après la naissance suite à l'affaissement de leurs poumons. Par conséquent, le facteur de transcription YY1 agit de manière dose dépendante pour contrôler une multitude de gènes impliqués dans la croissance cellulaire, la prolifération, la cytokinèse, l'apoptose, le développement et la différenciation.

La protéine YY1 joue un rôle important dans plusieurs processus biologiques essentiels au développement normal de l'embryon. Par exemple, le stockage à long terme des ARNm maternels, au niveau d'un complexe appelé particules ribonucléoprotéiques messagères (mRNPs), afin d'être utilisés lors de la maturation ainsi que de la fertilisation, nécessite la présence de la protéine YY1 au niveau du cytoplasme (Belak et Ovsenek, 2007; Belak *et al.*, 2008). De plus, elle est également impliquée dans la régulation de l'expression du lncRNA *Xist* lors de l'inactivation du chromosome X chez la souris (Donohoe *et al.*, 2007; Jeon et Lee, 2011), dans l'établissement de l'empreinte génomique des gènes *Peg3* et *H19* (Kim et Kim., 2008), de même que dans la régulation transcriptionnelle des gènes codant pour les histones (Gokhman *et al.*, 2013). De façon intéressante, ces trois dernières

fonctions nécessitent également l'interaction du facteur de transcription ubiquitaire possédant une activité isolatrice CTCF. D'ailleurs, puisque la protéine YY1 est ubiquitaire, elle nécessite généralement la coopération d'une ou de plusieurs autres protéines afin de conférer sa spécificité d'action. Ces différentes protéines peuvent dicter une fonction spécifique à la protéine YY1 en la régulant de manière post-traductionnelle, ou alors grâce à une interaction physique comme cofacteur de liaison à l'ADN (Golebiowski *et al.*, 2012).

Bien que plusieurs modifications post-traductionnelles semblent affecter la structure/fonction de la protéine YY1, les effets de ces différentes modifications sur le rôle de la protéine YY1 demeurent peu connus. Premièrement, plusieurs sites de phosphorylation consensus Sérine/Thréonine ont été identifiés dans sa séquence d'acides aminés (Rizkallah et Hurt, 2009; Rizkallah *et al.*, 2011a; Riman *et al.*, 2012). De plus, le traitement à la phosphatase entraîne la perte de sa liaison avec la séquence LTR du virus leucémique murin (Becker *et al.*, 1994). D'ailleurs, plusieurs kinases sont maintenant connues comme étant en mesure de phosphoryler YY1: Polo 1 (Rizkallah *et al.*, 2011b), Raf (Baritaki *et al.*, 2007), de même que la kinase Aurora B (Kassardjian *et al.*, 2012). La protéine YY1 est également sujette à la S-nitrosylation (Hongo *et al.*, 2005), la O-glycosylation (Hiromura *et al.*, 2003), la sumoylation (Deng *et al.*, 2007), la poly(ADP-ribosyl)ation (Oei *et al.*, 1997, 1998) et à l'acétylation (Yao *et al.*, 2001).

Les exemples d'interactions protéiques entre YY1 et différents partenaires sont nombreux dans la littérature. Parmi ceux-ci, les exemples impliquant différents membres appartenant au PcG sont les plus connus. C'est pourquoi le rôle de répresseur joué par la protéine YY1 dans la régulation transcriptionnelle des gènes *Hox* est le plus documenté. Par exemple, l'expression du gène *Hoxb4* au niveau du tube neural est contrôlée par une séquence activatrice/amplificatrice, où les facteurs de transcription NFY ainsi que YY1 se compétitionnent entre eux pour lier l'ADN et remplir respectivement leur fonction activatrice ou répressive (Gilthorpe *et al.*, 2002). De plus, des expériences d'immunoprécipitation de la chromatine de cellules P19 non différenciées ont démontré que la protéine MEL18, membre du complexe PRC1, est recrutée au niveau de la chromatine du gène *Hoxd4* conjointement avec la protéine YY1 (Kobrossy *et al.*, 2006). Suite au traitement au RA, la protéine MEL18 se dissocie de la chromatine du gène *Hoxd4*,

alors que la protéine YY1 demeure liée à celle-ci. Ensuite, l'inhibition de l'expression du gène *Hoxa11* contrôlée par la liaison du facteur de transcription YY1 au niveau du promoteur nécessite le recrutement de HDAC2 (Luke *et al.*, 2006). Dans les cellules souches mésenchymateuses (MSC), la protéine YY1 cible le complexe PRC2 au niveau de régions spécifiques de certains gènes appartenant au complexe *HoxD* (Woo *et al.*, 2010). De la même façon, l'association de la protéine YY1 au complexe protéique PRC2 réprime la voie de différenciation *Cdx-Hox* dans les différentes lignées hématopoïétiques chez le poisson zèbre (Yue *et al.*, 2009). Chez la souris, le complexe PRC0/PRC2/PRC1 composé des protéines YY1/EED/BMI1 est détecté au niveau du promoteur des gènes *Hoxa5* et *Hoxc8* (Kim *et al.*, 2006). Par contre, bien que les souris mutantes HT pour le gène *Yyl* montrent plusieurs transformations d'antériorisation du squelette axial, aucune variation dans l'expression des gènes *Hox* n'a été détectée (Lorente *et al.*, 2006).

À l'opposé, peu d'exemples impliquant la régulation positive de l'expression des gènes *Hox* par le facteur de transcription YY1 ont été répertoriés. Par exemple, l'interaction protéique entre PHO, la protéine BRM membre du complexe TrX (voir section 1.3.9) et DSP1 a été observée au niveau d'un élément régulateur du gène *Src* chez la drosophile (Mohd-Sarip *et al.* 2002; Rappailles *et al.*, 2005). De plus, les facteurs de transcription SP1/SP3, NFY, YY1 et USF1 sont tous importants pour l'activation du gène *Hoxb7* humain dans un contexte d'essais luciférase (Meccia *et al.*, 2003). Enfin, plusieurs coactivateurs transcriptionnels interagissent également avec la protéine YY1, comme AP1/2 (Artishevsky *et al.*, 1987; Klug et Beato, 1996; Wang *et al.*, 2005, 2007; Allouche *et al.*, 2008), P300 (Shrivastava *et al.*, 1993; Inouye et Seto, 1994; Lee *et al.*, 1995; Yang *et al.*, 1996; Wang *et al.*, 2007), GATA (Bhalla *et al.*, 2001) et INO80 (Cai *et al.*, 2007; Wu *et al.*, 2007). D'ailleurs, les mutants chez la drosophile où il y a perte de fonction pour le gène *dINO80* meurent durant l'embryogenèse et présentent des transformations homéotiques associées à la diminution d'expression des gènes homéotiques *Scr*, *Antp*, *Ubx* et *AbdB*, suggérant la liaison de complexes protéiques PHO/dINO80/TrX au niveau des TRE présents dans les séquences des gènes *Hom* (Bhatia *et al.*, 2010).

1.3.9 Les protéines membres du groupe Trithorax (TrxG)

Tel que mentionné précédemment, les protéines du groupe Trithorax (TrxG) remplissent la fonction antagoniste à celle des gènes codant pour les protéines PcG, soit de maintenir la chromatine des gènes *Hox* ouverte pour qu'elle puisse être accessible aux activateurs transcriptionnels (Kennison, 1993; Peterson et Tamkun, 1995). Cette fonction s'effectue grâce à leurs activités de remodelage de la chromatine, telles les activités histone méthyltransférase (HMT) et histone acétyltransférase (HAT) (Mohan *et al.*, 2010; Schuettengruber *et al.*, 2007, 2011). Le gène *Trx*, membre fondateur du groupe du même nom, a été isolé à partir d'une mutation naturelle qui cause plusieurs transformations homéotiques d'antériorisation sur différents segments de la drosophile (Ingham 1983, 1998), reflétant ainsi la perte de fonction de divers gènes *Hox*. Par la suite différents complexes TrX ont été identifiés chez la drosophile. Par exemple, le complexe SWI/SNF, également appelé le complexe BRM, contient les protéines BRM, MOIRA, OSA et SNR1 (Pazin et Kadonaga, 1997; Papoulas *et al.*, 1998). Le complexe TAC1 regroupe les protéines TRX, SBF1 et dCBP, l'homologue de CBP/ P300 (Petruk *et al.*, 2001). De la même façon, le complexe impliquant la protéine ASH1 (Papoulas *et al.*, 1998) s'associe également avec dCBP (Bantignies *et al.*, 2000). L'analyse moléculaire de la protéine ASH1 a mis en évidence l'activité méthyltransférase de son domaine SET qui dépose la marque activatrice H3K4me3 (Beisel *et al.*, 2002, Byrd et Shearn, 2003; Smith *et al.*, 2004; Eissenberg et Shilatifard, 2010). Il existe également un complexe qui contient la protéine ASH2. Chez les mammifères, on retrouve 4 complexes TrX possédant un domaine SET : ASHL1, ASHL2, MLL1 et MLL2 (Wang *et al.*, 2009). Les souris déficientes pour le domaine SET (Δ SET), générées par recombinaison homologue en ESC, présentent une perte de la marque H3K4me3 au niveau des loci *Hox* (Milne *et al.*, 2002; Terranova *et al.*, 2006) de même que des défauts du squelette axial. D'ailleurs, les souris mutantes pour le gène *Mill* présentent une perte d'expression du gène *Hoxa7* (Yu *et al.*, 1998). De plus, il a été démontré pour la protéine ASH2 que son domaine à doigt PHD (Plant Homeo Domain) est important pour sa localisation au niveau des loci *Hox* (Chen *et al.*, 2011). Enfin, d'autres protéines du groupe TrxG, comme les déméthylases UTX et JMJD3, sont essentielles à l'activation ainsi qu'au maintien de l'expression des gènes *Hox* (Soshnikova et Duboule, 2008; Copur et Müller, 2013).

1.3.10 Transcrits multiples

Un autre mécanisme permettant l'expression d'un gène aux niveaux tissulaire, cellulaire, ou même lors d'un stade développemental précis, consiste à initier la transcription à partir de promoteurs alternatifs. Tout comme la duplication de gènes ou de complexes géniques, les transcrits multiples, qu'ils soient issus de promoteurs distincts, ou bien de variants d'épissage, permettent d'ajouter un niveau supplémentaire à la régulation spatio-temporelle génique (Ayoubi et Van De Ven, 1996; Keren *et al.*, 2010). Comme l'expression des gènes *Hox* doit être finement régulée dans le temps et l'espace afin que le développement embryonnaire se déroule normalement, la présence de transcrits multiples est aussi observée dans leur cas. Par exemple, le gène *Ubx* code pour six isoformes protéiques générées par épissage alternatif (de Navas *et al.*, 2011), où le transcrit correspondant à tous les exons du gène est plus efficace dans la formation des haltères et dans l'activation des gènes en aval que ses versions plus courtes. Toujours chez la drosophile, plusieurs promoteurs sont utilisés pour la transcription des gènes *Antp* et *AbdB* (Laughon *et al.*, 1986; Zavortink et Sakonju, 1989). En ce qui concerne les mammifères, les gènes *Hoxb3* et *Hoxb5* génèrent également de multiples transcrits exprimés de façon spécifique dans l'embryon (Krumlauf *et al.*, 1987). De plus, l'expression de chacun des gènes *Hox* appartenant au quatrième groupe paralogue est initiée à partir de deux promoteurs différents (Galliot *et al.*, 1989; Gutman *et al.*, 1994; Folberg *et al.*, 1997). Ils répondent différemment à la stimulation par le RA et respectent le principe de colinéarité. De plus, les gènes *Hoxb4* et *Hoxd4* codent respectivement pour un transcrit non coiffé contenant les miRNA miR-10a et miR-10b (Phua *et al.*, 2011). Toutefois, leurs fonctions biologiques demeurent inconnues. Quant au gène *Hoxc6*, il génère quatre transcrits différents à partir de deux promoteurs différents (Shimeld *et al.*, 1993). L'épissage alternatif est également responsable de production de transcrits multiples pour les gènes *Hoxa9* et *Hoxa10* (Benson *et al.*, 1995; Fujimoto *et al.*, 1998). D'ailleurs, les deux transcrits du gène *Hoxa9* codent respectivement pour la protéine pleine longueur (HOXA9) et pour un variant sans domaine homéo (HOXA9T), lesquels se compétitionnent pour la liaison au coactivateur transcriptionnel CBP/P300 (Dintilhac *et al.*, 2004). Des transcrits anti-sens exprimés de façon spécifique ont également été observés pour les gènes *Hoxb3*, *Hoxb4*, *Hoxb5* et *Hoxa11* (Hogan *et al.*, 1988; Whiting *et al.*, 1991; wiatek et Gridley, 1993; Hsieh-

Li *et al.*, 1995; Potter et Branford, 1998; Dubrulle *et al.*, 2001; Zákány *et al.*, 2001; Brend *et al.*, 2003). Ces transcrits s'hybrideraient à leur transcrit sens respectif et mèneraient à leur dégradation (voir aussi section 1.3.13.1 sur les miRNAs). Enfin, tout comme le reste des données concernant notre modèle d'étude, les transcrits alternatifs du gène *Hoxa5* feront l'objet d'une section à part (Section 1.4).

1.3.11 Longs ARN non codants (lncRNAs)

Selon les données du consortium international de séquençage du génome humain publiées en 2004, notre matériel génétique contient approximativement 20,000 gènes codant pour des protéines, ce qui correspond à un peu moins de 2% de notre génome. Comme la presque totalité du génome des eukaryotes est transcrit, cela suggère qu'un réseau complexe de transcrits incluant des milliers de longs ARNs non codants (lncRNAs), de même que des micro-ARNs (section 1.3.13.1). Bien qu'ils ont été longtemps perçus comme étant du bruit de fond transcriptionnel, leur expression spécifique à certains types cellulaires, leur localisation au niveau de compartiments subnucléaires ainsi que leur association à différentes pathologies confirment maintenant que les lncRNAs sont des joueurs clés dans la régulation génique en général, et des gènes *Hox* en particulier.

Chez la drosophile, plusieurs lncRNAs ont été localisés au niveau des deux complexes *HOM*. Certains de ceux-ci traversent des PRE lors de leur transcription entraînant la dérégulation du gène *Hom* associé (Bender et Fitzgerald, 2002; Hogga et Karsch, 2002; Rank *et al.*, 2002), suggérant que l'action de transcrire expulse les protéines PcG et ouvre la chromatine (Schmitt *et al.*, 2005). De plus, le lncRNA *dxd-Pre*, transcrit à partir du locus *Ubx*, est en mesure d'interagir avec la protéine méthyltransférase ASH1 et ainsi de maintenir l'expression du gène *Ubx* (Sanchez-Elsner *et al.* 2006).

Plusieurs données associant les lncRNAs à la régulation transcriptionnelle des gènes *Hox* ont également été recueillies chez les mammifères. Par exemple, une analyse informatique comparative des complexes *HoxA* humain et murin révèle la conservation de lncRNAs anti-sens en 5' du gène *Hoxa13*, de même que dans les séquences intergéniques *Hoxa6-Hoxa7*, *Hoxa3-Hoxa4* et *Hoxa1-Hoxa2* (Sasaki *et al.*, 2007). Concernant le

lncRNA *Mistral* (*Mira*) situé entre les gènes *Hoxa6* et *Hoxa7*, il permet de recruter la protéine méthyltransférase MLL1 au niveau de la chromatine de ces deux gènes et d'activer leur transcription suite au traitement de ESC murines avec du RA (Bertani *et al.*, 2011). Quant à celui initié entre les gènes *Hoxa3* et *Hoxa4* (*HOXA-AS2*), il est également induit par le traitement des cellules humaines NB4 au RA et semble jouer un rôle dans le peaufinage de l'apoptose lors de la différenciation myéloïde (Zhao *et al.*, 2013). Le lncRNA HOTTIP, situé en 5' du complexe *HoxA*, se rapproche de ses gènes cibles en 3' grâce à la formation d'une boucle chromosomique générée lors de l'activation du complexe par le RA et peut maintenir leur expression en formant un complexe avec WDR5, ASH2 et MLL1 (Chen *et al.*, 2011; Wang *et al.*, 2011). À l'inverse, le lncRNA de 2.2 kb *HOTAIR*, transcrit à partir de la région intergénique *Hoxc11-Hoxc12*, recrute des protéines membres du complexe PRC2 au niveau de plusieurs gènes dont certains situés dans le complexe *HoxD* (Rinn *et al.*, 2007; Gupta *et al.*, 2010). Il est également en mesure de recruter la protéine déméthylase ISD1 qui efface spécifiquement la marque active H3K4me3 (Tsai *et al.*, 2010). De plus, il lie le même site unique présent dans le complexe *HoxD* de différentes cellules même si elles n'ont pas le même profil d'inactivation (Chu *et al.*, 2011). Toutefois, la délétion du transcrit *HOTAIR* chez la souris n'a pas d'incidence majeure sur la répression des gènes postérieurs du complexe *HoxD*, ce qui peut être attribuable au très faible niveau de conservation des séquences du transcrit entre l'humain et la souris (Schorderet et Duboule, 2011). Chez l'humain, l'expression du transcrit *HOTAIR* est induite par l'estradiol dans les carcinomes mammaires (Bhan *et al.*, 2013). De plus, les niveaux élevés du transcrit *HOTAIR* et de la protéine EZH2, membre du complexe PRC2, sont généralement associés à de mauvais pronostics dans plusieurs types de cancers (Kogo *et al.*, 2011; Chisholm *et al.*, 2012; Niinuma *et al.*, 2012; Li *et al.*, 2013).

1.3.12 Régulation post-transcriptionnelle

1.3.12.1 MicroARNs (miRNAs)

Les microARNs (miRNAs) sont des ARN non-codants simple brin d'environ 18 à 24 nt de longueur. Ils sont initialement transcrits par l'ARN polymérase II en temps que miRNAs primaires (pri-miRNAs) (Lee *et al.*, 2004; Zeng *et al.*, 2005). Ceux-ci sont ensuite

pris en charge à l'intérieur du noyau par l'enzyme RNase III Drosha, ainsi que la protéine se liant à l'ADN double brin DGCR8/Pasha, pour générer des structures en boucle d'approximativement 70 nt que l'on appelle miRNAs précurseurs (pre-miRNAs) (Gregory *et al.*, 2005). Avec l'aide du facteur nucléaire d'exportation Exportin 5, ces pre-miRNAs sont transportés dans le cytoplasme où ils servent de substrat à l'enzyme RNase Dicer qui génère des duplexes d'ARN d'environ 22 nt (Hutvagner, 2004; Kurihara et Watanabe, 2004). Finalement, la molécule d'ARN double brin se dissocie et l'un des brins est chargé au niveau du complexe induisant l'interférence à l'ARN (RISC), lequel cible l'ARNm correspondant qui sera ultimement dégradé (Bartel, 2009). Ce mécanisme de régulation post-transcriptionnelle a été conservé durant l'évolution des espèces et est impliqué dans la régulation de l'expression des gènes *Hox*.

Chez le nématode, modèle animal avec lequel les miRNAs ont été découverts, l'expression ectopique de miR-57 entraîne de sévères défauts postérieurs, lesquels rappellent les phénotypes observés chez les mutants du gène homéotique *Nobl*, l'homologue du gène *AbdB* (Zhao *et al.*, 2010). Chez la drosophile, un site de liaison pour miR-10-5p est prédit au niveau du 3'UTR de l'ARNm du gène *Scr* et serait conservé au sein d'autres ordres d'insectes (Enright *et al.*, 2003; Brennecke *et al.*, 2005; Miura *et al.*, 2011). À l'inverse, miR-10-3p, un miRNA spécifique à la drosophile (Griffiths-Jones *et al.*, 2011), ciblerait potentiellement l'ARNm des gènes *Ubx* et *AbdB* (Stark *et al.*, 2007). De plus, le locus de miR-lab-4 est transcrit à partir des deux brins d'ADN générant ainsi à la fois un miRNA sens et antisens, lesquels causeraient des transformations homéotiques lorsque exprimés de façon ectopique (Stark *et al.*, 2008). Toutefois, des résultats plus récents mettent en évidence que les miRNAs miR-10-3p, miR-10-5p et miR-lab-4-5p, transcrits à partir des deux complexes *Hox* de la drosophile, ne semblent pas avoir d'effets significatifs sur l'expression de leurs gènes *Hox* cibles durant l'embryogenèse. Il est possible que l'effet biologique de la mutation de ces miRNAs ne soit observable qu'en condition de stress ou en combinaison avec des mutations sur d'autres miRNAs (Lemons *et al.*, 2012). Chez le poisson zèbre, on observe l'altération de l'expression de différents gènes *Hox*, de même que celle du récepteur au RA *Rarab*, associée à des défauts du bourgeon de nageoire pectorale, du sixième arc pharyngien et des somites, suite à un gain ou une perte de fonction de miR-196 (He *et al.*, 2011). En ce qui concerne les mammifères, les complexes des gènes *Hox*

contiennent deux familles de miRNAs: les familles miR-10 et miR-196. Les paralogues de miR-10 sont exprimés de façon précoce dans le développement et possèdent une frontière antérieure d'expression similaire à celle des gènes *Hox4* (Mansfield *et al.*, 2004; Wienholds *et al.*, 2005; Darnell *et al.*, 2006; Kloosterman *et al.*, 2006). De plus, ces limites antérieures d'expression au niveau du SNC diffèrent entre les paralogues, où miR-10a et miR-10c sont exprimés caudalement au rhombomère 6/7, alors que les miR-10b et miR-10d se retrouvent caudalement au cerveau postérieur primitif (Woltering et Durston, 2006, 2008). Quant à la famille miR-196, elle forme un groupe postérieur pour les gènes *Hox* impliqués dans la morphologie du thorax/tronc, où elle cible 75% des gènes *Hox* des groupes paralogues 6 à 8 (Hornstein *et al.*, 2005; McIntyre *et al.*, 2007; McGlinn *et al.*, 2009; Asli et Kessel, 2010). De plus, tel que mentionné à la section 1.3.3 à la Figure 1.8, puisque les présumés gènes *Hox* cibles sont généralement situés en 3' des deux familles de miRNAs, cela suggère que ces miRNAs répriment les programmes d'expression antérieure des gènes *Hox*, participant à la prévalence postérieure (Yekta *et al.*, 2008). Enfin, la perturbation de l'expression des familles miR-10 et miR-196 est associée à certaines pathologies et cancers (Plummer *et al.*, 2012; Mahdipour, 2013; Planchon *et al.*, 2013).

1.3.12.2 Protéines ribosomales

Historiquement, le ribosome était vu comme jouant un rôle constitutif, plutôt qu'un rôle régulateur spécifique, lors de la traduction des ARNm. Toutefois, il a été démontré que certaines mutations au niveau de la protéine ribosomale RPL38 entraînent des défauts spécifiques à certains tissus de même que des transformations homéotiques du squelette axial (Kondrashov *et al.*, 2011). En effet, ces mutants présentent une synthèse protéique globale inchangée, mais une traduction perturbée d'un groupe spécifique d'ARNm de gènes *Hox*. Par conséquent, les protéines ribosomales constituent potentiellement un nouveau palier de contrôle de l'expression génique lors du développement.

1.3.13 Interprétation des phénotypes et cassette de sélection

À la lumière de tous ces mécanismes différents orchestrant de façon précise l'expression spatio-temporelle des gènes *Hox* chez les mammifères, il est clair que l'analyse

des phénotypes observés chez certaines souris mutantes, où la perte de fonction a été obtenue par l'insertion maintenue d'une cassette de sélection, doit être faite avec certaines réserves. En d'autres termes, il faut différencier les phénotypes propres à la mutation du gène *Hox* d'intérêt et ceux inhérents à la perturbation des éléments régulateurs agissant sur les gènes *Hox* à proximité du gène ciblé (Olson *et al.*, 1996). Plusieurs exemples dans la littérature font état de souris mutantes pour un gène *Hox* présentant des phénotypes situés à l'extérieur des domaines d'expression endogène du gène ciblé. Par exemple, l'insertion de la cassette de sélection à la néomycine dans le gène *Hoxa2* entraîne un retard de croissance et une mortalité approximativement 20 jours après la naissance des individus *Hoxa2*^{floxneo/floxneo}. Avant même l'action de la recombinaise Cre, une diminution de l'expression du gène *Hoxa2* au niveau de r2 associée à une dérégulation du gène *Hoxa1* au niveau de r1 est observée chez ces souris (Ren *et al.*, 2002). De la même façon, le maintien de la cassette de sélection au niveau du gène *Hoxd10* mène à une mortalité périnatale ainsi qu'à plusieurs défauts cranio-faciaux et axiaux (Rijli *et al.*, 1994; Carpenter *et al.*, 1997). Enfin, comme il y a partage d'éléments régulateurs entre les gènes *Hox* voisins, certaines interactions entre le promoteur de la cassette de sélection et les éléments régulateurs présents dans les séquences avoisinantes peuvent entraîner l'altération du profil d'expression de gènes *Hox* voisins. Des exemples de ces interactions non spécifiques ont été observés chez les souris mutantes pour les gènes *Hoxb2* (Barrow et Capecchi, 1996) *Hoxa4* (Aubin *et al.*, 1998), *Hoxc4* (Boulet et Capecchi, 1996), *Hoxb8* (Chen et Capecchi, 1997; van den Akker *et al.*, 1999; Greer et Capecchi, 2002) et *Hoxa10* (Rijli *et al.*, 1995).

1.4 Le gène *Hoxa5*

Tel que mentionné, cette partie de l'introduction sera consacrée au gène sur lequel j'ai centré mes travaux, soit le gène *Hoxa5*. Membre du complexe *HoxA* situé sur le chromosome 6 chez la souris (chromosome 7 chez l'humain), il code pour un facteur de transcription de 270 aa liant l'ADN grâce à sa boîte homéo. Toutefois, bien que sa masse moléculaire théorique soit de 29.3 kDa, la protéine HOXA5 se comporte sur gel comme une protéine entre 35 et 43 kDa selon son niveau de phosphorylation (Odenwald *et al.*, 1987). Tout comme les autres gènes *Hox*, non seulement sa structure protéique, mais

également plusieurs de ses fonctions ont été conservées durant l'évolution des espèces, notamment avec son orthologue drosophilien *Scr*. En effet, l'expression exogène du gène *Hoxa5* chez la mouche atténue considérablement la perte de fonction du gène *Scr* (Zhao *et al.*, 1993). De façon intéressante, malgré le très haut degré de conservation entre les domaines homéo des différentes protéines HOX, la protéine SCR dimérise *in vivo* via ce domaine, contrairement à son plus proche parent la protéine ANTP (Papadopoulos *et al.*, 2012). Ce type d'interaction n'est pas connu chez les mammifères.

Chez la souris, trois promoteurs différents sont responsables de la synthèse des cinq différents transcrits *Hoxa5*. Tout d'abord, on retrouve le promoteur proximal à partir duquel est initié le transcrit de 1.8 kb qui correspond aux deux exons du gène *Hoxa5* (Jeannotte *et al.*, 1993). De plus, il s'agit du transcrit *Hoxa5* détecté en plus grande quantité ainsi que le plus précocement, soit au niveau des somites 5-7 d'un embryon au stade E8.25. Quatre longs transcrits *Hoxa5*, dont deux de 5kb, un de 9.5 kb et un de 11 kb, sont initiés lors du stade E8.75 à partir de deux promoteurs distaux: le promoteur du gène *Hoxa6* et un promoteur situé en 3' du gène *Hoxa7* (Zákány *et al.*, 1988; Larochelle *et al.*, 1999). D'ailleurs, la relation de colinéarité spatiale entre le court et les longs transcrits est observée au stade E12.5. En effet, les frontières antérieures d'expression des longs transcrits sont détectées à la jonction tube neural/myencéphale ainsi qu'au niveau de la prévertèbre (pv) 10, soit de façon plus postérieure que celles du transcrit de 1.8 kb, situées au niveau du plancher du myencéphale et de la pv3. Par conséquent, les résultats indiquent que le transcrit de 1.8 kb possède un domaine d'expression spécifique entre les prévertèbres 3 et 10, de même qu'au niveau de la scapula (Aubin *et al.*, 1998; Larochelle *et al.*, 1999). Le transcrit de 1.8kb est également exprimé au niveau de la composante mésenchymale des poumons, de l'estomac, de l'intestin, de la rate, des reins, des ovaires, de la glande mammaire et du thymus (Jeannotte *et al.*, 1993; Aubin *et al.*, 1997, 1999, 2002a; Garin *et al.*, 2006; Gendronneau *et al.*, 2012; Boucherat *et al.*, 2013). De plus, la présence d'un gradient d'expression rostrocaudale croissant au niveau de l'estomac est aussi détectée. Quant à l'expression des longs transcrits *Hoxa5* au niveau des organes, on ne les détecte que dans l'intestin en développement (E10.5-E13.5). Enfin, le chevauchement entre le profil d'expression du transcrit *Hoxa5* de 1.8 kb et celui de la protéine HOXA5, notamment au niveau du cerveau postérieur, des pv3 à pv10, ainsi que de la composante mésenchymale

des systèmes respiratoire et digestif, indique qu'il s'agit du transcrit fonctionnel codant pour la protéine (Joksimovic *et al.*, 2005). L'expression des transcrits *Hoxa5*, de même que celle de la protéine, au niveau de l'embryon, de même qu'au niveau des systèmes respiratoire et digestif, sont présentées aux figures 1.12 et 5.1 respectivement. J'ai participé à la caractérisation moléculaire des différents transcrits du gène *Hoxa5* et ce travail sera décrit au chapitre 5.

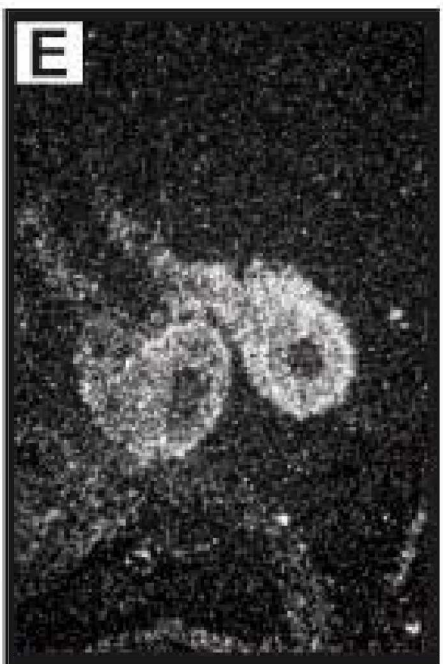
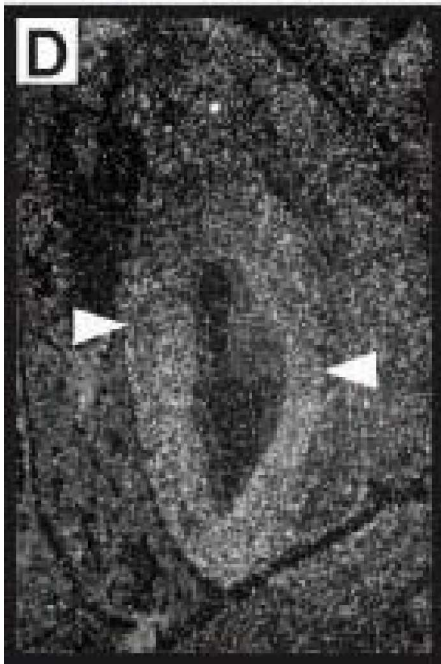
1.4.1 Fonction du gène *Hoxa5*

Deux différentes lignées de souris mutantes pour le gène *Hoxa5* ont été produites dans le laboratoire, respectivement par insertion d'une cassette de sélection à la néomycine au niveau du domaine homéo (Jeannotte *et al.*, 1991; 1993) et par excision conditionnelle de l'exon 2 qui contient ce domaine (Tabariès *et al.*, 2007). Comme le pourcentage de souris mutantes *Hoxa5*^{-/-} recensé lors du sevrage ne correspondait pas au pourcentage prédit par les lois de Mendel, un suivi quotidien des portées a mis en évidence la mort des nouveau-nés. Les individus *Hoxa5*^{-/-} présentent des poumons affaissés, de même que de l'air dans le système digestif (Jeannotte *et al.*, 1993; Aubin *et al.*, 1997). De façon intéressante, à l'opposé de ce qui a été observé pour les squelettes, ce sont les souris "outbred" qui possèdent le taux de mortalité périnatale le plus bas avec 50%, alors que les souches 129/SvEv et C57BL/6 les suivent avec 72% et 100% de mortalité respectivement. Ces souris mutantes ont permis l'étude des différentes fonctions du gène *Hoxa5* lors du développement.

1.4.1.1 Établissement du squelette axial

Les souris mutantes *Hoxa5*^{-/-} présentent diverses malformations au niveau du squelette axial qui sont localisées essentiellement entre les vertèbres C3 et T2, ce qui est en accord avec l'expression du transcrit de 1.8 kb ainsi que celle de la protéine HOXA5 qui sont situées entre la troisième et la dixième prévertèbre. De plus, tout comme la mutation de son gène orthologue *Scr* chez la drosophile, la perte de la fonction du gène *Hoxa5* entraîne à la fois des transformations homéotiques d'antériorisation et de postériorisation

Figure 1.12: Expression des transcrits du gène *Hoxa5* au niveau de l'embryon de souris au stade développemental E12.5. (A) Profil d'expression obtenu à l'aide d'une sonde détectant les 5 transcrits du gène *Hoxa5* au niveau du SNC, de la colonne prévertébrale et des membres antérieurs. (B) Profil d'expression obtenu à l'aide d'une sonde détectant les 4 longs transcrits du gène *Hoxa5*. Par analyse différentielle soustractive des deux profils d'expression, on déduit une limite d'expression antérieure du transcrit de 1.8 kb au niveau du plancher du myencéphale, comparativement à la jonction tube neural/myencéphale pour les longs transcrits (tête de flèche ouverte). Au niveau de la colonne prévertébrale, la frontière antérieure d'expression du transcrit de 1.8kb se situe au niveau de la prévertèbre (pv) 3, alors qu'elle est au niveau de pv10 en ce qui concerne les longs transcrits (tête de flèche noire). Enfin, l'expression scapulaire est spécifique au transcrit de 1.8 kb (flèche noire) (Larochelle *et al.*, 1999). (C) Expression spécifique au niveau du mésenchyme des voies respiratoires (trachée et poumon) (Aubin *et al.*, 1997), (D) de l'estomac (Aubin *et al.*, 2002a), et (E) de l'intestin (Aubin *et al.*, 1999). (Autorisation reçue de la part de(s) auteur(s).)



(Jeannotte *et al.*, 1993). D'ailleurs, un résumé des différents phénotypes répertoriés chez les souris mutantes *Hoxa5*^{-/-} est présenté à la Figure 1.13. Toutefois, la présence d'une grande variabilité au niveau des différents phénotypes est observée en fonction de la souche génétique des souris qui portent la mutation, suggérant que la mutation du gène *Hoxa5* est hypostatique et que des gènes modificateurs accentuent certains phénotypes (Jeannotte *et al.*, 1993; Aubin *et al.*, 1998; Garin *et al.*, 2006). Ceci étant dit, la transformation homéotique d'antériorisation de la vertèbre C6 en C5, caractérisée par la perte bilatérale du *tuberculum* antérieur, est pénétrante chez 100% des souris *Hoxa5*^{-/-} (Figure 1.13) (Aubin *et al.*, 1998). De la même façon, puisque l'expression du gène *Hoxa5* au niveau du squelette axial est maximale en C6 et C7, la transformation homéotique de postériorisation qui cause une côte ectopique à C7, la changeant ainsi en T1, est également très pénétrante. De plus, des expériences chez le poulet ont démontré que l'expression du gène *Hoxa5* au niveau du sclérotome est importante pour établir adéquatement la transition cervico-thoracique (Chen *et al.*, 2013). Enfin, d'autres phénotypes moins pénétrants sont également observés au niveau du squelette axial suite à la mutation du gène *Hoxa5* : des transformations d'antériorisation des vertèbres C3 et C4 en C2, caractérisées par l'apparition d'un processus dorsal, de T2 en T1, où il y a perte du processus dorsal spinal, de L1 en T13, sur laquelle se greffe une côte ectopique, de même qu'une transformation de postériorisation de T1 en T2, où dans ce cas il y a apparition d'un processus dorsal spinal (Aubin *et al.*, 1998).

1.4.1.2 Établissement du squelette paraxial et appendiculaire

En plus des altérations au niveau du squelette axial, les souris mutantes *Hoxa5*^{-/-} présentent également des défauts au niveau du squelette paraxial ainsi que du squelette appendiculaire. En effet, certains de ces individus *Hoxa5*^{-/-} possèdent une sternèbre additionnelle générée par la fission du manubrium (Jeannotte *et al.*, 1993). De plus, la quatrième sternèbre est plus courte et est attachée à la troisième sternèbre. Par conséquent, les sites d'attachement des côtes 5 à 7 se retrouvent pratiquement juxtaposés. Dans les cas les plus affectés, on observe la fusion de sternèbres. Enfin, le processus xyphoïde de certaines souris mutantes pour le gène *Hoxa5* est troué ou fissuré.

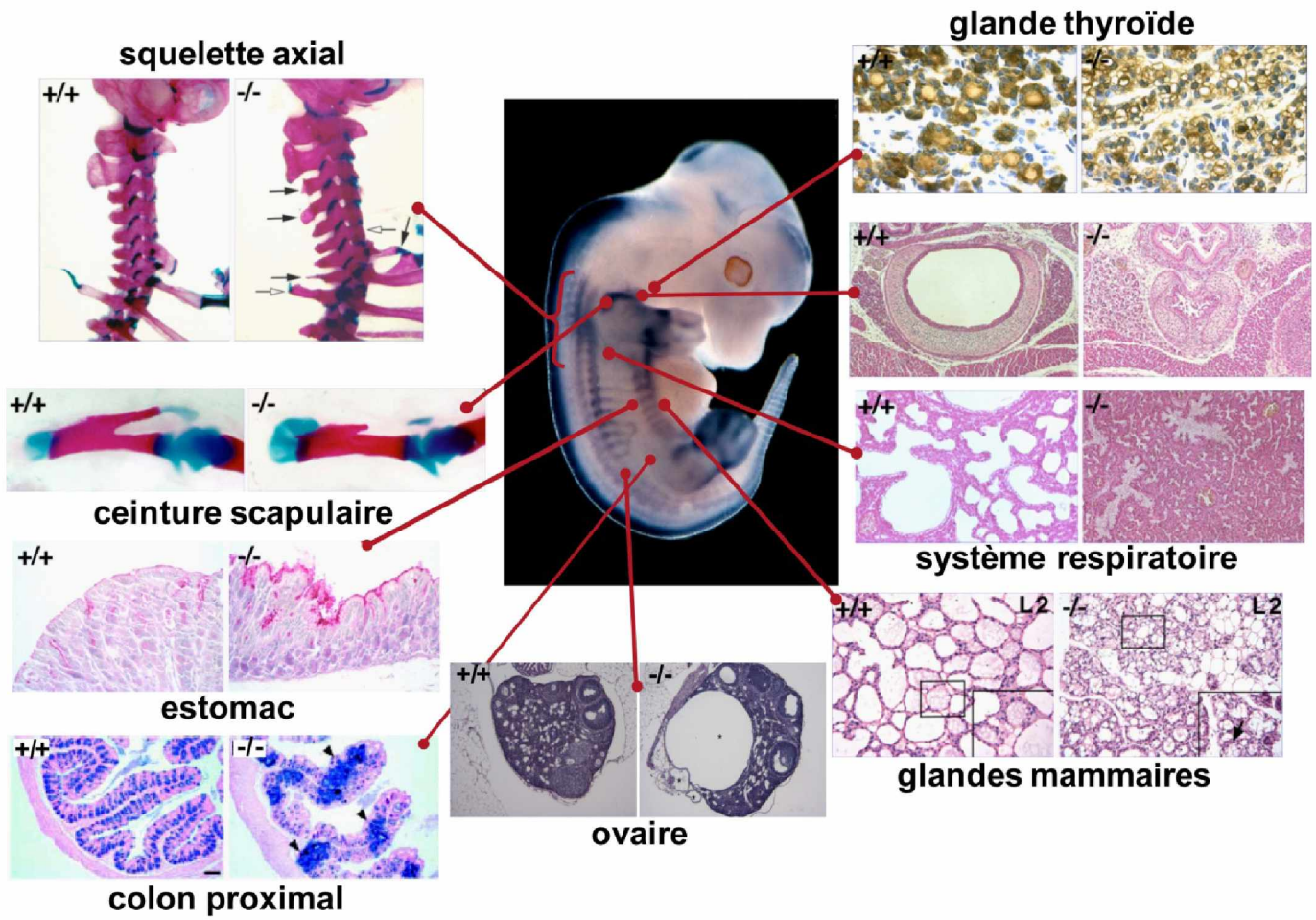
Concernant le squelette appendiculaire, la perte de fonction du gène *Hoxa5* chez les souris "outbred" se traduit par le rétrécissement de la projection osseuse émergeant de la ceinture scapulaire, appelée aussi acromion, et qui peut même aller jusqu'à son absence complète dans certains cas (Figure 1.13) (Aubin *et al.*, 1998). De plus, comme la lignée de souris *undulated (un)*, porteuses d'une mutation sur le gène *Pax1*, présente un phénotype au niveau de l'acromion qui est semblable à celui des souris mutantes *Hoxa5^{-/-}*, la synergie potentielle entre ces deux gènes a été étudiée (Aubin *et al.*, 2002b). Les résultats obtenus suggèrent un rôle de recrutement des cellules précurseurs pour le gène *Pax1*, tandis que le gène *Hoxa5* serait impliqué dans le positionnement des cellules positives pour le gène *Pax1*. De plus, le gène *Hoxa5* jouerait un rôle dans la chondrogenèse. (Aubin *et al.*, 2002b).

1.4.1.3 Système nerveux central (SNC) et cellules de la crête neurale (CCN)

Bien que la majorité des malformations du SNC associées aux gènes *Hox* découle majoritairement de mutations au sein des gènes *Hox* des trois premiers groupes paralogues, le gène *Hoxa5* fait exception à la règle. En effet, une approche par transgénèse avec des souris surexprimant le gène *Hoxa5* au niveau de la partie brachiale du tube neural, grâce à l'élément BSC (section 1.4.4.2), génère des anomalies de la partie cervicale du tube neural (Krieger *et al.*, 2004). Par conséquent, les neurones de la corde dorsale superficielle présentent divers problèmes de différenciation, de migration, d'apoptose ainsi qu'une expression ectopique du marqueur ventral neuronal *BRN3A* (Krieger *et al.*, 2004; Abbott *et al.*, 2005). Ces défauts se traduisent par des problèmes de motricité et de sensibilité au niveau des membres antérieurs.

Enfin, l'expression du gène *Hoxa5* est essentielle à la formation de deux colonnes différentes de neurones moteurs: les LMC des membres antérieurs et ceux de la colonne phrénique (PMC). En ce qui concerne les neurones LMC, le gène *Hoxa5* s'occupe d'établir correctement la connexion scapulo-humérale postérieure des neurones LMC rostraux (Dasen *et al.*, 2005; Dasen et Jessell, 2009; Lacombe *et al.*, 2013). Concernant les neurones PMC, ils innervent le diaphragme et expriment les gènes *Hoxa5* et *Hoxc5*. De plus, il s'agit des deux seuls gènes *Hox* exprimés dans ces neurones moteurs. La délétion combinée de

Figure 1.13: Résumé des différents phénotypes observés chez les souris *Hoxa5*^{-/-}. Dans le sens horaire: Désorganisation des follicules de la glande thyroïde doublée d'une distribution anormale de la thyroglobuline au stade développemental E18.5 (Meunier *et al.*, 2003). À la naissance, on observe une trachée au diamètre rétréci et qui peut même être complètement obstruée dans les cas les plus sévères. Quant aux poumons, ils sont collapsés (Aubin *et al.*, 1997). L'analyse des glandes mammaires de souris lactantes montrent une désorganisation caractérisée par des alvéoles plus petits et la présence de gouttelettes lipidiques (fenêtre) (Garin *et al.*, 2006). Au niveau des ovaires de femelles nullipares de 10 mois, des kystes d'inclusion se développent, rappelant l'endosalpingiose humaine (Gendronneau *et al.*, 2012). L'estomac présente un épithélium aminci ainsi qu'une hypertrophie de la sous-muqueuse (Aubin *et al.*, 2002a). Quant au colon proximal, il présente des villi plus courts, une distribution anormale des cellules productrices de mucus (tête de flèche noir) ainsi qu'une hypertrophie de la sous-muqueuse (Aubin *et al.*, 1999). En ce qui concerne la ceinture scapulaire, une réduction de la taille de l'acromion, pouvant aller jusqu'à l'absence complète, est observé chez ces individus mutants (Aubin *et al.*, 2002b). Le squelette axial des souris mutantes présente plusieurs anomalies, dont l'absence du tuberculum antérieur sur C6 (flèche blanche à droite), une côte sur C7 (flèche noire à droite), présence de processus spineux dorsaux à la fois sur C3, C4 et T1 (flèches noires à gauche) et perte de processus spineux sur T2 (flèche blanche). (Autorisation reçue de la part de(s) auteur(s).)



ces deux gènes *Hox* spécifiquement au niveau des neurones moteurs entraîne la diminution du nombre de neurones PMC, ce qui résulte en une détresse respiratoire causant la mort de certains nouveau-nés peu de temps après la naissance (Philippidou *et al.*, 2012). Le diaphragme de ces souris mutantes possède un nerf phrénique plus mince et les axones des PMC qui atteignent le diaphragme sont incapables de former les synapses responsables du mouvement respiratoire. De la même façon, le diaphragme des souris *Hoxa5*^{-/-} montre un phénotype similaire (Boucherat *et al.*, 2013). De plus, des résultats préliminaires de notre laboratoire montrent que la délétion du gène *Hoxa5* dans les neurones moteurs cause également une mortalité périnatale résultant d'une détresse respiratoire, démontrant l'importance de son expression dans les neurones PMC.

1.4.1.4 Organes glandulaires

Bien que certaines souris mutantes *Hoxa5*^{-/-} survivent et soient fertiles, elles présentent néanmoins un retard d'ouverture des yeux et de redressement des oreilles, de même qu'une taille plus petite comparativement à leurs congénères contrôles, caractéristiques d'une hypothyroïdie (Aubin et Jeannotte, 2001; Meunier *et al.*, 2003). En effet, les follicules thyroïdiens des glandes mutantes sont plus petits, désorganisés et contiennent des dépôts anormaux de thyroglobuline, précurseur des hormones T3 et T4 à partir de E16.5 (Figure 1.13). Ces phénotypes sont également associés à des changements d'expression de gènes régulateurs du développement de la glande thyroïde tels que *Nkx2.1*, *Titf2* et *Pax8* (Meunier *et al.*, 2003).

1.4.1.5 Système respiratoire

L'étude du phénotype associé à la détresse respiratoire observée chez les individus mutants *Hoxa5*^{-/-} révèle une trachée désorganisée, caractérisée par une fusion du cricoïde avec les premiers anneaux trachéaux, par la diminution du nombre d'anneaux, combinée à une apparence anormale. De plus, on observe la réduction du diamètre trachéal pouvant aller jusqu'à l'occlusion complète dans les cas plus sévères (Figure 1.13), la désorganisation de l'épithélium et l'épaississement de la lamina propria, laquelle origine du mésenchyme (Aubin *et al.*, 1997). D'ailleurs, bien que le gène *Hoxa5* soit exprimé uniquement dans la

composante mésenchymale, la perte de sa fonction entraîne plusieurs phénotypes affectant la composante épithéliale de différents organes (Figure 1.13) suggérant l'implication de ce dernier dans la communication mésenchyme-épithélium.

Au niveau pulmonaire, la désorganisation du mésenchyme perturbant la formation de tubules épithéliaux normaux est observée dès le jour E12.5, de même que la diminution de l'expression des gènes *Nkx2.1* ainsi que *Foxa2*, qui sont impliqués dans la régulation des protéines du surfactant, et l'augmentation du gène *N-myc*, qui joue un rôle dans la ramification des bronches (Aubin *et al.*, 1997). La réduction de ramifications des bronches cause une diminution du nombre d'acini pulmonaires, détectable à E15.5, ce qui mène à la compaction anormale des poumons à E18.5. À la naissance, les poumons sont affaissés, plus petits et présentent une diminution de la production de la protéine de surfactant SP-B, des pneumocytes de type I, de même qu'une diminution des cellules endothéliales (Boucherat *et al.*, 2013).

Bien que le gène *Hoxa5* soit essentiel au développement pulmonaire embryonnaire, certaines souris mutantes pour ce gène ne meurent pas à la naissance. Toutefois, tel que mentionné à la section 1.2.8.3, le stade de l'alvéogenèse pulmonaire prend place après la naissance entre les jours 5 et 30, durant lesquels il y a augmentation de la surface pulmonaire nécessaire aux échanges gazeux. Les souris *Hoxa5*^{-/-} échappant à la mortalité périnatale présentent une surface pulmonaire diminuée en raison d'un défaut de septation, causé par un défaut de migration des myofibroblastes alvéolaires et qui mène à des espaces aériens élargis (Mandeville *et al.*, 2006; Boucherat *et al.*, 2013). Pour pallier à une moins bonne oxygénation, les souris *Hoxa5*^{-/-} doivent augmenter leur apport d'air, notamment en augmentant leur fréquence respiratoire (Kinkead *et al.*, 2004). De plus, l'analyse des différentes populations de cellules épithéliales tapissant les voies aériennes a permis d'observer une augmentation du nombre des cellules de goblet (sécrétrices de mucus) au niveau des voies respiratoires proximales, de même qu'une présence ectopique de ces mêmes cellules de façon plus distale (Mandeville *et al.*, 2006; Boucherat *et al.*, 2012). D'ailleurs, cette métaplasie des cellules productrices de mucus résulte de la transdifférenciation anormale des cellules de Clara/club, présumées progéniteurs pulmonaires, en cellules de goblet. Cette transdifférenciation est aussi associée à une

augmentation de l'activité de la voie de signalisation NOTCH, ce qui s'apparente aux caractéristiques des poumons de patients atteints d'une maladie pulmonaire obstructive chronique (MPOC) (Boucherat *et al.*, 2012, 2013). La métaplasie des cellules à mucus retrouvée chez les souris mutantes *Hoxa5*^{-/-} se développe indépendamment d'un changement d'expression du gène *Foxa2* (Boucherat *et al.*, 2012). De plus, il y aurait une tendance à la diminution de l'expression du gène *Hoxa5* au niveau des poumons de patients emphysémateux, alors que ceux souffrant d'hypertension pulmonaire primaire le surexprimerait (Golpon *et al.*, 2001).

1.4.1.6 Système digestif

L'expression du gène *Hoxa5* est détectable au niveau de l'intestin antérieur primitif à partir de E9.0, soit au moment où les CCN envahissent le mésenchyme. Par la suite, son expression s'organise de manière à former un gradient rostro-caudal croissant dans l'estomac vers E11.0 (Aubin *et al.*, 2002a). L'abolition de l'expression du gène *Hoxa5* chez la souris entraîne des changements morphologiques au niveau de l'estomac ainsi que fonctionnels au niveau de l'intestin. Premièrement, on remarque un amincissement de l'épithélium gastrique ainsi qu'une hypertrophie de la sous-muqueuse (Figure 1.13). La fonction digestive est également altérée, puisqu'une diminution des cellules entéroendocrines ainsi que des cellules zymogéniques, responsables de la synthèse du pepsinogène, est observée. Bien que tous les types cellulaires présents dans l'estomac soient détectés, la présence ectopique de cellules productrices de mucus au niveau du pylore suggère que l'absence du gène *Hoxa5* génère une transformation homéotique de postériorisation de l'estomac en un devenir intestinal (Aubin *et al.*, 2002a). Enfin, les altérations morphologiques de l'épithélium gastrique observées chez les souris *Hoxa5*^{-/-} sont associées à une restriction du domaine d'expression du gène *Shh* et une extension de celui du gène *Ihh*.

Tout comme pour l'estomac, l'initiation de la transcription du gène *Hoxa5* est détectable au niveau de la composante mésenchymale du futur intestin dès E9.0. Elle se confine par la suite dans le système nerveux entérique (SNE) vers E17.5 et s'installe ultimement en gradient proximo-distal décroissant en s'éloignant du duodénum en direction

du colon (Aubin *et al.*, 1999). Tel que mentionné précédemment, les longs transcrits du gène *Hoxa5* sont exprimés plus tardivement soit à E10.5, alors qu'ils ne sont plus détectables dès E14.5, ce qui consolide à nouveau le rôle fonctionnel du transcrit de 1.8kb. D'ailleurs, l'intestin est le seul organe où l'expression des longs transcrits a été observée. Le phénotype intestinal associé à la perte de fonction du gène *Hoxa5* chez la souris consiste en une réduction de la taille des villi, une distribution anormale des cellules productrices de mucus ainsi qu'une hypertrophie de la sous-muqueuse (Figure 1.13) (Aubin *et al.*, 1999). Tout comme dans le cas de l'estomac, les souris *Hoxa5*^{-/-} présentent une fonction digestive intestinale altérée. En effet, on observe un délai dans l'établissement du mode digestif adulte avec une diminution de l'activité enzymatique de la sucrase, de la tréhalase et de la maltase. Toutefois, les niveaux enzymatiques redeviennent normaux lors du sevrage des individus mutants (Aubin *et al.*, 1999). De la même façon que pour les poumons et l'estomac, tous les types cellulaires sont présents au niveau de l'intestin des souris *Hoxa5*^{-/-}, ce qui s'accorde avec le rôle du gène *Hoxa5* dans la propagation d'un signal pro-différenciation du mésenchyme vers l'épithélium.

1.4.1.7 Système urogénital

Le profil d'expression du gène *Hoxa5* au niveau des ovaires est plutôt dynamique, puisqu'il varie en fonction du cycle oestral, du stade gestationnel, de même qu'en fonction de l'âge des souris, ce qui laisse croire que *Hoxa5* pourrait remplir diverses fonctions physiologiques. Chez les femelles vierges, la perte de fonction du gène *Hoxa5* entraîne une puberté précoce ainsi qu'un dérèglement du cycle œstral (Gendronneau *et al.*, 2012). En effet, ces femelles présentent un cycle œstral prolongé avec une augmentation de la durée du stade diœstrus, un phénotype qui s'exacerbe lors du vieillissement et qui mène à une diminution de fécondité. De plus, les souris *Hoxa5*^{-/-} vieillissantes sont prédisposées à développer des kystes épithéliaux d'inclusion ovariens rappelant l'endosalpingiose humaine. D'ailleurs, ces kystes possèdent les caractéristiques de lésions préneoplasiques, puisqu'ils expriment des marqueurs de carcinomes épithéliaux ovariens tels que les gènes *Pax8* et *Wt1*. (Gendronneau *et al.*, 2012). Enfin, le dérèglement du cycle oestral ainsi que la présence de kystes ayant comme origine l'épithélium de surface ovarien suggèrent un problème au niveau de la EMT. D'ailleurs, la présence de kystes corrèlent avec la

diminution de l'expression de différents membres de la cascade signalétique de l'EGF, responsable de la EMT dans l'OSE, comme les gènes *Egfr*, *Areg* et *Btc*.

1.4.1.8 Glande mammaire

Le développement de la glande mammaire débute chez la souris vers E10.5 et se poursuit après la naissance jusqu'à la puberté. Des modifications structurelles additionnelles surviennent lors de la gestation afin de préparer la glande à la lactation. La perte de fonction du gène *Hoxa5* cause un défaut de lactation observé initialement par l'absence de lait dans l'estomac des nouveau-nés. Une caractérisation plus exhaustive a révélé par la suite que le taux anormalement bas de lait produit par les souris *Hoxa5*^{-/-} est causé par l'accélération du développement lobulo-avéolaire ainsi que l'accumulation de gouttellettes lipidiques (Figure 1.13) (Garin *et al.*, 2006). De plus, le fait que le gène *Hoxa5* soit exprimé au niveau du stroma, alors que le phénotype des souris mutantes soit principalement épithélial, et que ce phénotype soit rétabli lorsqu'un épithélium mutant est greffé dans un stroma sauvage, mais non l'inverse, confirme à nouveau le rôle du gène *Hoxa5* dans la communication mésenchyme-épithélium.

Des études effectuées par un autre groupe ont suggéré que l'expression du gène *p53*, de même que celui du récepteur à la progestérone (*PR*), seraient positivement régulés au niveau transcriptionnel par la liaison de la protéine HOXA5 à leur promoteur respectif dans des lignées cellulaires épithéliales humaines de cancer du sein (Raman *et al.*, 2000a, 2000b). Toutefois, la régulation directe de l'expression de ces deux gènes par la protéine HOXA5 chez la souris a été écartée (Garin *et al.*, 2006; Gendronneau *et al.*, 2010).

1.4.1.9 Cellules hématopoïétiques

Les cellules souches hématopoïétiques se différencient en mégakaryoblastes, pro-érythroblastes, monoblastes, myéloblastes ainsi qu'en lymphoblastes. Cette dernière classe de progéniteurs est à l'origine des cellules B et des cellules T. Comme certaines voies de signalisation connues pour agir sur les gènes *Hox*, telle que la voie de signalisation NOTCH par exemple (section 1.3.4.4; Figure 1.9), jouent également un rôle dans le développement des lymphocytes T (Radtke *et al.*, 1999; Weng *et al.*, 2004), une analyse

de gènes *Hox* candidats a été entreprise. Bien que l'on retrouve le gène *Hoxa5* parmi les gènes identifiés, aucun changement dans les populations cellulaires dérivées des lignées progénitrices lymphoblastiques (cellules B, T et NK), de même que myéloblastiques, n'a été observé chez les individus *Hoxa5*^{-/-}, suggérant une certaine redondance fonctionnelle entre les différents gènes *Hox* lors du développement hématopoïétique (Weerkamp *et al.*, 2006).

Toutefois, des études plus récentes chez des patients souffrant de T-ALL, chez lesquels la transformation leucémique est causée par la protéine de fusion CALM-AF10, ont démontré que l'expression du gène *Hoxa5* est augmentée et contribue au processus leukémogénique (Okada *et al.*, 2006; Lin *et al.*, 2009). En effet, cette surexpression serait causée par le dépôt spécifique de la marque activatrice H3K79 sur le locus *Hoxa5* par la méthyltransférase hDOT1L, laquelle interagit normalement avec la protéine AF10.

1.4.1.10 Angiogenèse et microenvironnement tumoral

Le gène *Hoxa5* est également exprimé au sein des cellules endothéliales quiescentes, mais ne l'est pas dans celles qui sont activées (Rhoads *et al.*, 2005). D'ailleurs, les cellules endothéliales exprimant le gène *Hoxa5* ne répondent pas aux stimuli angiogéniques générés par VEGFA et FGF2. De plus, l'expression des gènes pro-angiogéniques *Vegfr2*, *EfnA1*, *Hif1* et *Cox2* est diminuée, celle du gène anti-angiogénique *Tsp2* est augmentée, alors que les jonctions adhérentes sont stabilisées dans ces cellules (Arderiu *et al.*, 2007). L'expression du gène *Hoxa5* est quant à elle régulée négativement par le miRNA miR-130A durant l'angiogenèse (Chen et Gorski, 2008). De façon intéressante, l'expression du gène paralogue *Hoxb5* entraîne à la fois l'augmentation de la vasculogenèse (Wu *et al.*, 2003) et de l'angiogenèse (Winnik *et al.*, 2009), ce qui va à l'encontre de l'hypothèse de redondance fonctionnelle. Enfin, l'expression du gène *Hoxa5* semble être dérégulée dans plusieurs types de cancers tels que certains carcinomes oraux à cellules squameuses (Rodini *et al.*, 2012), cancers du poumon non à petites cellules (Kim *et al.*, 2009; Liu *et al.*, 2012), cancers du sein (Bagadi *et al.*, 2008; Gendronneau *et al.*, 2010) et certains hémangiomes (Zhu *et al.*, 2009).

1.4.2 Régulation de l'expression du gène *Hoxa5*

Mis à part les travaux de caractérisation des éléments qui régulent transcriptionnellement l'expression du gène *Hoxa5*, effectués essentiellement avec une approche de transgénèse chez la souris, qui vous seront présentés dans les quatre prochaines sections (sections 1.4.2.1 à 1.4.2.4), des informations ont également été collectées en utilisant d'autres modèles d'études. Tout d'abord, des études chez la drosophile ont permis de regrouper certains gènes sous la bannière d'une nouvelle famille appelée activateurs/amplificateurs des gènes *Trithorax* et *Polycomb* (ETP). Cette nouvelle famille contient des gènes tels que *Pho*, l'homologue du gène *Yy1* abordé à la section 1.3.8.3, de même que *Corto* et *Dsp1* (Decoville *et al.*, 2001; Lopez *et al.*, 2001). D'ailleurs, ces deux derniers membres ETP se retrouvent liés à une séquence régulant en *cis* l'expression du gène *Scr*, l'homologue du gène *Hoxa5* (Salvaing *et al.*, 2006). D'autre part, l'utilisation de différentes lignées cellulaires comme modèles d'étude a permis de révéler qu'un traitement au RA entraîne la dérégulation du promoteur *Hoxa5*, caractérisée par la perte de la marque épigénétique de H3K27me3 ainsi que l'apparition des marques permissives H3K9/K14 ac et H3K4me3 (Laursen *et al.*, 2013). De plus, ce même traitement stabilise l'ARNm du gène *Hoxa5* via la liaison de la protéine se liant aux ARN HuR de même qu'en réprimant le miRNA *miR-130a* (Yang *et al.*, 2013).

Dans le laboratoire, la délétion successive des séquences flanquantes du gène *Hoxa5* dans un contexte de transgénèse chez la souris a permis d'identifier quatre grandes régions régulatrices dirigeant l'expression du transgène *Hoxa5/lacZ*, qui sont représentées dans la Figure 1.14. Il s'agit des éléments Temporel (section 1.4.2.1), BSC ou Brachial Spinal Cord (section 1.4.2.2), MES ou Mesodermal Enhancer Sequence (section 1.4.2.3), ainsi que l'élément Poumon/Intestin (section 1.4.2.4) (Tuggle *et al.*, 1990; Larochelle *et al.*, 1999; Nowling *et al.*, 1999; Moreau et Jeannotte, 2002; Tabariès *et al.*, 2005). La capacité des éléments régulateurs du gène *Hoxa5* à diriger l'expression d'une recombinase Cre sera décrite aussi au Chapitre 3.

1.4.2.1 Élément Temporel

La microinjection d'une construction d'ADN génomique de 11.1 kb, comprenant 3.8 kb de séquence en amont du promoteur du gène *Hoxa5* ainsi que 7.3 kb en aval, de même que les séquences d'ADN codant pour la β -Galactosidase, permet de récapituler en partie l'expression spatio-temporelle du transcrit fonctionnel de 1.8kb (Larochelle *et al.*, 1999). En effet, suite aux délétions successives en transgénèse, les colorations au X-gal sur des embryons entiers ont permis de déduire que l'élément Temporel est localisé dans le fragment d'ADN *EcoRI-AccI* de 648 pb (-2.1 kb à - 1.5 kb). Cette séquence régulatrice est en mesure de diriger l'expression du transgène *Hoxa5/lacZ* aussi tôt qu'au stade E8.25, au niveau des somites 5-7 (Figure 1.14). De plus, le fragment d'ADN de 11.1 kb englobe également les éléments BSC et MES qui seront présentés dans les deux prochaines sections (section 1.4.2.2 et 1.4.2.3 respectivement).

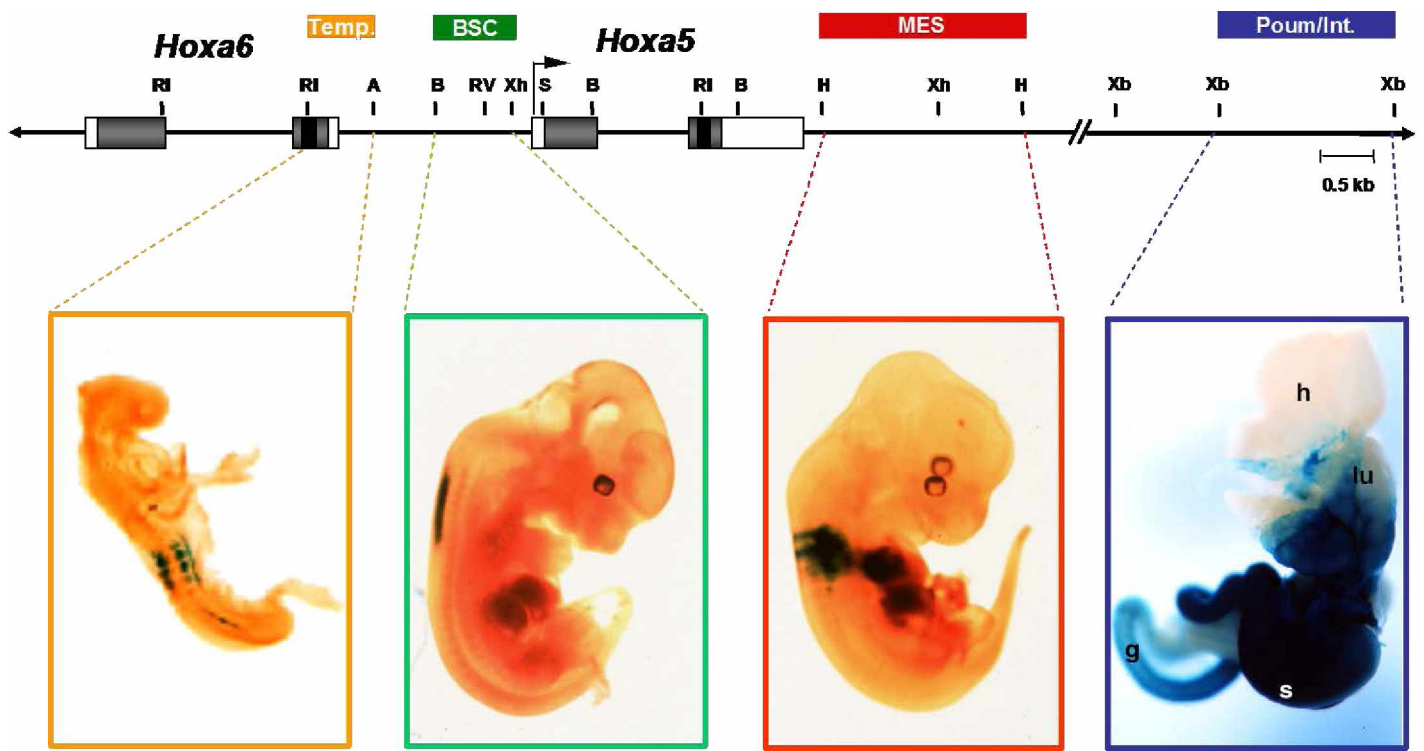
1.4.2.2 Élément BSC (Brachial Spinal Cord)

L'étude de l'élément BSC a été initiée par un autre groupe qui démontra qu'une séquence d'ADN génomique de 4.3 kb comprise entre les positions -1.5 kb et +2.9 kb contrôle l'expression du gène rapporteur *lacZ* à E11.0 au niveau de la partie brachiale du tube neural (Figure 1.14) (Zákány *et al.*, 1988; Tuggle *et al.*, 1990). En collaboration avec l'équipe du Dr Tuggle, des délétions successives du fragment de 4.3 kb ont été réalisées et un fragment de 912 pb capable de diriger l'expression du transgène *Hoxa5/lacZ* au niveau du tube neural, dans un domaine localisé au niveau des vertèbres C4 et T2, a été circonscrit. L'élément régulateur fut par la suite associé à un fragment d'ADN de 604 pb (-839 pb à - 235 pb) sans altérer ses propriétés régulatrices. De plus, trois sites de liaison AAATAA ont été identifiés au niveau des séquences de l'élément BSC comme étant essentiels à l'expression du transgène *Hoxa5/lacZ* dans la partie brachiale du tube neural (Nowling *et al.*, 1999).

1.4.2.3 Élément MES (Mesodermal Enhancer Sequence)

Tel que mentionné à la section 1.4.4.1, le fragment d'ADN génomique de 11.1 kb est en mesure de récapituler fidèlement le profil d'expression endogène du transcrit

Figure 1.14: Résumé des régions régulatrices contrôlant l'expression du gène *Hoxa5* identifiées par transgénèse chez la souris. (Haut) Représentation schématique des loci *Hoxa5* et *Hoxa6*, de même que de certains sites de restriction. Les boîtes blanches correspondent aux exons des gènes, les grises aux régions codant pour les protéines et les noires représentent les domaines homéo. Temp., Temporal; BSC, Brachial Spinal Cord; MES, Mesodermal Enhancer Sequence; Poum./Int., Poumon/Intestin. A, *AccI*; B, *BglII*, RI, *EcoRI*; RV, *EcoRV*; H, *HindIII*; S, *SacI*; Xb, *XbaI*; Xh, *XhoI*. (Bas) De gauche à droite: l'élément Temporel (orange) dirigeant l'expression du transgène *Hoxa5/lacZ* dès le jour E8.5 au niveau des somites 5-7 (Larochelle *et al.*, 1999), le BSC (vert) au niveau de la partie brachiale du tube neural à E12.5 (Tuggle *et al.*, 1990; Nowling *et al.*, 1999), le MES (rouge) dans les dérivés mésodermiques entre pv3 et pv10, les bourgeons de membres et le tractus urogénital (Larochelle *et al.*, 1999; Tabariès *et al.*, 2005) ainsi que l'élément poumon/intestin qui dirige l'expression dans les systèmes respiratoire et digestif (Moreau et Jeannotte, 2002). g, intestin; h, cœur; lu, poumon; s, estomac (tirée de la thèse de Ph.D. de S. Tabariès, 2006). (Autorisation reçue de la part de(s) auteur(s).)



fonctionnel *Hoxa5* de 1.8 kb. D'ailleurs, une construction d'ADN contenant un fragment *HindIII-HindIII* de 2.1 kb, situé en 3' de l'exon 2 (+2.9 kb à 5 kb), contrôle l'expression du transgène *Hoxa5/lacZ* au niveau des dérivés mésodermiques entre les pv 3 et 10, les bourgeons de membres antérieurs ainsi que le tractus urogénital (Figure 1.14) (Larochelle *et al.*, 1999). Une analyse par délétions successives de cette séquence régulatrice, appelée MES pour Mesodermal Enhancer Sequence, a mis en évidence la présence de deux sous-régions régulatrices, soit le fragment d'ADN de 164 pb *AvrII-Eco47III* (AvEc) et le fragment de 331pb *PstI-HindIII* (PsHi). L'analyse approfondie du fragment AvEc a fait ressortir des sites de liaison consensus pour les facteurs de transcription CDX, qui lient plus fortement la protéine CDX4 et qui sont nécessaires à l'établissement de la frontière postérieure d'expression du transgène *Hoxa5/lacZ* au niveau de pv10 (Tabariès *et al.*, 2005). Quant au fragment PsHi, bien qu'il soit suspecté de posséder une propriété activatrice/amplificatrice de l'élément MES, la confirmation de son rôle régulateur dans l'expression du gène *Hoxa5* reste à faire. Des précisions sur les rôles de ces deux sous-régions régulatrices seront présentées au chapitre 4.

1.4.2.4 Élément Poumon/Intestin

La séquence d'ADN génomique de 11.1 kb n'est pas en mesure de diriger adéquatement l'expression de la β -galactosidase dans les systèmes respiratoire et digestif, où le gène *Hoxa5* est normalement exprimé (Larochelle *et al.*, 1999). De façon intéressante, certaines données dans la littérature suggèrent une association entre le traitement au RA et l'activation de l'expression du gène *Hoxa5* dans les poumons, l'estomac et l'intestin. Par exemple, le traitement de poumons embryonnaires en culture au RA induit l'expression des gènes *Hoxa5*, *Hoxb5* et *Hoxb6* (Bogue *et al.*, 1994; Volpe *et al.*, 1997), alors que l'inactivation du gène *Raldh2* entraîne la diminution du gène *Hoxa5* dans les poumons et l'estomac (Wang *et al.*, 2006). De plus, un RARE nécessaire à l'expression du gène *Hoxa4* au niveau des poumons, de l'intestin et du métanéphros a été localisé à l'intérieur de la séquence intergénique *Hoxa4-Hoxa5* (Packer *et al.*, 1998). Comme les éléments régulateurs sont couramment partagés entre les gènes *Hox* voisins (section 1.3.5), l'hypothèse voulant que ce site RARE soit aussi important pour l'expression du gène *Hoxa5* qu'il l'est pour l'expression du gène *Hoxa4* a été mise de l'avant. D'ailleurs, ce site RARE

est localisé à l'intérieur d'un fragment d'ADN *XbaI-XbaI* de 1.5 kb, circonscrit entre les positions +9351 et +10816, qui récapitule le profil d'expression du gène *Hoxa5* dans ces organes (Figure 1.14) (Moreau et Jeannotte, 2002). La caractérisation approfondie de cet élément régulateur, que l'on appelle l'élément Poumon/Intestin (Poum./Int.), sera présentée au chapitre 2.

1.5 Projet de ma thèse de doctorat

Tel que mentionné, la régulation aberrante de l'expression du gène *Hoxa5* est associée ou directement reliée à certains types de tumeurs, dont les mieux documentés sont les leucémies (Okada *et al.*, 2006; Lin *et al.*, 2009) et les cancers du sein (Raman *et al.*, 2000a, 2000b; Bagadi *et al.*, 2008; Gendronneau *et al.*, 2010), ainsi que certaines pathologies, comme les MPOC (Golpon *et al.*, 2001; Boucherat *et al.*, 2012, 2013). De plus, les phénotypes observés chez les souris *Hoxa5*^{-/-} ont tous quelque chose en commun, à savoir qu'ils sont majoritairement confinés à la composante épithéliale de différents organes (Aubin *et al.*, 1997, 1999, 2002a; Meunier *et al.*, 2003; Garin *et al.*, 2006; Mandeville *et al.*, 2006; Gendronneau *et al.*, 2012; Boucherat *et al.*, 2012, 2013), alors que le gène *Hoxa5* est exprimé au niveau du mésenchyme, ce qui suggère qu'il joue un rôle dans la communication mésenchyme-épithélium. À la lumière de ces résultats, je me suis fixé comme but de mieux comprendre la régulation transcriptionnelle complexe du gène *Hoxa5*, puisqu'une meilleure compréhension des différentes cascades développementales impliquant le gène *Hoxa5* permettra ultimement de développer des avenues diagnostiques et/ou thérapeutiques potentielles pour différents cancers et pathologies.

À cet effet, comme la mortalité périnatale associée à la dysmorphogénèse des voies respiratoires correspond au phénotype le plus important retrouvé chez les souris *Hoxa5*^{-/-}, l'objectif principal de ma thèse consista en la caractérisation de l'élément poumon/intestin situé au sein du fragment d'ADN *XbaI-XbaI* de 1465 pb, localisé dans la séquence intergénique *Hoxa4-Hoxa5* et pouvant diriger l'expression d'un transgène *Hoxa5/lacZ* au niveau des systèmes respiratoire et digestif (Moreau et Jeannotte, 2002). J'ai entrepris d'identifier les séquences d'ADN agissant en *cis* sur l'expression du gène *Hoxa5* au niveau des organes à l'aide de délétions successives dans un contexte de transgénèse chez la souris.

Ensuite, j'ai identifié des facteurs de transcription susceptibles de se lier à ces séquences par EMSA et par ChIP. L'une des hypothèses préliminaires était qu'un site RARE essentiel à l'expression du gène *Hoxa4* dans les organes se voulait également important pour l'expression du gène *Hoxa5* à ce niveau. Cet objectif de recherche est décrit dans le cadre du chapitre 2 de ma thèse.

Une autre caractéristique des phénotypes observés chez les souris *Hoxa5*^{-/-} est qu'ils sont majoritairement regroupés au niveau cervico-thoracique entre les pv3 et pv10, soit le domaine d'expression spécifique au transcrit de 1.8 kb. Comme cette région correspond au domaine d'expression dirigé par l'élément MES (Larochelle *et al.*, 1999; Tabariès *et al.*, 2005), je me suis également fixé comme objectif de caractériser ces séquences régulatrices. Tout d'abord, je me suis penché sur la nature de l'interaction entre le fragment d'ADN AvEc-164 pb et les facteurs de transcription CDX4 et HOXB9, lors du positionnement de la limite postérieure d'expression au niveau de pv10, en utilisant essentiellement la même approche que pour l'objectif précédent. J'ai également tenté de localiser la séquence activatrice/amplificatrice à l'intérieur du fragment PsHi-331. Les résultats obtenus concernant cet objectif sont présentés au chapitre 4 de ma thèse.

Grâce aux connaissances acquises en tentant de répondre aux deux objectifs précédents, j'ai généré deux constructions d'ADN dans lesquelles les séquences encodant la recombinase Cre sont placées sous le contrôle de différentes combinaisons des séquences régulatrices du gène *Hoxa5*. La première englobe la totalité des séquences régulatrices du gène *Hoxa5* identifiées jusqu'à maintenant (Temporel, BSC, MES et Poumon/Intestin), alors que j'ai retiré l'élément Poumon/Intestin de la deuxième. J'ai effectué la caractérisation des lignées de souris obtenues, notamment en les utilisant pour déléter le gène *Hoxa5* et comparé les phénotypes obtenus avec ceux de la lignée de souris mutantes générée précédemment (Jeannotte *et al.*, 1993), notamment en effectuant des préparations squelettiques et des analyses histologiques du système respiratoire. Les lignées transgéniques *Hoxa5Cre* serviront d'outils pour confirmer des nouveaux membres potentiels de la cascade développementale impliquant le gène *Hoxa5*. La présentation de ces nouvelles lignées fait l'objet du chapitre 3 de ma thèse.

Enfin, bien que la majorité des phénotypes générés par la perte de fonction du gène *Hoxa5* chez la souris correspondent à des régions où seulement le transcrit de 1.8 kb est présent, il est improbable que les 4 longs transcrits non codants du gène *Hoxa5* soient uniquement du bruit de fond transcriptionnel, communément appelé transcrits poubelles. Par conséquent, j'ai participé à l'étude des rôles de ces longs transcrits du gène *Hoxa5*. Pour ce faire, j'ai vérifié le potentiel de traduction des transcrits *Hoxa5* dans un contexte *in vitro* (trousse de transcription/traduction) et *in vivo* (expression transitoire dans des HEK 293T). Cette étude est présentée au chapitre 5.

CHAPITRE II

YY1 acts as a transcriptional activator of *Hoxa5* gene expression in mouse organogenesis

Félix-Antoine Bérubé-Simard, Christelle Prudhomme and Lucie Jeannotte[#]

Department of Molecular Biology, Medical Biochemistry and Pathology, Université Laval, Québec, Canada; Centre de recherche sur le cancer de l'Université Laval; Centre de recherche du Centre Hospitalier Universitaire de Québec, L'Hôtel-Dieu de Québec, Québec, Canada

[#] Corresponding author Lucie Jeannotte
and reprint requests: Centre de recherche sur le cancer de l'Université Laval
CRCHU-Q, L'Hôtel-Dieu de Québec
9, rue McMahan
Québec, QC
Canada G1R 2J6
Phone: (418) 525-4444 ext. 15552
Fax: (418) 691-5439
e-mail: lucie.jeannotte@crhdq.ulaval.ca

Running title: Regulation of *Hoxa5* lung and gut expression by YY1
Key words: *Hoxa5*, YY1, gene regulation, lung and gut development
Grant support: This work was supported by grants from the Canadian Institutes of Health Research (MOP-15139; to L.J.) and the National Sciences and Engineering Research Council of Canada (to L.J.)
Disclosure statement: The authors have nothing to disclose.

Article publié en 2014 dans: Plos One 9(4): e93989

AVANT-PROPOS

La cartographie des différents éléments régulateurs responsables de diriger l'expression du gène *Hoxa5* a mené à l'identification d'une séquence d'ADN essentielle à son expression au niveau des systèmes respiratoire et digestif (Moreau et Jeannotte, 2002). La majeure partie de ma thèse a été consacrée à la poursuite de la caractérisation de l'élément Poumon/Intestin (P/I), situé dans la séquence intergénique *Hoxa4-Hoxa5* entre deux sites reconnus par l'enzyme de restriction *XbaI* qui sont respectivement localisés aux positions +9351 et + 10816.

Initiées tour à tour par Julie Moreau et Christelle Prudhomme durant leurs études graduées, j'ai poursuivi les délétions successives de cette région dans un contexte de transgénèse chez la souris générant ainsi près de la moitié des constructions (11/23) et identifiant deux sous-régions régulatrices : un fragment d'ADN de 259 pb *XbaI-BssHII* (positions +9351 et +9610) dirigeant essentiellement l'expression du transgène *Hoxa5/lacZ* au niveau du système digestif, de même qu'un fragment de 163 pb *NcoI-SacI* (positions +10198 et +10361), possédant des propriétés activatrices/amplificatrices et contrôlant l'expression au niveau des poumons, de l'estomac et de l'intestin. Par conséquent, la délétion de cette dernière séquence d'ADN entraîne l'absence de coloration X-Gal dans les systèmes respiratoire et digestif. L'analyse bioinformatique (TESS/TFSEARCH) des fragments de 259 pb et de 163 pb a souligné la présence de sites de liaison potentiels pour le facteur de transcription YY1. Par la suite, j'ai confirmé l'importance de ces sites *in vitro* par essais de retard sur gel (EMSA), puisque un anticorps dirigé contre la protéine YY1 abolit la liaison sonde-protéine. De plus, j'ai montré par immunoprécipitation de la chromatine (ChIP) que ces mêmes sites sont occupés *in vivo* et que leurs mutations dans un contexte de transgénèse est suffisantes pour abolir l'expression du transgène *Hoxa5/lacZ*.

Enfin, j'ai confirmé le lien génétique entre les gènes *Yy1* et *Hoxa5* dans le développement du poumon. Puisque les souris mutantes *Yy1*^{-/-} meurent au jour E6.5 (Donohoe *et al.*, 1999), alors que les souris hypomorphes *Yy1*^{Floxneo/-} meurent à la naissance de détresse respiratoire (Affar *et al.*, 2006), j'ai délété spécifiquement le gène *Yy1* au niveau

du mésenchyme en utilisant une lignée de souris exprimant la *Dermo1/Cre* (*Twist2/Cre*) (Yu *et al.*, 2003). J'ai ainsi démontré qu'en plus d'avoir un phénotype pulmonaire comparable à celui des souris *Hoxa5*^{-/-}, la délétion mésenchymale du gène *Yy1* entraîne une diminution de l'expression du gène *Hoxa5* au niveau des poumons d'embryons au stade E18.5, confirmant ainsi l'importance de ce facteur de transcription dans la régulation positive de l'expression du gène *Hoxa5* au niveau des systèmes respiratoire et digestif. Il s'agit d'une découverte très importante puisque la protéine YY1 est principalement connue pour son rôle de régulateur négatif dans l'expression des gènes *Hox*.

J'ai collaboré à l'écriture du manuscrit avec le Dre Lucie Jeannotte.

ABSTRACT

The *Hox* gene family encodes homeodomain-containing transcriptional regulators that confer positional information to axial and paraxial tissues in the developing embryo. The dynamic *Hox* gene expression pattern requires mechanisms that differentially control *Hox* transcription in a precise spatio-temporal fashion. This implies an integrated regulation of neighbouring *Hox* genes achieved through the sharing and the selective use of defined enhancer sequences. The *Hoxa5* gene plays a crucial role in lung and gut organogenesis. To position *Hoxa5* in the regulatory hierarchy that drives organ morphogenesis, we searched for *cis*-acting regulatory sequences and associated *trans*-acting factors required for *Hoxa5* expression in the developing lung and gut. Using mouse transgenesis, we identified two DNA regions included in a 1.5-kb *XbaI-XbaI* fragment located in the *Hoxa4-Hoxa5* intergenic domain and known to control *Hoxa4* organ expression. The multifunctional YY1 transcription factor binds the two regulatory sequences *in vitro* and *in vivo*. Moreover, the mesenchymal deletion of the *Yy1* gene function in mice results in a *Hoxa5*-like lung phenotype with decreased *Hoxa5* and *Hoxa4* gene expression. Thus, YY1 acts as a positive regulator of *Hoxa5* expression in the developing lung and gut. Our data also support a role for YY1 in the coordinated expression of *Hox* genes for correct organogenesis.

INTRODUCTION

Hox genes encode evolutionarily conserved transcription factors that control the formation of body segment-specific structures by regulating the transcription of downstream effectors that, in turn, direct the morphogenetic events leading to the complex body forms along the embryonic axes in metazoan [1-2]. Consequently, mutations in *Hox* genes alter segmental identity and cause morphological defects. In mammals, 39 *Hox* genes are distributed over four clusters, each containing 9 to 11 genes closely packed in less than 150-kb of sequences (*HoxA* to *D*). Their spatio-temporal expression profile during embryogenesis reflects their arrangement in the clusters: the 3' most genes are expressed earlier and their expression domain reaches a more anterior limit than those occupying 5' positions. As a result, members of the *Hox* complexes are expressed in nested and overlapping domains along the developing body suggesting that specific combinations of HOX proteins provide a unique address to defined regions [3]. Based on sequence homology and location within clusters, *Hox* genes are also classified into 13 paralog groups.

The *Hox* clustered organization is fundamental for the precise regulation and the function of each gene and hence for the correct formation of the embryo. Analysis of *Hox* mutant mice endorses the collinear relationship between the position of individual genes within *Hox* clusters and the structural defects observed along the anterior-posterior (A-P) axis [4]. For example, the mutation of the *Hoxa5* gene, located in the middle of the *HoxA* multigenic complex, affects axial specification at the cervico-thoracic level [5]. A high percentage of *Hoxa5*^{-/-} pups die at birth from impaired respiratory tract development [6]. Moreover, the loss of *Hoxa5* function results in a panoply of phenotypes indicative of the broad range of *Hoxa5* actions throughout life [7-11]. Most defects in *Hoxa5*^{-/-} mutants are confined to the cervico-thoracic region corresponding to the *Hoxa5* rostral-most expression domain, where the major *Hoxa5* transcript of 1.8-kb encoding the HOXA5 270-amino-acid protein, is specifically expressed [12]. Thus, *Hoxa5* appears as a critical determinant in the specification and the development of a subset of structures at the cervico-thoracic level.

While the developmental role of *Hox* genes is well established, the regulation of

Hox gene expression in the embryo remains incompletely understood. A complex array of different modes of regulation governs the precise *Hox* expression [13-14]. Regulation primarily occurs at the transcriptional level via the combinatorial interplay of several signaling pathways and transcriptional factors that interact with positive and negative *cis*-acting sequences to differentially control *Hox* expression in a spatio-temporal and tissue-specific fashion. The proximity of *Hox* genes in clusters implies the integrated regulation of adjacent *Hox* promoters through the sharing, the competition and/or the selective use of defined enhancers [15]. In parallel, global regulatory elements located outside the *Hox* clusters and able of long-distance action coordinate the expression of several genes along the *Hox* complexes [14]. Large-scale chromatin remodeling events also participate to the regulation of the collinear expression of *Hox* genes [16].

Transcriptional regulators of *Hox* gene expression have been identified [17]. They include developmentally regulated factors like the CDX homeodomain-containing proteins that integrate retinoic acid (RA), FGF and Wnt signaling for the setting of the correct expression domain of *Hox* genes [18-21]. *Hox* genes are also directly responsive to RA, which activates retinoic acid receptors that then interact with retinoic acid response elements (RARE) identified near *Hox* genes mainly from paralog groups 1 to 5 [22-23]. *Hox* expression is under the control of HOX proteins themselves involved in auto- and cross-regulation [24-25]. Finally, ubiquitously expressed transcription factors such as the multifunctional Yin Yang 1 (YY1) protein can modulate *Hox* expression in specific contexts [26-29].

A complex organization of overlapping transcriptional units encompassing the *Hoxa5* locus exists, which results from alternative splicing and the use of three promoters, one proximal producing the 1.8-kb transcript and two distal ones giving rise to long noncoding RNAs [12]. Using a transgenic approach, we have identified regulatory elements directing the developmental expression of the *Hoxa5* proximal promoter. An 11.1-kb genomic fragment can recapitulate the temporal expression and substantially reconstitute the spatial pattern of the 1.8-kb *Hoxa5* transcript in mouse embryos. It includes DNA control sequences, such as the 604-bp brachial spinal cord (BSC) enhancer and a 650-bp temporal control region, both contained in the *Hoxa5* 5' flanking sequences [30-

31]. A 2.1-kb mesodermal (MES) enhancer important for *Hoxa5* paraxial and lateral plate mesoderm expression in the cervico-upper thoracic region of the A-P axis is positioned 3' of the *Hoxa5* gene. CDX proteins bind this sequence acting as potential regulators for the regionalization of *Hoxa5* gene expression along the axis [32]. A 1.5-kb DNA region that targets *Hoxa5* lung and gut developmental expression was also identified in the *Hoxa4-Hoxa5* intergenic sequences [33].

Several *Hox* genes, mainly from paralog groups 1 to 8, are expressed along the respiratory tract [34]. However, except for *Hoxa5*, the lack of overt lung phenotype in single *Hox* mutants indicates that *Hoxa5* plays a predominant function in lung ontogeny [35]. The prevalent role of *Hoxa5* in organ development prompted us to further characterize *Hoxa5* lung and gut regulatory sequences. Here, we present evidence that *Hoxa5* lung and gut expression is under the control of several DNA elements. Some are shared with the flanking *Hoxa4* gene and they bind the transcription factor YY1, which acts as a positive regulator of *Hoxa5* gene expression in the developing lung and gut.

MATERIALS AND METHODS

Ethics statement

All animal experimentations were performed according to the guidelines of the Canadian Council on Animal Care and they were specifically approved by the institutional animal care committee (Comité de protection des animaux du Centre de recherche du Centre Hospitalier Universitaire de Québec (CPAC); Permit Number: 2012013-2).

Design of *Hoxa5/lacZ* transgenes

All *Hoxa5/lacZ* constructs contain the bacterial *lacZ* gene inserted into the *SacI* site of the first exon of the *Hoxa5* gene, which allows translation of the *lacZ* ORF from the HOXA5 AUG [31]. Construct 1 (also named pLJ272) was obtained by adding a 5.20-kb *EcoRI-ApaI* DNA fragment at the 3' end of construct 2 from ref. [33]. Construct 2 (pLJ123) was previously described (construct 6 in ref. [31]). It was used as the backbone for most *Hoxa5/lacZ* constructs in the present study. Constructs 3, 4, 8-12, 14-15 and 17-21 were

obtained by ligation at the 3' end of pLJ123 of genomic fragments from the *Hoxa4-Hoxa5* intergenic region. Construct 5 was produced by cloning the 1.5-kb *XbaI-XbaI* fragment upstream an *XhoI-HindIII Hoxa5* genomic fragment carrying a *Hoxa5* minimal promoter that cannot direct β -galactosidase expression by itself (pLJ143 in ref. [32]). Enhancer activity was tested in constructs 6, 7, 13, 16 and 22 by cloning *Hoxa5* regulatory regions in front of the *hsp68lacZpA* fragment, which contains an *hsp68* minimal promoter that cannot direct *lacZ* developmental tissue-specific expression by itself [31, 36]. For constructs 17 to 22, nucleotide substitutions in RARE and YY1 binding sites were produced by an overlapping PCR strategy using synthetic oligonucleotide primers carrying appropriate base changes (Table 1A). Introduced mutations were confirmed by sequencing and disruption of the binding sites was validated with the TFSEARCH (<http://molsun1.cbrc.aist.go.jp/research/db/TFSEARCH.html>) and TESS (<http://www.cbil.upenn.edu/cgi-bin/tess/tess>) softwares.

Production of transgenic mice

The microinjected *Hoxa5/lacZ* and *hsp68lacZpA* fragments were obtained by a *SalI+NotI* digestion and an *XhoI+XbaI* digestion, respectively, to eliminate plasmid sequences. Transgene microinjection into fertilized eggs from C57BL/6 X CBA F1 hybrid intercrosses were done according to standard procedures [37]. Transgenic founder embryos were recovered from foster mothers, genotyped by Southern analysis of yolk sac DNA using a *lacZ* specific probe to assess the integrity of the microinjected construct, and analyzed for *lacZ* expression by β -galactosidase staining. Permanent mouse lines were also obtained for construct 1 and genotyped with tail DNA [31].

Mice, genotyping, tissue collection, histology, immunohistochemistry (IHC) and β -galactosidase staining

To assess the effect of YY1 on *Hoxa5* expression, mutant mouse lines for the *Yy1*^{flox} and *Yy1*^{Floxneo} alleles were used (provided by Y. Shi) [38]. For the mesenchymal ablation of *Yy1*, we used *Dermo1*^{Cre} mice provided by D. Ornitz [39]. *Dermo1*^{Cre} mice express the Cre recombinase protein from the endogenous *Dermo1* (*Twist2*) locus. *Yy1*^{flox/flox}*Dermo1*^{+Cre} mice were generated by matings between *Yy1*^{flox/+}*Dermo1*^{+Cre} and

Yy1^{fl^{ox}/fl^{ox}} mice.

Age of the embryos was estimated by considering the morning of the day of the vaginal plug as embryonic (E) day 0.5. Experimental specimens were genotyped by Southern blot and PCR analyses. Organs were collected from E13.5, E15.5 and E18.5

Table 2.1: Oligonucleotide sequences

Fragment	Sequence (5'-3')	Amplicon (bp)
A) Mutagenesis		
259-bp <i>XbaI</i> - <i>BssHII</i> (mRARE)	F-GCGCAAAGTCCAAGGCCG <u>TGGTGC</u> ACTT CTAGTCGGCGCGTTCTAA	46
259-bp <i>XbaI</i> - <i>BssHII</i> (mYY1)	F-GACCCTGTCGGCCGCCAGGGC	22
163-bp <i>NcoI</i> - <i>SacI</i> (mYY1)	F-CAGTACCCTGTCACCCTGTCGG	22
B) EMSA on the 259-bp <i>XbaI</i>-<i>BssHII</i> DNA fragment (C3)		
Oligo RARE	F-GCGCAAAGTCCAAGGCCGAGGTGAACTT CAGGTCAGCGCGTTCTAA	46
Oligo mRARE	F-GCGCAAAGTCCAAGGCCG <u>TGGTGC</u> ACTT CTAGTCGGCGCGTTCTAA	46
Fragment A	F-CTAGAGAAGTGTGGATGAGATAG R-CATTATGGGATGTATTGACTG	85
Fragment B	F-CAGTCAATACATCCATAATG R-CAGTGTCTTTTAGAGAGCTGG	104
Fragment C	F-CCAGCTCTCTAAAAGACACTG R-CCCTCACCAGCCGACATTTTC	110
Oligo C1	F-CCCAGCTCTCTAAAAGACACTGTATAGA CCTTTTAGAAGCG	41
Oligo C2	F-AAGCGCAAAGTCCAAGGCCGAGGTGAACT TCAGGTCAGCG	40
Oligo C3	F-CAGCGCGTCTAACAAATATGAAAATGTC GGCTGGTGAGGGCGCG	44
Oligo C3(m1)	F-GCTATATTATAACAAATATGAAAATGTC GGCTGGTGAGGGCGCG	44
Oligo C3(m2)	F-CAGCGCGTCGCCACCCGCTGAAAATGTC GGCTGGTGAGGGCGCG	44
Oligo C3(m3)	F-CAGCGCGTCTAACAAATAGTCCCCGTGC GGCTGGTGAGGGCGCG	44
Oligo C3(m4)	F-CAGCGCGTCTAACAAATATGAAAATGTA TTAGTTGTAGGGCGCG	44
Oligo C3(m5)	F-CAGCGCGTCTAACAAATATGAAAATGTC GGCTGGTGCTCTATAC	44
Oligo-18(C3)	F-TGAAAATGTCGGCTGGTG	18
Oligo C3(mYY1)	F-TCAGCGCGTCTAACAAATATGACCCTGTC GGCCGCCAGGGCGCG	45
Oligo C3(m1YY1)	F-TCAGCGCGTCTAACAAATATGACCCTGTC GGCTGGTGAGGGCGCG	45
Oligo C3(m2YY1)	F-TCAGCGCGTCTAACAAATATGAAAATGTC GGCCGCCAGGGCGCG	45

Consensus YY1	F-GGGGATCAGGGTCTCCATTTTGAAGCGG GATCTCCC	36
C) EMSA on the 433-bp <i>Mef1-SacI</i> DNA fragment (G3)		
Oligo G1	F-CATGGGATTTTTGCTATGGCTTGCTTGCA AAGGGAGGCTGTGGAA	45
Oligo G2	F-TGGAAGGTTTCAGGAAAGGTACTGAGATT GTTTATTACAGCCATAA	45
Oligo G3	F-CATAAATCTTGCAGTAAAATGTCAAAAT GTCGGTGTGTGAGATAA	45
Oligo G4	F-GATAACACTTGGTGGTCTGGCTCCGTTT GTGTTTATGATAGGAGCT	47
Oligo G3(m1)	F- <u>GCGCCCCAGT</u> GCAGTAAAATGTCAAAAT GTCGGTGTGTGAGATAA	45
Oligo G3(m2)	F-CATAAATCTGTACTGCCGATGTCAAAAT GTCGGTGTGTGAGATAA	45
Oligo G3(m3)	F-CATAAATCTTGCAGTAAAGGTAAGTCCT GTCGGTGTGTGAGATAA	45
Oligo G3(m4)	F- CATAAATCTTGCAGTAAAATGTCAAAACTGA TTGTGGTGTGAGATAA	45
Oligo G3(m5)	F-CATAAATCTTGCAGTAAAATGTCAAAAT GTCGGTGTGTGAGATAA	45
Oligo-18(G3)	F-ATGTCAAAATGTCGGTGT	18
Oligo G3(mYY1)	F-CATAAATCTTGCAGTACCCTGTCAACCCTG TCGGTGTGTGAGATAA	45
Oligo G3 (m1YY1)	F-CATAAATCTTGCAGTACCCTGTCAAAATG TCGGTGTGTGAGATAA	45
Oligo G3 (m2YY1)	F-CATAAATCTTGCAGTAAAATGTCAACCCTG TCGGTGTGTGAGATAA	45
D) ChIP		
C3	F-ACTTCAGGTCAGCGGTCTAACAA R-CACTTGAATGCAACCCTGTCCCAA	161
G3	F-TTGCTATGGCTTGCTTGCAAAGGG R-CCTGTAACACGTCTTTGAGCTCCT	173
Ctl locus	F-TGAAGTGTGAAGGTGCAGCAAACG R-AAAGGTGAGACCTTGCAGAGCAGA	140
<i>Sfrs10</i>	F-TTTCTCCGCTTACCCTTGGGA R-AACGGTATCTTCTTTCCCGTTGGA	175
<i>Rcor3</i>	F-GTCTCAGCTGAAGGAATTTGGCCT R-GCCATCCTCATAGCTCCTGTCAA	157
E) qRT-PCR		
<i>Hoxa4</i>	F-CGCCGTCAACTCCAGTTAT R-AGTGAATTCCTTCTCCAGTTC	95
<i>Hoxa5</i>	F-CCCAGATCTACCCTGGATG R-GGCATGAGCTATTTTCGATCCT	173
<i>Rpl19</i>	F-GATCATCCGCAAGCCTGTGA R-GCATCCGAGCATTGGCAGTA	122
<i>Scgb1a1</i>	F-AAGCCTCCAACCTCTACCATG R-ATGTCCGAAGAAGCTGAGCTG	85
<i>T1α</i>	F-TGGCAAGGCACCTCTGGTA R-GGTGGACAGTTCCTCTAAGGGA	70
<i>Yyl</i>	F-CATGTGGTCTCGGATGAAA R- GGGAGTTTCTTGCTGTGATA	117

wild-type (wt), *Hoxa5/lacZ* transgenic and mutant embryos. For RNA extraction, tissues were snap-frozen in N₂. For histology and immunostaining, tissues were fixed in 4% ice-cold paraformaldehyde (PFA), paraffin-embedded and sectioned at 4µm [40]. Organ sections were stained according to standard hematoxylin and eosin procedure for morphology analysis. Immunohistochemistry was performed as described [11]. The primary antibodies used were a goat antibody against CC10 (1/500; gift for Dr. G. Singh), a syrian hamster antibody against podoplanin (T1α; 1/75; DSHB), and a rabbit antibody against YY1 (1/500; Santa Cruz Biotechnology). The biotinylated secondary antibodies used were a goat anti-rabbit antibody (1/300; Vector Laboratories), a swine anti-goat antibody (1/300; Cedarlane), and a goat anti-syrian hamster antibody (1/300; Jackson Immuno Research).

For whole-mount detection of β-galactosidase activity, E8.5 to E13.5 entire embryos and dissected organs from E13.5 embryos were fixed in ice-cold 0.25% glutaraldehyde in PBS for 15 min to one hr according to the age and/or size of the sample. Detection of β-galactosidase activity was performed as described [31]. In parallel, organs from E13.5 embryos were fixed one hr in ice-cold 4% PFA, 0.2% glutaraldehyde in PBS, put overnight in a 30% sucrose solution in 0.1M phosphate buffer pH8.0, and embedded in tissue freezing medium (Triangle Biomedical Sciences, Inc.). Ten µm cryosections were processed for *lacZ* expression and counterstained with nuclear Fast Red [41].

Electrophoretic mobility shift assay (EMSA) and supershift assay

Whole cell extracts (WCE) were prepared from organs (lung, stomach and intestine) of E13.5 wt mouse embryos as described [19, 32]. The expression vector pCDNAI-YY1 (provided by G. Blanck) was used with the TnT7 Quick coupled transcription-translation system (Promega) to produce *in vitro* YY1 protein that was analyzed on sodium dodecyl sulfate-polyacrylamide gel electrophoresis. Both WCE and YY1 protein were tested in EMSA.

The 259-bp *XbaI-BssHII* and the 433-bp *MefI-SacI* DNA fragments were radiolabelled from digested plasmids with the Klenow fragment and purified on a G-50 Sephadex column. These DNA probes were separated on non-denaturing 6%

polyacrylamide (29:1) gels, from which bands of interest were cut and electroeluted. Binding reactions containing 0.5 to 2ng of probe (2,000 cpm), 5µg of WCE, 1µg of poly(dI-dC), and 10µg of BSA were prepared in 5mM Hepes pH7.9, 10% glycerol, 25mM KCl, 0.05mM EDTA and 0.125mM PMSF. Samples were equilibrated for 10 min at room temperature and separated by electrophoresis through a 6% polyacrylamide (29:1) gel prepared in 0.25X Tris-borate-EDTA buffer. Binding specificity was assessed by addition of a 100-fold excess of unlabelled probe or competitor prior to addition of the radiolabelled probe.

Oligonucleotides were also used to generate radiolabelled probes and cold competitors (Table 1B-C). Radiolabelling was achieved with the T4 polynucleotide kinase (NEB). Binding reactions using these probes contained 0.5 to 2ng (50,000 cpm), 2µl of YY1 protein produced *in vitro* or 5µg of WCE, 1µg of poly(dI-dC), and 10µg of BSA (NEB) prepared in 15mM Hepes pH7.9, 50mM NaCl, 80 µM ZnCl₂, 800µM dithiothreitol, 0.5% NP-40, 2mM MgCl₂, and 3% Ficoll. Samples were equilibrated for 5 min at room temperature. In supershift assays, 2µg of YY1 antibody was added for an additional 15 min. Reaction products were separated by electrophoresis through a 6% polyacrylamide (29:1) gel prepared in 0.25X Tris-borate-EDTA and 2.5% glycerol buffer. Binding specificity was validated using only rabbit reticulocyte extract as negative control. The supershift specificity was assessed using as negative control a CDX2 antibody (provided by D. Lohnes).

Chromatin immunoprecipitation (ChIP) assays

Lung and stomach from E13.5 wt embryos were collected separately and mechanically disrupted to produce a single cell suspension prior cross-linking with 1% formaldehyde in PBS for 15 min at room temperature. Cross-linking was stopped by adding glycine to a final concentration of 0.125M. Extracts were then disrupted into crude lysates in 2ml of swelling buffer (5mM Pipes pH8.0, 85mM KCl, 1% NP-40 and protease inhibitors), equilibrated 20 min on ice, and centrifuged for 5 min at 3000 rpm at 4°C. The pellets were eluted in 1ml of nuclear lysis buffer (50mM Tris-HCl pH8.0, 10mM EDTA, 1% SDS) and protease inhibitors (Complete Mini-EDTA-free; Roche Diagnostics) and submitted to two successive rounds of sonication using a Bioruptor (Diagenode) for a total

of 10 cycles of a 30-sec pulse interrupted by a 30-sec pause at the highest setting in order to obtain an average DNA size of 300-600-bp. One hundred μg of fragmented chromatin was incubated overnight at 4°C with Dynabeads linked to protein G (Invitrogen) and 2 μg of rabbit anti-YY1 antibody, rabbit anti-H3 antibody (ab1791; Abcam), or control rabbit IgG (sc2027; Santa Cruz Biotechnology). Immunocomplexes were washed three times respectively with 0.5ml of low salt immune complex wash buffer (0.1% SDS, 1% Triton X-100, 2mM EDTA, 20mM Tris-HCl pH8.0, 150mM NaCl), high salt immune complex wash buffer (0.1% SDS, 1% Triton X-100, 2mM EDTA, 20mM Tris-HCl pH8.0, 500mM NaCl) and LiCl immune complex wash buffer (0.25M LiCl, 1% IGEPAL-CA630, 1% deoxycholic acid (sodium salt), 1mM EDTA, 10mM Tris-HCl pH8.0). A final wash in 0.5ml of TE buffer (10mM Tris-HCl pH8.0, 1mM EDTA) was followed by centrifugation at 3000 rpm at 4°C for 2 min. Protein-DNA complexes were eluted by adding 100 μl of elution buffer (1% SDS, 50mM Tris pH8.0, 10mM EDTA) and then equilibrated at 65°C for 15 min. Cross-links were reversed by adding 200mM NaCl and equilibrated at 65°C overnight. DNA fragments were purified using a QIAquick gel extraction kit (Qiagen) after a two hr treatment with RNase followed by a two hr treatment with proteinase K. qPCR analyses were performed with specific primers listed in Table 1D. The values for the samples immunoprecipitated by anti-YY1, anti-histone H3 and control IgG were recorded as the percentage relative to input. ChIP results were confirmed by three independent experiments and qPCR was performed in triplicate for each sample. The ChIP efficiency was calculated by dividing the amount of PCR product obtained with the immunoprecipitated DNA by the amount obtained with the input DNA, as described in ref. [42].

Quantitative RT-PCR (qRT-PCR) experiments

Total RNA was isolated from lungs of E18.5 embryos. RNA extraction, cDNA synthesis and qRT-PCR were performed as described [41]. Samples were analyzed in triplicate. The *Rpl19* gene was used as control. Eight specimens were used for each genotype tested. Primer sequences are listed in Table 1E.

Statistical analyses

Student's *t* test was performed for gene expression and ChIP studies. A significance level inferior to 5% ($p < 0.05$) was considered statistically significant.

RESULTS

A 14.5-kb genomic fragment recapitulates the spatio-temporal expression from the *Hoxa5* proximal promoter

We have previously shown that the 11.1-kb *Hoxa5* genomic region located between positions -3767-bp and +7188-bp (relative to *Hoxa5* transcription start site (TSS)) reproduced the temporal expression and substantially reconstituted the spatial profile of *Hoxa5* gene. However, it did not recapitulate the correct expression in the central nervous system (CNS) and in developing organs, indicating that additional sequences were required [31]. A 1.5-kb *XbaI-XbaI* DNA fragment located in *Hoxa4-Hoxa5* intergenic sequences at ~3.0-kb upstream the *Hoxa4* gene (positions +9351-bp to +10816-bp) was able to target *Hoxa5* expression in lung and stomach [33]. To assess if a DNA region encompassing the *Hoxa5* regulatory sequences identified can recapitulate *Hoxa5* expression from the proximal promoter, a 14.5-kb *Hoxa5/lacZ* transgene (positions -2128-bp to +12384-bp) was tested (Construct 1; Fig. 2.1A-B). As reported for the 11.1-kb *Hoxa5/lacZ* transgene, the onset of expression of construct 1 was ~E8.0-8.25, corresponding to that of *Hoxa5* endogenous expression (not shown) [31, 43]. At E8.5, staining was mainly detected in the foregut region and in somites 5 to 8, consistent with the *Hoxa5* expression domain in the mesoderm along the A-P axis at this age (Fig. 2.1C) [31]. At E9.5-10.5, transgene expression extended more caudally along the axis and appeared in limb buds (Fig. 2.1D-E). At E12.5-13.5, staining was detected along the neural tube with an anterior boundary in the posterior hindbrain corresponding to the limit of *Hoxa5* endogenous expression in CNS. No staining was seen in the most caudal part of the embryo, which expressed the larger *Hoxa5* transcripts from the distal promoters (Fig. 2.1F-G) [12]. Construct 1 directed transgene expression in lung, stomach and intestine, and the staining was mainly restricted to the mesenchymal layer as *Hoxa5* endogenous expression (Fig. 2.1H-J). The *Hoxa5* rostro-caudal gradient present in the developing stomach was reproduced with construct 1

(Fig. 2.1J) [8]. *lacZ* staining was detected in the intestine, being stronger in the proximal part of midgut and vanishing towards the hindgut (not shown). Thus, the 14.5-kb *Hoxa5/lacZ* transgene contains regulatory sequences that largely reproduce the *Hoxa5* spatio-temporal expression driven by the *Hoxa5* proximal promoter.

***Hoxa5* organ-specific expression involves several regulatory sequences**

To delineate the DNA regions involved in *Hoxa5* organ-specific expression, we analyzed shorter versions of the *Hoxa4-Hoxa5* intergenic region in E13.5 transgenic embryos (Fig. 2.2A-B). Construct 2 (pLJ123) was used as control. It contains the BSC and MES enhancer sequences. It also drives expression in forelimbs and spleen but not in lung and gut (Fig. 2.2C, I) [31, 33]. Insertion of the 3.7-kb *KpnI-ApaI* genomic fragment downstream of pLJ123 reproduced the spatial profile of construct 1 at the same age (construct 3; Fig. 2.2D, J). A similar result was obtained when the 1.5-kb *XbaI-XbaI* DNA fragment was inserted at the 3' end of pLJ123 indicating that the 1.5-kb sequence contained the regulatory information necessary for proper *Hoxa5* expression in organs (construct 4; Fig. 2.2E, K). Histological analyses of stained organs from construct 4-positive embryos revealed that expression was restricted to mesenchyme as seen with construct 1 (not shown). To assess if the 1.5-kb DNA region possesses enhancer properties, we put it in front of a *Hoxa5/lacZ* minimal plasmid (construct 5) and of a heterologous *hsp68/lacZ* plasmid (constructs 6, 7), both plasmids being unable to direct transgene expression on their own [30, 32, 36]. Construct 5 targeted a strong transgene expression in organs and CNS reinforcing the notion that tissue-specific sequences for these structures were present in the 1.5-kb sequence (Fig. 2.2F, L). In contrast, when tested in the *hsp68/lacZ* context, faint or no organ expression was observed while CNS expression was reproduced (Fig. 2.2G-H, M-N). Thus, the 1.5-kb *XbaI-XbaI* region encloses several *Hoxa5* regulatory elements including a CNS-specific enhancer that directs the appropriate limit of expression in the posterior hindbrain regardless of the promoter used, and an organ-specific sequence that requires a *Hoxa5* minimal environment for effectiveness.

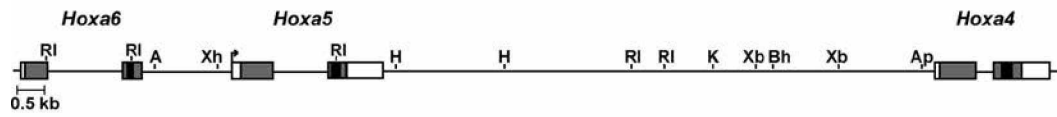
To narrow down the organ-specific regulatory sequences, a deletion analysis was undertaken (Fig. 2.3). A 1.0-kb *XbaI-SacI* fragment (construct 8) targeted organ and CNS expression similarly to construct 4 suggesting that the 455-bp located at the 3' end of the

Figure 2.1. β -galactosidase expression profile of construct 1 during embryogenesis

(A) Partial restriction map of the *Hoxa4-Hoxa6* genomic region. Black boxes represent homeoboxes, shaded boxes correspond to translated regions, and white boxes indicate transcribed regions. The arrow shows TSS from the *Hoxa5* proximal promoter. (B) Diagram of the *Hoxa5/lacZ* construct 1 (also named pLJ272) and summary of transgenic expression analysis at E13.5. The first column presents the number of *lacZ*-expressing F0 embryos out of the total number of F0 transgenic embryos generated. The number of positively stained embryos for each structure listed is indicated. A, *AccI*; Ap, *ApaI*; Bh, *BssHIII*; RI, *EcoRI*; H, *HindIII*; K, *KpnI*; Xb, *XbaI*; Xh, *XhoI*. (C-G) Representative transgenic embryos stained for β -galactosidase activity showed the expression pattern of construct 1 during development. Black arrowheads indicate the anterior limit of transgene expression in paraxial mesoderm at somite 5 and prevertebra 3. Open arrowheads point the anterior limit of transgene expression in the neural tube. (H) Whole-mount organs from E13.5 transgenic embryos showed β -galactosidase activity in lung, stomach and intestine. (I-J) Histological sections of lung and stomach, respectively, from E13.5 transgenic embryos stained for β -galactosidase activity. The inset in I shows that expression was restricted to mesenchyme. Black arrow in J shows the gradient of expression of the transgene in the developing stomach. c, caudal; e, epithelium; h, heart; i, intestine; l, lung; m, mesenchyme; nt, neural tube; pv, prevertebrae; r, rostral; s, stomach. Scale bar: 100 μ m.

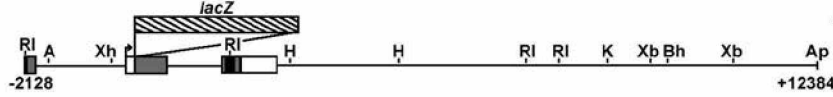
A

Hox loci



B

Construct 1 (pLJ272)



<i>lacZ</i> ⁺ /# Tg	Expression				
	nt	pv	l	s	i
5/6	5	5	5	5	5

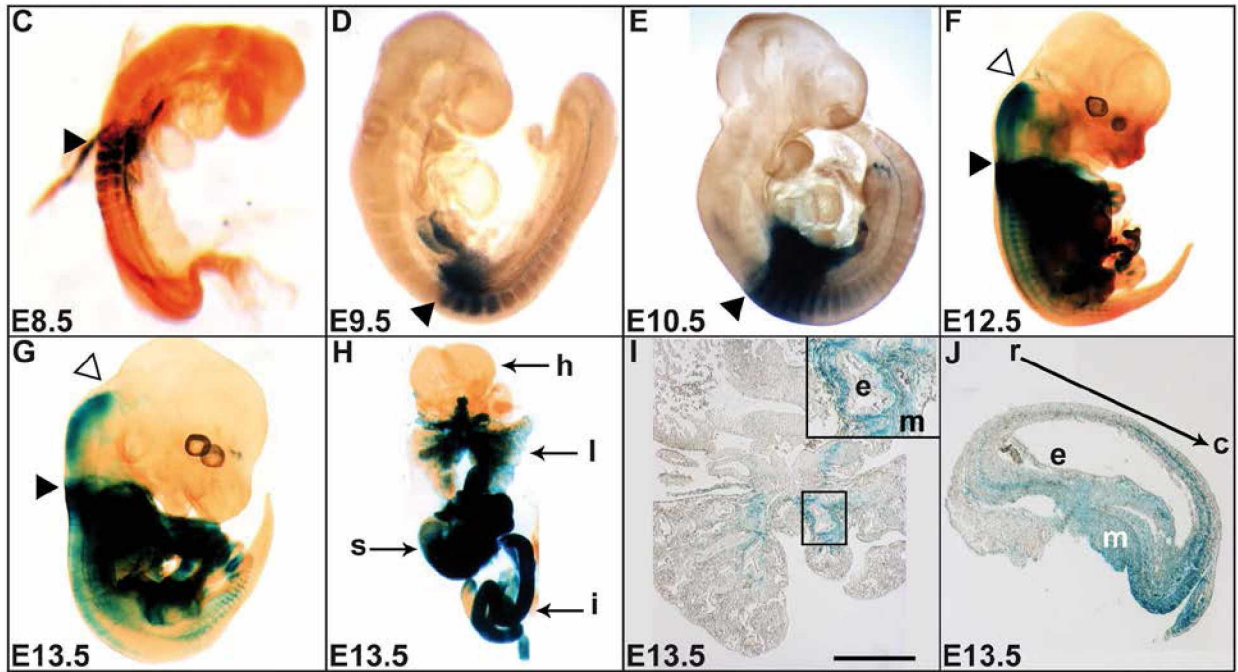
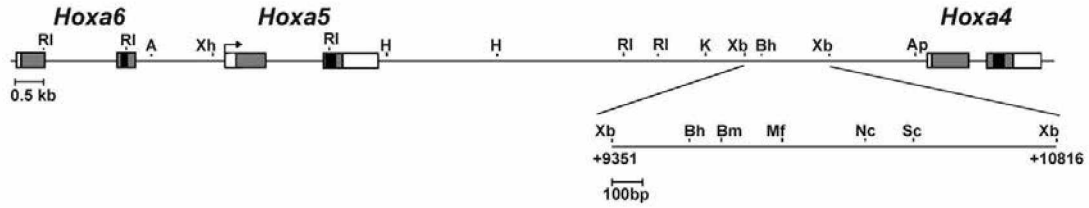


Figure 2.2. Characterization of the *Hoxa4-Hoxa5* intergenic region in E13.5 F0 transgenic mouse embryos

(A) Schematic representation of the 1.5-kb *XbaI-XbaI* DNA fragment in the *Hoxa4-Hoxa6* genomic region. (B) Diagram of the *Hoxa5/lacZ* constructs used to generate E13.5 F0 transgenic embryos and summary of transgenic expression analyses. A, *AccI*; Ap, *ApaI*; Bh, *BssHII*; Bm, *BsmI*; RI, *EcoRI*; H, *HindIII*; K, *KpnI*; Mf, *MfeI*; Nc, *NcoI*; Sc, *SacI*; Xb, *XbaI*; Xh, *XhoI*. (C-H) Carcass of representative E13.5 transgenic embryos and the associated organs (I-N) stained for β -galactosidase activity showed the effects of the different deletions on the expression pattern. Open arrowheads point the anterior limit of transgene expression in the neural tube. The number in the lower left corner of each panel corresponds to the transgene. h, heart; i, intestine; l, lung; nt, neural tube; pv, prevertebrae; s, stomach; sp, spleen.

A



B

#	lacZ ⁺ /# Tg	Expression				
		nt	pv	I	s	i
2 (pLJ123)	7/11	7	7	0	0	0
3 (pLJ155)	8/20	8	8	8	8	7
4 (pLJ160)	11/12	11	11	9	9	9
5 (pLJ172)	4/6	4	0	4	4	4
6 (pLJ424)	5/10	5	0	0	0	2
7 (pLJ516)	5/11	5	0	1	1	1

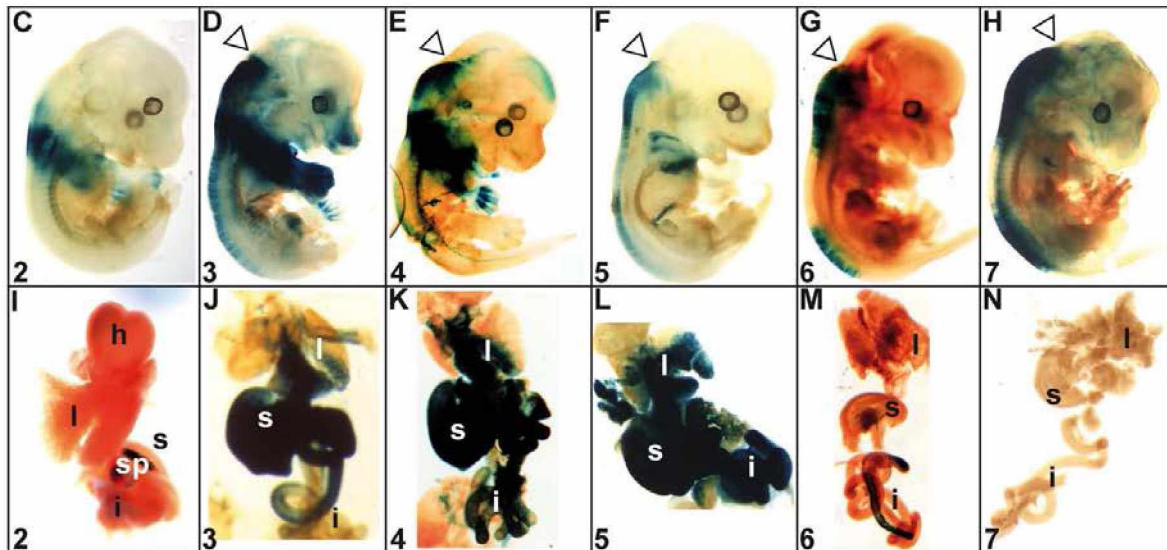
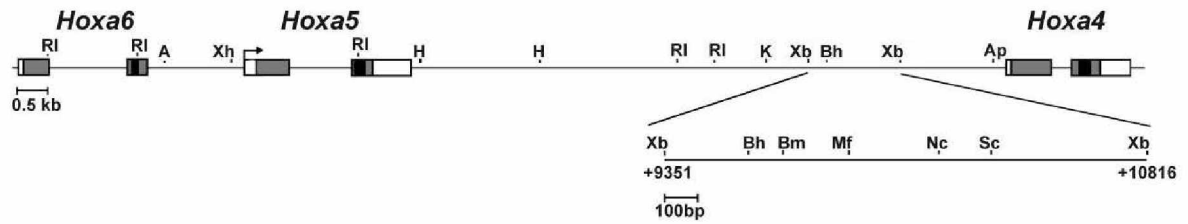


Figure 2.3. Characterization of the 1.5-kb *XbaI-XbaI* DNA fragment in E13.5 F0 transgenic mouse embryos

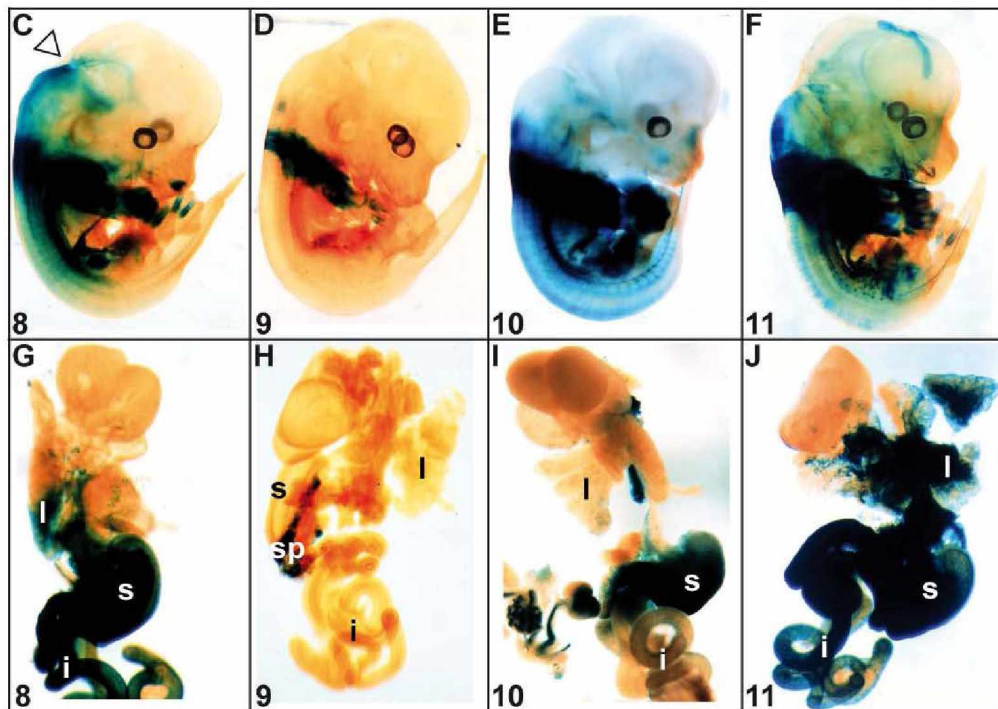
(A) Schematic representation of the 1.5-kb *XbaI-XbaI* DNA fragment in the *Hoxa4-Hoxa6* genomic region. (B) Diagram of the *Hoxa5/lacZ* constructs used to generate E13.5 F0 transgenic embryos and summary of transgenic expression analyses. (C-F) Carcass of representative E13.5 transgenic embryos and the associated organs (G-J) stained for β -galactosidase activity showed the effects of the different deletions on the expression pattern. Open arrowhead points the anterior limit of transgene expression in the neural tube. i, intestine; l, lung; nt, neural tube; pv, prevertebrae; s, stomach; sp, spleen.

A



B

#	lacZ ⁺ /# Tg	Expression				
		nt	pv	l	s	i
8 (pLJ307)	8/15	8	8	5	6	6
9 (pLJ292)	5/7	5	3	1	2	2
10 (pLJ173)	5/5	5	5	1	3	2
11 (pLJ252)	4/7	4	4	4	4	4



XbaI-XbaI sequence were not necessary (Fig. 2.3C, G). This was confirmed with construct 9 containing the 455-bp *SacI-XbaI* sequence. The expression profile of construct 9 was identical to that of construct 2 demonstrating the lack of regulatory activity in the 455-bp sequence (Fig. 2.3D, H). Additional deletion of sequences at the 3' end of the *XbaI-XbaI* sequence showed that the 259-bp *XbaI-BssHII* fragment (construct 10) directed expression in the stomach and intestine in few specimens but did not reproduce the anterior boundary in the hindbrain and the lung expression, suggesting the loss of tissue-specific elements (Fig. 2.3E, I). The reverse construct carrying the 1.2-kb *BssHII-XbaI* sequence (construct 11) targeted organ expression as construct 4. However, expression in CNS was not entirely recovered suggesting that neural-specific sequences were spread along the 1.0-kb *XbaI-SacI* region (Fig. 2.3F, J).

X-Gal staining in organs was similar in transgenic embryos carrying constructs 4, 8 and 11, suggesting that organ tissue-specific sequences were included into the 751-bp *BssHII-SacI* region. This was confirmed with construct 12 that showed expression in the respiratory and digestive tracts (Fig. 2.4B-C, H). Interestingly, when linked to the heterologous *hsp68/lacZ* plasmid, the 751-bp *BssHII-SacI* region targeted expression in organs in contrast to what was seen with constructs 6 and 7, suggesting the presence of repressive sequences in the 259-bp *XbaI-BssHII* and/or in the 455-bp *SacI-XbaI* DNA fragments (construct 13; Fig. 2.4D, I). Thus, the 751-bp *BssHII-SacI* fragment contains enhancer sequences that direct organ-specific expression.

Further deletions were performed. The 588-bp *BssHII-NcoI* sequence in construct 14 did not correctly reproduce the staining in organs pinpointing toward the 163-bp *NcoI-SacI* sequence as the one carrying the organ-specific enhancer (Fig. 2.4E, J). This was confirmed with the *NcoI-SacI* sequence in construct 15, which showed organ expression in transgenic embryos (Fig. 2.4F, K). A similar result was observed when the *NcoI-SacI* sequence was tested with the heterologous *hsp68/lacZ* plasmid (construct 16; Fig. 2.4G, L). Thus, organ-specific regulatory elements appear to be dispersed along the 1.5-kb *XbaI-XbaI* region, some located in the 259-bp *XbaI-BssHII* fragment, others in the 163-bp *NcoI-SacI* sequence.

The RARE sequence is not necessary for *Hoxa5* organ-specific expression

To gain insight into the critical DNA sequences involved in organ-specific expression, we investigated whether the 259-bp *XbaI-BssHII* and 163-bp *NcoI-SacI* DNA regions contain binding sites for putative transcriptional regulators. The *XbaI-BssHII* fragment includes a previously identified RARE-DR5 required for *Hoxa4* RA-responsiveness in CNS and for *Hoxa4* lung and stomach expression (Fig. 2.5A) [44]. Genomic sequence comparison with the zebrafish *HoxAa* cluster also revealed that this RARE was part of a DNA region of high homology (>70%) spanning ~100-bp, suggesting the presence of evolutionary conserved regulatory elements (not shown) [45]. To define if the RARE sequence was involved in *Hoxa5* regulation during organogenesis, we performed EMSA with WCE from lung, stomach and intestine of E13.5 mouse embryos and a 259-bp *XbaI-BssHII* radiolabelled probe. Binding was observed and the specificity was confirmed by competition with a 100-fold excess of unlabelled probe. However, competition was also detected when the RARE site was mutated (Fig. 2.5B, lanes 1-4). Moreover, no competition occurred with oligos containing wt or a mutated version of the RARE (Fig. 2.5B, lanes 5-6).

We assessed the RARE regulatory activity in E13.5 *Hoxa5/lacZ* transgenic embryos by mutating the RARE sequence in the 1.5-kb *XbaI-XbaI* DNA fragment (construct 17) and in the 259-bp *XbaI-BssHII* fragment (construct 18; Fig. 6B). Except for the CNS anterior boundary that was not reproduced, construct 17 targeted transgene expression in lung, stomach and intestine similarly as construct 4 (Fig. 2.6C, I). Likewise, construct 18 presented a comparable staining pattern than construct 10 with expression in the stomach and intestine but not in lung (Fig. 2.6D, J). Thus, the RARE sequence, found to be necessary to drive *Hoxa4* expression in embryonic lung and gut, does not play a key role in *Hoxa5* organ expression at E13.5.

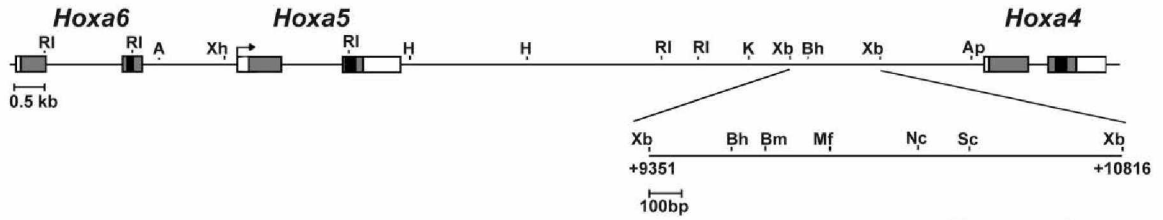
The YY1 transcription factor binds to *Hoxa5* organ-specific regulatory DNA regions

To identify which DNA region from the *XbaI-BssHII* radiolabelled probe bound proteins in WCE from organs of E13.5 embryos, the 259-bp sequence was divided into

Figure 2.4. Characterization of the 751-bp *Bss*HII-*Sac*I DNA fragment in E13.5 F0 transgenic mouse embryos

(A) Schematic representation of the 1.5-kb *Xba*I-*Xba*I DNA fragment in the *Hoxa4-Hoxa6* genomic region. (B) Diagram of the *Hoxa5/lacZ* constructs used to generate E13.5 F0 transgenic embryos and summary of transgenic expression analyses. (C-G) Carcass of representative E13.5 transgenic embryos and the associated organs (H-L) stained for β -galactosidase activity showed the effects of the different deletions on the expression pattern. i, intestine; l, lung; nt, neural tube; pv, prevertebrae; s, stomach.

A



B

#	<i>lacZ</i> ⁺ /# Tg	Expression				
		nt	pv	l	s	i
12 (pLJ336)	9/14	9	9	5	8	9
13 (pLJ354)	11/17	10	9	8	9	9
14 (pLJ401)	7/8	7	6	1	4	0
15 (pLJ524)	10/20	7	10	5	7	6
16 (pLJ525)	5/8	4	5	4	4	4

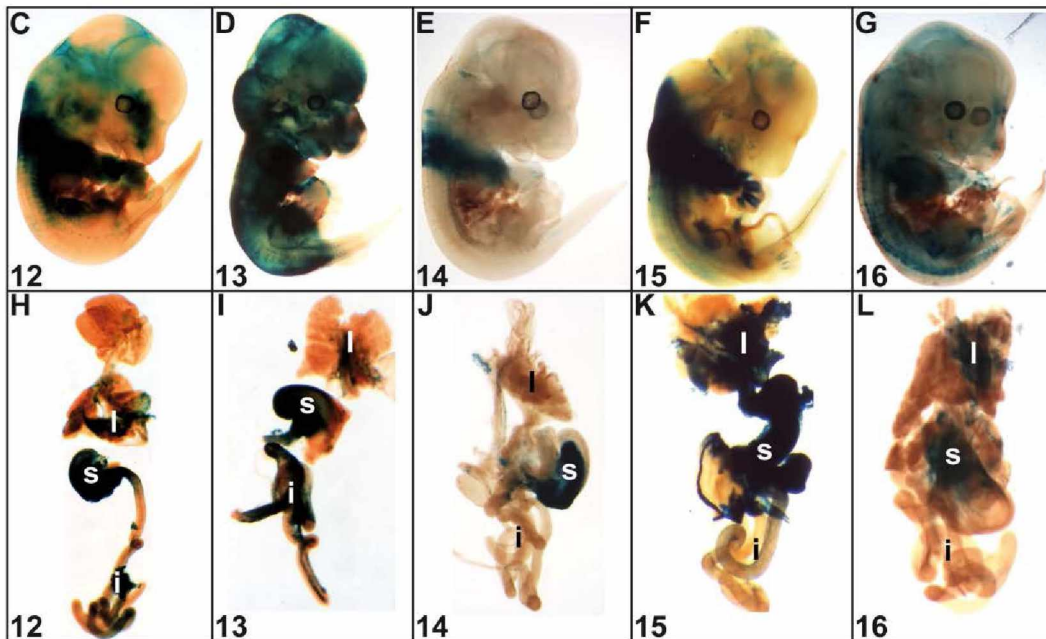
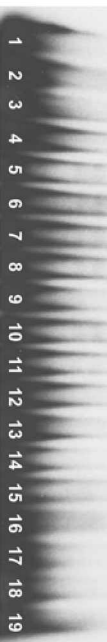


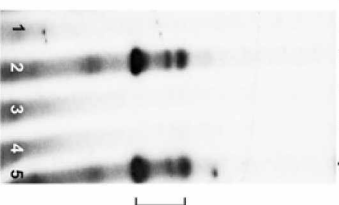
Figure 2.5. Characterization of the 259-bp *XbaI*-*Bss*HII DNA fragment by EMSA

(A) Restriction map of the 1.5-kb *XbaI*-*XbaI* sequence extending from +9.3-kb to +10.8-kb in the 3' half of the *Hoxa4*-*Hoxa5* intergenic region. The box denotes the location of the 259-bp *XbaI*-*Bss*HII regulatory region. Sequence of the 259-bp *XbaI*-*Bss*HII DNA fragment is indicated. Fragments A, B and C used as competitor in EMSA are underlined. Fragment C was further subdivided into Oligos C1, C2, C3 and Oligo RARE and Oligo-18(C3). Boxed nucleotides correspond to RARE-DR5 sequence and YY1 binding sites. Symbols * and @ indicate point mutations into RARE and YY1 binding sites, respectively (see Table 1B for nucleotide sequences). (B) EMSA with WCE from lung/stomach/intestine of E13.5 embryos and the 259-bp *XbaI*-*Bss*HII radiolabelled probe in presence of unlabelled competitors in 100-fold excess showed that protein binding occurred with the Oligo-18(C3) fragment via YY1 binding sites (lanes 9, 12, 18-19). No binding with the RARE site was observed (lanes 4-6). EMSA with *in vitro*-translated YY1 protein and the 259-bp *XbaI*-*Bss*HII probe showed specific binding that was competed by Oligo C3 (lanes 21-24). (C) The binding of WCE with YY1 consensus binding site and the loss of binding when the YY1 antibody was added confirmed the presence of YY1 protein in WCE (lanes 1-5). (D) EMSA with WCE and Oligo C3 radiolabelled probe showed binding that was competed by an excess of cold Oligo C3, Oligo-18(C3) sequence, the YY1 consensus sequence, localized mutations in YY1 sites, and the addition of the YY1-specific antibody (lanes 1-4, 6-9). No competition occurred when a non-specific probe was used (Oligo C2), when several mutations were distributed along the YY1 binding sites in Oligo C3 or when the CDX2 control antibody was used (lanes 5, 10-11). Arrows and brackets indicate the bands corresponding to YY1 binding. COMP, competitor; r. lysate, reticulocyte lysate.



C

PROBE	Oligo YY1 consensus				
	WCE	Oligo YY1	YY1Ab	CDX2 Ab	COMP
1	-	+	-	-	-
2	-	+	+	-	-
3	-	+	+	+	-
4	-	+	+	+	+
5	-	+	+	+	+



D

PROBE	Oligo C3					COMP
	WCE	Oligo C3	YY1	CDX2	Oligo C2	
1	-	+	-	-	-	
2	-	+	+	-	-	
3	-	+	+	+	-	
4	-	+	+	+	+	
5	-	+	+	+	+	
6	-	+	+	+	+	
7	-	+	+	+	+	
8	-	+	+	+	+	
9	-	+	+	+	+	
10	-	+	+	+	+	
11	-	+	+	+	+	

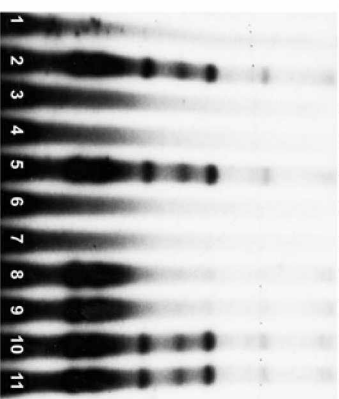
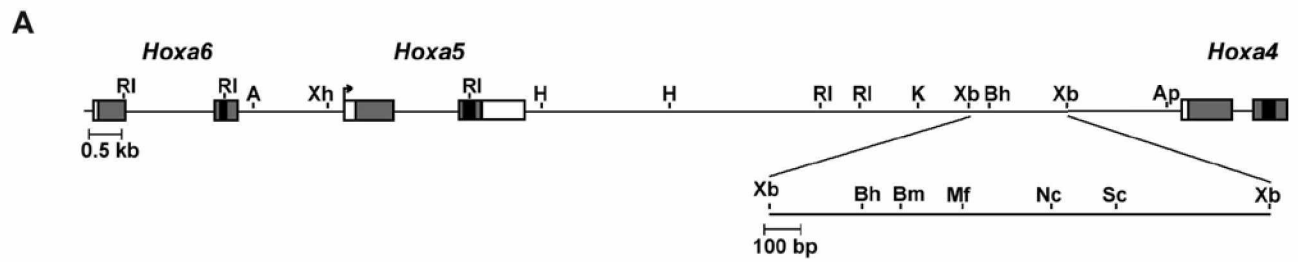


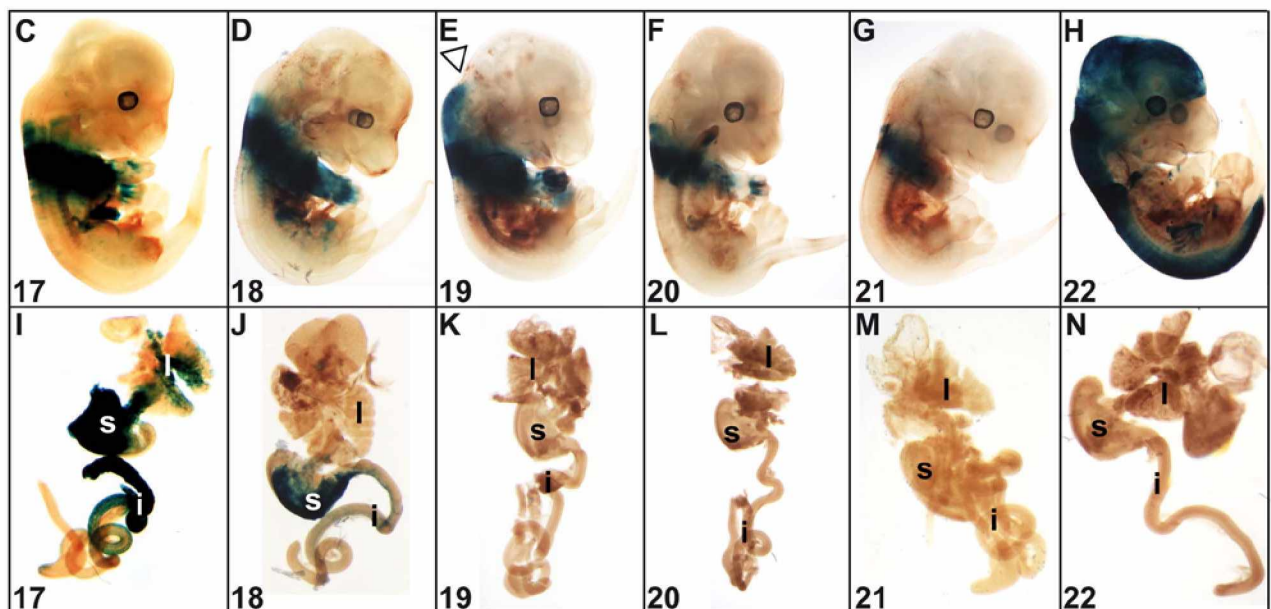
Figure 2.6. Characterization of the RARE and YY1 binding sites in E13.5 F0 transgenic mouse embryos

(A) Schematic representation of the 1.5-kb *XbaI-XbaI* DNA fragment in the *Hoxa4-Hoxa6* genomic region. (B) Diagram of the *Hoxa5/lacZ* constructs used to generate E13.5 F0 transgenic embryos and summary of transgenic expression analyses. The asterisks in constructs 17-18 and the @ symbols in constructs 19-22 correspond to mutations in the RARE and YY1 binding sites, respectively. (C-H) Carcass of representative E13.5 transgenic embryos and the associated organs (I-N) stained for β -galactosidase activity showed the effects of the mutations on the expression pattern. Open arrowhead points the anterior limit of transgene expression in the neural tube. i, intestine; l, lung; nt, neural tube; pv, prevertebrae; s, stomach.



B

#	lacZ	<i>lacZ</i> ⁺ /# Tg	Expression				
			nt	pv	l	s	i
17 (pLJ315)		5/7	5	5	5	5	3
18 (pLJ537)		5/10	4	5	1	3	3
19 (pLJ540)		6/13	6	6	1	1	1
20 (pLJ539)		7/14	7	7	0	1	0
21 (pLJ532)		7/12	7	7	1	1	1
22 (pLJ533)		10/12	10	9	2	2	2



three fragments (A to C) used as competitors. Only fragment C competed binding with the *XbaI-BssHII* probe (Fig. 2.5B, lanes 7-9). Moreover, only the 45-bp Oligo C3 located at the 3' end of fragment C could compete binding with the *XbaI-BssHII* probe (Fig. 2.5B, lanes 10-12). Using a linker scanning approach, we found that an 18-bp sequence in Oligo C3 (Oligo-18(C3)) was responsible for the binding (Fig. 2.5B, lanes 13-18). Sequence comparison with the TFSEARCH and TESS databases revealed YY1 binding sites in Oligo-18(C3). Mutations of the YY1 sites in Oligo C3 impaired its capacity to compete with the *XbaI-BssHII* probe (Fig. 2.5B, lane 19). To further establish that YY1 can bind the *XbaI-BssHII* fragment, we used a murine YY1 protein produced *in vitro*. Binding specificity to the *XbaI-BssHII* probe was confirmed and competition with Oligo C3 caused YY1 protein binding inhibition (Fig. 2.5B, lanes 20-24).

We then verified that our WCE contained YY1 protein by performing supershift assays. The specificity of the YY1 antibody was validated by testing an YY1 consensus sequence with WCE [28]. Addition of the YY1 antibody specifically resulted in the loss of DNA-protein complexes (Fig. 2.5C). A CDX2 antibody was used as negative control. In parallel, we tested the Oligo C3 sequence with WCE in EMSA. Binding was observed and it was competed by an excess of cold Oligo C3, Oligo-18(C3) and the YY1 consensus sequence (Fig. 2.5D, lanes 1-4, 8). No competition was observed when a non-specific probe was used (Oligo C2) or when several mutations distributed along the YY1 binding sites were inserted into Oligo C3 (Fig. 2.5D, lanes 5, 11). Presence of localized mutations in YY1 sites allowed competition to happen (Fig. 2.5D, lanes 6-7). Supershift assays were also performed with the YY1-specific antibody and the CDX2 control antibody. Only the addition of the YY1 antibody led to the loss of Oligo C3-protein complexes (Fig. 2.5D, lanes 9-10). Thus, the YY1 protein present in WCE from organs of E13.5 mouse embryos specifically binds a site located at the 3' end of the 259-bp *XbaI-BssHII* fragment.

A similar systematic approach was applied to decipher the regulatory elements of the 163-bp *NcoI-SacI* sequence. EMSAs combining the 433-bp *MfeI-SacI* radiolabelled probe with WCE demonstrated specific binding with the 163-bp *NcoI-SacI* portion (Fig. 2.7B, lanes 1-5). The latter was further divided into four oligos, leading to the identification of the 45-bp Oligo G3 as the binding sequence. Using the linker scanning approach, an 18-

bp sequence in oligo G3 (Oligo-18(G3)) was found to be necessary for protein binding (Fig. 2.7B, lanes 6-15). Sequence comparisons with databases revealed YY1 binding sites in Oligo-18(G3), which when mutated in Oligo G3 impaired its capacity to compete protein binding with the 433-bp *MfeI-SacI* sequence as the non-specific *BssHIII-BsmI* competitor (Fig. 2.7B, lanes 16-17). We also tested the Oligo G3 sequence with WCE, the YY1-specific antibody and various competitors and we confirmed that the YY1 protein present in WCE from organs of E13.5 mouse embryos can specifically bind the 163-bp *NcoI-SacI* sequence (Fig. 2.7C; lanes 1-11).

Functional YY1 binding sites are involved in *Hoxa5* organ-specific expression

To establish whether the YY1 binding sites identified in the 259-bp *XbaI-BssHIII* and 163-bp *NcoI-SacI* DNA regions are effective *in vivo*, a ChIP assay was performed on cross-linked chromatin isolated from either lung or stomach of E13.5 wt mouse embryos. DNA from the immunoprecipitate was subjected to qPCR analyses with specific primers for the 259-bp *XbaI-BssHIII* and 163-bp *NcoI-SacI* sequences, for a reported YY1 target gene used as positive control (*Srfs10*), and a known YY1 negative control (*Rcor3*) [46]. A downstream region located at 15-kb of *Hoxa5* TSS and devoid of YY1 binding sites was used as a negative control for the locus. In lung and stomach, YY1 was recruited to *Hoxa5* C3 and G3 regulatory sequences and to the positive *Srfs10* control whereas no binding was observed with the *Rcor3* and *Hoxa5* locus negative controls (Fig. 2.8; not shown). Thus, YY1 can bind to *Hoxa5* organ-specific regulatory DNA regions both *in vitro* and *in vivo*.

We then assessed the contribution of the YY1 binding sites to *Hoxa5* organ expression in E13.5 transgenic embryos. Construct 19, which contained mutations in the YY1 binding sites of the *XbaI-BssHIII* and the *NcoI-SacI* regions, correctly directed expression in the CNS as construct 4, while no organ expression was detected (Fig. 2.6E, K). Similarly, mutations of the YY1 binding sites in the *XbaI-BssHIII* sequence in construct 20 and in the *NcoI-SacI* sequence in construct 21 caused a loss of organ expression when compared to constructs 10 and 15, respectively (Fig. 2.6F-G, L-M). A comparable result was observed when the *NcoI-SacI* sequence with mutated YY1 sites was tested in front of the heterologous *hsp68/lacZ* plasmid (construct 22; Fig. 2.6H, N). In summary, YY1 acts as a positive regulator of *Hoxa5* lung, stomach and intestine expression during

Figure 2.7. Characterization of the 163-bp *NcoI-SacI* DNA fragment by EMSA

(A) Restriction map of the 1.5-kb *XbaI-XbaI* sequence extending from +9.3-kb to +10.8-kb in the 3' half of the *Hoxa4-Hoxa5* intergenic region. The box denotes the location of the 163-bp *NcoI-SacI* regulatory region. Sequence of the 163-bp *NcoI-SacI* DNA fragment is indicated. Oligos G1, G2, G3, G4 and Oligo-18(G3) are underlined. Boxed nucleotides correspond to YY1 binding sites. Symbols @ indicate point mutations into YY1 binding sites (see Table 1C for nucleotide sequences). (B) EMSA with WCE from lung/stomach/intestine of E13.5 embryos and the 433-bp *MfeI-SacI* radiolabelled probe in presence of unlabelled competitors in 100-fold excess showed that protein binding occurred with the Oligo-18(G3) fragment via YY1 binding sites (lanes 8, 15-16). (C) EMSA with WCE and Oligo G3 probe showed binding that was competed by an excess of cold Oligo G3, Oligo-18(G3) sequence, the YY1 consensus sequence, localized mutations in YY1 sites, and the addition of the YY1-specific antibody (lanes 1-4, 6-9). No competition occurred when a non-specific probe was used (Oligo G2), when several mutations were distributed along the YY1 binding sites in Oligo G3 or when the CDX2 control antibody was used (lanes 5, 10-11). Arrows and brackets indicate the bands corresponding to YY1 binding. COMP, competitor.

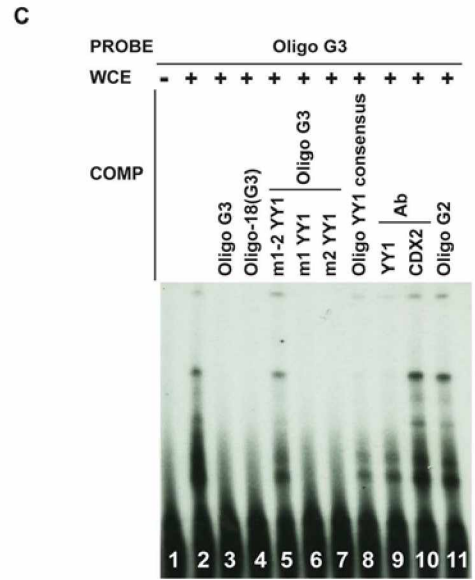
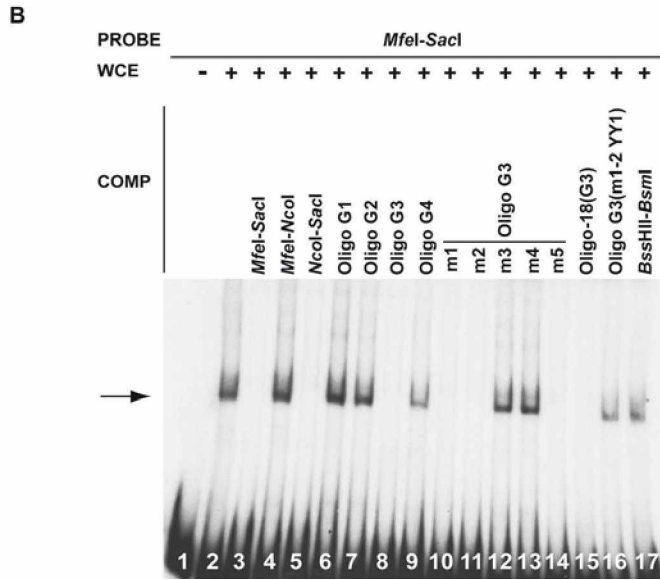
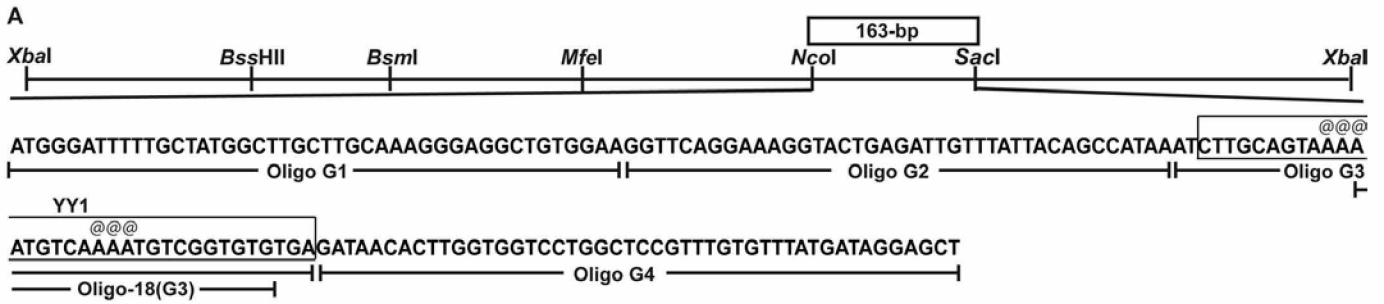
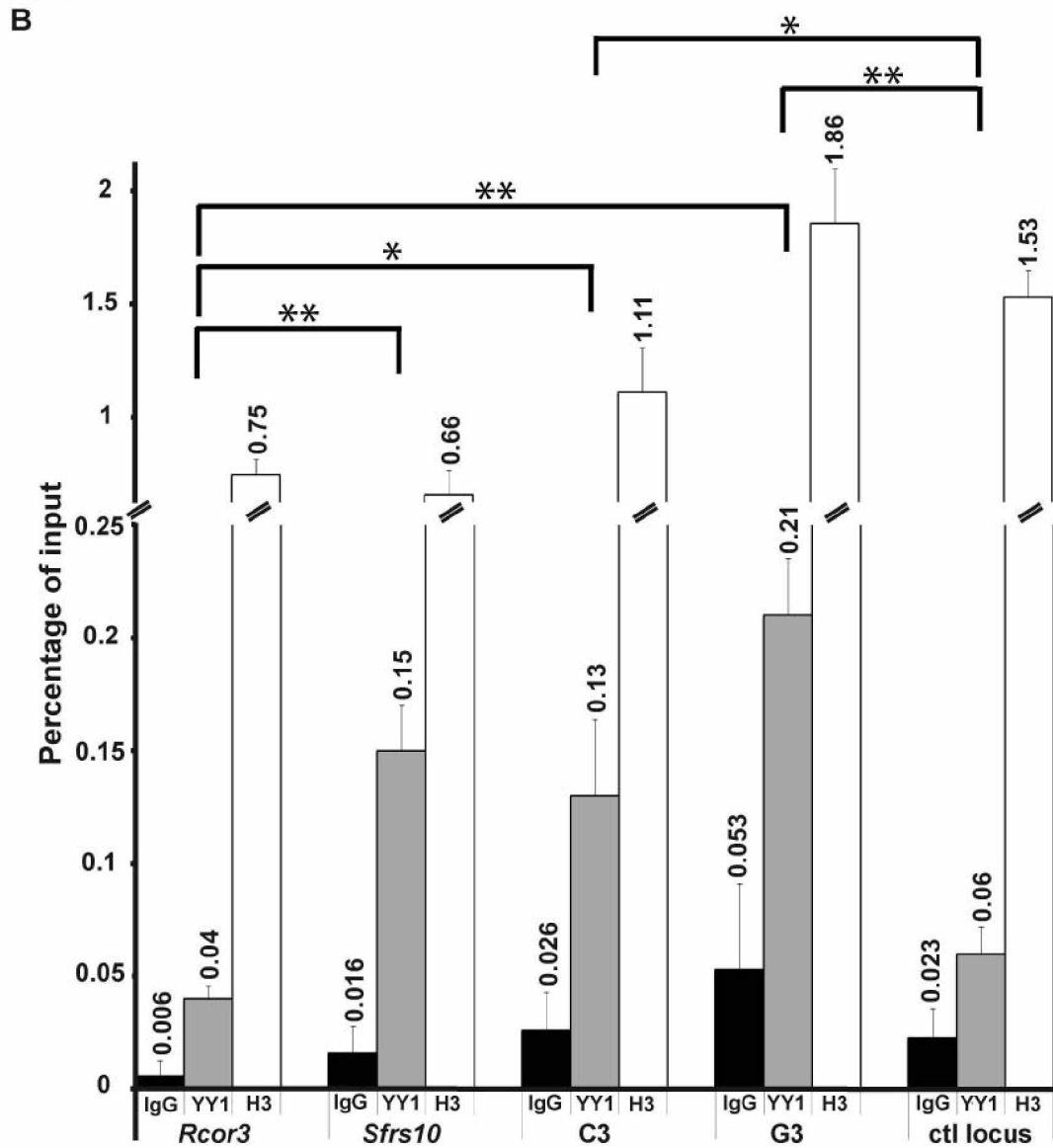
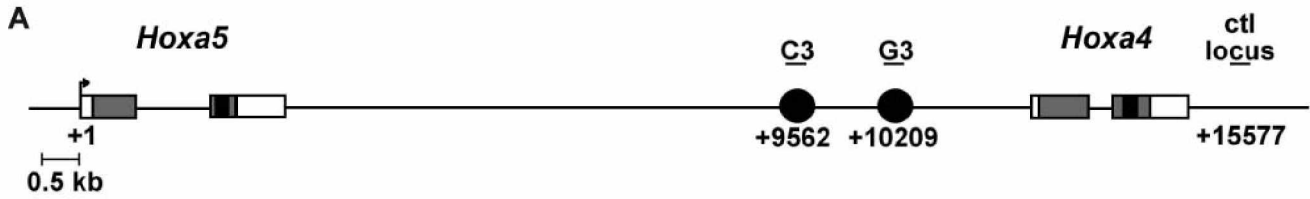


Figure 2.8. *In vivo* detection of YY1 protein binding to Oligo C3 and Oligo G3 regions by ChIP analysis

(A) Schematic representation and position relative to the *Hoxa5* TSS of the two YY1 binding sites in the *Hoxa4-Hoxa5* intergenic region. The black circles represent the Oligo C3 and Oligo G3 sequences identified by EMSA. The position of the qPCR fragments corresponding to Oligo C3, Oligo G3 and a *Hox* control locus located 15-kb downstream the *Hoxa5* TSS that does not contain YY1 binding sites (ctl locus) is indicated. (B) ChIP analysis of the *Hoxa4-Hoxa5* intergenic region in lungs from E13.5 mouse embryos. Chromatin was immunoprecipitated with rabbit IgG (negative control), anti-YY1 and anti-histone H3 (for chromatin integrity control) antibodies. Recruitment of YY1 and histone H3 on Oligo C3 and Oligo G3 sequences, a YY1 negative control (*Rcor3*), a YY1 positive control (*Sfrs10*), and the *Hox* control locus was evaluated by qPCR and is indicated as the percentage of input. The data are mean \pm SEM of three independent experiments. * $p < 0.05$, ** $p < 0.01$.



embryogenesis.

Inactivation of *Yy1* function in mesenchyme impacts on lung formation and on *Hoxa5* gene expression

In mice, the complete loss of *Yy1* gene function resulted in peri-implantation lethality, while the phenotypes associated with the *Yy1* conditional and hypomorphic alleles revealed a critical gene dosage requirement for YY1 during embryogenesis [38, 47]. Indeed, *Yy1*^{Floxneo/-} mice express 25% of normal YY1 protein levels and a high percentage of newborn pups die at birth from lung defects reminiscent to those observed in *Hoxa5*^{-/-} mice [6, 38]. To assess *in vivo* the requirement of *Yy1* gene function in *Hoxa5* lung expression, we generated *Yy1*^{Floxneo/-} mouse embryos carrying the *Hoxa5/lacZ* construct 1 transgene. No change in *lacZ* expression pattern and *Hoxa5* RNA levels was observed in the lungs from E13.5 and E18.5 *Yy1*^{Floxneo/-}*Tg*^{*Hoxa5/lacZ*#1} embryos (not shown). The lack of effect on *Hoxa5* expression might result from the remaining 25% of *Yy1* expression.

As *Hoxa5* is specifically expressed in lung mesenchyme, we inactivated *Yy1* gene function in the mesenchyme using the *Dermo1*^{Cre} deleter mouse line [35, 39]. No *Yy1*^{flox/flox}*Dermo1*^{+Cre} animals were recovered at weaning and monitoring of litters showed that all *Yy1*^{flox/flox}*Dermo1*^{+Cre} pups were found dead at birth (Fig. 2.9A). At E18.5, the distribution of genotypes was conformed to the expected mendelian ratio. Thus, mesenchymal *Yy1* deletion causes death at birth.

Comparative lung morphology of E18.5 embryos revealed that *Yy1*^{flox/+}*Dermo1*^{+/+} and *Yy1*^{flox/+}*Dermo1*^{+Cre} specimens presented a normal structure with dilated peripheral lung saccules and thin mesenchyme. However, lungs from *Yy1*^{flox/flox}*Dermo1*^{+Cre} embryos showed a collapsed appearance with narrow airspaces and thick mesenchyme similar to the *Hoxa5*^{-/-} lung phenotype, a likely explanation for the neonatal lethality of *Yy1* mutant pups (Fig. 2.9B-E). We also assessed the impact of the mesenchymal *Yy1* deletion on the integrity of the airway epithelium by looking at the expression of cell specific markers by qRT-PCR and IHC. A statistically decreased expression of the club cell (Clara cell) marker CC10 (encoded by the *Scgb1a1* gene) was observed in lung specimens from E18.5 *Yy1*^{flox/flox}*Dermo1*^{+Cre} embryos by qRT-PCR (Fig. 2.9Q). IHC analysis showed that the

decreased CC10 staining paralleled that seen in *Hoxa5*^{-/-} specimens (Fig. 9F-I) [41]. As well, expression of podoplanin (T1 α), a marker of type I pneumocytes, which are participating to gas exchanges with the underlying vascular endothelial cells, was decreased in *Yy1* mutant specimens similar to what was previously observed in *Hoxa5*^{-/-} mutants (Fig. 2.9J-M, Q) [35]. Together, these results indicated that the *Yy1* mesenchymal deletion perturbed epithelial cell differentiation along the respiratory tract, a phenotype reminiscent to that of *Hoxa5* mutants.

We validated the specific inactivation of *Yy1* in lung mesenchyme by IHC analysis. At E15.5 and E18.5, the YY1 protein showed ubiquitous expression in lung epithelial and mesenchymal compartments from control embryos, whereas in *Yy1*^{flox/flox}*Dermo1*^{+Cre} specimens, YY1 staining was greatly diminished in lung mesenchyme but unchanged in the epithelium (Fig. 2.9O-P; not shown). qPCR analysis confirmed the decreased *Yy1* expression in lungs from E18.5 *Yy1*^{flox/flox}*Dermo1*^{+Cre} embryos (Fig. 2.9R). We also looked at the impact of the *Yy1* mesenchymal deletion on *Hoxa5* and *Hoxa4* lung expression as the 1.5-kb *XbaI-XbaI* DNA fragment containing the YY1 binding sites was shown to be necessary for *Hoxa4* lung and gut expression [44]. qRT-PCR expression analysis revealed significantly reduced *Hoxa5* and *Hoxa4* expression levels in lungs from E18.5 *Yy1*^{flox/flox}*Dermo1*^{+Cre} embryos, indicating that an integrated regulation of *Hoxa4* and *Hoxa5* genes prevailed in lung expression through the sharing of YY1-responsive sequences.

DISCUSSION

Our search for *Hoxa5* transcriptional regulatory sequences led us to identify a 14.5-kb genomic fragment (construct 1) encompassing the *Hoxa5* gene and starting into the homeobox of the 5' flanking *Hoxa6* gene up to ~350-bp upstream of the *Hoxa4* TSS. The 14.5-kb transgene reproduces the spatio-temporal expression directed by the *Hoxa5* proximal promoter by targeting *lacZ* expression to the brachial region of the CNS and to the paraxial and lateral plate mesoderm at the cervico-thoracic level as reported for the shorter version of 11.1-kb in length [31]. In addition, the 14.5-kb transgene recapitulates

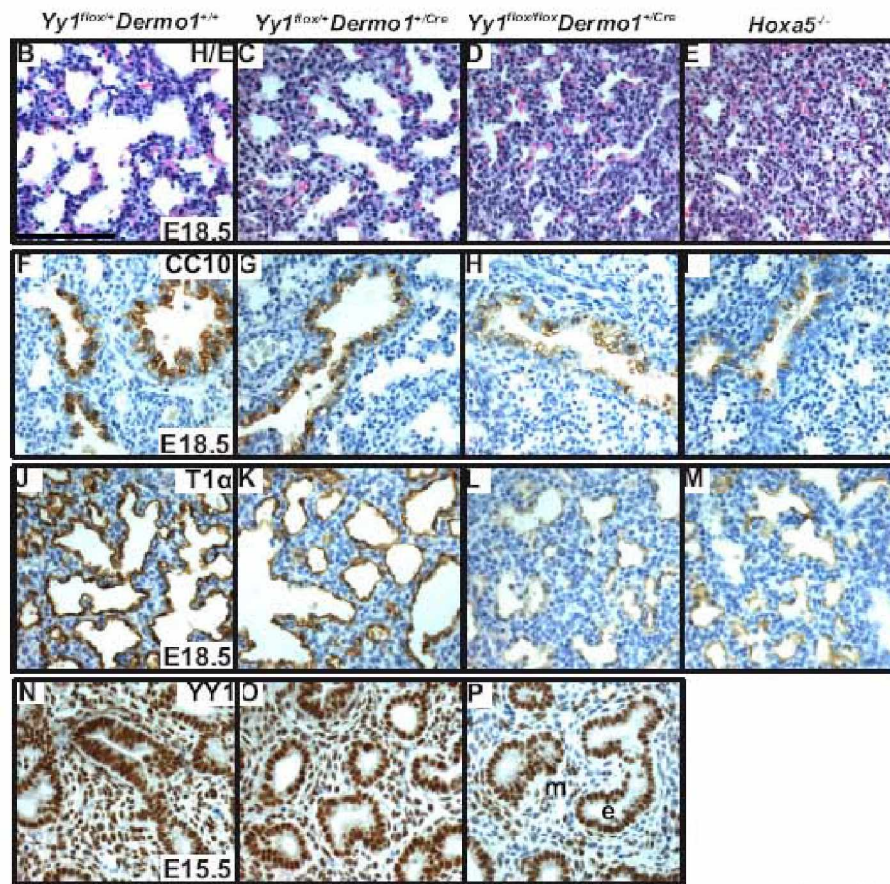
Figure 2.9. Analysis of the lung phenotype in *Yy1^{flox/flox}Dermo1^{+Cre}* mutants

(A) Ratios of genotypes of litters obtained from matings between *Yy1^{flox/+}Dermo1^{+Cre}* and *Yy1^{flox/flox}* mice. (B-E) Comparative lung histology of E18.5 *Yy1^{flox/+}Dermo1^{+/+}*, *Yy1^{flox/+}Dermo1^{+Cre}*, *Yy1^{flox/flox}Dermo1^{+Cre}* and *Hoxa5^{-/-}* embryos. *Yy1^{flox/+}Dermo1^{+/+}* and *Yy1^{flox/+}Dermo1^{+Cre}* specimens presented a normal lung structure, whereas lungs from *Yy1^{flox/flox}Dermo1^{+Cre}* and *Hoxa5^{-/-}* embryos were collapsed. (F-M) Characterization of the respiratory epithelium of E18.5 *Yy1^{flox/flox}Dermo1^{+Cre}* embryos. (F-I) Detection of club cells by CC10 immunostaining showed decreased labelling in lungs from *Yy1^{flox/flox}Dermo1^{+Cre}* and *Hoxa5^{-/-}* specimens. (J-M) Immunostaining with *Tia α* , a marker of type I pneumocytes, was reduced in lungs from *Yy1^{flox/flox}Dermo1^{+Cre}* and *Hoxa5^{-/-}* embryos. (N-P) YY1 immunostaining showed ubiquitous YY1 expression in lung epithelial and mesenchymal compartments in E15.5 control embryos, but an important decreased staining in lung mesenchyme from *Yy1^{flox/flox}Dermo1^{+Cre}* specimens. e, epithelium; m, mesenchyme. Scale bar: 200 μ m. (Q-S) qRT-PCR analysis for *Scgbl1a1*, *Tia α* , *Yy1*, *Hoxa5* and *Hoxa4* expression in lungs from E18.5 *Yy1^{flox/+}Dermo1^{+/+}* and *Yy1^{flox/flox}Dermo1^{+Cre}* embryos. Expression levels were significantly diminished for all genes tested in *Yy1^{flox/flox}Dermo1^{+Cre}* specimens. Values are expressed as means \pm SEM. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

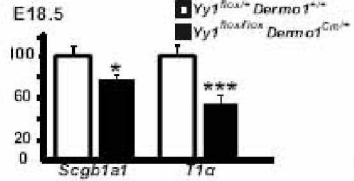
A

Age	# of litters	# of pups	<i>Yy1^{lox/+}Dermo1^{+/-}</i>	<i>Yy1^{lox/lox}Dermo1^{+/-}</i>	<i>Yy1^{lox/+}Dermo1^{+/-}Cre</i>	<i>Yy1^{lox/lox}Dermo1^{+/-}Cre</i>
E18.5	9	74	17(23%)	24(32%)	20(27%)	13(18%) [*]
D0	5	29	5(17%)	11(38%)	8(28%)	5(17%) [*]
D21	11	75	22(29%)	20(27%)	33(44%)	0(0%)
% expected			25	25	25	25

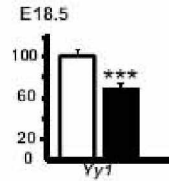
^{*}All found dead at birth



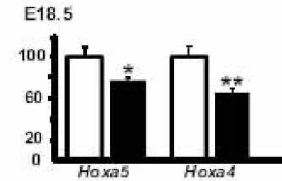
Q *Scgb1a1* and *T1a* relative expression



R *Yy1* relative expression



S *Hoxa5* and *Hoxa4* relative expression



Hoxa5 expression in the mesenchymal compartment of the respiratory and digestive tracts, the *Hoxa5* rostro-caudal gradient in the stomach, and the anterior limit of *Hoxa5* expression in the CNS at the level of the floor of myelencephalon [6-8, 31].

Hoxa5 expression in the developing neural tube is under the control of the BSC enhancer, which directs expression in the brachial region of the spinal cord from E11 to E13, a subset of the *Hoxa5* endogenous pattern in the neural tube [30]. We now report that sequences distributed in the 1.5-kb *XbaI-XbaI* DNA fragment replicate the anterior expression domain of *Hoxa5* in the neural tube. However, the onset of *Hoxa5* expression in the developing CNS up to E11 is not reproduced with the 14.5-kb transgene indicating that additional sequences are required [31]. Moreover, mutation of the RARE site located in the 1.5-kb *XbaI-XbaI* fragment causes a caudal shift of the anterior boundary in the CNS suggesting that RA signaling is involved in the establishment of the correct *Hoxa5* expression domain in the neural tube as shown for other *Hox* genes [48-49]. Interestingly, when tested in a *Hoxa4* context, mutation of the RARE sequence did not impact on the activation of *Hoxa4* embryonic expression in the developing neural tube but it affects *Hoxa4* response to exogenous RA in the neural tube [44]. Thus, the *Hoxa4* and *Hoxa5* genes share a RARE site that positively regulates their respective expression in the developing CNS. This RARE sequence is highly conserved between *Hox* clusters and between vertebrates. It was shown to possess neural enhancer activity for the *Hoxd4* gene [49-51]. In the *HoxB* cluster, the *Hoxb5* distal RARE was found to regulate the anterior expression boundary of 5' *Hoxb* genes in the posterior hindbrain raising the possibility that the RARE site located in the 1.5-kb *XbaI-XbaI* fragment may play a similar evolutionarily conserved role in the *HoxA* complex [23].

While the RARE site in the 1.5-kb *XbaI-XbaI* fragment appears important for *Hoxa5* expression in CNS, it is not required for *Hoxa5* expression in the developing lung and gut when tested in E13.5 transgenic embryos. In EMSA experiments with the 259-bp *XbaI-BssHII* fragment, the lack of binding between the RARE sequence and WCE prepared from lung, stomach and intestine of E13.5 mouse embryos further supports the absence of a role for the RARE sequence in *Hoxa5* organ expression. In contrast, *Hoxa4* expression in lung, gut and metanephros at the same embryonic age is dependent on the functional

RARE, suggesting that the action of the RARE site is *Hox* promoter-specific [44]. We cannot rule out the possibility that the RARE site may be functional in a time-dependent manner regulating *Hoxa5* lung and gut expression at earlier embryonic stages as suggested by studies showing that RA deficiency negatively impacts on *Hoxa5* expression in the developing lung and stomach at E10.5 [52]. Another RARE located at the 3' end of the human *HOXA5* gene was shown to mediate RA responsiveness of the gene in breast cancer cells [53]. This RARE site is conserved in the mouse genome at position +7.4-kb from the *Hoxa5* TSS. Even though our deletion studies did not identify the corresponding genomic region to be involved in *Hoxa5* gene regulation, we cannot exclude a role for this RARE in *Hoxa5* developmental expression.

Our combined approach of transgenesis and biochemistry revealed that YY1 is a positive regulator of *Hoxa5* gene expression in the developing respiratory and digestive tracts, while it does not play a major role in CNS expression. YY1 is a multifunctional zinc-finger-containing transcription factor, identified as the homolog of *Drosophila* Pleiohomeotic (PHO) protein, the latter recruiting Polycomb group (PcG) proteins to negatively regulate genes. YY1 plays crucial roles in numerous biological processes by selectively initiating, activating or repressing transcription, depending upon promoter contextual differences or specific protein interactions [54]. The role of YY1 in *Hox* gene regulation is mainly associated with repression as shown for the *Hoxb4*, *Hoxd4*, *HOXB13* and *HOXD11-HOXD12* genes. For instance, YY1 binds *Hoxb4* promoter and intron enhancer sequences through overlapping NFY/YY1 sites suggesting that the relative levels of binding of the transcriptional activator NFY and YY1 through the same site mediate opposing transcriptional effects on *Hoxb4* expression along the A-P axis [26]. YY1 represses *Hoxd4* expression in undifferentiated P19 cells by recruiting the PcG protein MEL18 to the *Hoxd4* proximal promoter, which maintains silencing at the *Hoxd4* locus [28]. YY1 also participates in the repression of *HOXB13* expression in prostate cancer cells through an epigenetic mechanism involving histone acetylation modification [56]. Finally, YY1 binding sites are present in the *HOXD11-HOXD12* intergenic region in human embryonic stem cells, YY1 recruiting PcG proteins for the transcriptional repression of the distal *HOXD* genes [29]. So far, only the *Hoxb7* gene expression is positively regulated by YY1 in tumor and transformed cell lines [27].

Here, our work established for the first time the positive role of YY1 in *Hox* gene expression in a normal context, the embryo. Our EMSA and ChIP data obtained with organ preparations from E13.5 embryos demonstrated that YY1 physically interacts with two *Hoxa5* regulatory sequences located in the *Hoxa5-Hoxa4* intergenic region. Mutations in these YY1 binding sites abolished *Hoxa5* expression in lung, stomach and intestine from transgenic embryos. Moreover, the conditional deletion of *Yy1* function in lung mesenchyme resulted in decreased levels of *Hoxa5* transcripts in the developing lung. Altogether, these data demonstrated that YY1 acts as a transcriptional activator of *Hoxa5* expression in lung and gut during embryogenesis.

In the trunk of E12.5 mouse embryos, YY1 binding sites are present in *Hoxa5* upstream sequences and they co-localize with EED and BMI1 binding sites, two PcG protein members of the PcG repressive complex 1. PcG binding is specific to the anterior domain of the trunk and it results in the transcriptional silencing of *Hoxa5* in this axial region demonstrating that PcG repression is involved in the establishment of the correct *Hoxa5* expression domain in the prevertebral column [56]. Thus, *Hoxa5* developmental expression is under the control of several YY1 binding sites distributed along the *Hoxa5* locus. Depending on the developmental context, YY1 can mediate PcG repression of *Hoxa5* expression via binding sites located in *Hoxa5* upstream sequences to finely define *Hoxa5* expression domain along the A-P axis [56]. Here, we showed that YY1 can also activate *Hoxa5* expression in the developing respiratory and digestive tracts acting via binding sites located in the intergenic *Hoxa5-Hoxa4* region and shared with the flanking *Hoxa4* gene. Studies have shown that YY1 acts by recruiting co-activators to function on YY1-activated targets [57-61]. This situation may also prevail for the *Hoxa5* gene. While some of these coactivators are expressed in the lung (INO80, PRMT1, BAP1, GATA-4, AP-1), no information is available about their potential role in lung development.

Genetic analyses have revealed the crucial role played by YY1 during embryogenesis. *Yy1*^{-/-} embryos die shortly after implantation [47]. The use of the *Yy1* conditional allele with lineage-specific *Cre* mouse lines circumvents the early embryonic lethality of *Yy1* null mutants revealing the large spectrum of YY1 actions throughout life [38, 62-65]. A dosage-dependent requirement for YY1 is essential for survival as newborn

mice expressing 25% of normal YY1 levels die at birth from respiratory failure [38]. Despite these observations, little is known about the *Yy1* function during lung morphogenesis. Here, we showed that the specific ablation of *Yy1* function in mesenchyme via the use of *Dermo1*^{+Cre} deleter mice causes neonatal death of mutant pups likely due to lung defects. Lungs from *Yy1*^{flox/flox}*Dermo1*^{+Cre} embryos were collapsed with narrow airspaces and thick mesenchyme. Epithelial cell differentiation along the respiratory tract was also affected as shown by the decreased expression of CC10 and T1 α , specific markers of club cells present in the respiratory airways and type I pneumocytes lining the alveolar epithelium, respectively. These lung phenotypes were similar to those observed in *Hoxa5*^{-/-} embryos, supporting the notion that mesenchymal YY1 action during lung formation is mediated, at least in part, by the control of *Hoxa5* expression [6, 35, 41]. This is further reinforced by the decreased *Hoxa5* expression in lungs from *Yy1*^{flox/flox}*Dermo1*^{+Cre} embryos. However, according to the broad transcriptional activity of YY1, the lung phenotype of *Yy1*^{flox/flox}*Dermo1*^{+Cre} mutants may also result from the deregulation of other genes.

YY1 was shown to be essential for the transcription of the *Scgb1a1* gene in endometrial cells [66]. In the present case, the decreased *Scgb1a1* expression observed in the respiratory epithelium of *Yy1*^{flox/flox}*Dermo1*^{+Cre} embryos is a non-cell autonomous phenotype since the *Yy1* mutation is specific to mesenchyme. *Scgb1a1* expression is also diminished in the respiratory epithelium of *Hoxa5*^{-/-} mutants [41]. Therefore, *Hoxa5* could be a downstream effector of YY1. Alternatively, *Yy1* and *Hoxa5* genes may act on *Scgb1a1* expression via distinct pathways.

Hoxa4 lung expression was reduced in lungs from *Yy1*^{flox/flox}*Dermo1*^{+Cre} embryos, suggesting that the YY1 binding sites present in the 1.5-kb *XbaI-XbaI* DNA fragment are involved in the regulation of both *Hoxa4* and *Hoxa5* genes. No lung phenotype was reported in *Hoxa4*^{-/-} mutants, but it is possible that *Hoxa4* exerts a role during lung development masked by functional redundancy. Even though the decrease in *Hoxa4* and *Hoxa5* expression was modest in lungs from *Yy1*^{flox/flox}*Dermo1*^{+Cre} embryos, the combined downregulation of *Hoxa4* and *Hoxa5* may participate to the lung phenotype observed in *Yy1*^{flox/flox}*Dermo1*^{+Cre} mutants.

In summary, our search for regulatory sequences that correctly reproduce *Hoxa5* developmental expression has led to the identification of additional *cis*-acting elements present in the *Hoxa4-Hoxa5* intergenic region that are important for *Hoxa5* expression in CNS, lung, stomach and intestine. Several of these sequences are shared with the flanking *Hoxa4* gene supporting the model that coordinated regulatory mechanisms between *Hox* genes are essential for the precise function of each gene and the correct development of the embryo [15]. We have also unveiled the crucial role of YY1 as a transcriptional activator of *Hoxa5* lung and gut expression and the repercussions of the *Yy1* conditional deletion in lung mesenchyme. Further studies will explore the extent of the regulatory role of YY1 in *Hoxa5* developmental expression.

ACKNOWLEDGEMENTS

We thank Drs. Jean Charron and Olivier Boucherat for comments on the manuscript, Drs. Yang Shi and David Ornitz for providing mouse lines, Drs. George Blanck, David Lohnes and Gurmukh Singh for sharing reagents, Drs. Jacques Côté and Benoit Leblanc for advices on ChIP and RARE experiments, respectively, and Julie Moreau and Kim Landry-Truchon for technical contribution. A special thank to Marcelle Carter for her skilled technical assistance.

REFERENCES

1. Akam ME (1987) The molecular basis for metamerism in the *Drosophila* embryo. *Development* 101: 1-22.
2. Pearson JC, Lemons D, McGinnis W (2005) Modulating *Hox* gene functions during animal body patterning. *Nat Rev Genet* 6: 893-904.
3. McGinnis W, Krumlauf R (1992) Homeobox genes and axial patterning. *Cell* 68: 283-302.
4. Wellik DM (2007) *Hox* patterning of the vertebrate axial skeleton. *Dev Dyn* 236: 2454-2463.
5. Jeannotte L, Lemieux M, Charron J, Poirier F, Robertson EJ (1993) Specification of axial identity in the mouse: role of the *Hoxa-5* (*Hox1.3*) gene. *Genes Dev* 7: 2085-2096.
6. Aubin J, Lemieux M, Tremblay M, Bérard J, Jeannotte L (1997) Early postnatal lethality in *Hoxa-5* mutant mice is attributable to respiratory tract defects. *Dev Biol* 192: 432-445.
7. Aubin J, Chailier P, Ménard D, Jeannotte L (1999) Loss of *Hoxa5* gene function in mice

- perturbs intestinal maturation. *Am J Physiol* 277: C965-C973.
8. Aubin J, Déry U, Lemieux M, Chailler P, Jeannotte L (2002) Stomach regional specification requires *Hoxa5*-driven mesenchymal-epithelial signaling. *Development* 129: 4075-4087.
 9. Meunier D, Aubin J, Jeannotte L (2003) Perturbed thyroid morphology and transient hypothyroidism symptoms in *Hoxa5* mutant mice. *Dev Dyn* 227: 367-378.
 10. Garin É, Lemieux M, Coulombe Y, Robinson GW, Jeannotte L (2006) Stromal *Hoxa5* function controls the growth and differentiation of mammary alveolar epithelium. *Dev Dyn* 235: 1858-1871.
 11. Gendronneau G, Boucherat O, Aubin J, Lemieux M, Jeannotte L (2012) The loss of *Hoxa5* function causes estrous acyclicity and ovarian epithelila inclusion cysts. *Endocrinology* 153: 1484-1497.
 12. Coulombe Y, Lemieux M, Moreau J, Aubin J, Joksimovic M, Bérubé-Simard FA, Tabariès S, Boucherat O, Guillou F, Larochelle C, Tuggle CK, Jeannotte L (2010) Multiple promoters and alternative splicing: *Hoxa5* transcriptional complexity in the mouse embryo. *PLoS ONE* 5(5): e10600. doi:10.1371/journal.pone.0010600.
 13. Deschamps J, van Nes J (2005) Developmental regulation of the *Hox* genes during axial morphogenesis in the mouse. *Development* 132: 2931-2942.
 14. Spitz F, Duboule D (2008) Global control regions and regulatory landscapes in vertebrate development and evolution. *Adv Genet* 61: 175-205.
 15. Sharpe J, Nonchev S, Gould A, Whiting J, Krumlauf R (1998) Selectivity, sharing and competitive interactions in the regulation of *Hoxb* genes. *EMBO J* 17: 1788-1798.
 16. Chambeyron S, Bickmore WA (2004) Chromatin decondensation and nuclear reorganization of the *HoxB* locus upon induction of transcription. *Genes Dev* 18: 1119-1130.
 17. Parrish M, Nolte C, Krumlauf R (2009) *Hox* genes expression. In Squire L, editor. *New Encyclopedia of Neuroscience*. 4th ed, Academic Press. pp. 1221-1231.
 18. Subramanian V, Meyer BI, Gruss P (1995) Disruption of the murine homeobox gene *Cdx1* affects axial skeletal identities by altering the mesodermal expression domains of *Hox* genes. *Cell* 83: 641-653.
 19. Charité J, de Graaff W, Consten D, Reijnen M J, Korving J, Deschamps J (1998) Transducing positional information to the *Hox* genes: critical interaction of *cdx* gene products with position-sensitive regulatory elements. *Development* 125: 4349-4358.
 20. Houle M, Prinos P, Iulianella A, Bouchard N, Lohnes D (2000) Retinoic acid regulation of *Cdx1*: an indirect mechanism for retinoids and vertebral specification. *Mol Cell Biol* 20: 6579-6586.
 21. Pilon N., Oh K, Sylvestre JR, Savory JG, Lohnes D (2007) Wnt signaling is a key mediator of *Cdx1* expression in vivo. *Development* 134: 2315-2323.
 22. Dupé V, Davenne M, Brocard J, Dollé P, Mark M, Dierich A, Chambon P, Rijli FM (1997) *In vivo* functional analysis of the *Hoxa-1* 3' retinoic acid response element (3'RARE). *Development* 124: 399-410.
 23. Oosterveen T, Niederreither K, Dollé P, Chambon P, Meijlink F, Deschamps J (2003) Retinoids regulate the anterior expression boundaries of 5' *Hoxb* genes in posterior hindbrain. *EMBO J* 22: 262-269.
 24. Pöpperl H, Featherstone MS (1992) An autoregulatory element of the murine *Hox-4.2* gene. *EMBO J* 11: 3673-3680.

25. Maconochie MK, Nonchev S, Studer M, Chan SK, Pöpperl H, Sham MH, Mann RS, Krumlauf R (1997) Cross-regulation in the mouse *HoxB* complex: the expression of *Hoxb2* in rhombomere 4 is regulated by *Hoxb1*. *Genes Dev* 15: 1885-1895.
26. Gilthorpe J, Vandromme M, Brend T, Gutman A, Summerbell D, Totty N, Rigby PW (2002) Spatially specific expression of *Hoxb4* is dependent on the ubiquitous transcription factor NFY. *Development* 129: 3887-3899.
27. Meccia E, Bottero L, Felicetti F, Peschle C, Colombo MP, Carè A (2003) HOXB7 expression is regulated by the transcription factors NF-Y, YY1, Sp1 and USF-1. *Biochim Biophys Acta* 1626: 1-9.
28. Kobrossy L, Rastegar M, Featherstone M (2006) Interplay between chromatin and trans-acting factors regulating the *Hoxd4* promoter during neural differentiation. *J Bio Chem* 281: 25926-25939.
29. Woo CJ, Kharchenko PV, Daheron L, Park PJ, Kingston RE (2010) A region of the human *HOXD* cluster that confers polycomb-group responsiveness. *Cell* 140: 99-110.
30. Zakany J, Tuggle CK, Patel M, Nguyen-Huu MC (1988) Spatial regulation of homeobox gene fusions in the embryonic central nervous system of transgenic mice. *Neuron* 1: 679-691.
31. Larochelle C, Tremblay M, Bernier D, Aubin J, Jeannotte L (1999) Multiple *cis*-acting regulatory regions are required for the restricted spatio-temporal *Hoxa5* expression. *Dev Dyn* 214: 127-140.
32. Tabariès S, Lapointe J, Besch T, Carter M, Woollard J, Tuggle CK, and Jeannotte L (2005) Cdx protein interaction with *Hoxa5* regulatory sequences contributes to *Hoxa5* regional expression along the axial skeleton. *Mol Cell Biol* 25: 1389-1401.
33. Moreau J, Jeannotte L (2002) Sequence analysis of *Hoxa4-Hoxa5* intergenic region including shared regulatory elements. *DNA Sequence* 13: 203-209.
34. Herriges JC, Yi L, Hines EA, Harvey JF, Xu G, Gray PA, Ma Q, Sun X (2012) Genome-scale study of transcription factor expression in the branching mouse lung. *Dev Dyn* 241: 1432-1453.
35. Boucherat O, Montaron S, Bérubé-Simard FA, Aubin J, Philippidou P, Wellik DM, Dasen JS, Jeannotte L (2013) Partial functional redundancy between *Hoxa5* and *Hoxb5* paralog genes during lung morphogenesis. *Am J Physiol Lung Cell Mol Physiol* 304: L817-L830.
36. Kothary R, Clapoff S, Darling S, Perry MD, Moran LA, Rossant J (1989) Inducible expression of an *hsp68-lacZ* hybrid gene in transgenic mice. *Development* 1989: 105: 707-714.
37. Hogan B, Beddington RS, Constantini F, Lacy E (1994) Manipulating the mouse embryo: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
38. Affar el B, Gay F, Shi Y, Liu H, Huarte M, Wu S, Collins T, Li E, Shi Y (2006) Essential dosage-dependent functions of the transcription factor yin yang 1 in late embryonic development and cell cycle progression. *Mol Cell Biol* 26: 3565-3581.
39. Yu K, Xu J, Liu Z, Susic D, Shao J, Olson EN, Towler DA, Ornitz DM (2003) Conditional inactivation of FGF receptor 2 reveals an essential role for FGF signaling in the regulation of osteoblast function and bone growth. *Development* 130: 3063-3074.
40. Jaffe L, Jeannotte L, Bikoff EK, Robertson EJ (1990) Analysis of beta 2-microglobulin gene expression in the developing mouse embryo and placenta. *J Immunol* 145: 3474-

3482.

41. Boucherat O, Chakir J, Jeannotte L (2012) The loss of *Hoxa5* function in mice promotes NOTCH-dependent goblet cell metaplasia in lung airways. *Biology Open* 1: 677-691.
42. Aparicio O, Geisberg JV, Sekinger E, Yang A, Moqtaderi Z, Struhl K (2005) Chromatin immunoprecipitation for determining the association of proteins with specific genomic sequences *in vivo*. *Current Protocols in Molecular Biology* 21.3.1-21.3.33.
43. Dony C, Gruss P (1987) Specific expression of the Hox1.3 homeo box gene in murine embryonic structures originating from or induced by the mesoderm. *EMBO J* 6: 2965-2975.
44. Packer AI, Crotty DA, Elwell VA, Wolgemuth DJ (1998) Expression of the murine *Hoxa4* gene requires both autoregulation and a conserved retinoic acid response element. *Development* 125: 1991-1998.
45. Santini S, Boore JL, Meyer A (2003) Evolutionary conservation of regulatory elements in vertebrate *Hox* clusters. *Genome Res* 13: 1111-1122.
46. Khang K, Chung JH, Kim J (2009) Evolutionary Conserved Motif Finder (ECMFinder) for genome-wide identification of clustered YY1- and CTCF-binding sites. *Nucl Acid Res* 37: 2003-2013.
47. Donohoe M, Zhang X, McGinnis L, Biggers J, Li E, Shi Y (1999) Targeted disruption of mouse Yin Yang 1 transcription factor results in peri-implantation lethality. *Mol Cell Biol* 19: 7237-7244.
48. Marshall H, Morrison A, Studer M, Pöpperl H, Krumlauf R (1996) Retinoids and *Hox* genes. *FASEB J* 10: 969-978.
49. Nolte C, Amores A, Nagy Kovács E, Postlethwait J, Featherstone M (2003) The role of a retinoic acid response element in establishing the anterior neural expression border of *Hoxd4* transgenes. *Mech Dev* 120: 325-335.
50. Pöpperl H, Featherstone MS (1993) Identification of a retinoic acid response element upstream of the murine Hox-4.2 gene. *Mol Cell Biol* 13: 257-265.
51. Mainguy G, In der Rieden PM, Berezikov E, Woltering JM, Plasterk RH, Durston AJ (2003) A position-dependent organisation of retinoid response elements is conserved in the vertebrate *Hox* clusters. *Trends Genet* 19: 476-479.
52. Wang Z, Dollé P, Cardoso WV, Niederreither K (2006) Retinoic acid regulates morphogenesis and patterning of posterior foregut derivatives. *Dev Biol* 297: 433-445.
53. Chen H, Zhang H, Lee J, Liang X, Wu X, Zhu T, Lo PK, Zhang X, Sukumar S (2007) HOXA5 acts directly downstream of retinoic acid receptor β and contributes to retinoic acid-induced apoptosis and growth inhibition. *Cancer Res* 67: 8007-8013.
54. Thomas MJ, Seto E (1999) Unlocking the mechanisms of transcription factor YY1: are chromatin modifying enzymes the key? *Gene* 236: 197-208.
55. Ren G, Zhang G, Dong Z, Liu Z, Li L, Feng Y, Su D, Zhang Y, Huang B, Lu J (2009) Recruitment of HDAC4 by transcription factor YY1 represses HOXB13 to affect cell growth in AR-negative prostate cancers. *Int J Biochem Cell Biol* 41: 1094-1101.
56. Kim SY, Paylor SW, Magnuson T, Schumacher A (2006) Juxtaposed Polycomb complexes co-regulate vertebral identity. *Development* 133: 4957-4968.
57. Rezai-Zadeh N, Zhang X, Namour F, Fejer G, Wen YD, Yao YL, Gyory I, Wright K, Seto E (2003) Targeted recruitment of a histone H4-specific methyltransferase by the transcription factor YY1. *Genes Dev* 17: 1019-1029.
58. Cai Y, Jin J, Yao T, Gottschalk AJ, Swanson SK, Wu S, Shi Y, Washburn MP, Florens

- L, Conaway RC, Conaway JW (2007) YY1 functions with INO80 to activate transcription. *Nat Struct Mol Biol* 14: 872-874.
59. Wang CC, Tsai MF, Dai TH, Hong TM, Chan WK, Chen JJW, Yang PC (2007) Synergistic activation of the tumor suppressor HLJ1 by the transcription factors YY1 and activator protein 1. *Cancer Res* 67: 4816-4826.
60. Yu H, Mashtalir N, Daou S, Hammond-Martel I, Ross J, Sui G, Hart GW, Rausher III FJ, Drobetsky E, Milot E, Shi Y, Affar el B (2010) The ubiquitin carboxyl hydrolase BAP1 forms a ternary complex with YY1 and HCF-1 and is a critical regulator of gene expression. *Mol Cell Biol* 30: 5071-5085.
61. Bhalla SS, Robitaille L, Nemer M (2001) Cooperative activation by GATA-4 and YY1 of the cardiac B-type natriuretic peptide promoter. *J Biol Chem* 276: 11439-11445.
62. Trask MC, Tremblay KD, Mager J (2012) Yin-Yang 1 is required for epithelial-to-mesenchymal transition and regulation of Nodal signaling during mammalian gastrulation. *Dev Biol* 368: 273-282.
63. Li J, Song JS, Bell RJ, Tran TN, Haq R, Liu H, Love KT, Langer R, Anderson DG, Larue L, Fisher DE (2012) YY1 regulates melanocyte development and function by cooperating with MITF. *PLoS Genet* 8(5): e1002688.
64. Gregoire S, Karra R, Passer D, Deutsch MA, Krane M, Feistritz R, Sturzu A, Domian I, Saga Y, Wu SM (2013) Essential and unexpected role of Yin Yang 1 to promote mesodermal cardiac differentiation. *Circ Res* 112: 900-910.
65. Rhee S, Guerrero-Zayas MI, Wallingford MC, Ortiz-Pineda P, Mager J, Tremblay KD (2013) Visceral endoderm expression of YY1 is required for VEGFA maintenance and yolk sac development. *PLoS One* 8(3): e58828. doi: 10.1371/journal.pone.0058828.
66. Klug J, Beato M (1996) Binding of YY1 to a site overlapping a weak TATA box is essential for transcription from the uteroglobin promoter in endometrial cells. *Mol Cell Biol* 16: 6398-6407.

CHAPITRE III

***Hoxa5/Cre* transgenic mice: novel tools for regional deletion along the anterior-posterior axis**

Félix-Antoine Bérubé-Simard and Lucie Jeannotte[#]

Department of Molecular Biology, Medical Biochemistry and Pathology, Université Laval, Québec, Canada; Centre de recherche sur le cancer de l'Université Laval; Centre de recherche du Centre Hospitalier Universitaire de Québec, L'Hôtel-Dieu de Québec, Québec, Canada

[#] Corresponding author: Lucie Jeannotte
Centre de recherche sur le cancer de l'Université Laval
CRCHU-Q, L'Hôtel-Dieu de Québec
9, rue McMahan
Québec, QC
Canada G1R 2J6
Phone: (418) 525-4444 ext. 15552
Fax: (418) 691-5439
e-mail: lucie.jeannotte@crhdq.ulaval.ca

Running head: *Hoxa5/Cre* deleter mouse lines

Key words: *Hoxa5*, Cre recombinase, lung and gut morphogenesis, anterior-posterior axis

Grant support: This work was supported by grants from the Canadian Institutes of Health Research (MOP-15139; to L.J.) and the National Sciences and Engineering Research Council of Canada (to L.J.)

Disclosure statement: The authors have nothing to disclose.

Article publié en 2014 dans: *Genesis* 52(2):149-56.

AVANT-PROPOS

L'étude des différentes séquences d'ADN orchestrant l'expression du gène *Hoxa5* a révélé la présence de quatre grandes régions régulatrices: l'élément temporel de 648 pb (positions -2128 et -1480) nécessaire à l'expression du gène *Hoxa5* dès le stade E8.25 au niveau des somites 5-7 (Larochelle *et al.* 1999), l'élément BSC de 604 pb (positions -839 à -235) nécessaire à l'expression dans la partie brachiale du tube neural au stade E12.5 (Zákány *et al.*, 1988 ; Tuggle *et al.*, 1990 ; Nowling *et al.*, 1999), l'élément MES de 1287 pb (positions +2851 et +4935) nécessaire à l'expression au niveau des dérivés mésodermiques entre les pv3-10, dans les bourgeons de membres et le système urogénital au stade E12.5 (Larochelle *et al.*, 1999; Tabariès *et al.*, 2005) et l'élément Poumon/Intestin de 1465 pb (positions +9351 et +10816) qui dirige l'expression dans les systèmes respiratoire et digestif au stade E13.5 (Moreau et Jeannotte, 2002).

Grâce à ces connaissances, j'ai élaboré deux constructions d'ADN où une recombinaison Cre est sous le contrôle de différentes combinaisons des éléments régulateurs du gène *Hoxa5*. La première construction d'ADN (pLJ447) englobe la totalité des séquences régulatrices connues du gène *Hoxa5* (14.5-kb *Hoxa5/Cre*), alors que la seconde (pLJ444) ne contient pas l'élément Poumon/Intestin (7.1-kb *Hoxa5/Cre*). Ensuite, j'ai entrepris la caractérisation des deux différentes lignées de souris transgéniques obtenues. Tout d'abord, j'ai croisé ces deux lignées avec des souris ROSA26^{R/R} (Friedrich et Soriano, 1991; Soriano, 1999) et observé que leur expression récapitulait majoritairement l'expression embryonnaire endogène dirigée par les éléments régulateurs en présence. De plus, l'expression au niveau des poumons (Aubin *et al.*, 1997), de l'estomac (Aubin *et al.*, 2002a), de l'intestin (Aubin *et al.*, 1999), de la glande mammaire (Garin *et al.*, 2006) et de l'ovaire (Gendronneau *et al.*, 2012) a également été validée chez l'adulte. J'ai ensuite vérifié l'efficacité de recombinaison inhérente à ces deux lignées en les croisant avec des souris *Hoxa5*^{flox/flox} et en comparant les phénotypes obtenus avec ceux des souris mutantes *Hoxa5*^{-/-} (Jeannotte *et al.*, 1991, 1993; Aubin *et al.*, 1997) et *Hoxa5*^{Δ/Δ} (Tabariès *et al.*, 2007), notamment au niveau du squelette et des poumons. Dans un premier temps, j'ai réalisé une étude de survie démontrant que la mortalité périnatale observée chez les individus mutants *Hoxa5*^{-/-} n'est présente que chez les souris *Hoxa5*^{flox/flox}, 14.5-kb *Hoxa5/Cre*. D'ailleurs, j'ai

confirmé la perte d'expression du gène *Hoxa5* dans les poumons de ces individus au stade E18.5 par PCR quantitatif, de même que leur apparence anormalement compacte par analyse histologique. Enfin, la délétion conditionnelle du gène *Hoxa5* générée par l'une ou l'autre des deux lignées de souris transgéniques *Hoxa5/Cre* entraîne différentes transformations au niveau du squelette qui sont caractéristiques de la perte de fonction du gène *Hoxa5*. Ces lignées de souris *Hoxa5/Cre* sont donc de nouveaux outils intéressants pour l'analyse fonctionnelle de différents gènes à l'intérieur de tissus définis le long de l'axe AP.

J'ai collaboré à l'écriture du manuscrit avec le Dre Lucie Jeannotte.

ABSTRACT

The *Hoxa5* homeobox gene encodes a transcription factor that plays a critical role in specifying the identity of the cervico-thoracic region along the anterior-posterior embryo axis and in orchestrating organ morphogenesis. The loss of *Hoxa5* function results in skeletal transformations, lethality at birth due to lung defects, and organ anomalies affecting the digestive tract, the mammary gland and the ovary. Study of *Hoxa5* gene regulation has revealed the interplay of several control regions that direct *Hoxa5* developmental expression. Enhancers targeting expression in the CNS, the paraxial and lateral plate mesoderm at the cervico-thoracic level, and in the mesenchymal compartment of the respiratory and digestive tracts have been identified. Using these molecular tools, we have generated two *Hoxa5/Cre* transgenic mouse lines carrying different combinations of *Hoxa5* regulatory enhancers and allowing site-specific recombination in subsets of *Hoxa5* expression sites as tested with the *Rosa26/lacZ* reporter mice. Further validation of the recombination efficiency of the *Hoxa5/Cre* transgenic lines was performed with mice carrying a *Hoxa5* conditional allele. *Hoxa5* deletion with the *Hoxa5/Cre* mouse lines recapitulates *Hoxa5* mutant phenotypes, such as skeletal defects, neonatal lethality and lung malformations. *Hoxa5/Cre* mouse lines provide novel genetic tools for gene function analysis in defined tissues along the anterior-posterior axis.

INTRODUCTION

The use of site-specific recombinase systems provides a powerful means to tailor the mouse genome. It allows cell lineage analysis, generation of subtle modifications mimicking human heritable mutations, and production of site- or time-specific mutations, that circumvent early embryonic lethality in numerous cases enabling a better comprehension of gene function *in vivo*. The power of the recombinase-based strategies relies mainly on the collection of mouse lines expressing the site-specific recombinase Cre from bacteriophage P1, which can delete any DNA sequence located between two *loxP* sites similarly oriented (Nagy *et al.*, 2009). Production of Cre mouse lines can be achieved by the insertion of Cre sequences by gene targeting into the locus of interest, which drives Cre expression by endogenous gene regulation. However, this approach can create a mutation in the targeted gene that may interfere with subsequent analyses. Classical transgenesis is also largely applied to generate Cre transgenic mice. Although it requires a good knowledge of the regulatory sequences used, it can usefully target Cre activity in specific subsets of expression sites depending on the enhancers included in the transgene.

Hox genes encode evolutionarily conserved transcription factors essential for the correct organization and morphogenesis of the body structures along the anteriorposterior (AP) body axis. Due to their crucial role during embryogenesis, an extensive characterization of *Hox* gene regulation has been undertaken revealing the interplay of transcriptional regulators interacting with positive and negative *cis*-acting sequences to differentially control *Hox* expression in a spatio-temporal and tissue-specific fashion. For instance, several regulatory elements directing the developmental expression of the *Hoxa5* gene are contained in a 14.5-kb genomic fragment that recapitulates the restricted spatio-temporal expression profile along the AP axis of the embryo directed by the *Hoxa5* proximal promoter (Bérubé-Simard, unpublished data). In embryos, the 14.5-kb *Hoxa5/lacZ* transgene targets *lacZ* expression to the brachial region of the central nervous system (CNS) with the correct anterior limit of *Hoxa5* expression at the level of the floor of myelencephalon (Zakany *et al.*, 1988). It reproduces *Hoxa5* expression in the paraxial and lateral plate mesoderm at the cervico-thoracic level (Larochelle *et al.*, 1999). Finally, it

replicates *Hoxa5* expression in the mesenchymal compartment of the respiratory and digestive tracts as well as the *Hoxa5* rostro-caudal gradient in the stomach (Aubin *et al.*, 1997; 1999; 2002; Moreau *et al.*, 2002). In agreement with the *Hoxa5* expression profile, *Hoxa5*^{-/-} mice present homeotic transformations affecting the cervico-thoracic region of the skeleton (Jeannotte *et al.*, 1993). Most *Hoxa5*^{-/-} mice die at birth from respiratory tract malformations (Aubin *et al.*, 1997; Boucherat *et al.*, 2013). The loss of *Hoxa5* function also affects the regionalization and the specification of the gastric epithelium as well as the functional maturation of the midgut (Aubin *et al.*, 1999; 2002). Surviving *Hoxa5*^{-/-} mice have mammary gland and ovary defects (Garin *et al.*, 2006; Gendronneau *et al.*, 2012).

Here, we report the characterization of two *Hoxa5/Cre* transgenic mouse lines carrying different combinations of *Hoxa5* regulatory enhancers and allowing site-specific recombination in subsets of *Hoxa5* expression sites. They differ from the Tg(*Hoxa5-cre*)1Fmr, MGI:5474192 mouse line, in which the entire *Hoxa* cluster was included into the transgene to drive the recombinase Cre sequences inserted into the *Hoxa5* locus (Di Meglio *et al.*, 2013). The 7.1-kb *Hoxa5/Cre* transgene contains the temporal (T) control region, the brachial spinal cord (BSC) enhancer, and the mesodermal (MES) enhancer directing expression in paraxial and lateral plate mesoderm from the cervico-upper thoracic domain (Fig. 1; Zakany *et al.*, 1988; Larochelle *et al.*, 1999). The 14.5-kb *Hoxa5/Cre* transgene contains an additional 7.4-kb of downstream sequences including the enhancer targeting *Hoxa5* lung and gut developmental expression (L/G; Moreau *et al.*, 2002). An IRES-eGFP^{Cre} cassette was cloned into the *Hoxa5* exon 1 downstream the HOXA5 AUG allowing translation of the eGFP^{Cre} fusion protein (Lee *et al.*, 2001). Nine and eight founder lines were obtained for the 7.1-kb *Hoxa5/Cre* and 14.5-kb *Hoxa5/Cre* transgenes, respectively.

RESULTS AND DISCUSSION

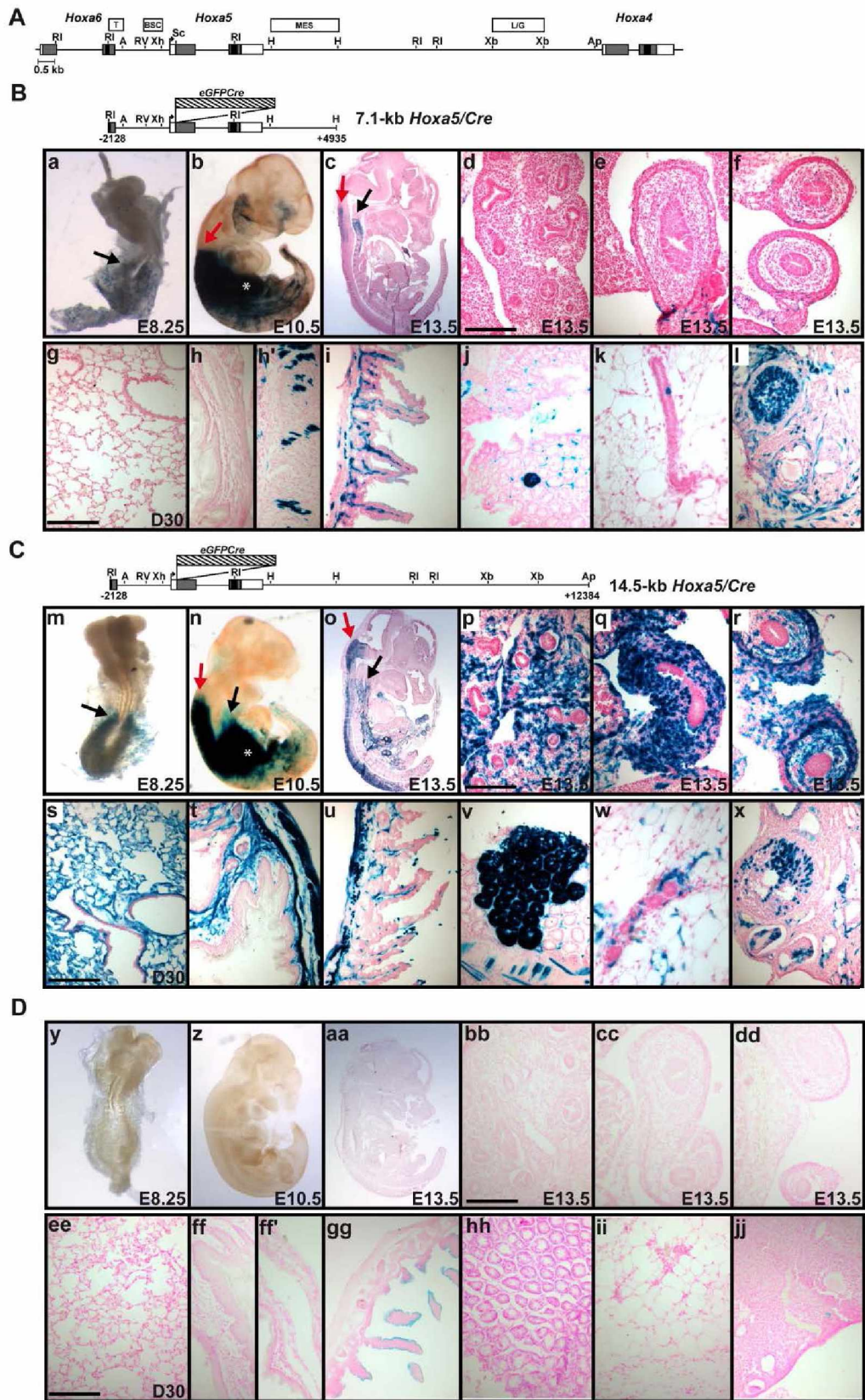
To assess the efficiency and the specificity of Cre-mediated recombination, all transgenic mouse lines were crossed with *Rosa26/lacZ* reporter (*R26R*) mice (Soriano, 1999). Based on the X-Gal staining results obtained at embryonic day (E) 13.5, two mouse lines were further characterized, the pLJ444C line for the 7.1-kb *Hoxa5/Cre* transgene, and

the pLJ447B line for the 14.5-kb *Hoxa5/Cre* transgene. Southern analysis with the *NheI* restriction enzyme cutting once into the transgenes indicated that both lines contained complete head-to-tail transgene integrations at one site (not shown). Quantification of the fragment detected with a *Cre* probe following tail DNA digestion with the *EcoRI* enzyme showed that the pLJ444C and pLJ447B lines contained ~2 and ~6 copies of their transgene, respectively, when compared to the single copy *Cre* gene present in the Twist2tm1(cre)Dor, MGI:3044412 mouse line (not shown). Our *Hoxa5/Cre* transgenic lines were viable and fertile with no overt morphological or behavioral defects. Breeding to homozygosity did not reveal apparent phenotype.

When bred with *R26R* mice, both lines targeted expression at E8.25 in somites 5-8, which corresponds to the onset of *Hoxa5* endogenous expression and to the correct AP expression domain at this age (Fig. 3.1a,m). At E10.5, the expression of both transgenes extended more caudally along the axis but remained confined to the cervico-thoracic region in a *Hoxa5*-like fashion. Expression was also detected in limb buds (Fig. 3.1b,n). In the neural tube, the 14.5-kb *Hoxa5/Cre* transgene recapitulated the correct *Hoxa5* anterior boundary at the posterior hindbrain level. At E13.5, the 7.1-kb *Hoxa5/Cre* transgenic line directed *Cre* expression in the CNS brachial region and in the cervico-upper thoracic domain of the prevertebral column, the latter being specific to the *Hoxa5* proximal promoter (Fig. 3.1c; Coulombe *et al.*, 2010). No staining was detected in the lung (Fig. 3.1d). Scarce punctuated *lacZ*-positive spots were observed in the developing gut that might be driven by the *Hoxa5* MES sequence (Fig. 3.1e-f; Larochelle *et al.*, 1999). The 14.5-kb *Hoxa5/Cre* transgene directed a broader expression pattern in the CNS and in the mesenchymal component of lung, stomach and intestine, reproducing *Hoxa5* endogenous expression in these structures (Fig. 3.1o-r).

Cre activity was also monitored in organs from adult mice at postnatal day (D) 30. *LacZ* staining was detected in lung and trachea mesenchyme only in the 14.5-kb *Hoxa5/Cre* transgenic line (Fig. 3.1g,s; not shown). In the stomach, a strong *Cre* activity was observed in the mesenchymal and muscular layers of the entire organ in 14.5-kb *Hoxa5/Cre* transgenic mice (Fig. 3.1t). In contrast, 7.1-kb *Hoxa5/Cre* specimens presented no activity

Figure 3.1. Characterization of *Hoxa5/Cre* mouse lines during development and at adult age. (A) Partial restriction map of the *Hoxa4-Hoxa6* genomic region. Black boxes represent homeoboxes, shaded boxes correspond to translated regions, and white boxes indicate transcribed regions. The arrow shows the transcription start site of the *Hoxa5* proximal promoter. The localization of the *Hoxa5* enhancer regions is indicated. (B-C) Schematic diagrams of the 7.1-kb *Hoxa5/Cre* and 14.5-kb *Hoxa5/Cre* transgenes. A, *AccI*; Ap, *ApaI*; RI, *EcoRI*; RV, *EcoRV*; H, *HindIII*; Sc, *SacI*; Xb, *XbaI*; Xh, *XhoI*. (a-b,m-n) Representative E8.25 (a,m) and E10.5 (b,n) transgenic embryos stained for β -galactosidase activity showed Cre-mediated recombination directed by the 7.1-kb *Hoxa5/Cre* (a-b) and 14.5-kb *Hoxa5/Cre* (m-n) transgenes. The position of limb buds is indicated by an asterisk. (c-f,o-r) Sagittal sections of E13.5 *R26R;Hoxa5/Cre+* transgenic embryos stained for β -galactosidase activity. Sections of complete embryos (c,o), lungs (d,p), stomach (e,q) and midgut (f,r) are shown. Black arrows indicate the anterior limit of *lacZ* staining in paraxial mesoderm at the level of somite 5 or prevertebra 3 for both transgenes. Red arrows point the anterior limit of expression in the neural tube. The 14.5-kb *Hoxa5/Cre* transgene reproduced the endogenous boundary at the posterior hindbrain level. (g-l,s-x) β -galactosidase activity in organs from D30 *R26R;Hoxa5/Cre+* mice. Sections of lungs (g,s), stomach (h,t: rostral part; h': caudal part), small intestine (i,u), colon (j,v), mammary gland (k,w) and ovary (l,x) are shown. (y-jj) No β -galactosidase activity was detected in *R26R* control embryos and in organs from D30 *R26R* control adults except at the luminal surface of the intestinal epithelium (gg). Scale bars: 100 μ m.



in the rostral portion of the stomach, but stained mesenchymal cells were detected in the caudal part, a pattern reminiscent of the endogenous *Hoxa5* rostro-caudal gradient previously reported (Fig. 3.1h,h'). The lack of gradient in 14.5-kb *Hoxa5/Cre* specimens may be hidden by the robust *lacZ* staining due to the high transgene copy number and/or the integration site. In the intestine, staining was detected in both mouse lines. In the 7.1-kb *Hoxa5/Cre* mouse line, *lacZ* expression was restricted to the mesenchymal cells of the villi, whereas staining was also detected in the underneath muscular layer of the intestine in 14.5-kb *Hoxa5/Cre* specimens (Fig. 3.1i,u). In the colon, staining of mesenchymal and muscular cells as well as clusters of X-Gal-stained epithelial cells were observed in both mouse lines, the clusters being larger and more abundant in 14.5-kb *Hoxa5/Cre* specimens (Fig. 3.1j,v). *Hoxa5* expression was not observed in the developing and adult gut epithelium, suggesting that the epithelial staining might result from ectopic expression of the transgene (Aubin *et al.*, 1999; not shown). Finally, Cre activity was detected in mammary glands and ovaries, two sites where *Hoxa5* expression was reported to be restricted to stroma (Garin *et al.*, 2006; Gendronneau *et al.*, 2012). *LacZ* staining was seen in few mammary epithelial cells in the 7.1-kb *Hoxa5/Cre* mouse line suggesting ectopic expression of the transgene (Fig. 3.1k). In the 14.5-kb *Hoxa5/Cre* mouse line, expression was also observed in mammary stroma, consistent with the presence of a *Hoxa5* mammary regulatory element in the 3' sequences contained in the transgene (Fig. 3.1w; Chen *et al.*, 2007). In ovaries, both transgenes targeted expression in the stroma but the ectopic expression seen in corpus lutea and follicles suggested that additional inhibitory sequences might be required to reproduce correct *Hoxa5* expression (Fig. 3.1l,x). No Cre activity was observed in *R26R* control specimens except in the adult small intestine where staining was detected at the luminal surface of the epithelium corresponding to the endogenous lactase activity (Fig. 3.1y-jj; Salehi *et al.*, 2009).

To further validate the recombination efficiency of the *Hoxa5/Cre* transgenic lines, they were bred with mice carrying the *Hoxa5* conditional allele (*Hoxa5^{flox}*; Tabariès *et al.*, 2007). Deletion with the 7.1-kb *Hoxa5/Cre* mouse line did not cause lethality of *Hoxa5^{flox/flox};7.1-kb Hoxa5/Cre⁺* pups. All possible genotypes were obtained at the expected Mendelian ratio at E18.5 and adult age (Fig. 3.2A). Conversely, 42% of the *Hoxa5^{flox/flox};14.5-kb Hoxa5/Cre⁺* pups were found dead at birth, a ratio similar to the lethal

phenotype observed for the *Hoxa5*^{-/-} pups (Fig. 3.2B; Jeannotte *et al.*, 1993). Lung comparative morphology indicated that *Hoxa5*^{flox/flox};14.5-kb *Hoxa5/Cre*⁺ pups displayed respiratory distress at birth as lungs from E18.5 *Hoxa5*^{flox/flox};14.5-kb *Hoxa5/Cre*⁺ embryos presented a collapsed appearance comparable to that reported for *Hoxa5*^{-/-} mutants (Fig. 3.2C-D; Aubin *et al.*, 1997). qRT-PCR analysis confirmed the decreased *Hoxa5* expression in lungs from E18.5 mutants (Fig. 3.2E). We also examined lungs from surviving D30 *Hoxa5*^{flox/flox};14.5-kb *Hoxa5/Cre*⁺ mutants and observed the enlargement of alveolar spaces as previously seen in *Hoxa5*^{-/-} adults (Fig. 3.2F-G; Mandeville *et al.*, 2006). Finally, we performed a skeletal analysis of E18.5 specimens carrying the different allelic variants. In concordance with the Cre activity pattern, *Hoxa5*^{flox/flox};7.1-kb *Hoxa5/Cre*⁺ embryos presented a loss of tuberculum anterior on the 6th cervical vertebra (C6), an abnormal xiphoid process at the caudal extremity of the sternum, and a normal trachea (Fig. 3.3A-D,G-I,L-N). Although Cre activity was reported in the paraxial mesenchyme of the cervico-upper thoracic region, the presence of ectopic ribs on C7, a transformation seen in most *Hoxa5*^{-/-} specimens, was not detected (Jeannotte *et al.*, 1993). Conversely, C7 extra ribs were observed in *Hoxa5*^{flox/flox};14.5-kb *Hoxa5/Cre*⁺ embryos as well as absent or shorter tuberculum anterior on C6, a misshapen xiphoid process and a profound disorganization of the trachea indicating the importance of the additional regulatory sequences present in the 14.5-kb *Hoxa5/Cre*⁺ transgene (Fig. 3.3A,E-F,J-K,O-P). Altogether these data demonstrated that the *Hoxa5* conditional deletion with the *Hoxa5/Cre* mouse lines can recapitulate several *Hoxa5* mutant phenotypes confirming the efficiency and specificity of the *Hoxa5/Cre* transgenes. In summary, these two *Hoxa5/Cre* transgenic lines provide to the scientific community novel useful tools for deleting floxed alleles in specific subsets of tissues along the AP axis.

METHODS

Design of *Hoxa5/Cre* transgenes and production of transgenic mice

A 2.7-kb *Bsu36I-HincII* fragment from the pIGCN21 vector (obtained from the

Figure 3.2. Impact of the conditional *Hoxa5* deletion with the *Hoxa5/Cre* transgenes. (A-B) Ratios of genotypes of litters obtained from matings between *Hoxa5^{flox/flox}* mice with *Hoxa5^{flox/+}*;7.1-kb *Hoxa5/Cre* and *Hoxa5^{flox/+}*;14.5-kb *Hoxa5/Cre* mice, respectively. (C-D,F-G) Comparative lung histology for E18.5 and D30 *Hoxa5^{flox/+}* and *Hoxa5^{flox/flox}*;14.5-kb *Hoxa5/Cre* specimens. Scale bar: 200 μ m. (E) qRT-PCR analysis for *Hoxa5* expression in lungs from E18.5 *Hoxa5^{+/+}*;14.5-kb *Hoxa5/Cre* and *Hoxa5^{flox/flox}*;14.5-kb *Hoxa5/Cre* embryos. Values are expressed as means \pm SEM. *** $p < 0.001$. (H) The *Hoxa5* alleles produced following the specific action of the Cre recombinase are indicated. The *Hoxa5* probe used for Southern analysis is represented. The expected fragments following *EcoRI* digestion are shown underneath each allele and transgene. Southern blot analysis of DNA from *Hoxa5/Cre* transgenic and *Hoxa5* mutant mice using an *EcoRI* digestion combined with the *Hoxa5* probe discriminates between the wild-type (3793-bp), *Hoxa5^{flox}* (3910-bp), *Hoxa5 Δ* (7679-bp) alleles and the *Hoxa5/Cre* transgenes (2232-bp). An additional band (*) was also detected in the *Hoxa5/Cre* specimens. For the left panel of the Southern blot presented at the right, the contrast was increased to reveal the *Hoxa5* flox and Δ alleles. Ap, *ApaI*; B, *BglIII*; Nh, *NheI*; RI, *EcoRI*; Sc, *SacI*; Xh, *XhoI*.

A

Age	# of litters	# of pups	<i>Hoxa5^{flox/flox} X Hoxa5^{flox/+}; 7.1-kb Hoxa5/Cre</i>			
			Cre-		Cre+	
			<i>Hoxa5^{flox/+}</i>	<i>Hoxa5^{flox/flox}</i>	<i>Hoxa5^{flox/+}</i>	<i>Hoxa5^{flox/flox}</i>
E18.5	8	55	14(25%)	14(25%)	13(24%)	14(25%)
D30	12	76	19(25%)	18(24%)	17(22%)	22(29%)
% expected			25	25	25	25

B

Age	# of litters	# of pups	<i>Hoxa5^{flox/flox} X Hoxa5^{flox/+}; 14.5-kb Hoxa5/Cre</i>			
			Cre-		Cre+	
			<i>Hoxa5^{flox/+}</i>	<i>Hoxa5^{flox/flox}</i>	<i>Hoxa5^{flox/+}</i>	<i>Hoxa5^{flox/flox}</i>
E18.5	8	64	16(25%)	17(26%)	14(23%)	17(26%)
D30	24	162	30(20%)	38(25%)	51(34%)	19(13%) 14*(9%)
% expected			25	25	25	25

* Found dead at birth

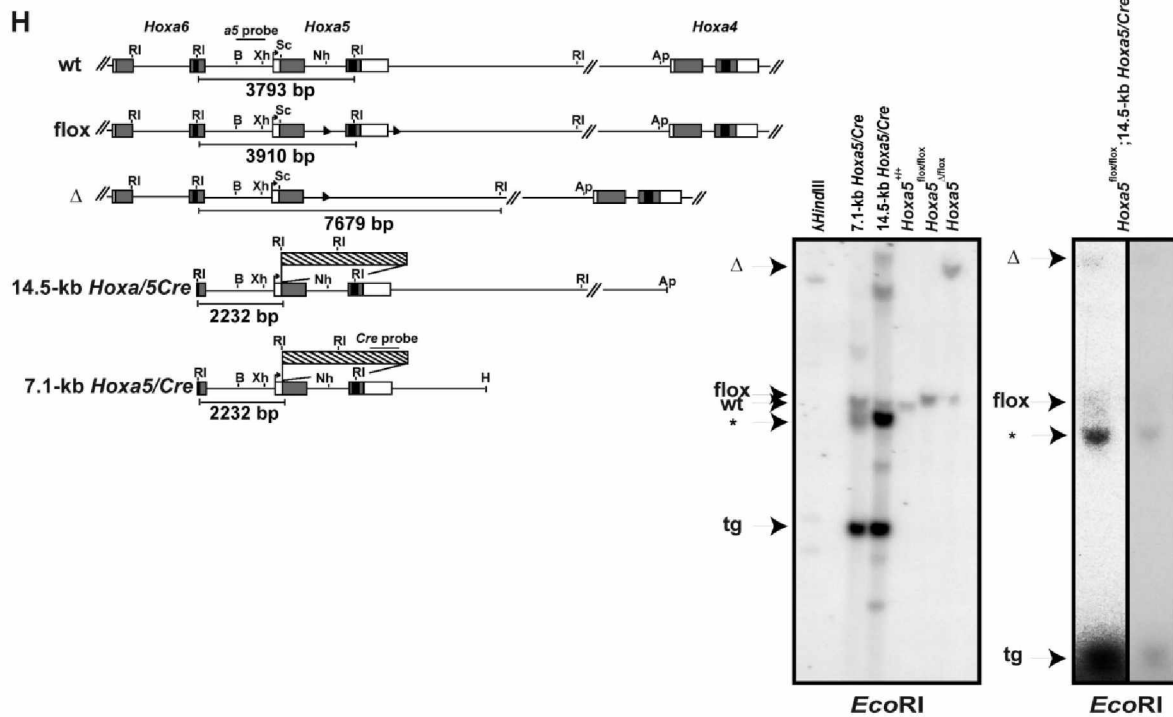
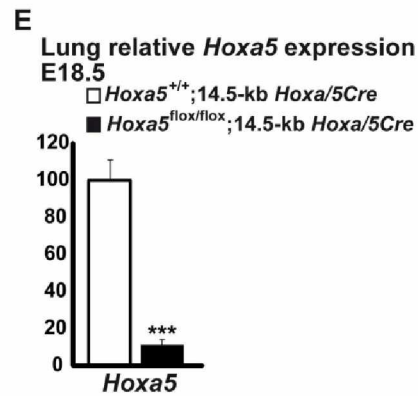
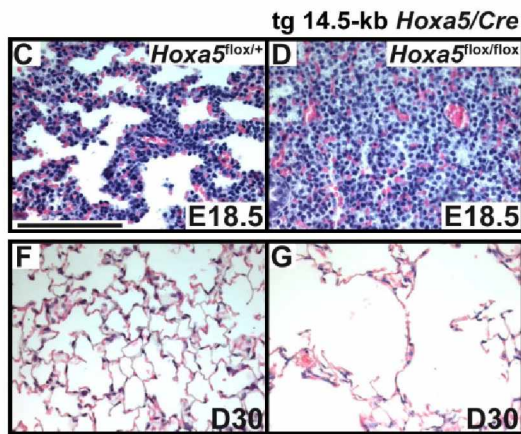
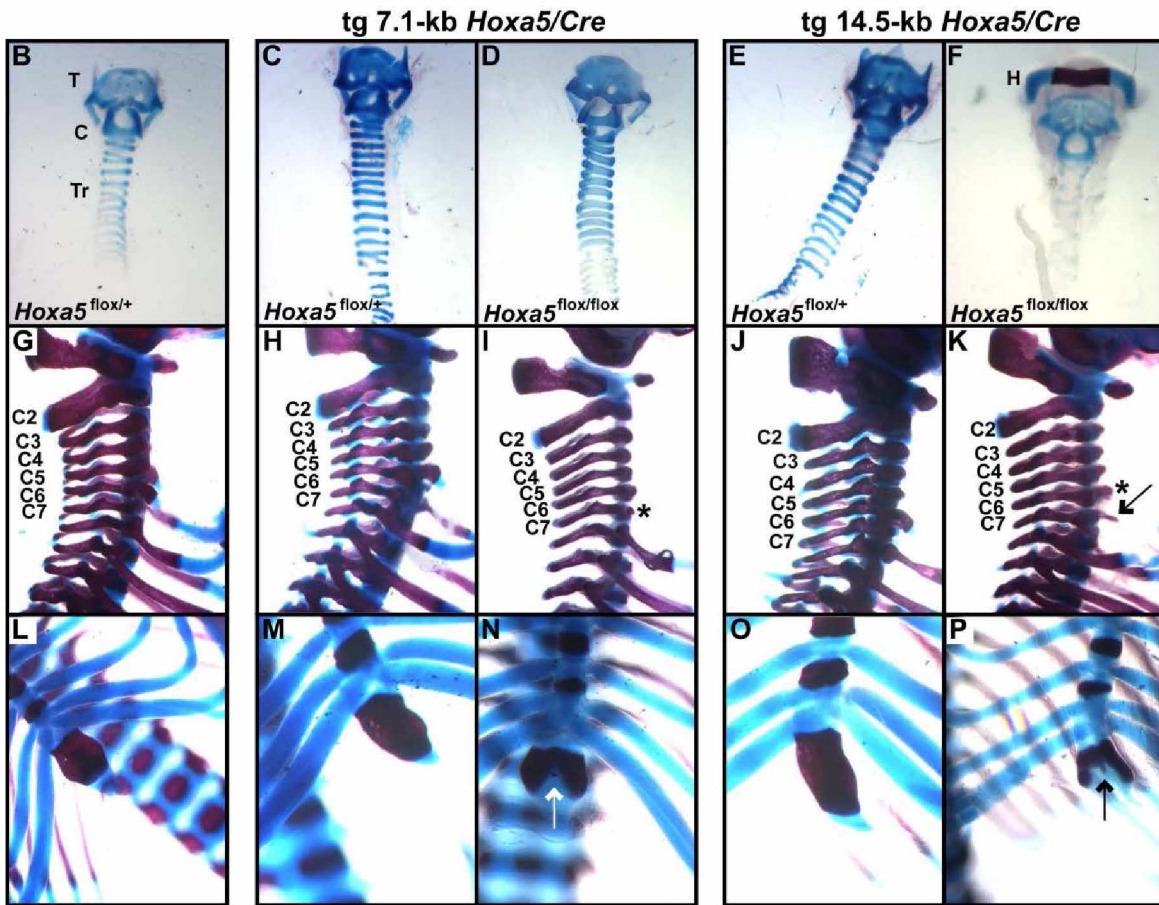


Figure 3.3. Comparative analysis of the skeletal morphology of *Hoxa5* conditional mutants. (A) List of skeletal transformations observed in E18.5 *Hoxa5*^{flox/+}, *Hoxa5*^{flox/+};7.1-kb *Hoxa5*/*Cre*, *Hoxa5*^{flox/flox};7.1-kb *Hoxa5*/*Cre*, *Hoxa5*^{flox/+};14.5-kb *Hoxa5*/*Cre* and *Hoxa5*^{flox/flox};14.5-kb *Hoxa5*/*Cre* embryos. (B-P) Representative specimens stained with alcian blue and alizarine red according to their genotype. (B-F) Ventral view of the trachea. C, Cricoid cartilage; H, Hyoid bone; T, Thyroid cartilage; Tr, Trachea cartilage. (G-K) Lateral view of the cervicothoracic transition region. Black arrow indicates ectopic rib on C7. Asterisks show absent or shorter tuberculum anterior on C6. The C2-C7 cervical vertebrae are identified. (L-P) Ventral view of the caudal part of the sternum region. Arrows indicated fissured xiphoid processes.

A

	<i>Hoxa5</i> ^{flox/+}	tg 7.1-kb <i>Hoxa5</i> /Cre		tg 14.5-kb <i>Hoxa5</i> /Cre	
		<i>Hoxa5</i> ^{flox/+}	<i>Hoxa5</i> ^{flox/flox}	<i>Hoxa5</i> ^{flox/+}	<i>Hoxa5</i> ^{flox/flox}
Trachea					
normal	10	10	10	9	2
abnormal	0	0	0	1	8
Tuber. ant. on C6					
present	10	9	1	9	2
absent/short	0	1	9	1	8
Ribs on C7					
absent	10	10	10	9	5
small	0	0	0	1	5
extensive	0	0	0	0	0
unilateral	0	0	0	1	4
bilateral	0	0	0	0	1
Xiphoid process					
normal	10	9	5	9	5
perforated	0	1	4	1	3
fissured	0	0	1	0	2
# of animals	10	10	10	10	10



laboratory of Dr. Copeland; Lee *et al.*, 2001), which contains the IRES-eGFPCre cassette, was inserted into the *SacI* site of the 5.0-kb *EcoRI-HindIII Hoxa5* genomic region, the latter encompassing the temporal (T) control region, the brachial spinal cord (BSC) enhancer, the proximal *Hoxa5* promoter, both exons and the intron of the *Hoxa5* gene. Insertion of the IRES-eGFPCre cassette into the first *Hoxa5* exon downstream the HOXA5 AUG allows translation of the eGFPCre ORF. To produce the 7.1-kb *Hoxa5/Cre* and the 14.5-kb *Hoxa5/Cre* transgenes, a 2.1-kb *HindIII-HindIII* fragment, containing the MES enhancer, and a 7.4-kb *HindIII-ApaI*, containing the MES and the L/G enhancers, were added at the 3' end of the *Hoxa5/Cre* basic construct, respectively.

To eliminate plasmid sequences, the microinjected *Hoxa5/Cre* fragments were obtained by a *SalI+NotI* digestion. Transgene preparation and microinjection into fertilized eggs from C57BL/6 X CBA F1 hybrid intercrosses were done according to standard methods (Hogan *et al.*, 1998). The established transgenic lines were maintained by intercrosses. Animal experimentations were performed according to the guidelines of the Canadian Council on Animal Care and approved by the institutional animal care committee. The 7.1-kb *Hoxa5/Cre* (Tg(*Hoxa5-cre*)444CLjea, MGI:5518946) and 14.5-kb *Hoxa5/Cre* (Tg(*Hoxa5-cre*)447BLjea, MGI:5518947) mouse lines will be available to the research community.

Mice, genotyping, tissue collection, histology β -galactosidase staining

The establishment of the *Hoxa5* conditional mutant mouse line was previously reported (*Hoxa5*^{tm1.1Ljea}, MGI:3723622; Tabariès *et al.*, 2007). The *R26R* mouse line was obtained from The Jackson Laboratory (Gt(*ROSA*)26Sor^{tm1Sor}, MGI:1861932; Soriano, 1999). Age of the embryos was estimated by considering the morning of the day of the vaginal plug as E0.5. Experimental specimens were genotyped by Southern blot analysis for the *Hoxa5* locus and *Hoxa5/Cre* transgenes as shown on Figure 2H and by PCR analysis for the *R26R* locus (Soriano, 1999). For the *Hoxa5*^{flox},*Hoxa5/Cre*⁺ specimens, purified genomic DNA was digested with *EcoRI* and the blots were hybridized to a 606- bp *BgIII-XhoI* ³²P-labeled genomic probe located in the *Hoxa5-Hoxa6* intergenic region (Fig. 2H; Coulombe *et al.*, 2010). Organs were collected from E18.5 and D30 *R26R,Hoxa5/Cre*⁺ and *Hoxa5*^{flox},*Hoxa5/Cre*⁺ specimens. For RNA extraction, tissues were snap-frozen in N₂. For

histology, tissues were fixed in 4% ice-cold paraformaldehyde (PFA), paraffin-embedded and sectioned at 4 μ m (Jaffe *et al.*, 1990). For adult mice, lungs were fixed by tracheal instillation of cold fixative until fully inflated under direct observation. Organ sections were stained according to standard hematoxylin and eosin procedure for morphology analysis. For whole-mount detection of β -galactosidase activity, E8.25 to E13.5 entire embryos were fixed in ice-cold 0.25% glutaraldehyde in PBS for 15 min to one hr according to the age of the embryo. Detection of β -galactosidase activity was performed as described (Larochelle *et al.*, 1999). In parallel, E13.5 embryos and D30 organs were fixed one hr in ice-cold 4% PFA, 0.2% glutaraldehyde in PBS, put overnight in a 30% sucrose solution in 0.1M phosphate buffer pH8.0, and embedded in tissue freezing medium (Triangle Biomedical Sciences, Inc.). Ten μ m cryosections were processed for *lacZ* expression and counterstained with nuclear Fast Red (Boucherat *et al.*, 2012).

Skeletal analysis

Whole-mount skeletons were prepared from E18.5 *Hoxa5^{fllox};Hoxa5/Cre⁺* compound mutants with alcian blue for staining the cartilage and alizarin red for staining the bone. Briefly, embryos were skinned, eviscerated and fixed overnight at 4°C in EtOH 95%. The next day, specimens were put into an alcian blue staining solution (alcian blue 0.015% in acetic acid 20% prepared in EtOH 95%) for 7 days at 37°C. Skeletons were then rinsed one hr in EtOH 95%, clarified 6 hrs in KOH 2% and stained overnight at room temperature in a solution of alizarin red 0.0075% prepared in KOH 1%. Specimens were transferred in KOH 1% for 7 days, and then kept at room temperature in a solution of glycerol/EtOH 1:1 until analysis. Skeletons were observed, and left and right sides of each vertebra were scored independently for bilateral markers.

Quantitative RT-PCR (qRT-PCR) experiments

Total RNA was isolated from lungs of E18.5 embryos. RNA extraction, cDNA synthesis and qRT-PCR were performed as described (Boucherat *et al.* 2012). Samples were analyzed in triplicate. The *Rpl19* gene was used as control. Five specimens were used for each genotype tested. The primer sequences are: *Hoxa5*-F:

CCCAGATCTACCCTGGATG, *Hoxa5*-R: GGCATGAGCTATTTTCGATCCT, *Rpl19*-F: GATCATCCGCAAGCCTGTGA and *Rpl19*-R: GCATCCGAGCATTGGCAGTA.

Statistical analyses

Student's *t* test was performed for gene expression studies. A significance level inferior to 5% ($p < 0.05$) was considered statistically significant.

ACKNOWLEDGEMENTS

This work was supported by grants from the Canadian Institutes of Health Research (MOP-15139; to L.J.) and the National Sciences and Engineering Research Council of Canada (194559; to L.J.). We thank Drs. Jean Charron and Olivier Boucherat for their scientific input and comments on the manuscript, Dr. Neil Copeland for sharing reagent, and Marcelle Carter for technical assistance.

REFERENCES

- Aubin J, Lemieux M, Tremblay M, Bérard J, Jeannotte L. 1997. Early postnatal lethality in *Hoxa-5* mutant mice is attributable to respiratory tract defects. *Dev Biol* 192: 432-445.
- Aubin J, Chailier P, Ménard D, Jeannotte L. 1999. Loss of *Hoxa5* gene function in mice perturbs intestinal maturation. *Am J Physiol* 277: C965-C973.
- Aubin J, Déry U, Lemieux M, Chailier P, Jeannotte L. 2002. Stomach regional specification requires *Hoxa5*-driven mesenchymal-epithelial signaling. *Development* 129: 4075-4087.
- Boucherat O, Chakir J, Jeannotte L. 2012. The loss of *Hoxa5* function in mice promotes NOTCH-dependent goblet cell metaplasia in lung airways. *Biology Open* 1: 677-691.
- Boucherat O, Montaron S, Bérubé-Simard FA, Aubin J, Philippidou P, Wellik DM, Dasen JS, Jeannotte L. 2013. Partial functional redundancy between *Hoxa5* and *Hoxb5* paralog genes during lung morphogenesis. *Am J Physiol Lung Cell Mol Physiol* 304: L817-L830.
- Chen H, Zhang H, Lee J, Liang X, Wu X, Zhu T, Lo PK, Zhang X, Sukumar S. 2007. HOXA5 acts directly downstream of retinoic acid receptor β and contributes to retinoic acid-induced apoptosis and growth inhibition. *Cancer Res* 67: 8007-8013.

- Coulombe Y, Lemieux M, Moreau J, Aubin J, Joksimovic M, Bérubé-Simard FA, Tabariès S, Boucherat O, Guillou F, Larochelle C, Tuggle CK, Jeannotte L. 2010. Multiple promoters and alternative splicing: *Hoxa5* transcriptional complexity in the mouse embryo. PLoS ONE 5(5): e10600. doi:10.1371/journal.pone.0010600.
- Di Meglio T, Kratochwil CF, Vilain N, Loche A, Vitobello A, Yonehara K, Hrycaj SM, Roska b, Peters AHFM, Eichmann A, Wellik D, Ducret S, Rijli FM. 2013. *Ezh2* orchestrates topographic migration and connectivity of mouse precerebellar neurons. Science 339: 204-207.
- Garin E, Lemieux M, Coulombe Y, Robinson GW, Jeannotte L. 2006. Stromal *Hoxa5* function controls the growth and differentiation of mammary alveolar epithelium. Dev Dyn. 235: 1858-1871.
- Gendronneau G, Boucherat O, Aubin J, Lemieux M, Jeannotte L. 2012. The loss of *Hoxa5* function causes estrous acyclicity and ovarian epithelila inclusion cysts. Endocrinology 153: 1484-1497.
- Hogan B, Beddington RS, Constantini F, Lacy E. 1994. Manipulating the mouse embryo: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Jaffe L, Jeannotte L, Bikoff EK, Robertson EJ. 1990. Analysis of beta 2-microglobulin gene expression in the developing mouse embryo and placenta. J Immunol 145: 3474-3482.
- Jeannotte L, Lemieux M, Charron J, Poirier F, Robertson EJ. 1993. Specification of axial identity in the mouse: role of the *Hoxa-5 (Hox1.3)* gene. Genes Dev 7: 2085-2096.
- Larochelle C, Tremblay M, Bernier D, Aubin J, Jeannotte L. 1999. Multiple *cis*-acting regulatory regions are required for the restricted spatio-temporal *Hoxa5* expression. Dev Dyn 214: 127-140.
- Lee EC, Yu D, Martinez de Velasco J, Tessarollo L, Swing DA, Court DL, Jenkins NA, Copeland NG. 2001. A highly efficient Escherichia coli-based chromosome engineering system adapted for recombinogenic targeting and subcloning of BAC DNA. Genomics. 73: 56-65.
- Mandeville I, Aubin J, LeBlanc M, Lalancette-Hébert M, Janelle MF, Tremblay GM, Jeannotte L. 2006. Impact of the loss of *Hoxa5* function on lung alveogenesis. Am J Pathol 169: 1312-1327.
- Moreau J, Jeannotte L. 2002. Sequence analysis of *Hoxa4-Hoxa5* intergenic region including shared regulatory elements. DNA Sequence 13: 203-209.

Nagy A, Mar L, Watts G. 2009. Creation and use of a Cre recombinase transgenic database. *Methods Mol Biol* 530: 365-378.

Salehi S, Eckley L, Sawyer GJ, Zhang X, Dong X, Freund J-N, Fabre JW. 2009. Intestinal lactase as an autologous β -galactosidase reporter gene for *in vivo* gene expression studies. *Hum Gene Ther* 20: 21-30.

Soriano P. 1999. Generalized lacZ expression with the ROSA26 Cre reporter strain. *Nat Genet* 21: 70-71.

Tabariès S, Lemieux M, Aubin J, Jeannotte L. 2007. Comparative analysis of *Hoxa5* allelic series. *genesis* 45: 218-228.

Zakany J, Tuggle CK, Patel M, Nguyen-Huu MC. 1988. Spatial regulation of homeobox gene fusions in the embryonic central nervous system of transgenic mice. *Neuron* 1: 679-691.

Chapitre IV

***Hoxa5* regional expression along the developing anteroposterior axis requires several regulatory elements**

Félix-Antoine Bérubé-Simard[#], Sébastien Tabariès[#] and Lucie Jeannotte

Department of Molecular Biology, Medical Biochemistry and Pathology, Université Laval, Québec, Canada; Centre de recherche sur le cancer de l'Université Laval; Centre de recherche du Centre Hospitalier Universitaire de Québec, L'Hôtel-Dieu de Québec, Québec, Canada

[#]These authors contributed equally to this work.

Manuscrit en préparation

AVANT-PROPOS

Chez les vertébrés, le concept de prévalence postérieure a été émis pour expliquer le fait que lorsqu'une protéine, encodée par un gène exprimé plus postérieurement le long de l'axe AP, est exprimée ectopiquement dans des domaines plus antérieurs, ces domaines se trouvent transformés et présentent des caractéristiques plus postérieures. Cependant, quand des gènes normalement exprimés dans des régions antérieures sont exprimés dans des domaines plus postérieurs, aucune transformation d'antériorisation n'est observée au niveau du domaine d'expression ectopique (Duboule et Morata, 1994). Aucune démonstration expérimentale de cette prévalence postérieure n'a été rapportée.

Dans ce chapitre, nous poursuivons la caractérisation de l'élément MES initiée par Christian Larochelle et décrite dans la thèse de doctorat de Sébastien Tabariès (Tabariès *et al.*, 2005). Une approche par délétions successives dans un contexte de transgénèse entreprise par Sébastien Tabariès a démontré que la perte des sites de liaison aux protéines CDX n'est pas en mesure de reproduire l'effet de la délétion du fragment d'ADN *AvrII-Eco47III* de 164 pb (AvEc 164-bp), suggérant ainsi l'implication d'autres facteurs au sein de cette séquence. L'analyse bioinformatique (TRANSFAC) menée par Sébastien Tabariès a mis en évidence un site potentiel de liaison pour les protéines HOX. Il a ensuite confirmé, par essais de retard sur gel (EMSA) réalisés avec des protéines HOX de souris produites *in vitro*, que des protéines HOX orthologues à la protéine AbdB de drosophile sont en mesure de se lier sur la séquence d'ADN Oligo1 de 25 pb comprise à l'intérieur du fragment AvEc 164-bp. Or, ces protéines sont produites par des gènes exprimés dans des parties postérieures de l'embryon soulevant ainsi la possibilité d'un mécanisme de répression de l'expression d'un gène *Hox* antérieur par le produit des gènes *Hox* plus postérieurs, appuyant ainsi la notion de prévalence postérieure (Duboule et Morata, 1994). Par ailleurs, ces mêmes expériences menées en utilisant des extraits protéiques embryonnaires au stade E9.5 ont permis d'identifier la protéine HOXB9 comme candidate potentielle.

De mon côté, j'ai confirmé que la liaison de la protéine HOXB9, tout comme celle de CDX4, s'effectue *in vitro* via les sites de liaison CDX situés sur la séquence d'ADN Oligo1 et non sur le site potentiel de liaison aux protéines HOX (Fig 4.3C; Fig. 4.5A-B). De la même façon, j'ai démontré, par immunoprécipitation de la chromatine (ChIP), que

les protéines CDX (préférentiellement CDX4) et HOXB9 se fixent sur le fragment AvEc 164-bp (Fig. 4.4C) Toutefois, des expériences de co-immunoprécipitation de protéines (CoIP) que j'ai réalisées montrent que ces deux protéines ne semblent pas interagir physiquement (Fig. 4.5C) , suggérant qu'elles contrôlent indépendamment l'activité répressive du fragment AvEc 164-bp sur l'expression du transgène *Hoxa5/LacZ* dans la partie postérieure à pv10.

Parallèlement, Sébastien Tabariès a également démontré par transgénèse qu'un fragment d'ADN *PstI-HindIII* de 331 pb (PsHi 331-bp), localisé à l'extrémité 3' de l'élément MES, est également important pour positionner la frontière postérieure d'expression du transgène *Hoxa5/lacZ* au niveau de pv 10, et ce potentiellement avec la participation d'un motif de 16 pb identifié par EMSA.

La rédaction du manuscrit a été réalisée conjointement avec Sébastien Tabariès ainsi que Dre Jeannotte.

ABSTRACT

Hox gene functions are intimately linked to the correct developmental expression of the genes. In the laboratory, we are studying the *Hoxa5* gene, which plays a crucial role in axial patterning and organogenesis. The MES, a mesodermal enhancer sequence located downstream of the *Hoxa5* coding sequences targets expression in mesodermal derivatives at the cervico-upper thoracic level. It contains at least two regulatory sequences, identified as the AvEc 164-bp and the PsHi-331 bp fragments. The AvEc-164 bp sequence is necessary for the correct *Hoxa5* expression between prevertebra 3 and 10. In the present study, we demonstrate that the HOXB9 protein specifically interacts with the AvEc-164 bp sequence supporting the concept of posterior prevalence. Moreover, we show that HOXB9 protein binds the same 25-bp DNA fragment previously identified for its interaction with CDX proteins. In parallel, the PsHi 331-bp DNA fragment also participates in restricting *Hoxa5* expression at cervico-upper thoracic level via the binding of a homeodomain-containing protein to a 16-bp motif. Our analysis demonstrates that coordination between several *cis*-acting regulatory elements is needed to fully reproduce the correct spatial *Hoxa5* gene profile.

INTRODUCTION

The *Hox* genes, an evolutionarily conserved gene family first described in *Drosophila*, are master regulators involved in specifying positional information along the anterior-posterior (AP) axis of diverse organisms (Lewis, 1978; Krumlauf, 1994). In vertebrates, these homeodomain-containing transcription factors are organized in four clusters, *HoxA* to *HoxD* (Krumlauf, 1994). The genomic organization of the *Hox* genes has been conserved throughout evolution and is fundamental for the precise spatio-temporal regulation and function of each gene (Garcia-Fernandez, 2005). It reflects the colinear relationship existing between the position occupied by each *Hox* gene on the chromosome, the expression domain along the embryonic axes, and the temporal onset: the *Hox* genes located at the 3' extremity of the clusters are expressed earlier and in more anterior regions of the embryo than the genes located at the 5' end of the clusters (Duboule and Morata, 1994; Kmita and Duboule, 2003; Wacker et al. 2004). As a result, *Hox* gene expression domains are both spatially restricted and overlapping, extending from specific anterior boundaries down to the caudal end of the embryo. Axial patterning was proposed to be established by the combination of *Hox* genes expressed at a specific axial level, referred to the '*Hox* code' (Kessel and Gruss, 1991). However, mutational analyses of *Hox* genes in mice have shown that their domains of action are mainly restricted to the anterior-most region of their expression domain (Stein et al., 1996; Coulombe *et al.*, 2010), suggesting that at a given axial level, the information was mainly provided by the most posterior acting *Hox* gene, a phenomenon termed 'posterior prevalence' (Duboule, 1991; Bachiller et al., 1994; Duboule and Morata, 1994). In vertebrates, this concept was put forward to explain the lack of phenotype when *Hox* genes are ectopically expressed in regions posterior to their endogenous expression domain (Duboule and Morata, 1994). How posterior prevalence is achieved remains poorly understood. Possible explanations include that posterior *Hox* genes would determine posterior body patterns by repressing target genes activated by more anterior *Hox* genes or by binding more efficiently to shared sets of target genes (Lamka et al., 1992; Duboule and Morata, 1994; Capovilla and Botas, 1998; Caronia et al., 2003; Williams et al., 2005; Williams et al., 2006). Other possibilities imply negative regulation mechanisms by non-coding RNAs. For example, *miR-10a/miR-10b*, respectively transcribed from *Hoxb4-Hoxb5* and *Hoxd4-Hoxd5* intergenic regions,

potentially target mRNA of the *Hox1-3* paralogous, whereas *miR-196b/miR-196a1/miR196a2*, respectively transcribed from *Hoxa9-Hoxa10*, *Hoxb9-Hoxb13* and *Hoxc9-Hoxc10* intergenic regions, would disrupt *Hox4-Hox9* mRNA (Yekta et al., 2004, 2008; Woltering and Durston, 2008). Finally, as lncRNAs transcription from both orientation are now associated with genes regulation and since transcripts that form complementary pairs are expressed in reciprocal profiles, it is possible that both sense and anti-sense lncRNAs transcribed from the Hox Clusters participate in posterior prevalence (Mainguy et al., 2007; Coulombe et al., 2010; Li et al., 2013)

A better comprehension of the controls governing Hox gene expression is needed to decipher the mechanistic basis responsible for Hox gene function in axial patterning. Knowledge of the regulatory mechanisms involved in mammalian Hox gene expression mainly comes from transgenic mice studies. Hox gene regulation is thought to be achieved by a combination of strategies involving: (i) chromatin remodeling (Chambeyron and Bickmore, 2004; Wurtele and Cartrand, 2006; Morey et al., 2007; Noordermeer et al., 2011; Rousseau et al., 2014); (ii) global enhancer sequences located outside the cluster and acting in a relatively promoter-unspecific manner (Zákány et al., 2004; Di-Pöi et al., 2007; Tschopp et al., 2012; Andrey et al., 2013); and (iii) integrated regulation of neighbouring genes through the sharing, the competition and/or the selective use of defined local cis-acting sequences (Kmita and Duboule, 2003; Peljto et al., 2010; Kashyap et al., 2011; Mahony et al., 2011; Mazzone et al., 2013; Ahn et al., 2014). The endogenous expression pattern of some Hox genes can be recapitulated partly or entirely upon random integration in the genome of transgenes containing their DNA regulatory elements (Oosterveen et al., 2003; Gonzalez et al., 2007; Rouso et al., 2008; Tschopp et al., 2009; Montavon et al., 2011; Bérubé-Simard and Jeannotte, 2014). The common emerging picture is that the expression pattern of individual *Hox* gene is generated by the positional information transduced by transcriptional factors interacting with a combination of both positive and negative *cis*-acting sequences to restrict the enhancer activities to a defined domain with specific expression boundaries (Morrison et al., 1997; Gilthorpe *et al.*, 2002; Pearson et al., 2005; Kobrossy *et al.*, 2006; Lengerke et al., 2008). Among the upstream factors identified so far, the *caudal*-related family members have emerged as key regulators since CDX proteins are able to convey positional information to *Hox* genes (van den Akker

et al., 2002; Houle et al., 2003; Lohnes, 2003; Gaunt et al., 2005; Deschamps and van Nes, 2005; Tabariès et al., 2005; Davidson and Zon, 2006; van de Ven et al., 2011). HOX proteins themselves have also been reported as regulators of their own expression or that of other Hox genes (Manzanares et al., 2001; Lampe et al., 2004; Ferretti et al., 2005; Wong et al., 2011; Sheth et al., 2013).

To fully understand the regulatory events governing *Hox* gene expression, we are examining the *Hoxa5* gene. This gene provides a good model because of its central location in the *HoxA* cluster and its crucial role during embryogenesis as its loss-of-function mutation affects a well-defined subset of structures mainly located at the cervico-thoracic level (Coulombe et al., 2010). In the developing embryo, *Hoxa5* is expressed in the neural tube caudal to the posterior myelencephalon, in the axial skeleton up to the level of prevertebra (pv) 3 and in the mesenchymal component of several organs, including the trachea, the lung, the stomach, the intestine, the kidneys, the mammary glands and the ovaries (Dony and Gruss, 1987; Gaunt et al., 1990; Aubin et al., 1997; Aubin et al., 1998; Aubin et al., 1999; Aubin et al., 2002a, Garin et al., 2006; Gendronneau et al., 2012). Aside from morphological defects in foregut derivatives, mammary glands and ovaries, (Aubin et al., 1997; Aubin et al., 2002a; Meunier et al., 2003; Garin et al., 2006; Gendronneau et al., 2012), the targeted disruption of the *Hoxa5* gene perturbs axial skeleton identity in a region confined to the anterior-most region of the *Hoxa5* domain of expression along the prevertebral axis between pv3 and pv10 (Jeannotte et al., 1993; Aubin et al., 1998; Tabariès et al., 2007; Bérubé-Simard and Jeannotte, 2014). Four transcripts of 1.8, 5.0, 9.5 and 11 kb in length encompassing *Hoxa5* coding sequences originate from the use of different promoters and alternative splicing (Coulombe et al, 2010).

The presence of complex and overlapping transcriptional units at the *Hoxa5* locus implies dispersed and shared regulatory regions in the cluster. Moreover, *Hox* gene function is intimately linked to its correct developmental regulation as illustrated by the homeotic transformations observed when *Hox cis*-acting regulatory elements are mutated (Dupé et al., 1997; Zákány et al., 1997). This prompted us to characterize the regulatory elements directing *Hoxa5* developmental expression. Using a transgenic approach, we have demonstrated that an 14.5-kb genomic fragment, located between positions -3.8 kb and

+12.4 kb, can recapitulate the temporal expression and substantially reproduce the spatial pattern of the *Hoxa5* 1.8 kb transcript (Bérubé-Simard and Jeannotte, 2014), which is the only protein-coding *Hoxa5* transcript (Coulombe et al., 2010). Deletion analyses have revealed several *cis*-acting DNA control elements, including a 604-bp brachial spinal cord enhancer (BSC) in 5' flanking sequences (Tuggle et al., 1990; Nowling et al., 1999), a 2.08-kb mesodermal enhancer sequence (MES) located downstream of *Hoxa5* coding sequences, which is essential for paraxial and lateral plate mesoderm expression in the cervical and upper thoracic region (Larochelle et al., 1999), and a 1.5-kb lung/gut element responsible for respiratory and digestive tracts expression (Moreau et Jeannotte, 2002; Bérubé-simard et al., 2014) The MES contains DNA elements that limit the *Hoxa5* regional specific expression domain along the AP axis, including a 164-bp DNA sequence (AvEc 164-bp) that binds CDX proteins to correctly position the *Hoxa5* expression domain (Tabariès et al., 2005). In pursuing the characterization of the MES sequence, we now present evidence that HOXB9 participates to caudally restrict *Hoxa5* expression by binding the AvEc-164 bp sequence. In addition, we demonstrate that a 331 bp *PstI-HindIII* fragment is also contributing to the MES activity for the establishment of the specific *Hoxa5* expression domain along the AP axis.

MATERIALS AND METHODS

Constructions of *Hoxa5/lacZ* transgenes

All the constructs were derived from construct 9 (pLJ133) described in Larochelle et al. (1999). Briefly, pLJ133 contains the bacterial *lacZ* gene inserted into the *SacI* site of the first exon of the *Hoxa5* gene, allowing translation of the *lacZ* open reading frame from the *Hoxa5* AUG. This construct was digested with *HindIII* to remove a 2.08 kb fragment from +2.85 kb to +4.93 kb and self-ligated to produce an intermediate construct (pLJ143), containing a continuous *XhoI-HindIII Hoxa5* genomic fragment extending from -235 bp to +2.85 kb that does not direct β -galactosidase expression (Zákány et al., 1988; Tabariès et al., 2005). Insertion of the *PstI-HindIII* fragment referred to as PsHi-331 bp (from +4.60 kb to +4.93 kb) at the *HindIII* site (+2.85 kb) of pLJ143 resulted in construct 1. Deletion of the last 104-bp (from +4.83 kb to +4.93 kb) located in the 3' part of the PsHi-331 bp

fragment from construct 1 resulted in construct 2. Deletion of the the 16-bp sequence (described in Fig.2) from construct 1 and adding the *AccI-XhoI* fragment (from -1480 bp to -235 bp) at the *XhoI* site (-235 bp) resulted in construct 3. This construct contained the BSC element previously reported by Tuggle et al. (1990). Construct 4 was derived from construct 8 in Tabariès et al. (2005) by deleting the PsHi 331-bp fragment and adding the *AccI-XhoI* fragment to its 5' end. Construct 5 was obtained by deleting the Oligo1 sequence from pLJ133 via a PCR strategy using synthetic oligonucleotide primers (Table 1) containing the appropriate deletion (Wang and Malcolm, 1999). Sequencing subsequently validated all introduced mutations. Nucleotide substitutions in the identified CDX binding sites used in Tabariès et al. (2005) were introduced into the 2.08 kb *HindIII* fragment extending from +2.85 kb to +4.93 kb. The plasmids containing the desired mutations were digested with *PstI* to remove the PsHi-331 bp DNA fragment and self-ligated. The resulting plasmid was digested with *HindIII*, and the fragment was subcloned into the *HindIII* site (+2.85 kb) of pLJ143 to obtain construct 6. Finally, construct 7 was derived from pLJ133 by introducing mutations in Cdx binding sites and deleting the 16 bp sequence. All the constructs were made in pBluescript SKII+ (Stratagene) and purified following a cesium chloride centrifugation.

Production and genotyping of transgenic mice

All experiments were performed according to the guidelines of the Canadian Council on Animal Care and approved by the institutional animal care committee. The *Hoxa5/lacZ* sequences from all constructs were isolated using a *KpnI-NotI* digestion to remove vector sequences and they were purified on agarose gel. They were then injected into the pronuclei of fertilized eggs derived from (C57BL/6 x CBA) F1 hybrid intercrosses following standard procedures (Hogan et al., 1994). Transgenic founder embryos were recovered from foster mothers at day 12.5 of gestation (E12.5), genotyped by Southern analysis of yolk sac DNA using a *lacZ* specific probe to verify the integrity of the microinjected construct, and analyzed for *lacZ* expression by β -galactosidase staining as described (Larochelle et al., 1999).

Electrophoretic mobility shift assays

The expression vectors PSG5-*Cdx1*, pSG5-*Cdx2*, pSG5-*Cdx4*, pCMV-*Hoxb4*, pBluescript-*Hoxd4*, pSG5-*Hoxa5*, pET28a-*Hoxa7*, pET28b-*Hoxb8*, pET28b-*Hoxa9*, pET28b-*Hoxb9*, pSG5-*Hoxd9*, pSG5-*Hoxd10*, pET28a-*Hoxa11*, pSG5-*Hoxd11*, pET28a-*Hoxd12*, pSG5-*Pbx1a*, pSG5-*Pbx1b*, pET28b-*Meis1a*, pET28b-*Meis1b* were used with the TnT7 Quick coupled transcription/translation system (Promega Corporation) to produce *in vitro* proteins that were analyzed by SDS-PAGE gel electrophoresis and directly assayed for electrophoretic mobility shift assay (EMSA) as described (Tabariès et al., 2005). The AvEc 164-bp, AvEc 164-bp Δ (HoxBS) and AvEc164-bp Δ (Oligo1) fragments were obtained by PCR reaction from plasmids carrying the MES sequence containing the appropriate deletion (pLJ122, pLJ409 and pLJ367, respectively) (Tabariès et al., 2005) with the Sc33 and Sc35 primers (Table 4.1) end-labelled with T4 polynucleotide kinase and used in the EMSA reactions. The Oligo1 sequence contained in the AvEc 164-bp fragment was also used as a double-stranded end labelled oligonucleotide probe (Tabariès et al., 2005).

Various EMSA probes were obtained by PCR from pLJ122 using end-labelled primers (Table 4.1). They include (i) the PsHi-331 bp fragment; (ii) the 5' half of PsHi 331-bp fragment; (iii) the 3' half of PsHi 331-bp fragment; (iv) the first third (PsHiA) of PsHi 331-bp; (v) the second third (PsHiB) of PsHi 331-bp (vi) the last third (PsHiC) of PsHi-331 bp; and (vii) the overlapping 67-bp sequence located in the PsHiB fragment. Double-stranded end-labelled oligonucleotides covering (i) the 5' half of the 67 bp (Oligo 34-bp, and (ii) the 3' half of the 67 bp sequence (Oligo 33-bp) encompassing protein binding site were also tested (Table 4.1). We focused on the Oligo 33-bp sequence that showed protein binding activity and EMSAs were performed with double-stranded end labelled oligonucleotides harbouring various mutations of the Oligo 33-bp, named Oligo 33-bp (mut1) to Oligo 33-bp (mut4), as well as Oligo 16-bp (Table 4.1).

These EMSA probes were tested for protein binding using E9.5 whole cell extracts prepared as described in Charité et al. (1998). Binding reactions were done as described (Tabariès et al., 2005). The specificity of protein binding was assessed by addition of 100-fold excess of unlabeled probe prior to that of the radiolabeled probe. A 23-bp doubled-

stranded DNA fragment of the sucrase-isomaltase promoter, SIF1 (Table 4.1), that contains two sequences interacting with the CDX proteins, was used as an unrelated competitor (Suh et al., 1994).

Supershift assays were done as previously described (Tabariès et al., 2005) using rabbit anti-HOXA9 (Millipore), rabbit anti-HOXB9 (Imagex), goat anti-HOXD10 (Santa Cruz Biotechnology) or rabbit anti-HOXD12 antibodies (Aviva).

Table 4.1: Oligonucleotide sequences

Fragment	Sequence (5'-3')	Amplicon (bp)
A) Mutagenesis		
AvEc 164-bp (Δ Oligo1)	F-GTTCTCACCAGCCTGTGAGTCTTTAACTTTAGTCTTTGAATC R-GATTCAAAGACTAAAGTTAAAGACTCACAGGCTGGTGAGAAC	42
B) EMSA on the 331-bp <i>Pst</i>I-<i>Hind</i>III (PsHi 331-bp) DNA fragment		
PsHi 331-bp	F-CCAAGTGCAGACAAATTCCTC R-ATATGCGGCCGCAAGCTTGGCTTAAGCACATGG	331
PsHi 5' half	F-CCAAGTGCAGACAAATTCCTC R-ATATGCGGCCGCAATTCCTACTGACAAAGGGG	160
PsHi 3' half	F-ATATCTGCAGTCCTCTTCTCTTTGTTATTTGTT R-ATATGCGGCCGCAAGCTTGGCTTAAGCACATGG	171
PsHiA	F-CCAAGTGCAGACAAATTCCTC R-ATATAAGCTTATGTGTCAAGAAAAAC	115
PsHiB	F-CACAAAGCTTCATAAAGGGTACTTTA R-ATATAAGCTTGGAAAAACATAGATTATG	112
PsHiC	F-CACAAAGCTTAAAAACAGTCTCCAAGATT R-ATATGCGGCCGCAAGCTTGGCTTAAGCACATGG	104
67-bp	F-ATATCTGCAGTCCTCTTCTCTTTGTTATTTGTT R-ATATAAGCTTGGAAAAACATAGATTATG	67
Oligo 34-bp	F-TCCTCTTCTCTTTGTTATTTGTTCTTTCTGTTT	34
Oligo 33-bp	F-GATCCCAGGAAATTTACATAATCTATGTTTTCC	33
Oligo 33-bp (mut1)	F- <u>ACCTGTGTG</u> AAATTTACATAATCTATGTTTTCC	33
Oligo 33-bp (mut2)	F-GATCCCAG <u>ATTCGGGC</u> ATAATCTATGTTTTCC	33
Oligo 33-bp (mut3)	F-GATCCCAGGAAATTTAAG <u>CCGCTC</u> TGTTTTCC	33
Oligo 33-bp (mut4)	F-GATCCCAGGAAATTTACATAATCT <u>GCACCGGTA</u>	33
Oligo 16-bp	F-GAAATTTACATAATCT	16
SIF1	F-GTGCAATAAACTTTATGAGTAA	23
C) EMSA on the 164-bp <i>Avr</i>II-<i>Eco</i>47III (AvEc 164-bp) DNA fragment		
AvEc 164-bp	F-CCTAGGAGGTACTCTGCTGT R-AGCGCTTTAACCCTTTCAA	164
Oligo1	F-TGAAGCAATAAACTTTATGGCCGGA	25
Oligo1 (mutCDXBS)	F-TGAAGCAAT <u>GCCCTTGCGGGCCGGA</u>	25
D) ChIP		
AvEc 164-bp	F-CTAGGCCCAACCCTGTTAGTTCTT R-CCCTTTCAATTACCCTGAGGATTC	127

Nucleotide substitutions are underlined.

Chromatin immunoprecipitation (ChIP) assays

Head, trunk and tail from E10.5 wild-type embryos were collected separately and

mechanically disrupted to produce a single cell suspension prior cross-linking with 1% formaldehyde in PBS for 15 min at room temperature. Cross-linking was stopped by adding glycine to a final concentration of 0.125M. Extracts were then disrupted into crude lysates in 2ml of swelling buffer (5mM Pipes pH8.0, 85mM KCl, 1% NP-40 and protease inhibitors), equilibrated 20 min on ice, and centrifuged for 5 min at 3000 rpm at 4°C. The pellets were eluted in 1ml of nuclear lysis buffer (50mM Tris-HCl pH8.0, 10mM EDTA, 1% SDS) and protease inhibitors (Complete Mini-EDTA-free; Roche Diagnostics) and submitted to two successive rounds of sonication using a Bioruptor (Diagenode) for a total of 10 cycles of a 30-sec pulse interrupted by a 30-sec pause at the highest setting in order to obtain an average DNA size of 300-600-bp. One hundred µg of fragmented chromatin was incubated overnight at 4°C with Dynabeads linked to protein G (Invitrogen) and 2µg of rabbit anti-CDX1 and anti-CDX2 antibodies (provided by Dr D. Lohnes), mouse anti-CDX4 (Abnova), rabbit anti-HOXA9, rabbit anti-HOXB9, goat anti-HOXD10, rabbit anti-HOXD12 antibodies, rabbit anti-H3 antibody (Abcam), or control rabbit IgG (Santa Cruz Biotechnology). Immunocomplexes were washed three times respectively with 0.5ml of low salt immune complex wash buffer (0.1% SDS, 1% Triton X-100, 2mM EDTA, 20mM Tris-HCl pH8.0, 150mM NaCl), high salt immune complex wash buffer (0.1% SDS, 1% Triton X-100, 2mM EDTA, 20mM Tris-HCl pH8.0, 500mM NaCl) and LiCl immune complex wash buffer (0.25M LiCl, 1% IGEPAL-CA630, 1% deoxycholic acid (sodium salt), 1mM EDTA, 10mM Tris-HCl pH8.0). A final wash in 0.5ml of TE buffer (10mM Tris-HCl pH8.0, 1mM EDTA) was followed by centrifugation at 3000 rpm at 4°C for 2 min. Protein-DNA complexes were eluted by adding 100µl of elution buffer (1% SDS, 50mM Tris pH8.0, 10mM EDTA) and then equilibrated at 65°C for 15 min. Cross-links were reversed by adding 200mM NaCl and equilibrated at 65°C overnight. DNA fragments were purified using a QIAquick gel extraction kit (Qiagen) after a two hours treatment with RNase followed by a two hours treatment with proteinase K. qPCR analyses were performed with specific primers listed in Table 4.1. The values for the samples immunoprecipitated by anti-CDX1, anti-CDX2, anti-CDX4, anti-HOXA9 anti-HOXB9, anti-HOXD10, anti-HOXD12 anti-histone H3 and control IgG were recorded as the percentage relative to input. ChIP results were confirmed by three independent experiments and qPCR was performed in triplicate for each sample. The ChIP efficiency was calculated

by dividing the amount of PCR product obtained with the immunoprecipitated DNA by the amount obtained with the input DNA, as described in Aparicio et al. (2005).

Co-immunoprecipitation (Co-IP) assays

HEK293 cells were transiently transfected by the calcium phosphate method in 60 mm petri dishes with 10 mg/dish of the pCDNA3-CDX1, pCDNA3-CDX4, pCDNA3-HOXD4-FLAG and pCDNA3-HOXB9-FLAG expression vectors, as described in Sambrook et al. 1989. Transfection of each plasmid was done in duplicate. Chloroquine was added at a final concentration of 50 mM. Three hours after transfection, the cells were shocked for 30 seconds at 37°C with 15% glycerol in HEPES-buffered saline. Forty-eight hours after transfection, protein extracts were obtained after cell lysis in 300 ml of ice-cold Co-IP buffer ((25mM Tris pH 7.5, 1mM EDTA, 0.1mM EGTA, 5mM MgCl₂, 150mM NaCl, 2mM Na₃VO₄, 10% Glycérol, 1% Nonidet P-40) and protease inhibitors (Complete Mini-EDTA-free; Roche Diagnostics)). After 15 minutes on ice, the extracts were centrifuged at 15,000 g at 4°C for 10 mins. Protein content of the supernatant was quantified using a Lowry based assay (DC Protein Assay, Bio-Rad) and samples were stored at -80°C to test extracts (inputs). For Co-IP experiment, 250 µl of proteins extracts were incubated 2 hours at 4°C with 20 µl of Dynabeads linked to protein G (Invitrogen) conjugated with 2 µg of rabbit anti-CDX1, mouse anti-CDX4, mouse anti-FLAG (Sigma) or rabbit anti-FLAG (Sigma). Immunocomplexes were washed three times with 1ml of Co-IP buffer, one time with Nonidet P-40 free Co-IP buffer and resolved on a denaturing 12% SDS-PAGE, electrotransferred onto nitrocellulose (PALL) and probed overnight at 4°C with either the (1/2,500), the CDX4 antibody (1/5,000), or the FLAG antibody (1/10,000) according to manufacturer's instructions. Membranes were also incubated with a rabbit anti-ERK antibody (1/20,000; Dr J. Charron) as a loading control. Membranes were then incubated with the appropriate secondary horseradish peroxidase conjugated antibody: a donkey anti-rabbit IgG (1/100,000; Jackson ImmunoResearch Laboratories) or a donkey anti-mouse IgG (1/80,000; Jackson ImmunoResearch Laboratories). Proteins were revealed by chemiluminescence using the Western Lighting Plus-ECL system (PerkinElmer) according to manufacturer's protocol.

RESULTS

The PsHi 331-bp fragment participates to the MES enhancer activity.

We have previously shown that removal of the *PstI-HindIII* DNA fragment of 331-bp, referred as PsHi 331-bp located at the 3' end of the MES caused a loss of expression in the mesenchymal condensations of the developing limb buds (Tabariès et al., 2005). To further define the contribution of the PsHi 331-bp fragment to the MES activity, a construct carrying this DNA sequence in 3' of a *Hoxa5/lacZ* minimal construct was designed and tested in E12.5 F0 transgenic embryos (construct 1; Fig. 4.1B). This *Hoxa5/lacZ* minimal vector included 3.09 kb of *Hoxa5* genomic sequences from -235 bp to +2.85 kb that on its own cannot direct transgene expression (Zákány et al., 1988). Construct 1 drove β -galactosidase expression in the cervical region of the prevertebral column with a well-established posterior limit in the upper thoracic region at the level of pv10 (Fig. 4.1C). This result suggested that the PsHi 331-bp possessed enhancer properties since this sequence was able to direct region-specific expression on its own in the context of the proximal *Hoxa5* promoter and thus may contribute to the MES enhancer activity.

To define the sequences responsible for the enhancer activity of the PsHi 331-bp fragment, we first investigated whether the PsHi 331-bp fragment contains binding sites for known *trans*-acting factors. By sequence comparison with TFsearch database or Transcription Element Search System (<http://www.cbil.upenn.edu/cgi-bin/tess/tess>), we found several binding sites of putative *Hox* transcriptional regulators, like CDX (Fig. 4.2A), PBX, MEIS et OCT (not shown). To establish whether these proteins can bind to the PsHi 331-bp fragment, CDX, PBX1A, PBX1B, MEISA and MEISB proteins were produced *in vitro* and EMSAs were performed. However, no binding was observed (not shown). We repeated EMSA experiments with the PsHi-331 bp fragment using E9.5 embryo protein extracts (Fig. 4.2B). Binding complexes were detected (Fig.2B, lane 2) and further characterization of the specific interactions of the protein extracts with each half and each third of the PsHi 331-bp fragment revealed that embryonic proteins bound specifically to the 3' half (not shown) and the second third of the fragment named PsHiB-112 bp (Fig. 4.2B, lane 6). Depending on the protein extract used, some binding was

sometimes detected on the first third named PsHiA-115 bp (not shown). No binding was observed on the last third of the PsHi 331-bp (PsHiC) sequence (data not shown). In agreement, deletion of the PsHiC sequence from construct 1 had no effect on the enhancer activity carried by the PsHi 331-bp fragment in E12.5 F0 *Hoxa5/lacZ* transgenic embryos and on its ability to drive *lacZ* expression in the pv3-pv10 region (construct 2; Fig. 4.1D). Taken together, these results suggested that the activity of the PsHi-331 bp fragment is carried by the first two thirds of the PsHi-331 bp and that protein can bind to a 67-bp on the PsHiB-112 bp (Fig. 4.2B, lane 10).

To delineate the binding sequence in the 67-bp fragment, synthetic oligonucleotides corresponding to each half of the 67 bp fragment were used in EMSAs experiments and revealed that proteins bound specifically to the 3' half named Oligo-33 bp (Fig. 4.2B, lane 14). Synthetic oligonucleotides containing successive mutations of the Oligo 33-bp, namely Oligo 33-bp (mut 1), Oligo 33-bp (mut2), Oligo 33-bp (mut3) and Oligo 33-bp (mut 4) (Table 4.1), were tested in EMSAs (Fig. 4.2C, lanes 2, 6, 10 and 14). The loss of binding with Oligo-33 bp (mut2) and Oligo-33 bp (mut 3) led us to identify a 16 bp double-stranded oligonucleotide, named Oligo 16-bp that spans one putative CDX consensus binding site, as the only target for protein binding identified *in silico* (Fig. 4.2C, lane 18). Specificity of binding was confirmed by competition studies with a 100-fold excess of unlabeled probe that resulted in a substantial reduction of protein binding (Fig. 4.2B-C, lanes 3, 7, 11, 15 and 19). However, addition of a 100-fold excess of an unrelated SIF1 primer containing a known CDX binding site did not compete (Fig. 4.2B-C, lanes 4, 8, 12, 16 and 20). Thus, despite the presence of a putative binding site for CDX proteins, these might not be involved in the protein binding to the Oligo-16 bp sequence since SIF1 cannot efficiently compete for binding. To test if other homeodomain-containing proteins could bind to the Oligo-16 bp sequence, we directly assessed by EMSA the ability of HOXA5, HOXD9, HOXD10, HOXD11, CDX1, CDX2, CDX4, PBX1A and PBX1B proteins produced *in vitro* to bind to Oligo-33 bp. Unfortunately, no binding was observed with any of these proteins (not shown).

To further establish the functional significance of the Oligo-16 bp motif, E12.5 F0 *Hoxa5/lacZ* transgenic embryos were produced to assess the effect of the 16 bp sequence

Figure 4.1. Characterization of MES regulatory element intergenic region in E12.5 F0 transgenic mouse embryos

(A) Partial restriction map of the *Hoxa4-Hoxa6* genomic region. Black boxes represent homeoboxes, shaded boxes correspond to translated regions and white boxes indicate transcribed regions. The arrow shows the transcriptional start site from the *Hoxa5* proximal promoter. *Hoxa5* regulatory elements are also indicated: T, Temporal; BSC, Brachial Spinal Cord; MES, Mesodermal; L/G, Lung/Gut (B) Diagram of the *Hoxa5/lacZ* constructs used to generate E12.5 F0 transgenic embryos and summary of transgenic expression analyses. Black triangle (constructs 3 and 7), open triangle (construct 5) and asterisks (constructs 6 and 7) respectively represent deletion of Oligo 16-bp, deletion of Oligo 1 and mutation of CDX binding sites. A, *AccI*; Ap, *ApaI*; Av, *AvrII*; Bh, *BssHIII*; E, *Eco47III*; H, *HindIII*; K, *KpnI*; RI, *EcoRI*; P, *PstI*; Xb, *XbaI*; Xh, *XhoI*. post. bound., posterior boundary; pv, prevertebrae (C-I) representative E12.5 transgenic embryos stained for β -galactosidase activity showing the effects of the various constructs on the transgene expression pattern. Black arrowheads point the posterior limit of transgene expression in the prevertebrae. The number in the lower left corner of each panel refers to the construct number.

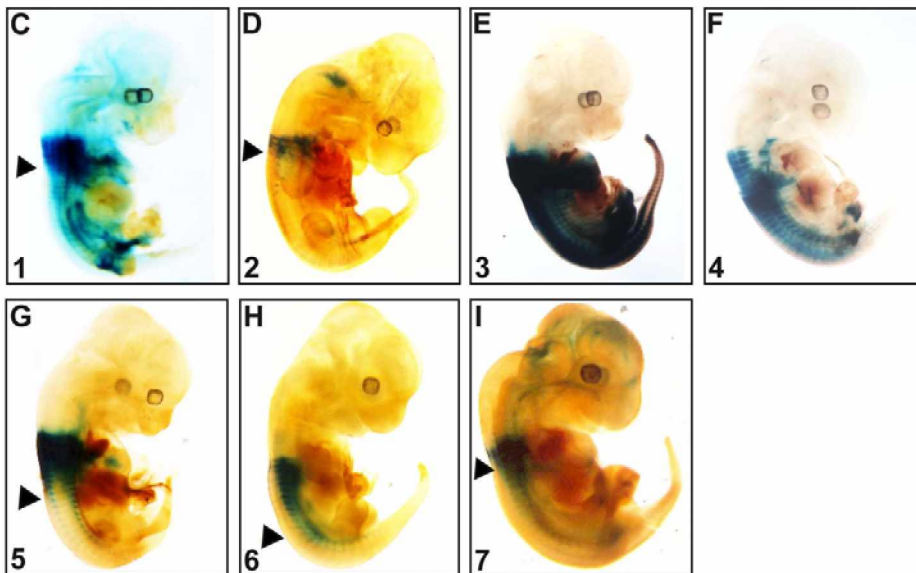
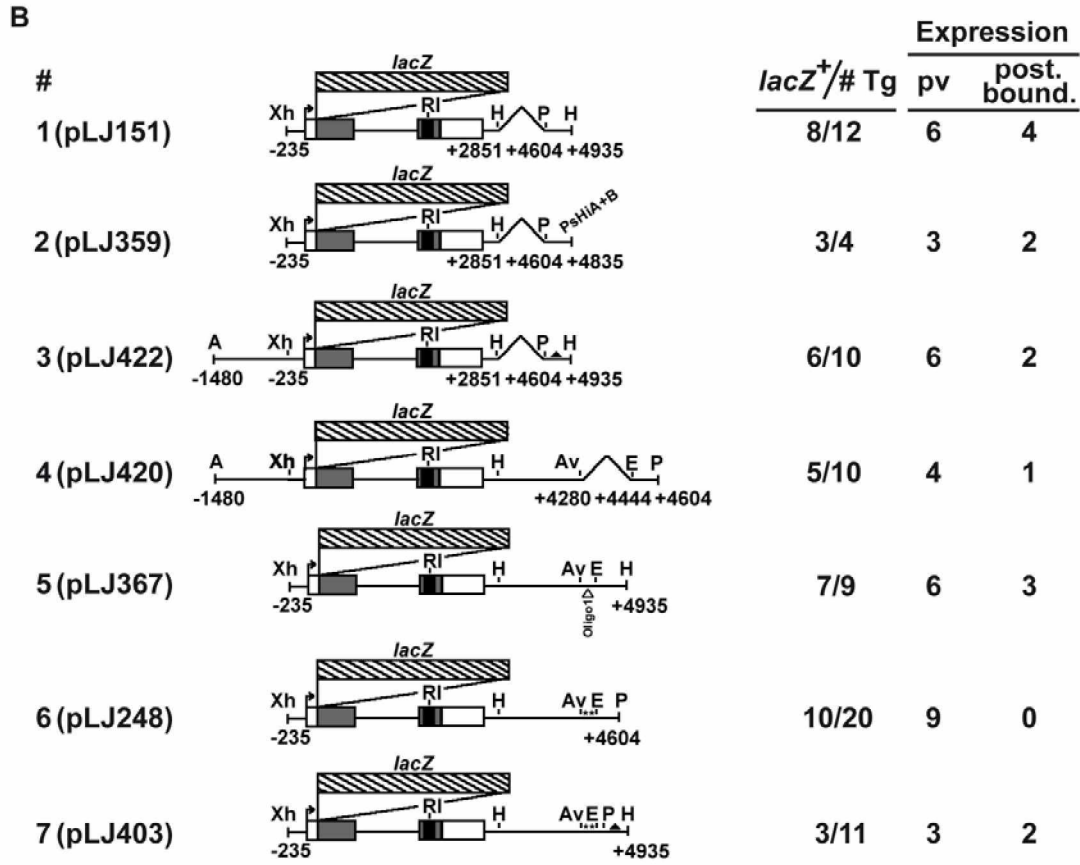
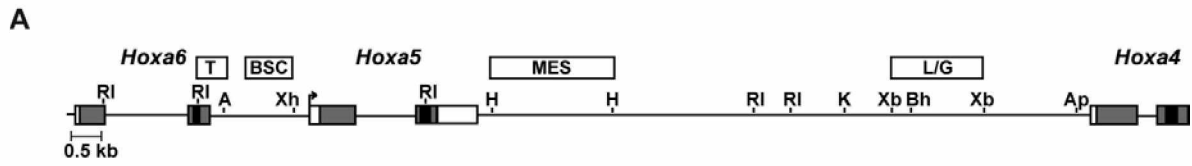
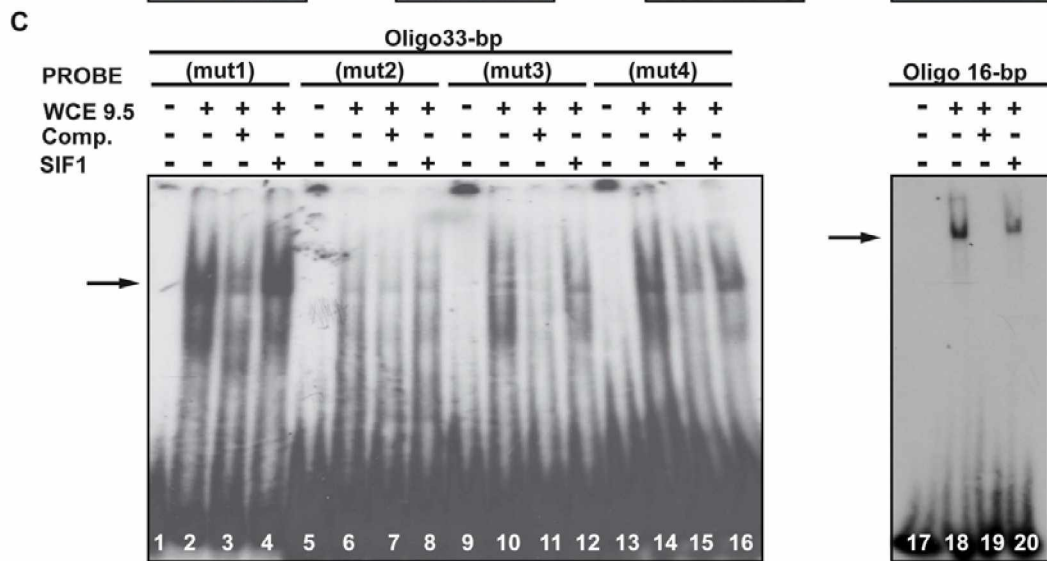
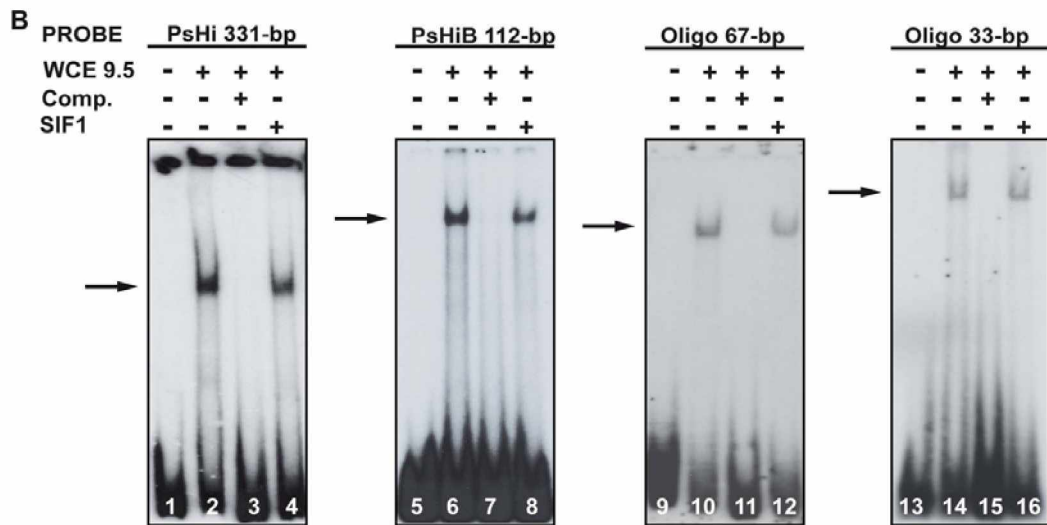
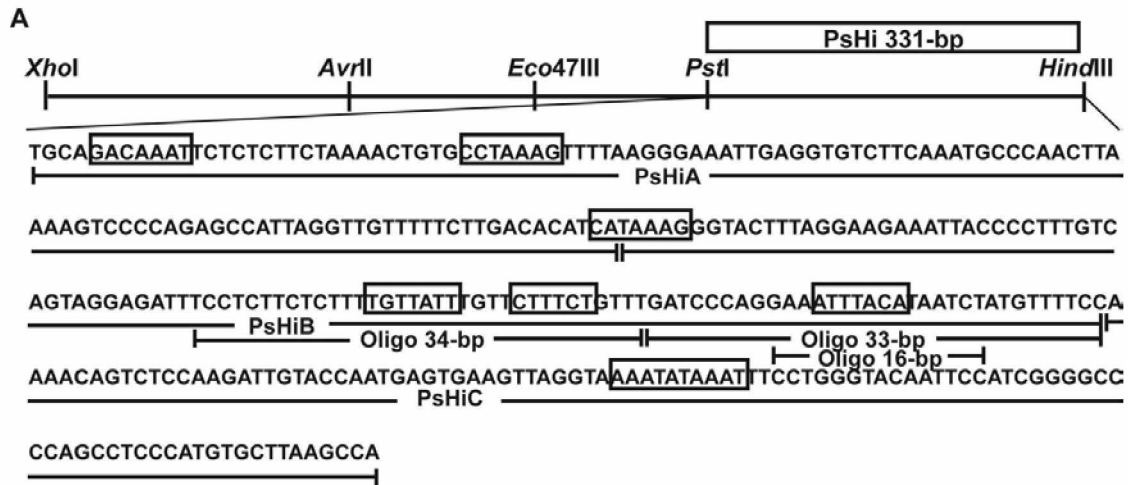


Figure 4.2. Characterization of the 331-bp *Pst*I-*Hind*III DNA fragment by EMSA

(A) Restriction map of the 892-bp *Xho*I-*Hind*III sequence extending from +4.04-kb to +4.93-kb in the 3' half of the MES regulatory region. The box denotes the location of the 331-bp *Pst*I-*Hind*III regulatory region as defined by transgenic studies. Sequence of the PsHi 331-bp DNA fragment is indicated. PsHiA, PsHiB and PsHiC used as probe/competitor in EMSA are underlined. Fragment PsHiB was further divided into 34-bp, 33-bp and 16-bp subfragments. Boxed nucleotides correspond to putative CDX binding sites. (B) The capacity of the PsHi 331-bp, PsHi 112-bp, Oligo- 67-bp and Oligo 33-bp to bind WCE E9.5 embryos extract was tested by EMSA (lanes 2, 6, 10 and 14). Specific binding complexes were competed by a 100-fold excess of unlabeled probe (lanes 3, 7, 11 and 15) or unrelated SIF1 (CDX binding site) (lanes 4, 8, 12 and 16) (C) The capacity of the Oligo 16-bp, identified by linker scanning mutagenesis of the Oligo 33-bp, to bind WCE E9.5 embryos extract was tested by EMSA (lanes 2, 6, 10, 14 and 18). Specific binding complexes were competed by a 100-fold excess of unlabeled probe (lanes 3, 7, 11, 15 and 19) or unrelated SIF1 (CDX binding site) (lanes 4, 8, 12, 16 and 20). Arrows indicate the bands corresponding to the binding of an unidentified transcription factor presents in the protein extract. Comp, competitor; SIF1, Sucrase-Isomaltase Factor 1; WCE, whole cell extract.



deletion on transgene expression (construct 3; Fig. 4.1B). Six embryos out of the 10 transgenic embryos obtained presented transgene expression in prevertebrae (Fig. 4.1E). Together, these results indicated that unknown protein (s) from E9.5 embryonic extracts could specifically bind to a 16-bp motif enclosed in the PsHi 331-bp fragment, but did not seem essential to direct *Hoxa5/lacZ* expression in the prevertebral column.

Multiple and distinct MES regulatory sequences mediate *Hoxa5/lacZ* regional expression

Our previous works have shown that a *AvrII*-*Eco47III* DNA fragment of 164 bp, named AvEc-164 bp, acts as a repressor in the caudal part of the embryo to define the posterior boundary of transgene expression at the level of pv10 (Tabariès et al., 2005). On the other hand, construct 1 (Fig. 4.1B), carrying the PsHi 331-bp DNA sequence in 3' of a *Hoxa5/lacZ* minimal construct, drove specifically *Hoxa5/lacZ* expression, between prevertebrae 3 and 10, in four *lacZ* positive transgenic embryos out of six even in the absence of the AvEc 164-bp caudal repressor element (Fig. 4.1B-C). Similar results were observed with construct 2 (Fig. 4.1D), suggesting that DNA sequences located in PsHiA-PsHiB 227-bp are also important to control *Hoxa5* regional expression. Moreover, deletion of the 16-bp DNA fragment tends to extend posteriorly *Hoxa5/lacZ* posterior limit of expression in the prevertebrae (Fig. 4.1E, black arrowhead). These results suggest that the binding site located in the 16-bp DNA fragment identified in Figure 4.2C could play a role in restricting *Hoxa5* expression between pv3 and pv10.

To determine if some cooperation exists between the AvEc 164-bp and the PsHi 331-bp fragments, deletion of both the AvEc-164 bp and PsHi-331 bp fragments was tested (Construct 4; Fig. 4.1B). This resulted in 75% of *lacZ* positive embryos presenting posterior extension of *Hoxa5/lacZ* expression in prevertebrae (Fig. 4.1F) compared to 33% and 29% respectively for individual deletions (construct 1; Fig. 4.1B) (construct 3; Tabariès et al., 2005), suggesting a potential cooperation between regulatory elements located in both the AvEc 164-bp and PsHi 331-bp fragments to control *Hoxa5* regional expression.

To gain further insight into the MES, deletions were made to delineate the DNA

elements enclosed. Deletion of the Oligo1 sequence, located in the AvEc 164-bp and that contains two CDX binding sites necessary to establish the *Hoxa5* expression posterior boundary at the pv10 level was undertaken (construct 5; Fig. 4.1B) (Tabariès et al., 2005). Construct 5 drove β -galactosidase expression with a loss of the posterior boundary at the level of pv10 (Fig. 4.1G) in a pattern similar to that observed when the two CDX binding sites were mutated (construct12; Tabariès et al., 2005). However, the effect was less important than the one observed with the complete deletion of the 164-bp fragment in which the expression domain extended to the caudal part of the embryo (referred to construct 8 in Tabariès et al., 2005). This result supported a cooperation between Oligo1 sequence and other DNA elements for the posterior restriction of the *Hoxa5* expression domain. These DNA elements could be located in the AvEc 164-bp, in the PsHi 331-bp, such as the 16-bp DNA fragment, and/or in another DNA fragment located in MES regulatory sequence. Similarly, the PsHi 331-bp deletion coupled with the nucleotide substitution in the CDX binding sites present in the Oligo1 sequence (construct 6; Fig 1B) posteriorly shifted the *Hoxa5/lacZ* expression boundary in prevertebra (Fig. 4.1H, black arrowhead), but not as much as with the complete deletion of both fragments (Construct 4; Fig. 4.1B, F), suggesting that not only CDX binding sites were important for the AvEc 164-bp function in *Hoxa5* regional expression regulation. Finally, the 16-bp deletion in the PsHi 331-bp coupled with the nucleotide substitution in the CDX binding sites present in the Oligo1 sequence (Construct 7; Fig. 4.1B) did not seem to have a synergistic action to specify the posterior boundary of the *Hoxa5/lacZ* transgene at the level of pv10 (Fig. 4.1I). Taken together, these results support the importance of distinct DNA regulatory elements located in the MES sequences, where some of them are in the AvEc 164-bp and the PsHi 331-bp fragments, to orchestrate proper *Hoxa5* regional expression in the cervico-thoracic region of the axial skeleton. Unlike for the CDX binding sites present in the Oligo1 sequence, we did not clearly demonstrate the significance of the 16-bp sequence located in the PsHi 331-bp fragment neither for the MES enhancer activity nor the *Hoxa5* regional expression control. Therefore, we focused on the AvEc 164-bp fragment for the rest of the study.

***Hox* gene products bind *in vitro* to the AvEc 164-bp fragment**

To identify any other critical sequences within the AvEc-164 bp sequence that would be involved in axial regional specification, we first searched for binding sites of known *trans*-acting factors that could transduce positional information to *Hox* genes. By sequence comparison with TFsearch database(<http://molsun1.cbrc.aist.go.jp/research/db/TFSEARCH.html>), we found one potential site with a high homology to the consensus-binding site for the *AbdB* homeotic gene product (Fig. 4.3A). To establish whether HOX proteins could bind the AvEc-164 bp fragment, we performed EMSA with *in vitro* produced murine HOX proteins from different paralogous groups (Fig. 4.3B). Among these proteins, HOXA9, HOXB9, HOXD9, HOXD10, HOXD11 and HOXD12 orthologs to the *Drosophila* AbdB protein formed binding complexes with the AvEc-164 bp fragment (Fig. 4.3B, lanes 9, 11, 13, 15, 19, and 21) Specificity of binding was confirmed by competition studies with a 100-fold excess of unlabeled probe that resulted in a reduction of protein binding (Fig. 4.3B, lanes 10, 12, 14, 16, 20 and 22). No binding was observed with HOXD4, HOXA5, HOXA7 and HOXA11 indicating the specificity of binding to defined AbdB-like murine HOX proteins (Fig. 4.3B, lanes 3, 5, 7 and 17). To delineate the DNA element that bind these HOX proteins, EMSAs were performed using either the wild-type AvEc-164 bp fragment (WT) or, versions of the AvEc-164 bp fragment carrying deletion of the ABDB binding site (Δ HoxBS) or the Oligo1 sequence (Δ Oligo1) (Fig. 4.3C). These experiments revealed that the above-mentioned binding did not occur to the HOX binding site found by TRANSFAC analysis (lanes 9-12) but rather via the Oligo1 sequence (lanes 15-18), since deletion of Oligo1 sequence prevented HOX proteins binding to the AvEc 164-bp in opposition to deletion of AbdB binding site. Together, these observations revealed that, in agreement with the concept of posterior prevalence, only HOX proteins expressed posteriorly to *Hoxa5* expression domain such as the AbdB group could bind the AvEc 164-bp and act as repressors of the *Hoxa5* gene expression.

HOXB9 protein binds *in vivo* to the AvEc 164-bp fragment

To establish the functional significance of the HOX protein binding, we performed supershift assays using E9.5 embryo protein extracts and the AvEc 164-bp as probe (Fig. 4.4A). Several proteins present in the extract were able to bind the AvEc 164-bp fragment

(Fig. 4.4, lane 10). To test if the E9.5 embryo protein complexes that bound the AvEc 164-bp sequence contained HOX proteins, supershift EMSAs were performed using HOX specific antibodies (Fig. 4.4A, lanes 11-14). The specificity of each antibody was assessed by EMSA with the different HOX proteins produced *in vitro*, as shown for the anti-HOXB9 antibody (Fig. 4.4A, lanes 4, 6 and 8; Fig.4.4B, lanes 4 and 6). While addition of the anti-HOXA9, anti-HOXD10 or anti-HOXD12 antibodies to the embryo protein extract-AvEc 164-bp reaction did not modify the pattern of bands (Fig. 4.4A, lanes 11, 13 and 14), addition of the anti-HOXB9 antibody to the embryo protein extract-AvEc 164-bp reaction led to a near-complete loss of complex formation (Fig. 4.4A, lane 12). Taken together, these results demonstrated that the HOXB9 protein present in E9.5 embryo protein extract could bind specifically the Oligo1 sequence. Thus, HOXB9 protein appears to be a potential regulator for the correct positioning of the posterior boundary of *Hoxa5* expression in the upper thoracic region along the axial skeleton.

To confirm whether the CDX and HOX binding sites identified in the AvEc 164-bp DNA regions occur *in vivo*, a ChIP assay was performed on cross-linked chromatin isolated from either head, trunk or stomach of E10.5 wild-type mouse embryos (Fig. 4.4C). DNA from the immunoprecipitate was subjected to qPCR analyses with specific primers for the AvEc 164-bp sequence. In both trunk and tail chromatin extracts, all CDX members were recruited to the AvEc 164-bp region, but the CDX4 protein bound with more affinity than CDX2 and CDX1, consistent to what was observed by EMSA in Tabariès et al, 2005. Furthermore, data obtained by supershift assays were confirmed since only the HOXB9 protein was recruited to the AvEc 164-bp region *in vivo*. Interestingly, proteins recruitment to the AvEc 164-bp region were more important using chromatin isolated from tails of E10.5 embryos compared to trunks extracts. This could be in part explained by the presence of a mix population of expressing and regionally restricted *Hoxa5* cells in the trunks chromatin extract, whereas all cells in the tails extract might need HOXB9 and CDX4 to negatively regulate *Hoxa5* expression. Finally, heads chromatin extract, which is not expected to contain neither CDX proteins, nor posterior HOX proteins, and do not express *Hoxa5*, was used as a negative control in our ChIP experiments.

Figure 4.3. Characterization of the 164-bp *AvrII-Eco47III* DNA fragment by EMSA

(A) Restriction map of the 892-bp *XhoI-HindIII* sequence extending from +4.04-kb to +4.93-kb in the 3' half of the MES regulatory region. The box denotes the location of the 164-bp *AvrII-Eco47III* regulatory region as defined by transgenic studies. Sequence of the AvEc 164-bp DNA fragment is indicated and sequence of the Oligo1 fragment is underlined. Boxed nucleotides correspond respectively to a putative AbdB-like HOX binding site, a PBX binding site and the two previously described CDX binding sites in Tabariès et al., 2005. (B) The capacity of the AvEc 164-bp fragment to bind HOX proteins produced *in vitro* was tested by EMSA (lanes 3, 5, 7, 9, 11, 13, 15, 17, 19 and 21). Specific binding complexes were competed by a 100-fold excess of unlabeled probe (lanes 10, 12, 14, 16, 20, 22). HOXA9, HOXB9, HOXD9, HOXD10, HOXD11 and HOXD12 proteins have the ability to bind the AvEc 164-bp sequence but not HOXD4, HOXA5, HOXA7 and HOXA11 proteins (lanes 3, 5, 7 and 17). (C) EMSA with HOX and CDX4 proteins produced *in vitro* using the AvEc 164-bp fragment (WT) (lanes 1-6), the AvEc 164-bp with a deletion of the HOX binding site (Δ HOXBS) (lanes 7-12) and the the AvEc 164-bp with a deletion of the Oligo1 sequence (Δ Oligo1) (lanes 13-18) as probes. HOX binding was only lost when the Oligo1 sequence was deleted suggesting that the binding of HOX proteins occurred via Oligo1. Comp, competitor; R.L., Reticulocyte lysate

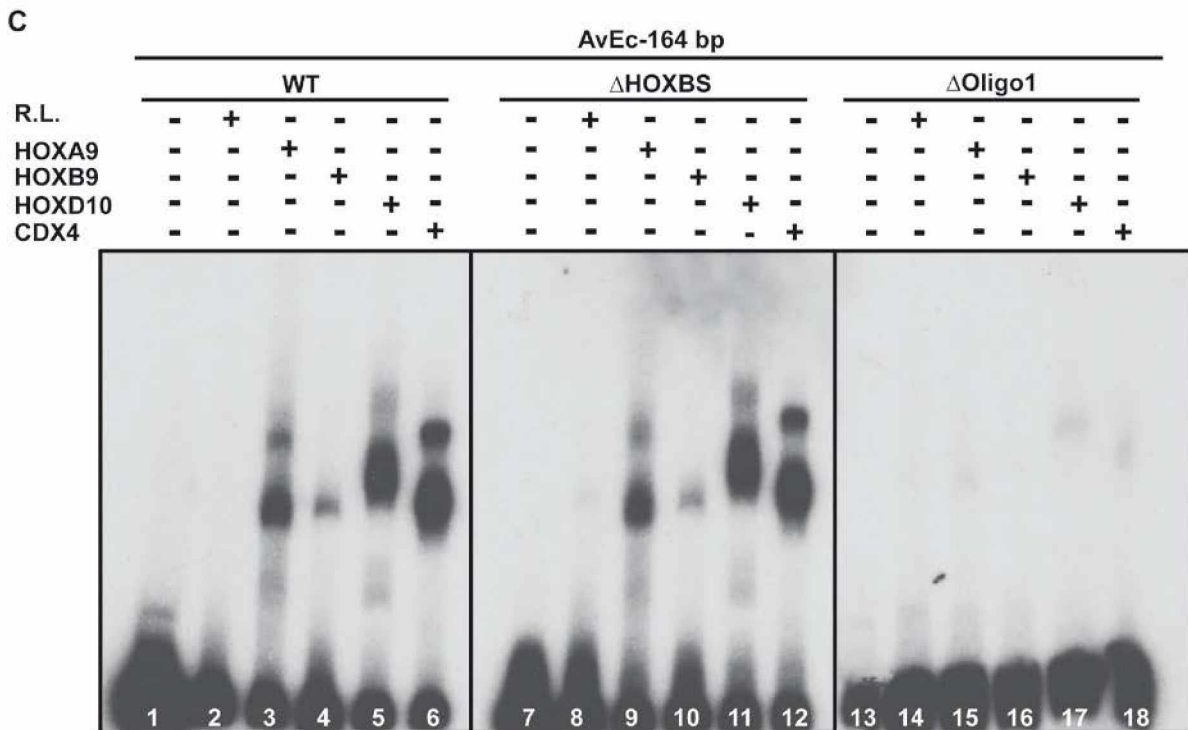
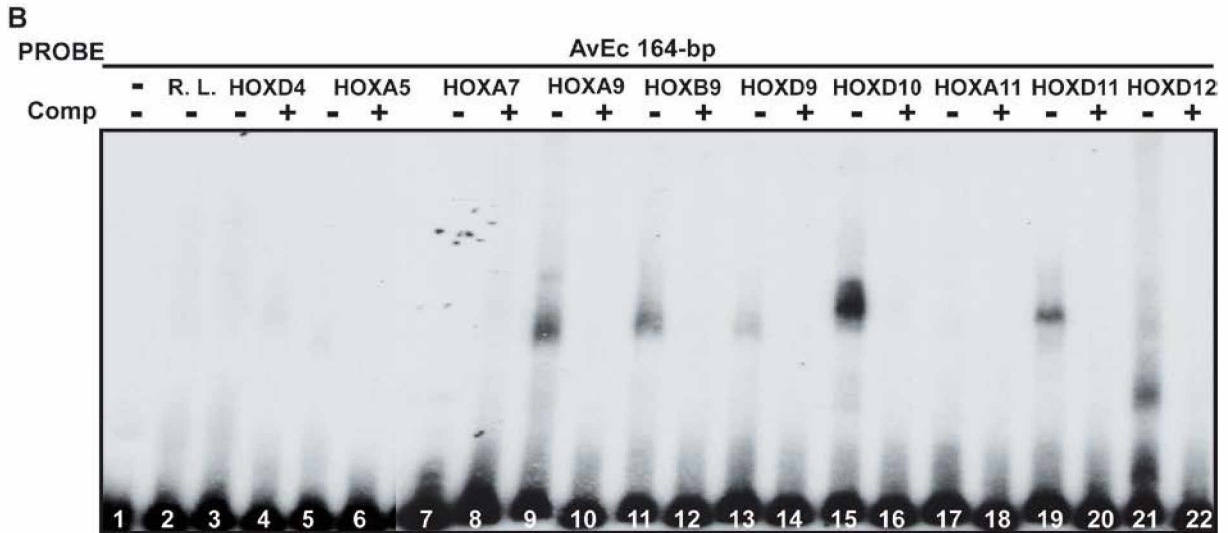
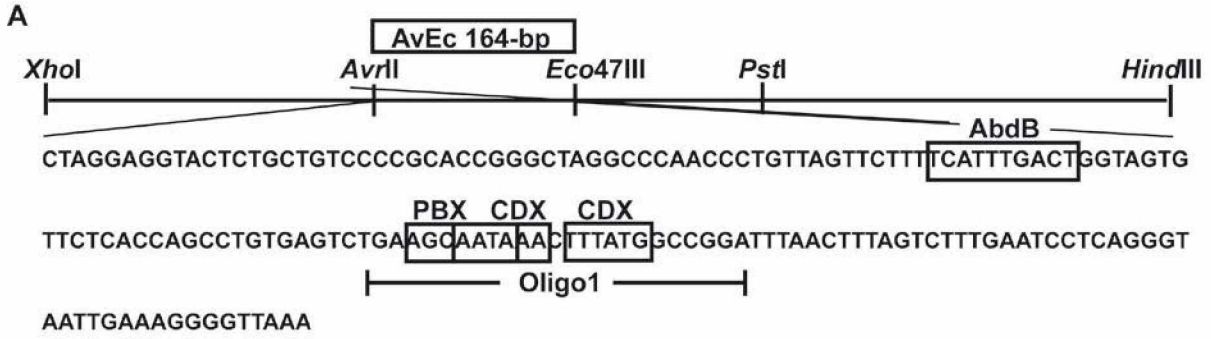
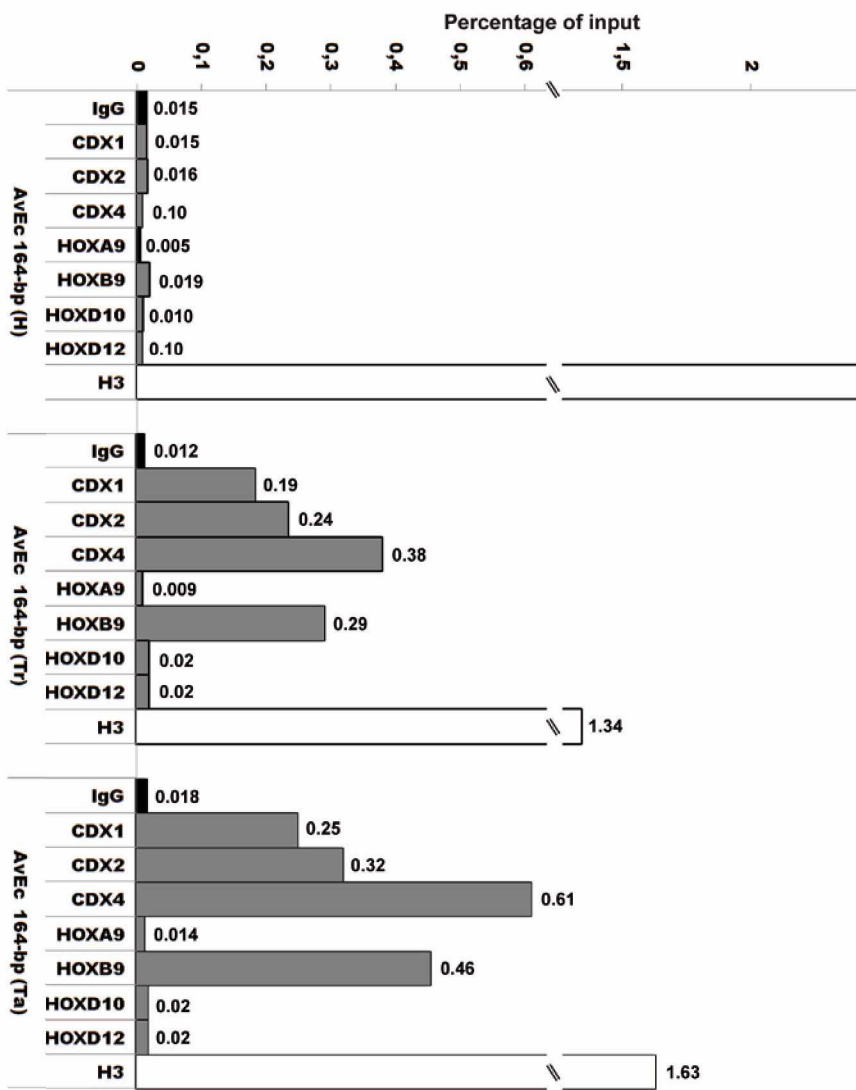


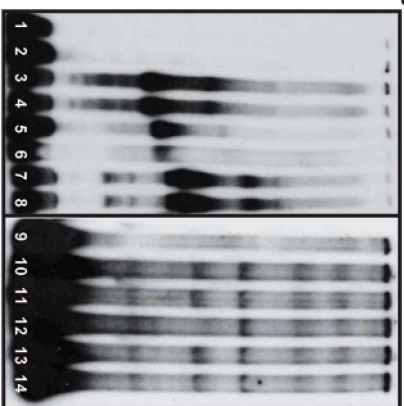
Figure 4.4. *In vivo* detection of HOXB9 protein binding to the AvEc 164-bp region

(A) EMSA with the AvEc 164-bp fragment and *in vitro*-translated proteins (HOXA9, HOXB9 and HOXD10) (lanes 3-8), or E9.5 embryonic protein extract (lanes 10-14). The specificity of the anti-HOXB9 antibody was confirmed by the observation of a decreased binding only in the presence of the HOXB9 protein (lanes 4, 6 and 8). A near-complete loss of one AvEc 164-bp-WCE E9.5 complex was observed in the presence of the anti-HOXB9 antibody (lane 12), but not when Anti-HOXA9, anti-HOXD10 and anti-HOXD12 antibodies were added (lanes 11, 13 and 14). (B) Anti-HOXD10 and anti-HOXD12 antibody also produced specific supershifts with their respective antigen (lanes 4 and 6). Ab, antibody; R. L., Reticulocyte lysate; WCE, whole cell extract. (C) Schematic representation and position relative to the *Hoxa5* TSS of the AvEc 164-bp fragment in the 3' end of the MES regulatory sequence. The black circle represents the AvEc 164-bp fragment identified by a transgenesis approach. The position of the qPCR fragment corresponding to AvEc 164-bp amplified sequence is indicated. ChIP analysis of the AvEc 164-bp DNA region in head (H), trunk (Tr) and tail (Ta) from E10.5 mouse embryos, chromatin was immunoprecipitated with rabbit IgG (negative control), anti-CDX1, anti-CDX2, anti-CDX4, anti-HOXA9, anti-HOXB9, anti-HOXD10, anti-HOXD12 and anti-histone H3 (for chromatin integrity control) antibodies. The recruitment of CDX1, CDX2, CDX4, HOXB9 and histone H3 on AvEc 164-bp DNA region was evaluated by qPCR and is indicated as the percentage of input.

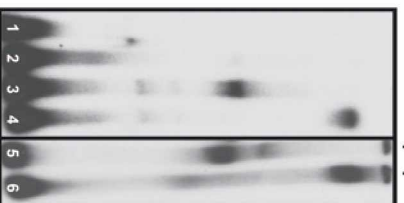
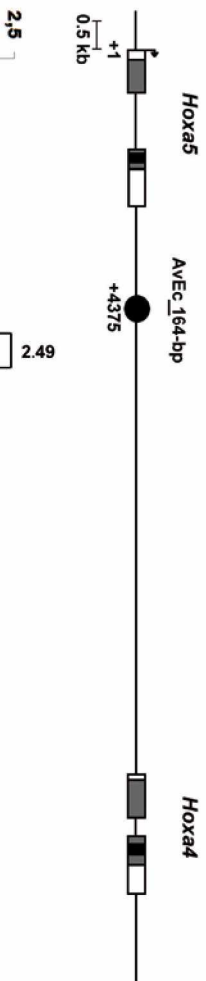


A

PROBE		AveC 164-bp	
<i>in vitro</i>	R. L.	-	-
HOXA9	+	-	-
HOXB9	+	-	-
HOXD10	+	-	-
WCE 9.5	-	-	-

**B**

PROBE		AveC 164-bp	
<i>in vitro</i>	R. L.	-	-
HOXD10	+	-	-
HOXD12	+	-	-

**C**

HOXB9 and CDX4 proteins do not physically interact

As HOX proteins binding on the AvEC 164-bp do not occur via the AbdB-like HOX binding site, but rather via the Oligo1 sequence (Fig. 4.3C; Fig. 4.5A, lanes 2-5), and since both HOXB9 and CDX4 proteins are the major players in AvEc 164-bp regional control of *Hoxa5* expression (Fig. 4.4; Tabariès et al., 2005), we have further characterized interactions between HOXB9, CDX4 and Oligo1 by EMSA. While the loss of CDX4 protein binding was already described (Tabariès et al, 2005), mutation of the two CDX binding sites in Oligo1 sequence (Oligo1 mutCDXBS, see also Table 4.1) led to binding disruption of homeodomain containing proteins (Fig. 4.5A, lanes 10-14). These results suggest that HOXB9 could physically interact with CDX4, physically interact with another protein that binds a DNA sequence overlapping CDX binding site or directly bind the CDX binding sites located in Oligo1 DNA sequences.

To further establish the relation between HOXB9 and CDX4, we investigated if these two proteins interact physically via the Oligo1 sequence. The absence of shift modifications observed by EMSA, such as an additional shift, a shift loss or a supershift, when *in vitro* HOXB9 and CDX4 proteins were added simultaneously to Oligo1 sequence, compared to individually, suggested that they do not heterodimerise to bind the Oligo1 DNA sequence (Fig. 4.5A, lanes 2, 4 and 7). Moreover, CDX4 *in vitro* protein seemed to preferentially bind Oligo1 sequence than HOXB9 did, as the CDX4-HOXB9-Oligo1 binding profile (lane7) observed is more similar to CDX4-Oligo1 (lane2) than to the HOXB9-Oligo1 (lane4) binding profile. On the other hand, both proteins were found bound to Oligo1 with the E9.5 embryos proteins extract (lane 8; black arrows), suggesting that CDX4 and HOXB9 compete for binding on Oligo1 sequence. This competition hypothesis for binding sites between CDX4 and HOXB9 was also tested using a gradient binding assay (Fig. 4.5B). At the same concentration (Fig. 4.5B, lanes 2-6 versus lanes 7-11), CDX4 binding was weaker in the presence of HOXB9 *in vitro* protein compared to what was observed in absence of HOXB9 (Fig. 4.5B, lanes 6 versus 11), suggesting that HOXB9 and CDX4 compete for binding sites to fulfill their role in controlling *Hoxa5* regional expression through an interaction with a 25-bp DNA sequence (Oligo1) located in AvEc 164-bp fragment.

Besides binding displacement hypothesis, in which CDX4 protein binds Oligo1 with higher affinity than HOXB9 protein, another possibility was that CDX4 tethered HOXB9 protein from its binding sites via physical interactions. To test this hypothesis, we co-transfected HEK293T cells with DNA constructs containing *Cdx4* and *Hoxb9-Flag* mRNA and undertook co-immunoprecipitation (CoIP) assays. Since HOXD4-CDX1 heterodimer was previously reported in the literature (Lafontaine et al., 2012), we initially reproduced that experiment as a positive control (Fig. 4.5C). We then assessed interaction between CDX4 and HOXB9. However, our CoIP experiments did not allow us to demonstrate any physical interaction between HOXB9 and CDX4 proteins in HEK 293T, suggesting that both transcription factors act independently as repressors of *Hoxa5* expression in prevertebra rather than synergically. This data also gives a new molecular perspective supporting the posterior prevalence concept, since *Cdx4* and *Hoxb9* anterior expression boundaries, located at somites 11-12 and 12-13 respectively (Chen and Capecchi, 1997; van den Akker et al., 2002; Gaunt et al., 2005), will restrict *Hoxa5* expression domain to the presumptive cervico-thoracic region by positioning its posterior expression limit to somite 10 in E9.5 embryo (Larochelle et al., 1999).

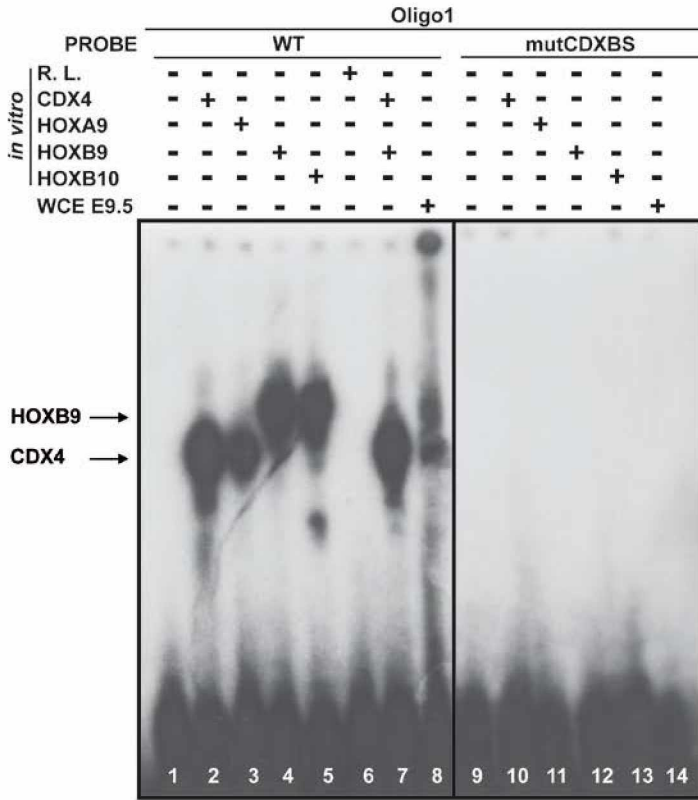
DISCUSSION

To better define the mechanisms regulating *Hoxa5* gene expression involved in the correct development of structures from the cervical and upper thoracic axial level, we have pursued our analysis of the MES region. Coordination between several *cis*-acting regulatory elements is needed to fully reproduce the correct spatial and temporal *Hoxa5* gene profile. In addition to the MES, regulatory sequences include the BSC sequence, which specifically drives *Hoxa5* gene expression in the brachial region of the spinal cord from E11.0 to E13.0 as well as the lung/gut sequence, located in the *Hoxa4-Hoxa5* intergenic region (Tuggle et al., 1990; Larochelle et al., 1999; Nowling et al., 1999; Moreau and Jeannotte, 2002; Tabariès et al., 2005; Bérubé-Simard et al., 2014).

Figure 4.5. HOXB9 and CDX4 proteins do not physically interact

(A) EMSA with HOX and CDX4 proteins produced *in vitro*, or protein extracts from E9.5 embryos, and Oligo1 (lanes 1-8) or Oligo1 mutCDXBS (lanes 9-14). The loss of HOX protein binding on oligo1 mutCDXBS (lanes 11-14) indicated that mutations of the two CDX binding sites disrupt HOX protein binding. The loss of CDX4 protein binding was already described in Tabariès et al, 2005. HOXB9 and CDX4 proteins did not interact physically on the Oligo1 sequence since no additional larger band (or shifted band) was observed when both was added simultaneously (lanes 2, 4 and 7). Moreover, CDX4 *in vitro* protein seemed to preferentially bind Oligo1 sequence than HOXB9 did, whereas both proteins were found bound to bind Oligo1 in the E9.5 embryos proteins extract (lane 8; black arrows) (B) EMSA with a increasing amount of CDX4 *in vitro* protein (lanes 2-6, 7-11) in presence of a fixed amount of HOXB9 *in vitro* protein (lanes 7-11) and Oligo1. HOXB9 and CDX4 proteins compete for the Oligo1 sequence since CDX4 binding was weaker when HOXB9 protein was added (lanes 5-6, 10-11). R.L., Reticulocyte lysate; WCE, Whole cells extract (C) CoIP by which interaction of CDX1 with HOXB4-FLAG protein in HEK293T cells was confirmed as described in Lafontaine et al., 2012. (D) CoIP by which no interaction of CDX4 with HOXB9-FLAG protein in HEK293T cells was observed. Transfection efficiency was tested by western blot on transfected cells (Input). Mock lanes consist of HEK293T transfected with empty pCDNA3 plasmid. IP, Immunoprecipitation; WB, Western blot.

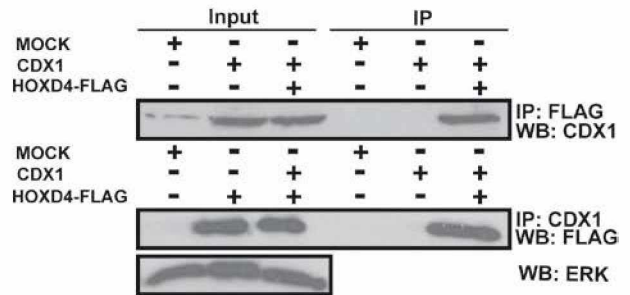
A



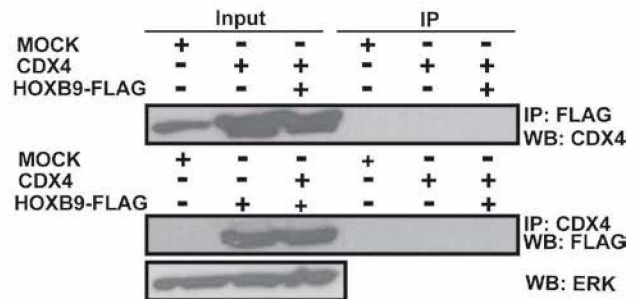
B



C



D



A homeodomain-containing protein binds to the PsHi 331-bp element

We have previously shown that the MES directs expression in the skeleton at the cervical and upper thoracic axial level and that the MES sequence encloses DNA elements in its 3' half able to target transgene expression in the limb buds and in the urogenital and intestinal tracts (Tabariès et al., 2005). We have now mapped another regulatory element, the PsHi-331 bp fragment that likely participates in MES restriction of *Hoxa5* activity in the pv3-pv10 region as well as in its enhancer activity (Fig. 4.1C). Proteins from E9.5 embryo extracts can physically interact with the PsHi 331-bp regulatory sequence via a 16-bp motif containing a homeodomain protein binding site. According to our assays, CDX and HOX proteins do not bind the sequence and additional work will be needed to identify the homeodomain-containing factor involved. However, construct 3 carrying deletion of the 16-bp sequence does not abolish *Hoxa5/LacZ* expression in prevertebral column (Fig. 4.1E), suggesting that the 16-bp sequence is not essential for enhancer activity. Interestingly, the 16-bp motif seems involved in MES restriction of *Hoxa5* expression to the cervico-thoracic level, since its deletion posteriorly extends *Hoxa5* expression domain in prevertebrae. Furthermore, depending on the protein extracts used in EMSA experiments, binding was sometimes observed on the PsHiA 115-bp fragment (not shown), suggesting that other motifs could be required for restricting *Hoxa5* expression. This hypothesis could also explain why deletion of the 16-bp motif in PsHi 331-bp combined to mutations of the two CDX binding sites in AvEc 164-bp fragment do not recapitulate the effect on *Hoxa5* expression posterior boundary generated by the complete deletion of both fragments combined to mutations of the two CDX binding sites in AvEc 164-bp fragment (Fig. 4.1I; Tabariès et al., 2005). Consequently, our analysis indicates that coordination between different *cis*-acting regulatory elements is needed to fully reproduce the correct spatial *Hoxa5* gene profile.

***Hoxb9* gene product binds to the AvEc-164 bp element**

Definition of expression domain boundaries is necessary for proper *Hox* gene action. The functional limits of *Hoxa5*, as delineated by the axial skeletal phenotype observed in *Hoxa5* mutant mice, are restricted to the pv3-pv10 region (Jeannotte et al., 1993; Aubin et al., 1998; Bérubé-Simard and Jeannotte, 2014). Posterior restriction at the

pv10 level is further delimited by the AvEc-164 bp fragment, enclosed within the MES, that acts as a repressor in the caudal part of the embryo (Tabariès et al., 2005). The pv3-pv10 region corresponds to the specific expression domain of the 1.8 kb transcript, which is the sole *Hoxa5* transcript expressed in the structures affected by the mutation at this axial level, such as the respiratory tract, the thyroid gland and the pectoral girdle (Aubin et al., 1997; Aubin et al., 1998; Larochelle et al., 1999; Aubin et al., 2002b; Meunier et al., 2003; Coulombe et al, 2010; Boucherat et al, 2013). As the 1.8 kb transcript encodes the functional form of the *Hoxa5* gene, its expression domain must be tightly regulated. The transgenic vectors used in our approach allow the regulation of the proximal promoter, which is a useful tool to investigate the regulatory mechanisms governing *Hoxa5* regional expression in its functional domain.

We have previously identified two functional CDX binding sites in the AvEc 164 - bp fragment involved in the positioning of the posterior boundary of *Hoxa5* gene expression domain (Tabariès et al., 2005). Moreover, we have shown that CDX4 specifically interacts with these sites. Among the murine *Cdx* genes, *Cdx4* is the most posteriorly expressed and its expression ceases at E10.5 (Gamer and Wright, 1993). Therefore, the X-gal staining pattern reported in E12.5 F0 transgenic embryos is likely due to an earlier regulatory event directed by CDX4. Indeed, *in vivo* detection of CDX4 protein binding occurred with E10.5 embryo extract but not with extracts from older specimens suggesting an involvement of CDX4 in the establishment of the correct *Hoxa5* posterior boundary at the upper thoracic level. Recently, the targeted inactivation of the murine *Cdx4* gene did not reveal homeotic transformation in mutants (van Nes et al., 2006). This result was unanticipated based on our hypothesis that CDX4 acts to restrict the *Hoxa5* gene expression domain. Indeed, an extension of the *Hoxa5* expression domain should have resulted from the *Cdx4* inactivation and caused homeotic transformations in the domain where *Hoxa5* would have been ectopically expressed. Furthermore, no effect on the *Hoxa5* expression pattern was observed in E8.5 *Cdx4* mutants (J. Deschamps and L. Jeannotte, unpublished data). One possibility is that CDX4 is involved in *Hoxa5* expression along the AP axis earlier than E8.5, while other homeobox-containing factors, such as HOXB9, could gradually take over later to regionally restrict *Hoxa5* expression in somites/prevertebrae. Thus, HOXB9 expression would prevail over potentially ectopically-

expressed *Hoxa5* gene due to *Cdx4* gene inactivation, as stated by the concept of posterior prevalence. However, we cannot exclude the possibility that CDX4 does not play a preponderant role in controlling *Hoxa5* expression in prevertebrae.

In the present study, we demonstrated that HOX proteins produced *in vitro* could bind the AvEc 164-bp repressive sequence restricting *Hoxa5* gene expression caudally to pv10. Interestingly, only proteins encoded by murine orthologs of the *Drosophila* *AbdB* gene can do so (Fig. 4.3B). Among these, HOXB9 emerges as the potential *in vivo* trans-acting factor in *Hoxa5* regulation. Interestingly, as reported with CDX protein, HOXB9 protein binding was disrupted by mutation of the two CDX binding sites enclosed in the Oligo1 sequence. From these results, it could be hypothesized that CDX4 and HOXB9 proteins cooperate, compete, or act independently for the binding to this sequence at specific stage during AP axis patterning. However, data collected using EMSA gradient assay and CoIP suggested that both proteins compete for the CDX binding site, or bind independently different sequence included in Oligo1, especially since a hypothetical PBX binding site predicted *in silico* by TRANSFAC analysis is overlapping the first CDX binding site (Fig. 4.3A). Thus, we cannot rule out the possibility that CDX4 binding prevent HOXB9-PBX heterodimer to bind Oligo1 sequence by steric hindrance. Although we have shown the importance of the Oligo1 sequence, the impact of the complete deletion of the AvEc-164 bp is more important than the one seen when the Oligo1 sequence is absent, suggesting that other factors could act on the AvEc 164-bp sequence (Fig 4.1B, G; Tabariès et al., 2005). In fact, other proteins specifically bind the AvEc-164 bp fragment (Fig. 4.4A, lanes 10-14) suggesting the possible cooperation between several factors for the establishment of the correct *Hoxa5* posterior boundary at the upper thoracic level.

The posterior prevalence concept

It was proposed that axial patterning relies upon a given partition of *Hox* functional domains along the developing AP axis (Kmita and Duboule, 2003). The concept of posterior prevalence states that posteriorly expressed *Hox* genes functionally prevail over more anterior genes (Duboule and Morata, 1994). This mechanism was put forward to explain the fact that ectopic expression of *Hox* genes generally results in homeotic transformations confined to region anterior to the normal expression domain of the

ectopically expressed gene (Gonzalez-Reyes and Morata, 1990; Mann and Hogness, 1990). However, not much is known about the way this posterior prevalence is achieved. Several strategies have been postulated to explain the suppressive mechanism by which a posterior *Hox* gene imposes its functional dominance (Kmita and Duboule, 2003). For instance, it was proposed that the posterior prevalence reflects a gradation in the binding activity of the *Hox* gene products, the posterior-expressed genes presenting a better efficiency to bind to common sets of target genes (LaRonde-LeBlanc and Wolberger, 2003). On the other hand, a competition could occur for the binding to shared co-factors (Zhao and Potter, 2002).

The HOXB9 binding to regulatory sequences of the *Hoxa5* gene is in agreement with the posterior prevalence concept. In fact, the *Hoxb9* gene possesses an anterior boundary of expression at the pv10/11 level in E12.5 embryos, which corresponds to *Hoxa5* posterior boundary (Chen and Capecchi, 1997; Larochelle et al., 1999; Tabariès et al. 2005; Coulombe et al. 2010). It could be hypothesized that CDX4 protein, whose anterior expression boundary corresponds to *Hoxa5* posterior expression boundary (somite 7), initiates and maintains *Hoxa5* expression domain restriction until E8.5. Since CDX proteins are known to activate *Hoxb9* gene expression along the AP axis (van den Akker et al, 2002), CDX4 could participate to *Hoxb9* gene expression between E8.5-10.5, where both proteins restrict *Hoxa5* expression domain to pv3-pv10. Then, HOXB9 could take over completely CDX4 task, whose expression is turned off around E10.75 (Gamer and Wright, 1993), and binds the AvEc-164 bp sequence to prevent the expression of the major 1.8 kb *Hoxa5* transcript caudally to pv10. The study of the *Hoxa5* gene expression in *Hoxb9* mutant mice will be of interest to confirm the impact of HOXB9 binding to restrict the *Hoxa5* gene expression domain. Nevertheless, because of the difficulty to obtain specific HOX antibodies working in supershift/ChIP experiments, it remains to be tested if other caudally expressed HOX proteins would act on the Oligo1 sequence to restrict the *Hoxa5* expression domain to the pv3-pv10 region. Among the HOX proteins expressed posteriorly to pv10 and not tested in this study, only HOXB7 has an anterior boundary in the vicinity of pv11 (Izpisua-Belmonte et al., 1991; Burke et al., 1995). Taken together, these results are a first evidence for a direct interaction of a posterior HOX protein with a repressive regulatory element to prevent an anteriorly-expressed *Hox* gene to be expressed caudally

of its functional domain. How the binding of HOXB9 protein represses the *Hoxa5* gene caudally to pv10 still remains unknown. However, it could be hypothesized that the HOXB9 binding acts at the transcriptional level preventing the *Hoxa5* gene transcription in the *Hoxb9* expression domain. On the other hand, post-transcriptional events whereby a *Hox* transcript is selectively destabilized in posterior domain may be used to confine *Hox* gene expression as demonstrated for the *Hoxb4* gene (Brend et al., 2003). These strategies may co-exist.

In summary, our work gives insights into the complex mechanisms by which *trans*-acting factors participate to the posterior delimitation of the *Hoxa5* functional domain. Furthermore, our data favour the concept of posterior prevalence as a strategy to define *Hox* gene functional domain at the cervical-upper thoracic axial level.

ACKNOWLEDGMENTS

This work was supported by grants from the Canadian Institutes of Health Research (MOP-15139; to L.J.) and the National Sciences and Engineering Research Council of Canada (194559; to L.J.). We thank Drs J. Charron, J. Deschamps, M. Featherstone, D. Lohnes and C. Largman for sharing antibodies and expressing vectors. A special thank to Marcelle Carter for her skilled technical assistance.

REFERENCES

- Ahn Y, Mullan HE, Krumlauf R. 2014. Long-range regulation by shared retinoic acid response elements modulates dynamic expression of posterior Hoxb genes in CNS development. *Dev Biol.* pii: S0012-1606(14)00066-9.
- Andrey G, Montavon T, Mascrez B, Gonzalez F, Noordermeer D, Leleu M, Trono D, Spitz F, Duboule D. 2013. A switch between topological domains underlies HoxD genes collinearity in mouse limbs. *Science* 340(6137):1234167.
- Aparicio O, Geisberg JV, Sekinger E, Yang A, Moqtaderi Z, Struhl K. 2005. Chromatin immunoprecipitation for determining the association of proteins with specific genomic sequences *in vivo*. *Current Protocols in Molecular Biology* 21.3.1-21.3.33.
- Aubin J, Chailier P, Ménard D, Jeannotte L. 1999. Loss of Hoxa5 gene function in mice perturbs intestinal maturation. *Am J Physiol.* 277:C965-973.

- Aubin J, Déry U, Lemieux M, Chailier P, Jeannotte L. 2002a. Stomach regional specification requires Hoxa5-driven mesenchymal-epithelial signaling. *Development*. 129:4075-4087.
- Aubin J, Lemieux M, Moreau J, Lapointe J, Jeannotte L. 2002b. Cooperation of Hoxa5 and Pax1 genes during formation of the pectoral girdle. *Dev Biol*. 244:96-113.
- Aubin J, Lemieux M, Tremblay M, Behringer RR, Jeannotte L. 1998. Transcriptional interferences at the Hoxa4/Hoxa5 locus: importance of correct Hoxa5 expression for the proper specification of the axial skeleton. *Dev Dyn*. 212:141-156.
- Aubin J, Lemieux M, Tremblay M, Bérard J, Jeannotte L. 1997. Early postnatal lethality in Hoxa-5 mutant mice is attributable to respiratory tract defects. *Dev Biol*. 192:432-445.
- Bachiller D, Macias A, Duboule D, Morata G. 1994. Conservation of a functional hierarchy between mammalian and insect Hox/HOM genes. *EMBO J* 13:1930-1941.
- Bérubé-Simard FA, Jeannotte L. 2014. Hoxa5/Cre transgenic mice: Novel tools for regional deletion along the anterior-posterior axis. *Genesis* 52(2). doi: 10.1002/dvg.22733.
- Boucherat O, Montaron S, Bérubé-Simard FA, Aubin J, Philippidou P, Wellik DM, Dasen JS, Jeannotte L. 2013. Partial functional redundancy between *Hoxa5* and *Hoxb5* paralog genes during lung morphogenesis. *Am J Physiol Lung Cell Mol Physiol* 304, L817-L830.
- Brend T, Gilthorpe J, Summerbell D, Rigby PW. 2003. Multiple levels of transcriptional and post-transcriptional regulation are required to define the domain of Hoxb4 expression. *Development* 130:2717-2728.
- Burke AC, Nelson CE, Morgan BA, Tabin C. 1995. Hox genes and the evolution of vertebrate axial morphology. *Development* 121:333-346.
- Capovilla M, Botas J. 1998. Functional dominance among Hox genes: repression dominates activation in the regulation of Dpp. *Development* 125:4949-4957.
- Caronia G, Goodman FR, McKeown CM, Scambler PJ, Zappavigna V (2003) An I47L substitution in the HOXD13 homeodomain causes a novel human limb malformation by producing a selective loss of function. *Development* 130, 1701–1712.
- Chambeyron S, Bickmore WA. 2004. Chromatin decondensation and nuclear reorganization of the HoxB locus upon induction of transcription. *Genes Dev* 18:1119-1130.
- Chen F, Capecchi MR. 1997. Targeted mutations in hoxa-9 and hoxb-9 reveal synergistic interactions. *Dev Biol* 181:186-196.
- Coulombe Y, Lemieux M, Moreau J, Aubin J, Joksimovic M, Bérubé-Simard FA, Tabariès S, Boucherat O, Guillou F, Larochelle C, Tuggle CK, Jeannotte L. 2010. Multiple promoters and alternative splicing: *Hoxa5* transcriptional complexity in the mouse embryo. *PLoS ONE* 5(5): e10600. doi:10.1371/journal.pone.0010600.
- Davidson AJ, Zon LI. 2006. The caudal-related homeobox genes *cdx1a* and *cdx4* act redundantly to regulate hox gene expression and the formation of putative hematopoietic stem cells during zebrafish embryogenesis. *Dev Biol*. 292(2):506-18.
- Deschamps J, van Nes J. 2005. Developmental regulation of the Hox genes during axial morphogenesis in the mouse. *Development*. 132:2931-2942.

- Di-Poï N, Zákány J, Duboule D. 2007. Distinct roles and regulations for HoxD genes in metanephric kidney development. *PLoS Genet.* 3(12):e232.
- Dony C, Gruss P. 1987. Specific expression of the Hox 1.3 homeo box gene in murine embryonic structures originating from or induced by the mesoderm. *EMBO J* 6:2965-2975.
- Duboule D. 1991. Patterning in the vertebrate limb. *Curr Opin Genet Dev* 1:211-216.
- Duboule D, Morata G. 1994. Colinearity and functional hierarchy among genes of the homeotic complexes. *Trends Genet* 10:358-364.
- Dupé V, Davenne M, Brocard J, Dollé P, Mark M, Dierich A, Chambon P, Rijli FM. 1997. In vivo functional analysis of the Hoxa-1 3' retinoic acid response element (3'RARE). *Development* 124:399-410.
- Ferretti E, Cambronero F, Tumpel S, Longobardi E, Wiedemann L, M Blasi F, Krumlauf R. (2005). Hoxb1 enhancer and control of rhombomere 4 expression: complex interplay between PREP1-PBX1-HOXB1 binding sites. *Mol Cell Biol* 25, 8541-52.
- Gamer LW, Wright CV. 1993. Murine Cdx-4 bears striking similarities to the Drosophila caudal gene in its homeodomain sequence and early expression pattern. *Mech Dev* 43:71-81.
- Garcia-Fernandez J. 2005. The genesis and evolution of homeobox gene clusters. *Nat Rev Genet.* 6:881-892.
- Garin E, Lemieux M, Coulombe Y, Robinson GW, Jeannotte L. 2006. Stromal *Hoxa5* function controls the growth and differentiation of mammary alveolar epithelium. *Dev Dyn* 235:1858-1871.
- Gaunt SJ, Coletta PL, Pravtcheva D, Sharpe PT. 1990. Mouse Hox-3.4: homeobox sequence and embryonic expression patterns compared with other members of the Hox gene network. *Development* 109:329-339.
- Gaunt SJ, Drage D, Trubshaw RC. 2005. cdx4/lacZ and cdx2/lacZ protein gradients formed by decay during gastrulation in the mouse. *Int. J. Dev. Biol.* 49: 901-908..
- Gendronneau G, Boucherat O, Aubin J, Lemieux M, Jeannotte, L. (2012) The loss of *Hoxa5* function causes estrous acyclicity and ovarian epithelila inclusion cysts. *Endocrinology* 153:1484-1497.
- Gonzalez F, Duboule D, Spitz F. 2007. Transgenic analysis of Hoxd gene regulation during digit development. *Dev Biol.* 306(2): 847-59.
- Gonzalez-Reyes A, Morata G. 1990. The developmental effect of overexpressing a Ubx product in Drosophila embryos is dependent on its interactions with other homeotic products. *Cell.* 61:515-522.
- Hogan BM, Beddington RS, Costantini F, Lacy E. 1994. Manipulating the mouse embryo, a laboratory manual. Cold Spring Harbor Press 2nd edition.
- Izpisúa-Belmonte JC, Falkenstein H, Dollé P, Renucci A, Duboule D. 1991. Murine genes related to the Drosophila AbdB homeotic genes are sequentially expressed during development of the posterior part of the body. *EMBO J* 10:2279-2289.
- Jeannotte L, Lemieux M, Charron J, Poirier F, Robertson EJ. 1993. Specification of axial identity in the mouse: role of the Hoxa-5 (Hox1.3) gene. *Genes Dev.* 7:2085-2096.

- Kashyap V, Gudas LJ, Brenet F, Funk P, Viale A, Scandura JM. 2011. Epigenomic reorganization of the clustered Hox genes in embryonic stem cells induced by retinoic acid. *J Biol Chem.* 286(5):3250-60.
- Kessel M, Gruss P. 1991. Homeotic transformations of murine vertebrae and concomitant alteration of Hox codes induced by retinoic acid. *Cell* 67:89-104.
- Kmita M, Duboule D. 2003. Organizing axes in time and space; 25 years of colinear tinkering. *Science* 301:331-333.
- Krumlauf R. 1994. Hox genes in vertebrate development. *Cell* 78:191-201.
- Lafontaine CA, Grainger S, Hess BL, Béland M, Lohnes D. 2012. Cdx1 interacts physically with a subset of Hox proteins. *Biochemistry* 51(48):9698-705.
- Lamka ML, Boulet AM, Sakonju S. 1992. Ectopic expression of UBX and ABD-B proteins during *Drosophila* embryogenesis: competition, not a functional hierarchy, explains phenotypic suppression. *Development* 116:841-854.
- Lampe X, Picard JJ, Rezsöházy R. (2004). The Hoxa2 enhancer 2 contains a critical Hoxa2 responsive regulatory element. *Biochem Biophys Res Commun* 316, 898-902.
- Larochelle C, Tremblay M, Bernier D, Aubin J, Jeannotte L. 1999. Multiple cis-acting regulatory regions are required for restricted spatio-temporal Hoxa5 gene expression. *Dev Dyn.* 214:127-140.
- LaRonde-LeBlanc NA, Wolberger C. 2003. Structure of HoxA9 and Pbx1 bound to DNA: Hox hexapeptide and DNA recognition anterior to posterior. *Genes Dev* 17:2060-2072.
- Lewis EB. 1978. A gene complex controlling segmentation in *Drosophila*. *Nature.* 276:565-570.
- Li L, Liu B, Wapinski OL, Tsai MC, Qu K, Zhang J, Carlson JC, Lin M, Fang F, Gupta RA, Helms JA, Chang HY. (2013) Targeted disruption of Hotair leads to homeotic transformation and gene derepression. *Cell Rep.* 5(1):3-12.
- Lohnes D. 2003. The Cdx1 homeodomain protein: an integrator of posterior signaling in the mouse. *Bioessays* 25:971-980.
- Mahony S, Mazzoni EO, McCuine S, Young RA, Wichterle H, Gifford DK. 2011. Ligand dependent dynamics of retinoic acid receptor binding during early neurogenesis. *Genome Biol.*12(1):R2
- Mainguy G, Koster J, Woltering J, Jansen H, Durston A. (2007) Extensive Polycistronism and Antisense Transcription in the Mammalian Hox Clusters. *PLoS One* 2(4): e356.
- Mann RS, Hogness DS. 1990. Functional dissection of Ultrabithorax proteins in *D. melanogaster*. *Cell* 60:597-610.
- Manzanares M, Bel-Vialar S, Ariza-McNaughton L, Ferretti E, Marshall H, Maconochie MM, Blasi F, Krumlauf R. 2001. Independent regulation of initiation and maintenance phases of Hoxa3 expression in the vertebrate hindbrain involve auto- and cross-regulatory mechanisms. *Development* 128:3595-3607.
- Mazzoni EO, Mahony S, Peljto M, Patel T, Thornton SR, McCuine S, Reeder C, Boyer LA, Young RA, Gifford DK, Wichterle H. 2013. Saltatory remodeling of Hox chromatin in response to rostrocaudal patterning signals. *Nat Neurosci.*16(9):1191-8.
- Meunier D, Aubin J, Jeannotte L. 2003. Perturbed thyroid morphology and transient hypothyroidism symptoms in Hoxa5 mutant mice. *Dev Dyn.* 227:367-378.

- Montavon T, Soshnikova N, Mascrez B, Joye E, Thevenet L, Splinter E, de Laat W, Spitz F, Duboule D. (2011) A regulatory archipelago controls Hox genes transcription in digits. *Cell* 147(5):1132-45.
- Moreau J, Jeannotte L. 2002. Sequence analysis of a Hoxa4-Hoxa5 intergenic region including shared regulatory elements. *DNA Seq.* 13:203-209.
- Morey C, Da Silva NR, Perry P, Bickmore WA. 2007. Nuclear reorganisation and chromatin decondensation are conserved, but distinct, mechanisms linked to Hox gene activation. *Development* 134: 909–919.
- Morrison A, Ariza-McNaughton L, Gould A, Featherstone M, Krumlauf R. 1997. HOXD4 and regulation of the group 4 paralog genes. *Development* 124:3135-3146.
- Noordermeer D, Leleu M, Splinter E, Rougemont J, De Laat W, Duboule D. 2011. The Dynamic Architecture of Hox Gene Clusters. *Science* 334(6053):222-5.
- Nowling T, Zhou W, Krieger KE, Larochelle C, Nguyen-Huu MC, Jeannotte L, Tuggle CK. 1999. Hoxa5 gene regulation: A gradient of binding activity to a brachial spinal cord element. *Dev Biol.* 208:134-146.
- Oosterveen T, van Vliet P, Deschamps J, Meijlink F. 2003. The direct context of a hox retinoic acid response element is crucial for its activity. *J Biol Chem* 278:24103-24107.
- Peljto M, Dasen JS, Mazzoni EO, Jessell TM, Wichterle H. 2010. Functional diversity of ESC-derived motor neuron subtypes revealed through intraspinal transplantation. *Cell Stem Cell* 7, 355–366.
- Rousseau M, Crutchley JL, Miura H, Suderman M, Blanchette M, Dostie J. 2014. Hox in motion: tracking HoxA cluster conformation during differentiation. *Nucleic Acids Res.* 42(3):1524-40.
- Rouso DL, Gaber ZB, Wellik D, Morrisey EE, Novitch BG. 2008. Coordinated Actions of the Forkhead Protein Foxp1 and Hox Proteins in the Columnar Organization of Spinal Motor Neurons. *Neuron* 59(2):226-40.
- Sambrook J, Fritsch EF, Maniatis T. 1989. Molecular cloning, A laboratory manual. New York: Cold Spring Harbour Laboratory Press.
- Sharpe J, Nonchev S, Gould A, Whiting J, Krumlauf R. 1998. Selectivity, sharing and competitive interactions in the regulation of Hoxb genes. *EMBO J* 17:1788-1798..
- Sheth R, Bastida MF, Kmita M, Ros M. (2013) ‘Self-Regulation’, a new facet of Hox genes’ function. *Dev Dyn.* doi: 10.1002/dvdy.24019.
- Stein S, Fritsch R, Lemaire L, Kessel M. 1996. Checklist: vertebrate homeobox genes. *Mech Dev* 55:91-108.
- Suh E, Chen L, Taylor J, Traber PG. 1994. A homeodomain protein related to caudal regulates intestine-specific gene transcription. *Mol Cell Biol* 14:7340-7351.
- Tabariès S, Lapointe J, Besch T, Carter M, Woollard J, Tuggle CK, Jeannotte L. 2005. Cdx protein interaction with Hoxa5 regulatory sequences contributes to Hoxa5 regional expression along the axial skeleton. *Mol Cell Biol.* 25:1389-1401.
- Tschopp P, Christen AJ, Duboule D. 2012. Bimodal control of Hoxd gene transcription in the spinal cord defines two regulatory subclusters. *Development* 139(5):929-39.
- Tschopp P, Tarchini B, Spitz F, Zakany J, Duboule D. 2009. Uncoupling Time and Space in the Collinear Regulation of Hox Genes. *PLoS Genet.* 5(3):e1000398.

- Tuggle CK, Zákány J, Cianetti L, Peschle C, Nguyen-Huu MC. 1990. Region-specific enhancers near two mammalian homeo box genes define adjacent rostrocaudal domains in the central nervous system. *Genes Dev* 4:180-189.
- van den Akker E, Forlani S, Chawengsaksophak K, de Graaff W, Beck F, Meyer BI, Deschamps J. 2002. Cdx1 and Cdx2 have overlapping functions in anteroposterior patterning and posterior axis elongation. *Development* 129: 2181-2193.
- van Nes J, de Graaff W, Lebrin F, Gerhard M, Beck F, Deschamps J. 2006. The Cdx4 mutation affects axial development and reveals an essential role of Cdx genes in the ontogenesis of the placental labyrinth in mice. *Development* 133:419-428.
- van de Ven C, Bialecka M, Neijts R, Young T, Rowland JE, Stringer EJ, Van Rooijen C, Meijlink F, Nóvoa A, Freund JN, Mallo M, Beck F, Deschamps J. 2011. Concerted involvement of Cdx/Hox genes and Wnt signaling in morphogenesis of the caudal neural tube and cloacal derivatives from the posterior growth zone. *Development* 138(16):3451-62.
- Wacker SA, Jansen HJ, McNulty C, Houtzager E, Durston AJ (2004) Timed interactions between the Hox expressing non-organiser mesoderm and the Spemann organiser generate positional information during vertebrate gastrulation. *Dev. Biol.* 268, 207–219.
- Wang W, Malcolm BA. 1999. Two-stage PCR protocol allowing introductions of multiple mutation deletions and insertions using quick-change site-directed mutagenesis. *Biotechniques* 26:680-682.
- Williams ME, Lehoczky JA, Innis JW (2006) A group 13 homeodomain is neither necessary nor sufficient for posterior prevalence in the mouse limb. *Dev Biol.* 297(2): 493-507.
- Williams TM, Williams ME, Innis JW (2005) Range of HOX/TALE superclass associations and protein domain requirements for HOXA13: MEIS interaction. *Dev. Biol.* 277, 457–471.
- Woltering JM, Durston A (2008) MiR10 represses HoxB1a and HoxB3a in Zebrafish. *PLoS ONE* 3(1): e1396.
- Wong EY, Wang XA, Mak SS, Sae-Pang JJ, Ling KW, Fritsch B, Sham MH. (2011) Hoxb3 negatively regulates Hoxb1 expression in mouse hindbrain patterning. *Dev. Biol.* 352(2): 382-92.
- Würtele H, Chartrand P. 2006. Genome-wide scanning of HoxB1-associated loci in mouse ES cells using an open-ended Chromosome Conformation Capture methodology. *Chromosome Res.* 14(5): 477-95.
- Yekta S, Shih H, Bartel DP (2004) MicroRNA-directed cleavage of HOXB8 mRNA. *Science* 304 (5670): 594–596.
- Yekta S, Tabin CJ., Bartel DP (2008) MicroRNAs in the Hox network: an apparent link to posterior prevalence. *Nat. Rev. Genet.* 9(10): 789–796.
- Zákány J, Gérard M, Favier B, Duboule D. 1997. Deletion of a HoxD enhancer induces transcriptional heterochrony leading to transposition of the sacrum. *EMBO J* 16:4393-4402.
- Zakany J, Kmita M, Duboule D. 2004. A dual role for Hox genes in limb anterior–posterior asymmetry. *Science* 304, 1669-1672.

Zhao Y, Potter SS. 2002. Functional comparison of the Hoxa 4, Hoxa 10, and Hoxa 11 homeoboxes. *Dev Biol* 244:21-36.

CHAPITRE V

Multiple Promoters and Alternative Splicing: Hoxa5 Transcriptional Complexity in the Mouse Embryo.

Yan Coulombe¹, Margot Lemieux¹, Julie Moreau¹, Josée Aubin¹, Milan Joksimovic², Félix-Antoine Bérubé-Simard¹, Sébastien Tabariès¹, Olivier Boucherat¹, François Guillou¹, Christian Larochelle¹, Christopher K. Tuggle², Lucie Jeannotte^{1*}

1 Centre de recherche en cancérologie de l'Université Laval, Centre Hospitalier Universitaire de Québec, L'Hôtel-Dieu de Québec, Québec, Québec, Canada,

2 Department of Animal Science, Iowa State University, Ames, Iowa, United States of America

Article publié en 2010 dans : *PLOS ONE* **5**(5): e10600

AVANT-PROPOS

Lors du développement embryonnaire, il est nécessaire que le principe de colinéarité d'expression des gènes *Hox* soit respecté afin que l'intégrité du plan d'organisation normal de l'embryon soit maintenue. En d'autres termes, l'organisation génomique au sein des complexes doit être conservée pour que la régulation spatio-temporelle des gènes *Hox* soit orchestrée précisément. Le gène *Hoxa5* ne fait pas exception à la règle, puisque son locus englobe plusieurs transcrits qui se chevauchent et qui ont été mis en évidence grâce à des expériences de buvardage Northern, de 5'RACE, de 3' RACE et de RT-PCR. L'unité transcriptionnelle du gène *Hoxa5* génère 5 différents transcrits susceptibles de coder pour la protéine HOXA5. Premièrement, on retrouve le transcrit majeur de 1.8 kb, initié à partir du promoteur proximal (P) du gène *Hoxa5* et qui porte les deux exons du gène. Ensuite, trois longs transcrits polyadénylés de 5.0, 9.5 et 11.0 kb sont initiés à partir d'un promoteur distal (D2) situé dans les séquences en aval du gène *Hoxa7*. Enfin, un transcrit de 5.0 kb initié à partir du promoteur du gène *Hoxa6* (D1) pourrait lui aussi être bicistronique et coder à la fois pour les protéines HOXA6 et HOXA5. De plus, un autre codon ATG, situé à 333 pb en amont du promoteur proximal du gène *Hoxa5*, est théoriquement en mesure de générer une protéine HOXA5 allongée (381 aa).

À la lumière de ces informations, nous avons entrepris la caractérisation de ces différents longs transcrits afin de comprendre leurs rôles dans la fonction et la régulation de l'expression des gènes *Hox*. L'efficacité des deux promoteurs distaux (D1 et D2) à initier une expression de type "Hox-like" a été confirmée par transgénèse au sein de constructions comprenant le gène rapporteur *lacZ* de même que l'élément de régulation MES (+2.9 kb à +4.9 kb par rapport au promoteur proximal du gène *Hoxa5*). Ensuite, des expériences d'hybridation *in situ* de l'ARN ont permis de constater que l'expression des différents transcrits respecte le principe de colinéarité, puisqu'une sonde reconnaissant seulement les longs transcrits dévoile un profil d'expression plus postérieur que celle reconnaissant tous les transcrits du gène *Hoxa5*. Par analyse différentielle, le profil d'expression du transcrit majeur de 1.8 kb semble correspondre à la partie brachiale du tube neural, aux dérivés mésodermiques entre pv3 et pv10 de même qu'à la ceinture scapulaire, ce qui correspond

aux sites dans lesquels la protéine HOXA5 est présente, tel qu'observé par immunofluorescence, ce qui suggère qu'il s'agit du transcrit fonctionnel du gène *Hoxa5*.

Pour ma part, j'ai voulu vérifier la capacité des différents transcrits à générer *in vitro* et *in vivo* les protéines HOXA5 ainsi que HOXA6. J'ai tout d'abord participé à l'élaboration des différentes constructions d'ADN consistant à l'apposition de l'étiquette MYC au niveau de la protéine HOXA5 et de l'étiquette FLAG au niveau de la protéine HOXA6. Ensuite, tirant avantage du système de transcription/traduction de protéines dans les réticulocytes de lapin, j'ai démontré le potentiel de tous les transcrits testés à générer la protéine HOXA5, de même que la protéine HOXA6 à partir des transcrits le permettant (Fig. 5.7A). Toutefois, lorsque transfectés dans des cellules HEK293T, seulement le transcrit de 1.8 kb est en mesure de produire la protéine HOXA5, alors que seul le transcrit de 5.0 kb initié à partir du promoteur du gène *Hoxa6* (D1) produit la protéine HOXA6 (Fig. 5.7B-C).

Enfin, la perturbation des longs transcrits engendrée par l'insertion de la cassette de sélection à la néomycine chez les individus *Hoxa6*^{-/-} ne semble pas avoir d'impact sur la spécification du squelette axial. Par contre, la diminution de l'expression du transcrit majeur de 1.8 kb chez les souris *Hoxa6*^{-/-} suggère soit une interférence transcriptionnelle liée à la cassette de sélection, soit des propriétés régulatrices propres aux longs transcrits. Cette dernière hypothèse est présentement testée dans le laboratoire.

ABSTRACT

Background: The genomic organization of *Hox* clusters is fundamental for the precise spatio-temporal regulation and the function of each *Hox* gene, and hence for correct embryo patterning. Multiple overlapping transcriptional units exist at the *Hoxa5* locus reflecting the complexity of *Hox* clustering: a major form of 1.8 kb corresponding to the two characterized exons of the gene and polyadenylated RNA species of 5.0, 9.5 and 11.0 kb. This transcriptional intricacy raises the question of the involvement of the larger transcripts in *Hox* function and regulation.

Methodology/Principal Findings: We have undertaken the molecular characterization of the *Hoxa5* larger transcripts. They initiate from two highly conserved distal promoters, one corresponding to the putative *Hoxa6* promoter, and a second located nearby *Hoxa7*. Alternative splicing is also involved in the generation of the different transcripts. No functional polyadenylation sequence was found at the *Hoxa6* locus and all larger transcripts use the polyadenylation site of the *Hoxa5* gene. Some larger transcripts are potential *Hoxa6/Hoxa5* bicistronic units. However, even though all transcripts could produce the genuine 270 a.a. HOXA5 protein, only the 1.8 kb form is translated into the protein, indicative of its essential role in *Hoxa5* gene function. The *Hoxa6* mutation disrupts the larger transcripts without major phenotypic impact on axial specification in their expression domain. However, *Hoxa5*-like skeletal anomalies are observed in *Hoxa6* mutants and these defects can be explained by the loss of expression of the 1.8 kb transcript. Our data raise the possibility that the larger transcripts may be involved in *Hoxa5* gene regulation.

Significance: Our observation that the *Hoxa5* larger transcripts possess a developmentally-regulated expression combined to the increasing sum of data on the role of long noncoding RNAs in transcriptional regulation suggest that the *Hoxa5* larger transcripts may participate in the control of *Hox* gene expression.

INTRODUCTION

Hox genes play a crucial role in specifying regional identity along the body axes and in regulating morphogenesis during animal development. Inappropriate expression and mutation of *Hox* genes can disrupt normal programs of growth and differentiation leading to malformations, tumor formation, and even death [1]. In mammals, 39 *Hox* genes are distributed over four clusters sharing a similar organization that reflects the relationship existing between the relative position of each *Hox* gene along the cluster, its expression domain in the embryo and its temporal onset. *Hox* genes have RNA expression domains extending from the caudal end of the embryo to a defined anterior limit. The resulting spatio-temporal profile of *Hox* gene expression during embryogenesis correlates with the arrangement of the clusters: the 3' most genes being expressed earlier and in more anterior domains than the 5' located ones [2]. Consequently, the clustered organization appears fundamental for the precise spatio-temporal regulation and the function of each Hox gene and hence for the correct patterning of the embryo.

How *Hox* gene expression is modulated along the developing axes still remains elusive. Our initial knowledge of the regulatory mechanisms governing *Hox* gene expression comes mostly from transgenic mice studies, which have shown that *Hox* dynamic expression patterns result from positional information transducing via transcription factors that interact with a combination of positive and negative cis-acting sequences to differentially control *Hox* gene expression in a spatio-temporal and tissue-specific fashion. However in most cases, only limited subsets of the proper spatial and temporal expression patterns are reconstituted by the transgenes. A likely explanation is the presence of complex and overlapping transcriptional units in *Hox* genes that implies dispersed regulatory regions in the clusters [3–5]. There is also evidence for the integrated regulation of neighboring *Hox* genes through the sharing, the competition and/or the selective use of defined cis-acting sequences [6–8]. Moreover, global enhancer sequences located outside the *Hox* clusters can coordinate the expression of several genes in a relatively promoter-unspecific manner [9–12]. Finally, large-scale chromatin remodeling events participate to the regulation of *Hox* loci [13,14].

Hox RNAs and HOX proteins can colocalize, which reinforces the notion that transcriptional control is a primary mechanism for *Hox* gene regulation. However in some instances, HOX proteins are detected in a subdomain of the RNA pattern suggesting the existence of post-transcriptional control [8,15]. The discovery of microRNAs that can mediate the targeted degradation of specific *Hox* transcripts has unveiled an additional level of regulation of *Hox* gene expression [16,17]. In addition, antisense transcripts and long noncoding RNAs are found throughout *Hox* clusters and they are proposed to be part of the epigenetic regulation of *Hox* gene expression [18–22]. Altogether, these data indicate that a complex array of different modes of regulation is essential for the proper spatio-temporal *Hox* gene expression.

To fully understand the regulatory events governing *Hox* gene expression, we are using as a model the *Hoxa5* gene. This gene plays a crucial role during embryogenesis as well as being involved in tumorigenesis [23–25]. In the developing embryo, *Hoxa5* is expressed in the neural tube caudal to the posterior myelencephalon, in the axial skeleton up to the level of prevertebra (pv) 3 and in the mesenchymal component of several organs, including the trachea, the lung, the stomach, the intestine and the kidneys [26–31]. We have shown that the loss of *Hoxa5* function in the mouse affects a well-defined subset of structures mainly located at the cervico-thoracic level [23,26,27,32,33]. Aside from morphological defects in foregut derivatives and mammary glands [26,29,34,35], the targeted disruption of the *Hoxa5* gene perturbs axial skeleton identity between pv3 and pv10, the anterior-most region of the *Hoxa5* domain of expression along the prevertebral axis [23,27].

Polyadenylated transcripts of 1.8, 5.0, 9.5 and 11.0 kb in length encompassing *Hoxa5* coding sequences are produced in the embryo. They are also detected after birth in a tissue-specific fashion [23,36]. The 1.8 kb transcript is the most abundant and it corresponds to the two characterized exons of the *Hoxa5* gene [37]. It encodes the 270 amino acid (a.a.) HOXA5 protein. Previous RNAase protection assays have shown that the larger forms initiate more upstream from sequences that remain to be identified [37]. Differences in the expression profile of these different transcripts are also observed: the 1.8 kb transcript is expressed as early as embryonic day (E) 8.0–8.25, whereas the larger

transcripts are first detected around E8.5–8.75 [32]. The larger transcripts are present in more posterior structures of the embryo with an anterior limit of expression in the pv column corresponding to pv10, while that of the 1.8 kb transcript is pv3. In the neural tube, a posterior shift was observed for the larger transcripts [32]. Similarities between the *Hoxa7* expression profile and that of the larger transcripts indicate that they may share regulatory mechanisms [38].

The presence of multiple overlapping transcriptional units at the *Hoxa5* locus suggests that *Hoxa5* gene regulation may be complex. Using a transgenic approach, we have shown that several DNA control elements located both upstream and downstream the *Hoxa5* coding sequences are involved in the expression of the 1.8 kb transcript [32,39–41]. An intricate situation prevails as some of these regulatory sequences are shared with the flanking *Hoxa4* gene, while others overlap with the *Hoxa6* coding sequences [32,39,42]. The presence of larger transcripts encompassing the *Hoxa5* coding sequences also implies that more DNA regions involved in *Hoxa5* gene regulation may be distributed along the cluster.

To assess the importance of the *Hoxa5* larger transcripts in *Hoxa5* gene function and regulation and to eventually define how they integrate in the developmental program, we have undertaken their molecular characterization. Our data revealed the complex organization of the different transcriptional units encompassing the *Hoxa5* and *Hoxa6* loci. It results from the use of three specific promoters and alternative splicing. Even though all these transcripts can potentially produce the genuine 270 a.a. HOXA5 protein, only the 1.8 kb form appears to generate the protein. Furthermore, the *Hoxa5* functional domain along the embryonic axis coincides with the expression region of the protein, where the larger transcripts are excluded. This pinpoints at the 1.8 kb form as the *Hoxa5* functional transcript in regional specification and leaves opened a role for the larger transcripts as long noncoding RNAs.

RESULTS

Molecular characterization of the *Hoxa5* alternate transcripts

Previous northern analysis of polyA⁺ RNA from mouse embryo using a DNA probe corresponding to the 3'-untranslated region of the second exon of the *Hoxa5* gene has shown that polyadenylated transcripts of approximately 1.8, 5.0, 9.5 and 11.0 kb in length contain sequences from the *Hoxa5* locus [23]. The 1.8 kb transcript corresponds to the putative *Hoxa5* transcript as demonstrated from cDNA sequence analyses [36,37]. We thus aimed to determine the molecular origin of the larger transcripts. To do so, we applied a series of molecular approaches, and by merging all the data obtained from northern, 3'- and 5'-RACE, RT-PCR and cDNA analyses, we established a schematic representation of the major *Hoxa5* transcripts produced in the e12.5 mouse embryo (Fig. 5.1).

First, we performed northern analyses with E12.5 mouse embryo polyA⁺ RNA using as antisense riboprobes several genomic fragments encompassing *Hoxa5* and flanking Hox genes (Fig. 5.2). The RNA was obtained from wild-type (wt), *Hoxa5*^{-/-} and *Hoxa6*^{-/-} embryos. We took advantage of the mutant forms of the *Hoxa5* and *Hoxa6* transcripts produced in *Hoxa5*^{-/-} and *Hoxa6*^{-/-} mice, respectively. These mutant transcripts are 1 kb larger than the endogenous ones due to the insertion of a 1 kb neo cassette into the homeobox sequence of each gene [23,43,44]. This difference in length allowed us to distinguish the alternate *Hoxa5* transcripts among the several products detected. Northern analyses of polyA⁺ RNA from wt and *Hoxa5*^{-/-} embryos revealed that the 1.8, 5.0, 9.5 and 11.0 kb transcripts contained the two *Hoxa5* exons and they were all affected by the insertion of the neo cassette into the *Hoxa5* mutant allele as previously shown (probes 13 and 16; Fig. 5.2) [23]. Moreover, these sense transcripts were all transcribed from the same DNA strand. The 1.8 kb transcript corresponded to the two *Hoxa5* exons. As shown by probes 2 to 12, the 5.0, 9.5 and 11.0 kb RNA species initiated in the *Hoxa6*-*Hoxa7* intergenic region further upstream from the identified 1.8 kb transcript start site (position +1; Figs. 5.1 and 5.2) [37]. These larger transcripts contained the *Hoxa6* sequences and they showed the expected shift in size in *Hoxa6*^{-/-} RNA due to the presence of the neo cassette [44]. A complex splicing pattern also prevailed explaining the difference

in length between the 5.0, 9.5 and 11.0 kb transcripts. The *Hoxa6* intron sequences were only detected in the 11.0 kb transcript (probe 8; Fig. 5.2), while some *Hoxa6-Hoxa7* intergenic sequences did not hybridize to the 5.0 kb band (probes 2, 4 and 5; Fig. 5.2). Faint bands approximately 1 kb larger than the expected transcripts were also distinguished in the wt specimens with probes 14 and 15 that correspond to *Hoxa5* intron sequences, suggesting that transcripts with unspliced *Hoxa5* intron sequences may exist at low abundance (Fig. 5.2). Additional bands of about 2.5 kb in length were detected with the *Hoxa6-Hoxa7* intergenic probe 3. Their origin was not investigated but they could correspond to RNA species initiating in the *Hoxa6-Hoxa7* intergenic region that skip the *Hoxa6* and *Hoxa5* loci to continue further downstream in the *Hoxa4-Hoxa5* sequence, like the GenBank mRNA AK051552, or extend towards the vicinity of the *Hoxa3* gene, as the Y11717 mRNA (GenBank). Finally, all 1.8, 5.0, 9.5 and 11.0 kb transcripts ended at the same polyA site at the 3' end of *Hoxa5* exon 2 (position +2637) as revealed by 3'-RACE (Rapid Amplification of cDNA Ends) experiments. This was also demonstrated by the lack of hybridization in northern analysis with a riboprobe located in genomic sequences 3' to the *Hoxa5* polyA site (Fig. 5.2; data not shown). To map the transcriptional start site of the larger *Hoxa5* transcripts, we designed a 5'-RACE strategy based on the data obtained from the northern analyses. We used different sets of primers specific either to the 9.5 and 11.0 kb transcripts (primer 1) or located in sequences shared by the 5.0, 9.5 and 11.0 kb transcripts (primers 2 and 3; Fig. S5.1A). Clones obtained with primer 1 indicated that the 9.5 and 11.0 kb transcripts initiate in the *Hoxa6-Hoxa7* intergenic region at position -8905 bp, which is about 2.3 kb downstream of the 3'-end of *Hoxa7* gene. With primers 2 and 3, 5'-RACE products revealed the presence of a 4.6 kb intron and an initiation site coinciding with that of the largest transcripts at position -8905 bp. A second population of clones was also obtained with primer 2 with a transcription start site at position -4409 bp, which corresponds to the putative first base of *Hoxa6* exon 1. Thus, two distal promoters, one related to the *Hoxa6* gene (promoter D1) and a more distal one located downstream the *Hoxa7* gene (promoter D2), participate in the production of the *Hoxa5* alternate transcripts (Fig. 5.1).

Figure 5.1. Schematic representation of the different transcripts encompassing *Hoxa5* sequences in the e12.5 mouse embryo.

Genomic organization of the *Hoxa5*, *Hoxa6* and *Hoxa7* genes along the HoxA cluster. Black, grey and open boxes indicate homeobox, translated and transcribed sequences, respectively. The two known exons of *Hoxa5* and the two in-frame ATG are represented. Position +1 corresponds to the transcription initiation site of *Hoxa5* exon 1. The 3' non-coding sequences of *Hoxa6* exon 2 extend further downstream into the *Hoxa6-Hoxa5* intergenic region and the adjacent *Hoxa5* coding sequences and they are indicated by dotted lines. The ATG of the putative HOXA6 protein is indicated. The promoters driving expression of the different transcripts are shown: proximal promoter, P; distal promoters D1 and D2. The transcripts are represented underneath based on northern, 5'-RACE, 3'-RACE and RT-PCR assays used to define their molecular structure. *Hoxa5* intron is represented by a dotted line to indicate the non-spliced isoforms. The longest ORFs deduced from the sequence of each transcript are represented by waved lines: the 270 a.a. HOXA5 protein, the 381 a.a. HOXA5 isoform and the HOXA6 protein. Ba, *Bam*HI; H, *Hind*III; RI, *Eco*RI; Xh, *Xho*I.

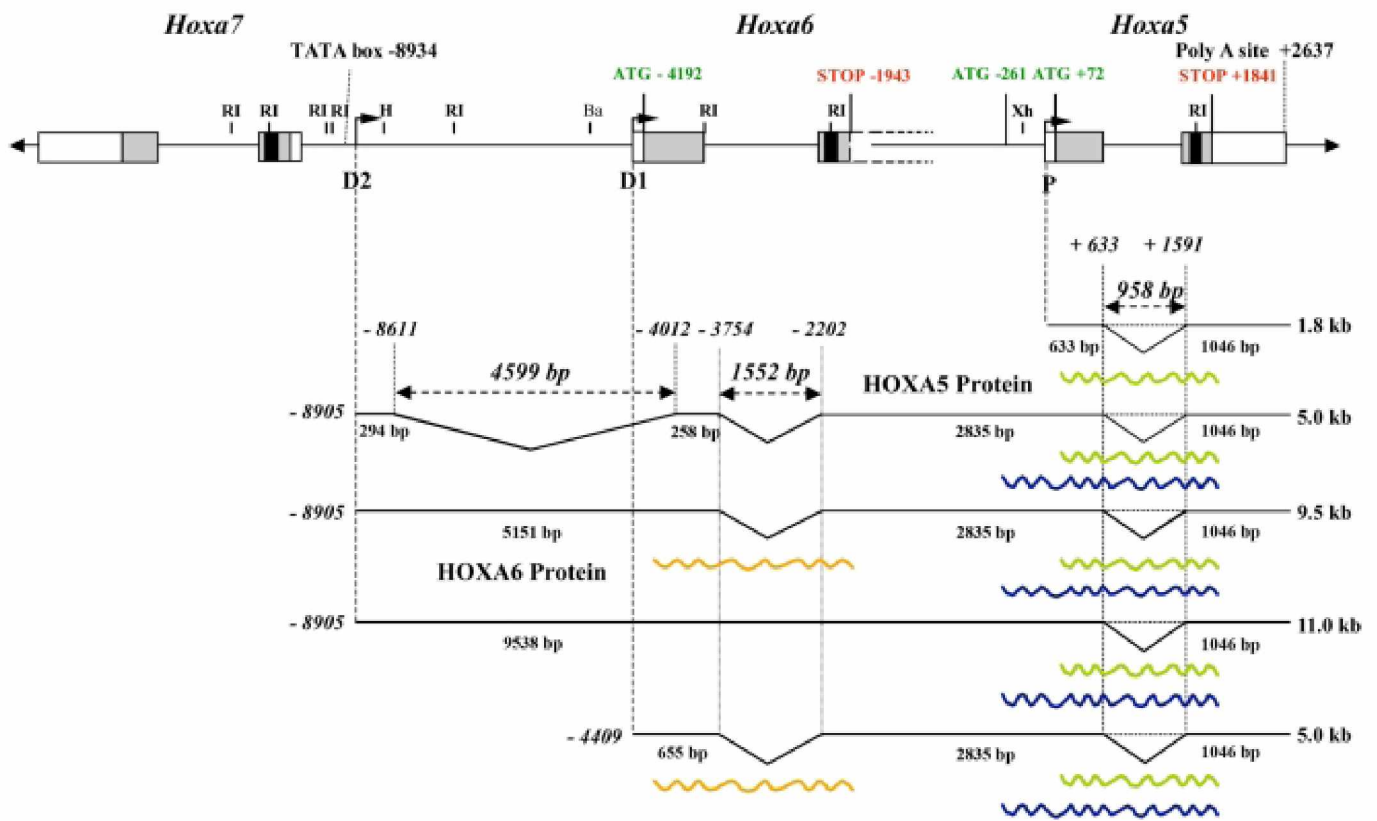
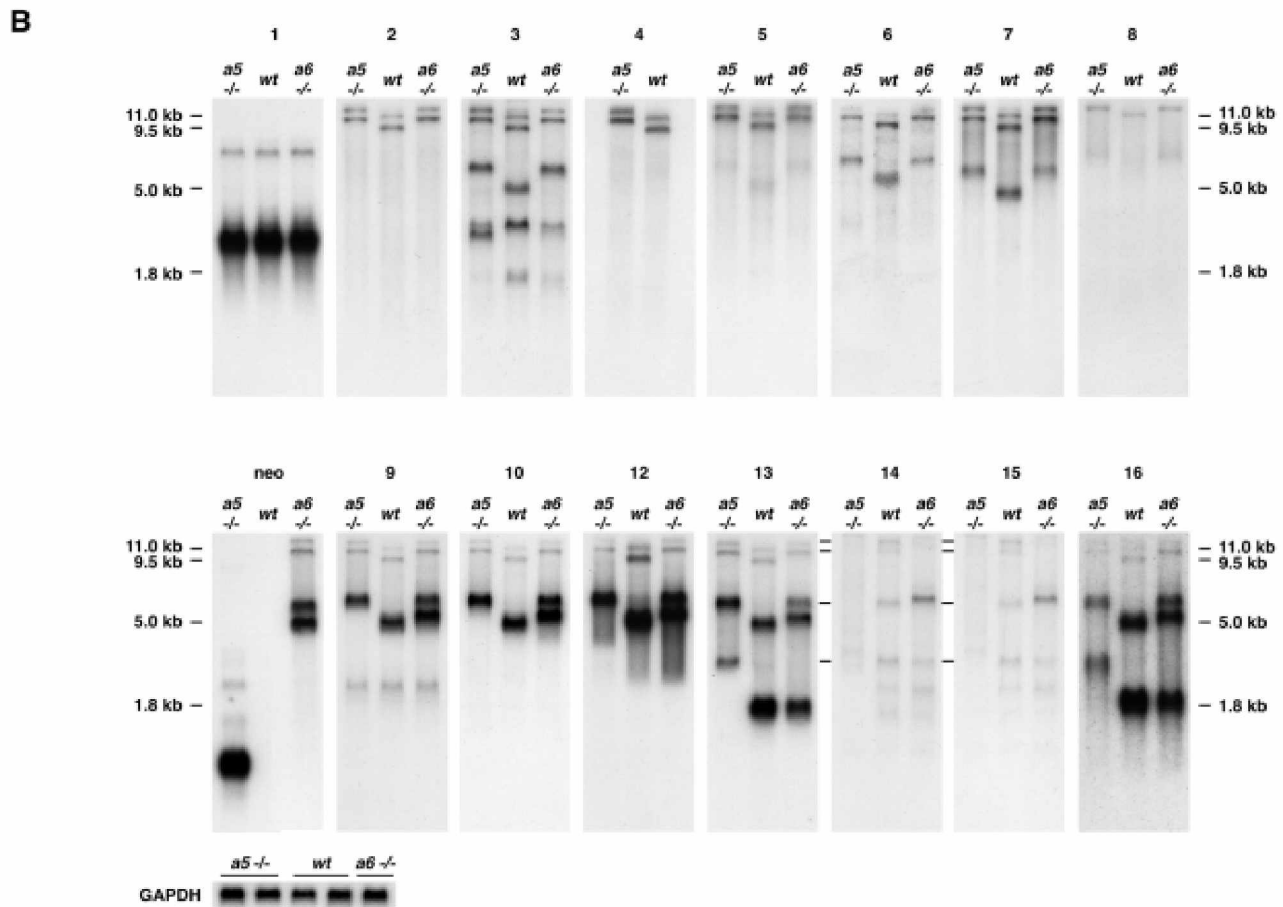
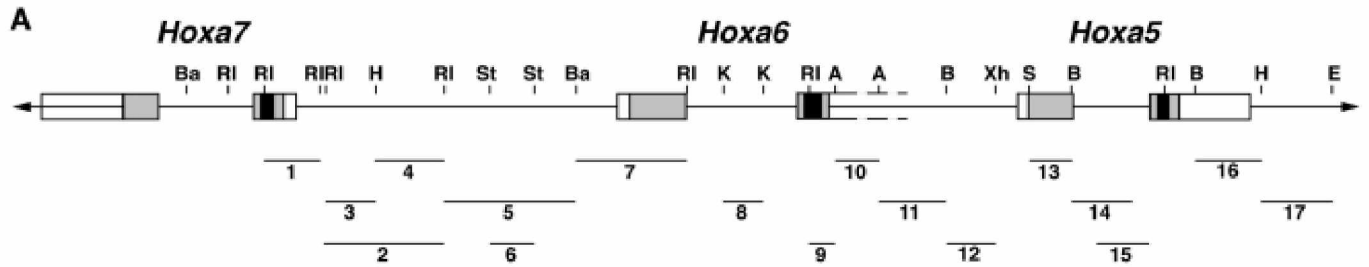


Figure 5.2. Molecular characterization of *Hoxa5* transcripts by northern analysis.

(A) Genomic organization of the *Hoxa5*, *Hoxa6* and *Hoxa7* genes along the cluster. Black, grey and open boxes indicate homeobox, translated, and transcribed sequences, respectively. Probes used for northern analyses are indicated by numbered lines (1–17).

(B) Northern analyses of polyA⁺ RNA from E12.5 wild-type, *Hoxa5*^{-/-} and *Hoxa6*^{-/-} mouse embryos were performed with probes covering the genomic region located between *Hoxa5* and *Hoxa7*. The mutated form of the *Hoxa5* transcripts produced in *Hoxa5*^{-/-} mice are 1 kb larger than the endogenous ones due to the insertion of the neo cassette into the homeobox sequence. A similar situation prevails for the *Hoxa6* mutation where the insertion of the neo cassette in the *Hoxa6* homeobox sequence also disrupts all transcripts encompassing the *Hoxa5* and *Hoxa6* loci, except the 1.8 kb form. The 5.0, 9.5 and 11.0 kb transcripts contain *Hoxa5* exons 1 and 2 sequences but initiated further upstream of the known 1.8 kb transcript start site in the *Hoxa6-Hoxa7* intergenic region. The difference in length between the 9.5 and 11.0 kb transcripts is due to the *Hoxa6* intron sequences present only in the 11.0 kb transcript (probe 8). All transcripts end at the same polyadenylation site at the 3' end of *Hoxa5* exon 2 since no hybridization was observed by northern analyses with probes covering more than 1 kb of genomic sequences 3' to the *Hoxa5* polyA site (probe 17; data not shown). The lines on the left of northern blots 14 and 15 indicate *Hoxa5* nonspliced isoforms that are 1 kb larger. Gapdh was used as a loading control. A, *AccI*; B, *BglIII*; Ba, *BamHI*; E, *EagI*; H, *HindIII*; K, *KpnI*; RI, *EcoRI*; S, *SacI*; St, *StuI*; Xh, *XhoI*.



Finally to resolve the molecular structure of the different transcripts, we used various combinations of primers in RT-PCR experiments (Fig. S5.1B). Sequencing data of the clones obtained confirmed the importance of alternative splicing in the production of the various *Hoxa5* transcripts and revealed other minor forms (clones pLJ282 and 284).

Figure 5.1 summarizes the molecular characterization of the different *Hoxa5* transcripts and from this, several observations were made. First, the sequences between the *Hoxa7* and the *Hoxa5* genes can be entirely transcribed to give rise to the 11.0 kb transcript. Second, the 5.0 kb band detected by northern analyses included two main RNA species, one initiating at position -8905 bp, like the larger forms of 9.5 and 11.0 kb, and containing a large intron of 4.6 kb (identified as the 5 kb-*Hoxa5* transcript), and a second starting at -4409 bp, from the putative *Hoxa6* promoter (identified as the 5 kb-*Hoxa6/a5* transcript). Third, no specific *Hoxa6* transcript corresponding solely to the two known *Hoxa6* exons was detected by northern analysis. Indeed, such signal was not observed with a probe including the putative *Hoxa6* exon 1 sequences (probe 7; Fig. 5.2). A weak band of about 2.4 kb in length was seen with the *Hoxa6* exon 2 probe containing part of the homeobox sequence (probe 9; Fig. 5.2). However, it could not correspond to a *Hoxa6* transcript since it did not produce a mutant form 1 kb larger in the *Hoxa6*^{-/-} RNA sample. Sequence blast of probe 9 against the mouse genome revealed homologies with some *Hox* genes, the highest being 94% homology with the *Hoxa7* homeobox sequence (data not shown). Since the 2.4 kb band matched the main transcript seen with probe 1, which included *Hoxa7* homeobox sequence, it is likely that this faint band may result from the cross-hybridization of probe 9 with the major *Hoxa7* transcript.

Our northern and 3'-RACE studies unveiled that all *Hoxa5* transcripts use the polyA site of the *Hoxa5* gene. Search for polyadenylation sequences at the *Hoxa6* locus did not reveal the presence of a consensus site nearby the presumptive 3'-end of the *Hoxa6* gene. Identified consensus motifs were either overlapping the end of the *Hoxa6* homeobox sequence (position -2006 bp relative to the start site of the 1.8 kb transcript), or located further downstream in the *Hoxa6-Hoxa5* intergenic region (positions -825 bp and -468 bp). The lack of a functional polyadenylation sequence at the *Hoxa6* locus was further confirmed by the presence of a neo transcript in *Hoxa6*^{-/-} RNA sample of about 5.5 kb in

length (probe neo; Fig. 5.2). This transcript initiated at the MC1 promoter of the MC1neo cassette, which does not contain a polyA addition signal [44]. Thus, transcription of the neo cassette must end at the nearest functional polyadenylation site, the latter being localized 3' of the *Hoxa5* gene. In summary, the use of different promoters and alternative splicing may account for the production of several transcriptional units containing sequences from both *Hoxa5* and *Hoxa6* loci.

Transcriptional activity in the *Hoxa6-Hoxa7* intergenic region

Our observation that the 5 kb-*Hoxa5*, 9.5 and 11.0 kb alternate transcripts initiated from a DNA region located downstream the *Hoxa7* gene prompted us to define the transcriptional activity of the sequences encompassing the potential distal promoters D1 and D2. We first performed comparison of the sequences encompassing the *Hoxa5*, *Hoxa6* and *Hoxa7* loci between divergent vertebrate species using the University of California, Santa Cruz (UCSC) Genome Browser (<http://genome.ucsc.edu/>; Mouse July 2007 assembly) [45]. As expected, Hox exon sequences showed very high homology (Fig. 5.3). A ~500 bp DNA region surrounding the transcription start site at position -8905 bp and the D2 promoter was also highly conserved among the species. Alignment of the nucleotide sequences indicated a DNA region of 160 bp, which includes a putative TATA box at position -8934 bp, very highly preserved, arguing for the presence of evolutionary conserved important regulatory DNA elements that may be involved in the production of the larger transcripts. Highly homologous sequences located just 5' from *Hoxa6* exon 1 and corresponding to the putative D1 promoter were also found.

The transcriptional activity of the distal promoters D1 and D2 was directly assessed by transgenesis (Fig. 5.4). For each promoter, a ~4 kb DNA fragment containing sequences flanking the transcription start site was fused to an IRES- β geo cassette used as a reporter. A MES enhancer sequence, known to drive *Hoxa5* regionalized expression along the embryonic axis, was added to each construct in order to improve the detection of a minimal promoter activity [32]. A *Hox*-like staining pattern was observed with the two transgenes indicating that the sequences upstream the distal transcription initiation sites D1 and D2 possess promoter activity.

Figure 5.3. Evolutionary conservation of the *Hoxa5-Hoxa7* genomic region among animal species.

(A) The mouse sequence of the *Hoxa5-Hoxa7* loci was compared to that of rat, human, orangutan, dog, horse, opossum, chicken, xenopus tropicalis, stickleback, zebrafish, fugu and medaka using the UCSC genome browser. The regions with vertical lines indicate conserved sequences. In addition to *Hox* exons, the DNA regions located upstream the *Hoxa5* distal transcription start sites at positions -4409 bp (D1) and -8905 bp (D2) show high homology between divergent species (boxes), suggesting the presence of evolutionary conserved important regulatory DNA elements. (B) Alignment of the nucleotides of a 160-bp DNA fragment from the D2 region indicates the presence of a consensus TATA box and a transcription initiation site (boxes) in most species.

doi:10.1371/journal.pone.0010600.g003

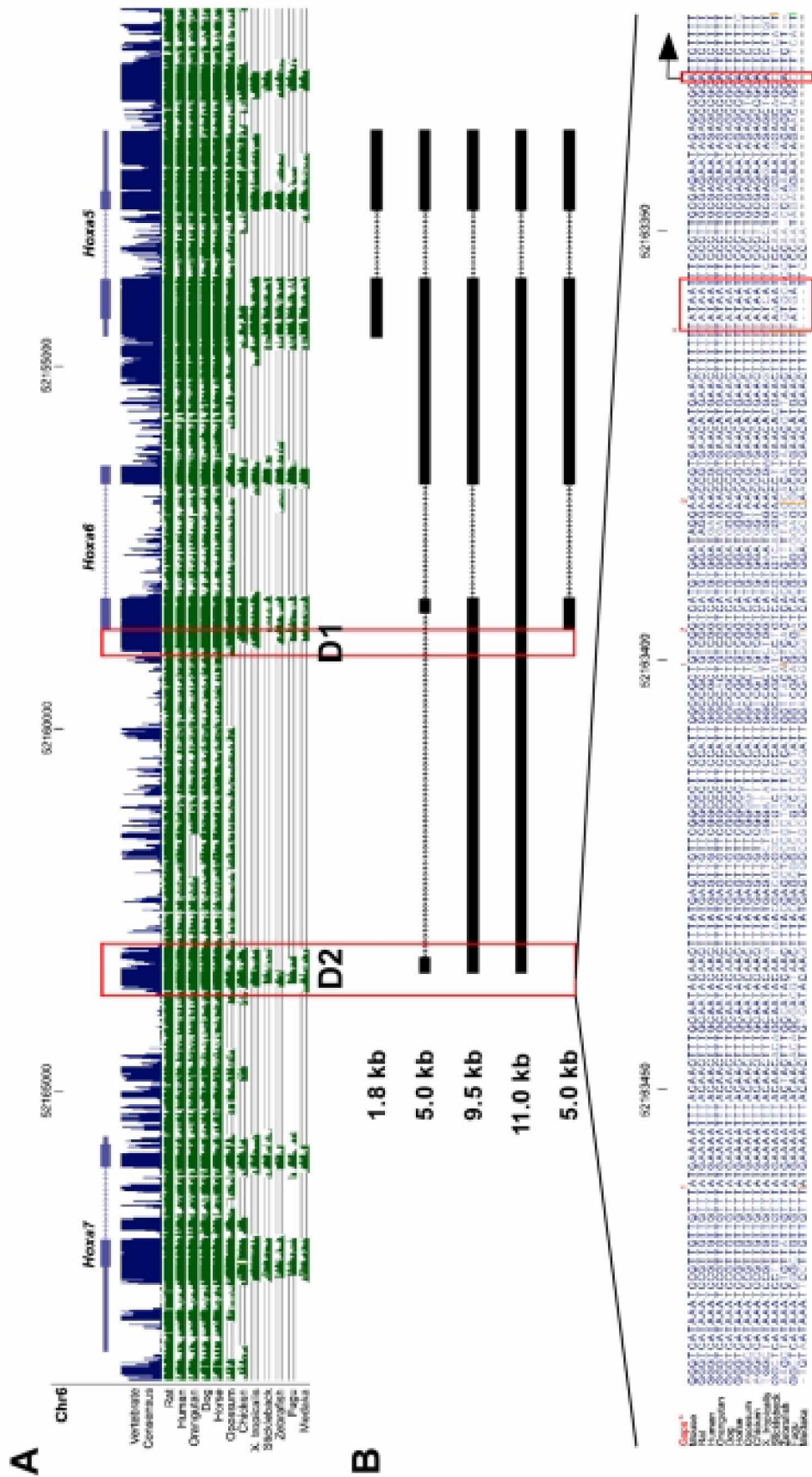
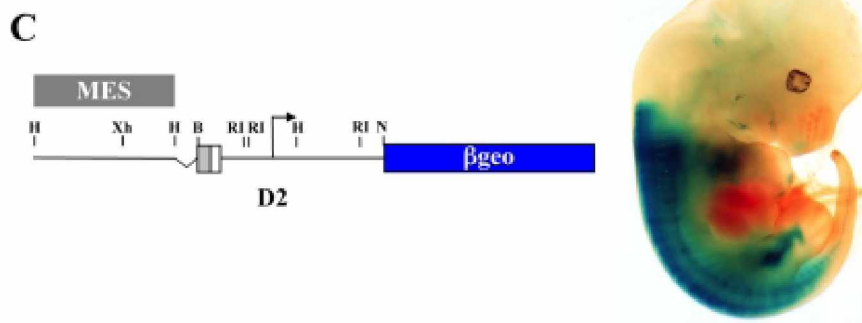
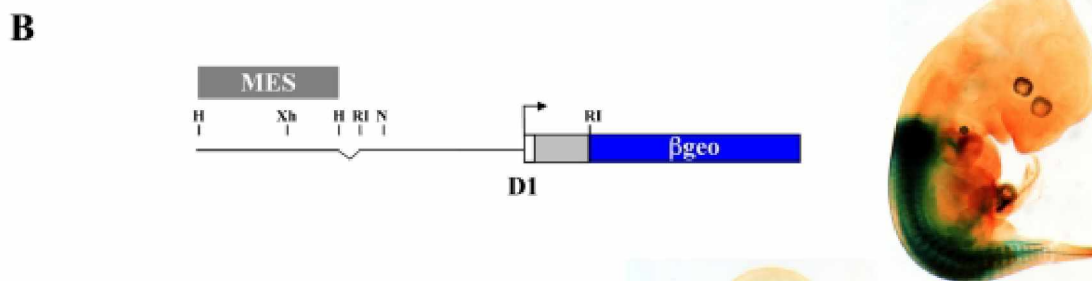
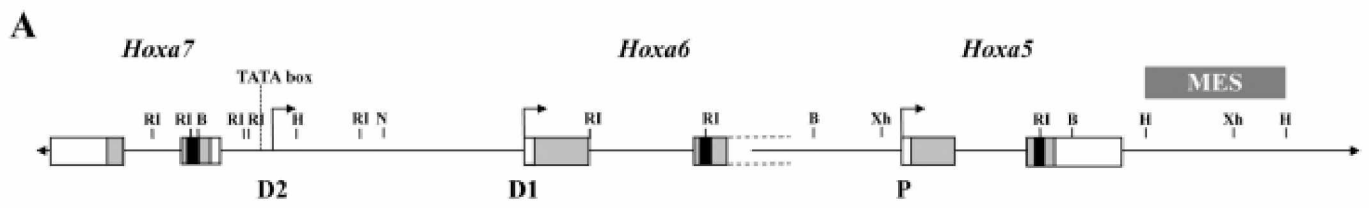


Figure 5.4. Transcriptional activity of the D1 and D2 putative promoters in E12.5 transgenic embryos.

(A) A schematic representation of the *Hoxa5*, *Hoxa6* and *Hoxa7* genes along the HoxA cluster. The proximal promoter P and the two distal promoters D1 and D2 are indicated. (B) Detection of β -galactosidase activity in D1-*lacZ* transgenic embryos in presence of the mesodermal enhancer sequence (MES) indicates that the 4 kb-DNA region encompassing the putative *Hoxa6* promoter (D1) can drive *Hox*-like expression along the antero-posterior axis. (C) As well, a 4 kb-DNA fragment containing the D2 putative promoter region of the larger transcripts possesses a similar transcriptional activity. B, *Bgl*III; H, *Hind*III; N, *Nru*I; RI, *Eco*RI; Xh, *Xho*I.

doi:10.1371/journal.pone.0010600.g004



Differential expression pattern of the *Hoxa5* transcripts

Our previous studies have demonstrated that the 1.8 kb transcript is expressed earlier during embryogenesis and in more anterior structures than the larger transcripts [32]. To gain information on the potential function of the alternate transcripts during embryogenesis, we performed comparative whole-mount in situ hybridization analyses at E12.5. Since all transcripts included the two *Hoxa5* exons corresponding to the 1.8 kb transcript and shared most of their sequences, we used probes that recognize either all transcripts or combinations of the larger forms. As shown on figure 5.5A and B, probe “a” contains *Hoxa5* exon 2 sequences common to all transcripts; probe “b” corresponds to the *Hoxa5-Hoxa6* intergenic region recognizing the 5.0, 9.5 and 11.0 kb forms; probe “c” is localized in the *Hoxa6-Hoxa7* intergenic region and detects the 9.5 and 11.0 kb transcripts; and probe “d” includes *Hoxa6* intron sequences hybridizing only to the 11.0 kb transcript. The expression profile detected with probe “a”, but not with probes “b”, “c” and “d”, revealed structures that exclusively express the 1.8 kb transcript (Fig. 5.5C). In the pv column, the anterior limit of expression of the 1.8 kb transcript corresponded to pv3, while the 5.0, 9.5 and 11.0 transcripts shared the same boundary at pv10. In the neural tube, a posterior shift was observed for the larger transcripts, and the shift was more caudal with probes “c” and “d”. In the future pectoral girdle, hybridization signal was detected only with probe “a”. Previous in situ hybridization experiments on E12.5 embryo sections have shown a strong expression with probe “a” in the mesenchymal component along the entire respiratory tract while probe “b” produced a weak signal restricted to the distal tip of the lungs [26,27]. Similarly, hybridization in the thyroid gland region was observed only with probe “a” (data not shown) [35]. In the developing gastrointestinal tract, a dynamic *Hoxa5* expression pattern prevails. Probe “a” detected expression in the gut mesenchyme as early as E9.0, while the onset of expression with probe “b” was delayed to E12.5 in the foregut, and E10.5 in the midgut [28,29]. In the hindgut, expression of the larger transcripts was detected as early as E9.5 indicating that they shared the same onset as the 1.8 kb transcript (Fig. S5.2). After E12.5, the expression profile in the developing stomach was similar for probes “a” and “b” with a widespread distribution throughout the gastric mesenchyme until E17.5, followed by restriction to the submucosa and muscular layers and extinction at postnatal day 15 (data not shown) [29]. In the mid- and hindgut, expression was detected

with both probes in the mesenchyme up to E14.5 and E17.5, respectively. Expression of the larger transcripts then extinguished in the midgut and the hindgut, while that of the 1.8 kb transcript got restricted to the enteric nervous system and was maintained after birth (Fig. S5.2) [28].

The differential expression of the transcripts encompassing *Hoxa5* sequences along the antero-posterior axis of the skeleton and the rostro-caudal axis of the developing gut reflects the organization of the different promoters (P, D1 and D2) along the cluster and respects the relationship of colinearity characterizing the Hox complexes. It also raises questions about the role played by each of these transcripts during development. Even though the *Hoxa5* mutation perturbs all *Hoxa5* transcripts, most of the defects observed in the *Hoxa5*^{-/-} mutant mice are confined to the cervico-thoracic region and they affect structures and organs that solely express the 1.8 kb transcript. Using a HOXA5-specific antibody, we looked at the HOXA5 protein distribution in the E12.5 mouse embryo [46]. HOXA5 immunoreactivity was observed along the pv column in the pv3-pv10 region and in the mesenchyme of the trachea, lung, stomach and intestine (Fig. 5.6A, C–F). No immunostaining was detected in *Hoxa5*^{-/-} specimens (Fig. 5.6B). Except for the gastrointestinal tract where all *Hoxa5* transcripts were detected (Fig. S5.2), the expression seen in the pv column and the respiratory tract matched that of the 1.8 kb transcript, raising the possibility that only the 1.8 kb transcript produces the HOXA5 protein.

Translational capability of the *Hoxa5* transcripts

There are multiple ORFs predicted from the sequence of the different *Hoxa5* transcripts and the longest ones are represented in figure 5.1. All *Hoxa5* transcripts include the HOXA5 ORF suggesting that they can potentially produce the genuine 270 a.a. HOXA5 protein. Sequence analysis also revealed the presence of a distal in-frame ATG codon located 333 nucleotides upstream of the proximal promoter (P; Fig. 5.1) that can produce a larger HOXA5 isoform of 381 a.a. Moreover, a HOXA6 protein of 232 a.a. can potentially be translated from the 5 kb-*Hoxa6/a5* and the 9.5 kb transcripts, raising the possibility of bicistronic transcriptional units.

Figure 5.5 Differential expression pattern of *Hoxa5* transcripts.

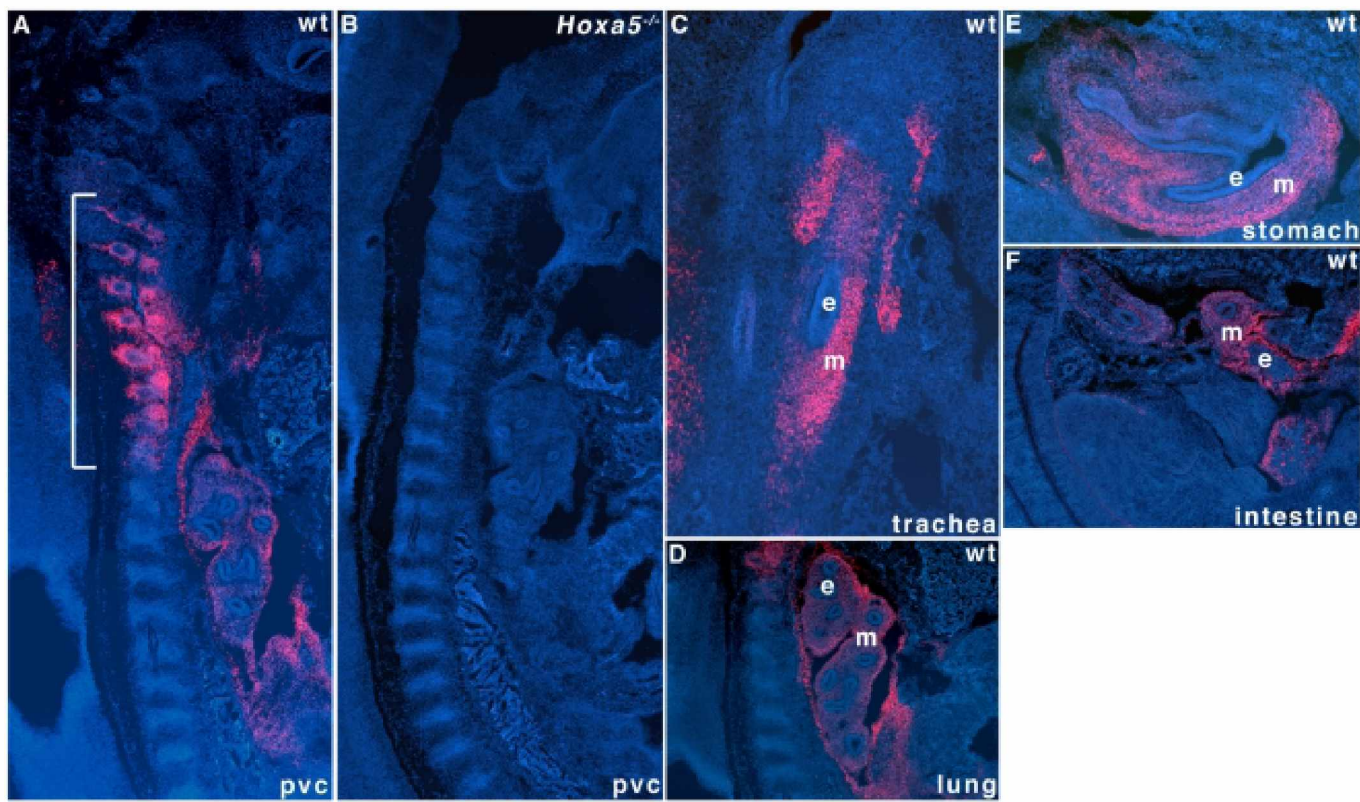
(A) Genomic organization of the *Hoxa5*, *Hoxa6* and *Hoxa7* genes along the cluster. Probes a, b, c and d used for northern blot analyses and whole-mount in situ hybridization are indicated below. (B) Northern blots of polyA⁺ RNA extracted from E12.5 wild-type and *Hoxa5*^{-/-} embryos were hybridized with each probe. (C) Whole-mount in situ hybridization of E12.5 wildtype embryos with probes a-d. Anterior limits of expression are indicated for the neural tube (white arrowheads) and the prevertebral column (black arrows). The different profiles reveal the specific expression of the transcripts: probe a allows to identify the structures that exclusively express the 1.8 kb transcript. These structures include the pv3-pv10 axial domain and the pectoral girdle (black arrowhead). The larger transcripts are expressed in more posterior structures than the 1.8 kb transcript.

B, *Bgl*III; H, *Hind*III; K, *Kpn*I; RI, *Eco*RI; Xh, *Xho*I.

doi:10.1371/journal.pone.0010600.g005

Figure 5.6. Restricted spatial distribution of HOXA5 protein along the antero-posterior axis and in the respiratory and digestive tracts.

HOXA5 immunoreactivity is detected in the pv3-pv10 region of the prevertebral column (pvc) of e12.5 wild-type mouse embryos by immunofluorescence (bracket; A). No immunoreactivity is seen in the *Hoxa5*^{-/-} specimens confirming the absence of the protein in the null mutant mouse line (B). HOXA5 is also detected in the mesenchymal component of the trachea (C), lung (D), stomach (E) and intestine (F) of e12.5 mouse embryo. e, epithelium; m, mesenchyme.



To define the capacity of the larger transcripts to produce the HOXA5 protein, we made expression vectors containing the entire 5 kb-*Hoxa5*, 5 kb-*Hoxa6/a5* or 9.5 kb cDNA sequence with a MYC tag at the carboxy terminus of the HOXA5 protein. We also added a FLAG tag at the carboxy terminus of the HOXA6 protein for the 5 kb-*Hoxa6/a5* and 9.5 kb cDNA vectors. As a control, we used a MYC-tagged version of the 1.8 kb cDNA. In addition, we made a vector containing an extended version of the 1.8 kb transcript (up to position -555 relative to the start site of the 1.8 kb transcript) that includes the distal ATG codon. The plasmids were first tested in vitro using a coupled transcription/translation assay with incorporation of radio-labeled methionine (Fig. 5.7A). As expected, vectors containing the 1.8 kb cDNA sequence with or without the MYC tag produced radioactive products of about 38 kD in size, slightly larger for the vector carrying the MYC tag and which corresponded to the 270 a.a. HOXA5 protein. For the extended version of the 1.8 kb transcript, the 38 kD HOXA5 protein was produced as well as a protein of about 50 kD, compatible to the 381 a.a. isoform. For the vector carrying the 5 kb-*Hoxa5* cDNA, the HOXA5-MYC isoform of 270 a.a. was the unique band observed. In the case of the vector containing the 5 kb-*Hoxa6/a5* cDNA sequence, the 270 a.a. HOXA5-MYC protein was detected as well as a smaller protein of about 36 kD, likely corresponding to the HOXA6-FLAG protein. Similar observations were made for the 9.5 kb cDNA, even though the two bands were faint. No band related to the larger HOXA5 isoform was detected with the 5 kb-*Hoxa5*, the 5 kb-*Hoxa6/a5* and the 9.5 kb cDNAs. Thus when tested in vitro, all vectors can produce the genuine HOXA5 protein. The 5 kb-*Hoxa6/a5* and 9.5 kb transcripts can also generate the HOXA6 protein, acting as bicistronic units.

Transfection assays in HEK293 cells followed by western analyses with the MYC or FLAG antibodies showed that only the 1.8 kb-MYC vector produces the 270 a.a. HOXA5-MYC protein (Fig. 5.7B). The HOXA5-MYC isoform of 381 a.a. was neither produced from the extended version of the 1.8 kb transcript nor from the 5 kb-*Hoxa5*, the 5 kb-*Hoxa6/a5* and the 9.5 kb vectors. Moreover, only the 5 kb-*Hoxa6/a5* vector produced the HOXA6-FLAG protein of 232 a.a. As verified by RT-PCR analysis, all *Hoxa5* cDNA expression vectors were transcribed in HEK293 cells (Fig. 5.7C). In summary, all *Hoxa5* transcripts can be efficiently translated in in vitro assays. However in cell cultures, only

the 1.8 kb transcript can encode the 270 a.a. HOXA5 protein and the HOXA6 protein can solely be produced from the 5 kb-*Hoxa6/a5* transcript.

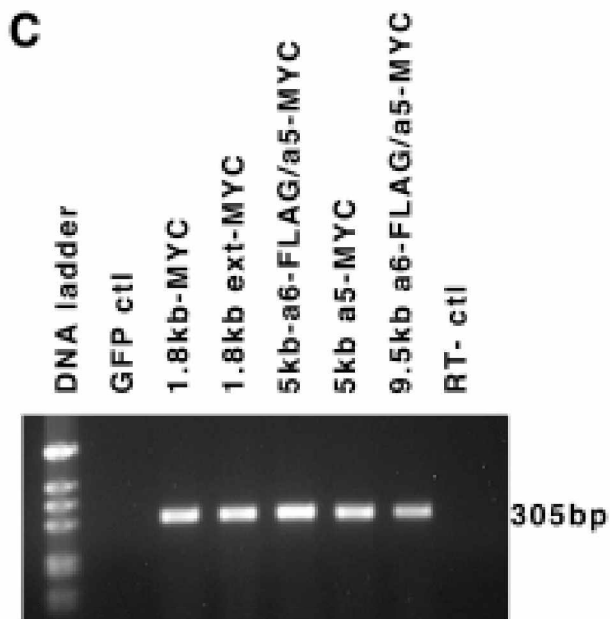
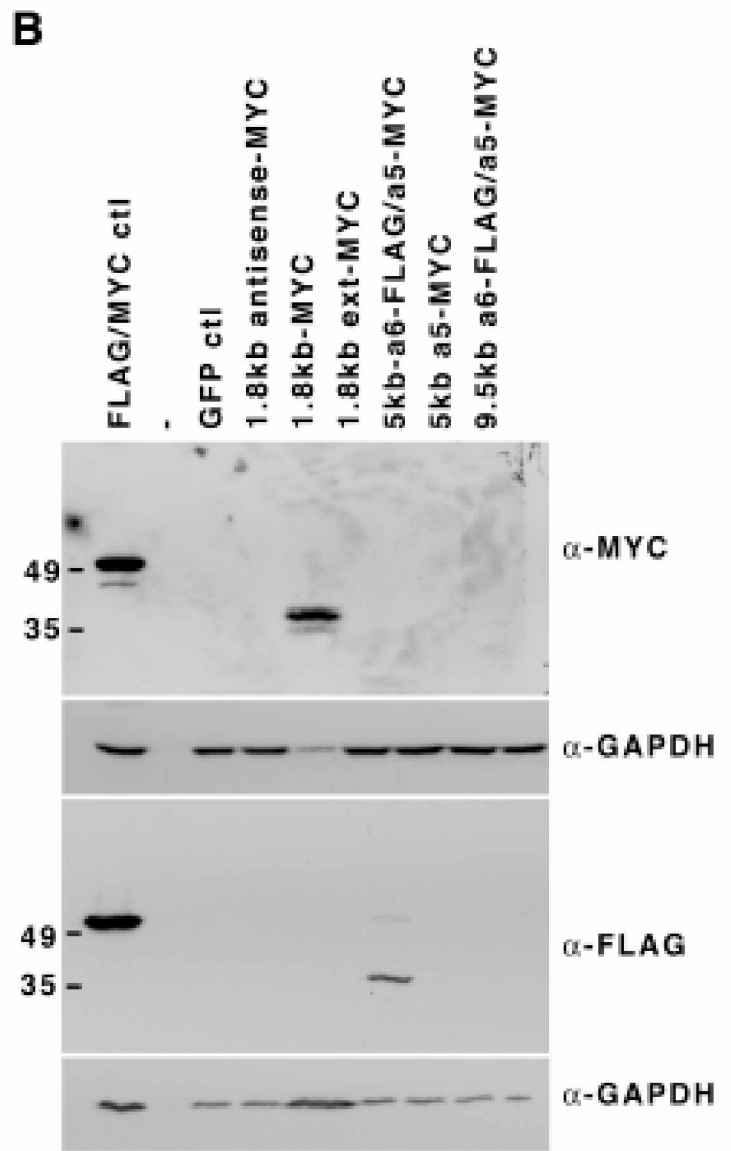
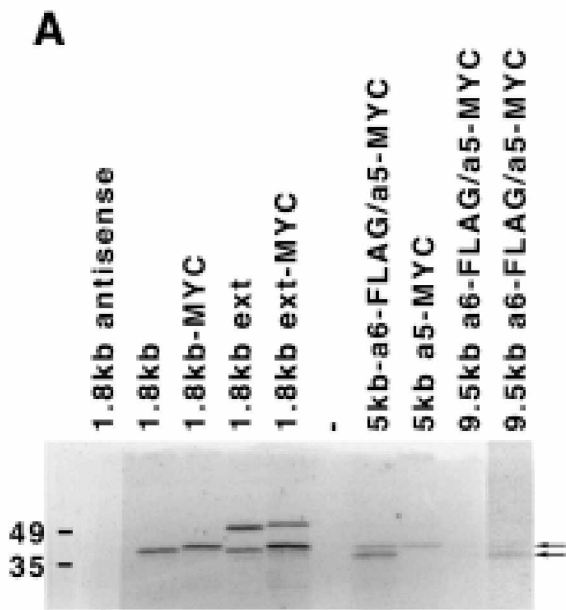
Skeletal transformations in *Hoxa6* and *Hoxa5*; *Hoxa6* transheterozygous mutant mice

The *Hoxa6* mutant mouse line provides a valuable tool for investigating the role of the larger *Hoxa5* transcripts since the insertion of the MC1neo cassette into the *Hoxa6* homeobox sequences disrupts the 5.0, 9.5 and 11.0 RNA species encompassing the *Hoxa5* and *Hoxa6* loci (Fig. 5.2). The *Hoxa6* mutation causes a relatively mild phenotype, which consists in the presence of ectopic ribs on the 7th cervical vertebra (C7) in less than half of the mutants [44]. One puzzling aspect of the *Hoxa6*^{-/-} phenotype is that it is incompatible with the pv10 anterior expression boundary of the gene, as established by the expression analysis using probes specific for the larger transcripts (Fig. 5.5C). Two possibilities could account for this discrepancy. The integrity of the larger transcripts including *Hoxa6* sequences is necessary for the correct patterning at the pv7 axial level. Alternatively, the presence of the neo cassette in the *Hoxa6* locus may interfere with the expression of the nearby *Hoxa5* gene, which then can impact on the skeletal phenotype. Ectopic ribs on C7 are a hallmark of the *Hoxa5* mutation, as they are found in most *Hoxa5* mutants [23,27]. Moreover, transcriptional interference is not unusual in *Hox* mutations, and we have previously reported the deleterious long-range cis effect of the *Hoxa4* mutation on *Hoxa5* expression [27]. To discriminate between these options and to define the respective role of the *Hoxa5* and *Hoxa6* genes in the specification of the cervico-upper thoracic region, we generated *Hoxa5*; *Hoxa6* transheterozygous animals (*Hoxa5*^{+/-}; *Hoxa6*^{+/-}), which are heterozygotes for both genes on different chromosomes. First, we examined the skeleton of new cohorts of single mutants and transheterozygous newborn pups (Table 5.1). In this mixed genetic background, *Hoxa5*^{-/-} mutants displayed the skeletal transformations previously reported: the lack of tuberculum anterior on C6 in 84% of the specimens analyzed; the presence of ectopic ribs on C7 (87%), most being present on both sides of the vertebra; abnormal acromion (42%) and fused tracheal rings (100%). *Hoxa5*^{+/-} pups also presented the C7 homeotic transformation at a lesser frequency (57%) than the *Hoxa5*^{-/-}

Figure 5.7. HOXA5 protein production from *Hoxa5* transcripts.

(A) Expression vectors carrying cDNAs corresponding to the 1.8 kb (with and without a MYC-tag), the extended 1.8 kb (MYC-tagged and non-tagged), the 5.0 kb-*Hoxa6* FLAG/*a5*-MYC, the 5.0 kb-*Hoxa5*-MYC and the 9.5 kb *Hoxa6*-FLAG/*a5*-MYC transcripts were tested in vitro using a coupled transcription/translation system. A [³⁵S]-radiolabeled protein corresponding to a ~38 kD genuine HOXA5 protein is translated from all vectors. The larger HOXA5 isoform of ~50 kD is only translated from the extended 1.8 kb cDNA version. The 5.0 kb-*Hoxa6*-FLAG/*a5*-MYC and the 9.5 kb *Hoxa6*-FLAG/*a5*-MYC vectors also produce a ~36 kD band likely corresponding to the putative HOXA6 protein. Translation of the HOXA5 and HOXA6 proteins from the 9.5 kb *Hoxa6*-FLAG/*a5*-MYC vector is weakly detected and a longer exposure is shown with arrows to indicate the position of both proteins. (B) The HEK293 cells were transfected with the MYC-tagged version of the 1.8 kb and extended 1.8 kb expression vectors and with the 5.0 kb-*Hoxa6*-FLAG/*a5*-MYC, the 5.0 kb-*Hoxa5*-MYC and the 9.5 kb *Hoxa6*-FLAG/*a5*-MYC vectors. In parallel, control plasmids expressing either the green fluorescent protein (GFP ctl), the 1.8 kb *Hoxa5*-MYC in the antisense orientation or the pMEK1-MYC-FLAG plasmid (MYC/FLAG ctl) were transfected. Protein lysates were western-blotted with anti-MYC or anti-FLAG antibodies. Solely the 1.8 kb-MYC vector produces a genuine HOXA5 protein whereas the HOXA6 protein is only detected with the 5.0 kb-*Hoxa6*-FLAG/*a5*-MYC plasmid. GAPDH was used as a loading control. (C) RNA expression of each expression vector transfected in HEK293 cells was tested by RT-PCR. A 305 bp fragment is observed in RNA samples from HEK293 cells transfected with the *Hoxa5* cDNA vectors. No expression is detected in the GFP ctl specimen.

doi:10.1371/journal.pone.0010600.g007



mutants, but with a much higher incidence than the *Hoxa6*^{-/-} mutants (35%; Table 5.1). Ectopic ribs on C7 were observed in 58% of the *Hoxa5*^{+/-}; *Hoxa6*^{+/-} pups analyzed, a frequency similar to that of *Hoxa5*^{+/-} mutants, suggesting that the *Hoxa6* contribution to the C7 skeletal specification was weak.

We also monitored *Hoxa5* expression in *Hoxa6*^{-/-} and *Hoxa5*^{+/-}; *Hoxa6*^{+/-} e12.5 embryos by in situ hybridization using the riboprobes “a” and “b” described above (Figs. 5.5A and 5.8). The signal detected with probe “a” was specifically less intense in the pv3-pv10 region of the pv column from *Hoxa6*^{-/-} and *Hoxa5*^{+/-}; *Hoxa6*^{+/-} embryos when compared to wt specimens. No change in expression was observed with probe “b”. Thus, the disruption of the large transcripts by the *Hoxa6* mutation does not have a major impact on axial specification as no skeletal anomaly was seen for the vertebrae localized caudally of pv10. Moreover, the *Hoxa6* mutation alters expression of the 1.8 kb transcript in the pv3-pv10 region. One possible explanation may be that the insertion of the neo cassette into the *Hoxa6* locus impairs the activity of the *Hoxa5* proximal promoter, which consequently alters the specification of the C7 vertebra. On the other hand, the larger transcripts may also be involved in the regulation of the *Hoxa5* proximal promoter by a mechanism that remains to be defined and their disruption by the *Hoxa6* mutation may affect the expression of the 1.8 kb transcript that in turn impacts on skeletal patterning in the cervico-thoracic region.

DISCUSSION

Multiple transcriptional units are present at the *Hoxa5* locus

This study aimed at characterizing the different transcriptional units encompassing the *Hoxa5* locus. Presence of multiple transcripts is not unique to the *Hoxa5* gene. Other Hox genes have been reported to produce several RNAs and in few cases, the molecular nature of these transcripts has been analyzed [45, 47, 48]. In the case of the *Hoxa5*, *Hoxd4* and *Hoxb3* genes, the additional transcripts are expressed according to a *Hox*-like pattern

Table 5.1. Newborn skeletal morphology according to the *Hoxa5* and *Hoxa6* genotypes.

	Genotype						
	wt	<i>Hoxa5</i> ^{+/-} <i>Hoxa6</i> ^{+/+}	<i>Hoxa5</i> ^{-/-} <i>Hoxa6</i> ^{+/+}	<i>Hoxa5</i> ^{+/-} <i>Hoxa6</i> ^{-/+}	<i>Hoxa5</i> ^{+/+} <i>Hoxa6</i> ^{+/-}	<i>Hoxa5</i> ^{+/+} <i>Hoxa6</i> ^{-/-}	<i>Hoxa5</i> ^{+/-} <i>Hoxa6</i> ^{-/-}
Tuberculum anterior on C6 ^a							
Absent	-	-	32	-	-	-	-
Present	14	44	6	38	20	26	
Ribs on C7 ^a							
Absent	12	16	5	16	7	17	
Present	2	28	33	22	13	9	
Unilateral	-	6	3	6	1	5	
Bilateral	1	11	15	8	6	2	
Acromion ^a							
Normal	14	44	22	38	19	26	
Abnormal	-	-	16	-	1	-	
Trachea							
Normal	7	22	-	19	10	13	
Abnormal	-	-	19	-	-	-	
Number of animals	7	22	19	19	10	13	

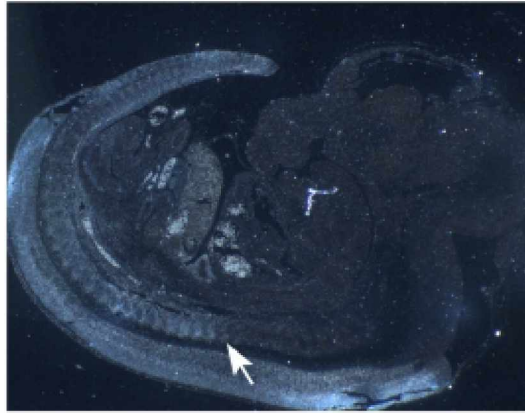
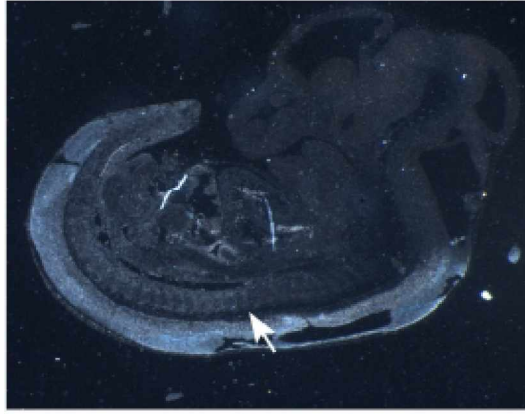
^aleft and right sides were scored independently.
doi:10.1371/journal.pone.0010600.t001

Figure 5.8. Comparative expression patterns of Hoxa5 transcripts in e12.5 wild-type, Hoxa6 homozygous and Hoxa5; Hoxa6 transheterozygous mutants.

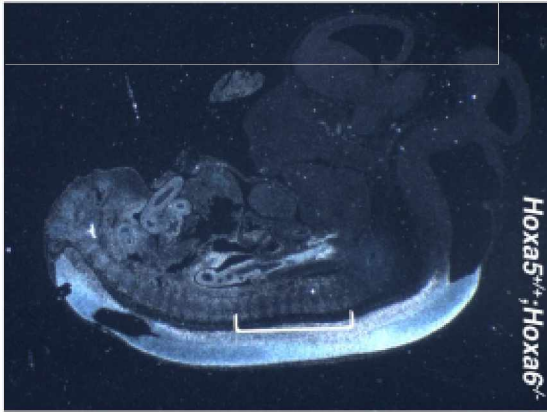
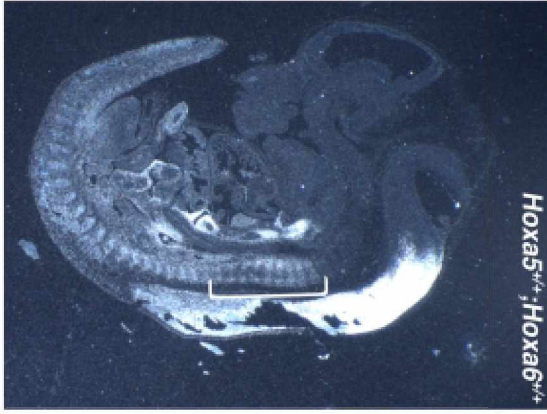
In situ hybridization experiments were performed on comparable sagittal sections. Representative specimens are shown. Genotype is indicated on the top right of each column and probe on the left of each row. The bracket indicates the pv3-pv10 domain, which hybridizes with probe a in wild-type specimens, but not with probe b. In *Hoxa6*^{-/-} and *Hoxa5*^{+/-}; *Hoxa6*^{+/-} mutants, the signal in this region is significantly decreased with probe a. With probe b, the anterior limit of expression corresponds to pv10 in all samples regardless of the genotype (arrows) with no major change in signal intensity.

doi:10.1371/journal.pone.0010600.g008

probe b



probe a



but they have distinct boundaries of expression. Their generation involves multiple promoters, alternative splicing and/or the use of different polyadenylation sites. We have identified three promoters for the *Hoxa5* transcripts. The proximal one corresponds to the genuine *Hoxa5* promoter driving expression of the major 1.8 kb transcript. The larger transcripts are generated from either the distal D1 promoter, which is in fact the putative *Hoxa6* promoter, or the D2 promoter located downstream the 3' extremity of the *Hoxa7* gene. Alternative splicing is also an important process contributing to the diversity of *Hoxa5* transcripts. Furthermore, the *Hoxa5* intron is not rigorously spliced in a small proportion of all *Hoxa5* transcripts adding to the number of RNA species detected. In contrast, the *Hoxa6* intron is systematically spliced in all transcripts that include the *Hoxa6* locus with the exception of the 11.0 kb form, which retains the intron sequences. Finally, only one polyadenylation site located at the end of the *Hoxa5* locus is utilized by all transcripts encompassing the *Hoxa5* and *Hoxa6* sequences. This data reveals the lack of a functional polyadenylation site at the *Hoxa6* locus. Thus in the E12.5 mouse embryo, the *Hoxa6* transcript exists only as a bicistronic gene product that can potentially generate two HOX proteins, HOXA5 and HOXA6.

Several studies have revealed the high transcriptional activity occurring along the mammalian *Hox* clusters [19–22]. Many of these transcriptional units are antisense to *Hox* genes. They are highly conserved between mouse and human, and some are polycistronic. For the HoxA cluster, a search with the UCSC genome browser indicates the existence of numerous sense and antisense transcripts covering the mouse *Hoxa5-Hoxa7* genomic region. Except for the *Hoxa5* 1.8 kb transcript, none of the larger transcripts characterized in the present work have been reported in databanks. However, portions of some of the transcripts listed include sequences of the *Hoxa5* larger forms. For instance, the first exon of the GenBank mRNA AK051552 corresponds exactly to the 5kb-*Hoxa5* transcript first exon. Twelve kb downstream in the *Hoxa4-Hoxa5* intergenic region, the second exon of the AK051552 transcript encompasses sequences corresponding to a distal *Hoxa3* exon annotated as the *Hoxa3* Y11717 mRNA (GenBank). The latter also initiates within the first exon of the 5 kb-*Hoxa5* transcript. In fact, the portrait of the different RNA species initiating in the *Hoxa6-Hoxa7* intergenic region reveals extensive transcription and production of large transcripts starting in the D2 promoter region and extending towards

the vicinity of the *Hoxa3* gene. These data combined to ours argue for an intricate transcriptional activity at the D2 promoter. Interestingly, the integration of the Mouse Moloney Leukemia Virus (MMLV) nearby the D2 promoter in the *Hoxa6-Hoxa7* intergenic sequence impacts dramatically on the expression of the *Hoxa3* to *Hoxa10* genes [49]. Even though we cannot rule out the possibility of a long-distance perturbing effect from the MMLV enhancer on the *HoxA* promoters, it is tempting to speculate that the MMLV insertion predominantly affects the D2 promoter activity revealing the important role of the latter in the control of a subset of *HoxA* genes.

Production of the HOXA5 and HOXA6 proteins

In in vitro assays, all the cDNAs corresponding to the larger *Hoxa5* RNA species tested can produce the 270 a.a. HOXA5 protein but not the 381 a.a. isoform initiating at the distal ATG (Fig. 1). However in HEK293 cultured cells, only the 1.8 kb transcript can be translated into the genuine HOXA5 protein. Moreover, expression of the HOXA5 protein along the mouse embryonic axis was not detected caudally of pv10, the anterior boundary of the expression domain of the larger transcripts. The concordance between the *Hoxa5* mutant phenotype and the expression domains of the HOXA5 protein and the 1.8 kb transcript support the notion that the 1.8 kb RNA is the functional *Hoxa5* transcript.

In the case of the HOXA6 protein, the bicistronic 5 kb-*Hoxa6/a5* and 9.5 transcripts generate both HOXA5 and HOXA6 proteins in vitro but only the 5 kb-*Hoxa6/a5* cDNA produces the HOXA6 protein in cultured cells. Whether the 5 kb-*Hoxa6/a5* transcript can generate the HOXA6 protein in the embryo remains to be defined via the development of a specific antibody. The absence of HOXA5 protein from the bicistronic 5 kb-*Hoxa6/a5* unit is in accordance with the previous observations that polycistronic translation is a rare phenomenon in eukaryotes. In the case of polycistronic transcripts, the 5' proximal cistron is usually the translated one as observed here [50]. Polycistronic transcription is not unusual in Hox clusters [19]. For instance, the *Hoxc4*, *-c5* and *-c6* genes are transcribed from a common promoter producing a primary transcript alternatively spliced to produce mature messengers encoding different proteins [51]. Finally, the presence of long 5' untranslated region (UTR) sequences in the 5 kb-*Hoxa5* and 9.5 kb transcripts can explain

the absence of translation as it can greatly reduce translational efficiency [50]. Thus, the molecular characterization of the different *Hoxa5* RNA species has unveiled an unexpected transcriptional organization and the promiscuity between the *Hoxa5* and *Hoxa6* loci in protein production.

Role of the *Hoxa6* locus

Our studies raise questions about the role of the *Hoxa6* gene during development. Interestingly in teleosts, the *Hoxa6* gene is not present, most likely lost during the duplication process [52]. In mice, the *Hoxa6* mutation results in a mild skeletal phenotype resembling that of the *Hoxa5* mutants and occurring at an axial level located outside the *Hoxa6* expression domain. Indeed, the *Hoxa6* phenotype can be attributed to transcriptional interference that hinders transcription from the *Hoxa5* proximal promoter. In fact, the neo cassette used to mutate the *Hoxa6* gene was inserted in a temporal regulatory sequence responsible for the correct onset of *Hoxa5* expression supporting the notion that the presence of exogenous sequences nearby control regions impact on their efficiency [32,44]. We cannot rule out the possible implication of the larger transcripts in the regulation of the *Hoxa5* proximal promoter and the effect their disruption by the *Hoxa6* mutation may have. To directly address this possibility would require the specific abolition of the larger transcripts in mice and the analysis of the phenotypic and molecular consequences. Taken together, our data suggest that it seems unlikely that the *Hoxa6* gene plays a role in axial specification, although it may serve other functions yet to be defined.

Functional role of the *Hoxa5* 1.8 kb transcript

Along the antero-posterior axis, the HOXA5 protein is detected in the most-rostral subdomain of expression of the *Hoxa5* gene, which corresponds to the exclusive axial expression domain of the 1.8 kb transcript. In *Hoxa5* mutant mice, most of the defects observed lie within the HOXA5 protein expression domain. This further supports the importance of the 1.8 kb transcript as the biological effector of the *Hoxa5* gene during development.

Restricted HOX proteins expression was also reported for the *Hoxb4* and *Hoxb5* genes, and in both cases, the proteins were similarly localized in the anterior part of the gene expression domain [8, 15]. Comparatively to the RNA distribution, vertebrate HOX proteins may be more confined to precise axial levels, a situation comparable to what is observed for homeotic proteins in *Drosophila* embryos. Lots of efforts have been put on the regulatory mechanisms establishing the anterior boundary of *Hox* expression domains in vertebrates. Our findings enlighten the relevance of examining in more details how posterior boundaries may be fixed along the embryonic axes as well.

Implication of the *Hoxa5* long noncoding RNAs

In E12.5 mouse embryo, the different polyA⁺ transcripts covering the *Hoxa5* coding sequences reported in this study originate from the DNA coding strand. Moreover, the larger transcripts have similar Hox-like expression profiles. They are expressed later during embryogenesis and in more posterior structures than the 1.8 kb transcript. Similarities between the expression profile of the larger transcripts and that of the *Hoxa7* gene indicate that they may share regulatory elements [38]. These long and interspersed transcripts also imply that DNA regions involved in *Hoxa5* gene regulation may be distributed along the cluster and emphasize the importance of the *Hox* cluster organization for the correct expression of Hox genes.

The larger *Hoxa5* transcripts cannot generate the HOXA5 protein. However, the 5 kb-*Hoxa6/a5* transcript can produce the HOXA6 protein in HEK293 cells. Thus, the 5kb-*Hoxa5*, the 9.5 and the 11.0 kb transcripts, all transcribed from the distal D2 promoter, can be considered as long noncoding RNAs. Both *Hoxa5* and *Hoxa6* null mutations disrupt the 5.0, 9.5 and 11.0 kb transcripts without any phenotypic consequence in the domain where they are expressed. These two mutations do not preclude the transcription of the transcripts but produce mutant versions 1 kb larger due to the presence of a neo cassette in each mutated locus. Thus, disruption of the larger transcripts does not impact on axial specification. Transcription of intergenic regions or upstream promoter sequences can affect the expression of adjacent genes, either by producing transcriptional interference, promoter competition for a limiting factor or by altering chromatin structure, leading to the hypothesis that the act of transcription per se of long noncoding RNAs is responsible for

the regulatory effect [53]. Alternatively, long noncoding RNAs may regulate in transgene expression as shown for HOTAIR, which participates to the Polycomb Repressive Complex 2 [20]. Our results in conjunction with the increasing sum of data on the potential role of long noncoding RNAs in transcriptional regulation now raise the following questions: do the *Hoxa5* larger transcripts represent “transcriptional noise” or do they contribute by themselves to the control of Hox gene expression [54].

New genome-wide technologies have unveiled the complex architecture of the eukaryotic transcriptome. The extensive overlap between transcriptional units, the existence of nonco-linear transcripts and the multifunctional roles of genomic sequences have even led to a re-evaluation of the current concept of the nature of the gene [55–57]. In this context and due to the fact that several of these features occur in *Hox* clusters, the latter appear as a paradigm from which we may learn more about the link existing between transcriptional complexity, functionality and genome organization.

MATERIALS AND METHODS

Mouse strains and genotyping

The establishment of the *Hoxa5* mutant mouse line in the MF1-129/SvEv C57BL/6 mixed background and the genotype by Southern analysis has been previously reported [23]. The *Hoxa6* mouse line in the 129/SvEv-C57BL/6 genetic background was provided by Dr. Mario Capecchi and genotyped by Southern blot analysis as described [44].

Hoxa5 mutant mice were intercrossed with *Hoxa6* mutant mice to produce transheterozygous animals (*Hoxa5*^{+/-}; *Hoxa6*^{-/+}). For genetic background homogeneity, transheterozygous animals were interbred to generate mice carrying the possible *Hoxa5*; *Hoxa6* allelic combinations for subsequent skeletal analyses.

Embryonic age was estimated by considering the morning of the day of the vaginal plug as E0.5. All experiments were performed according to the guidelines of the Canadian Council on Animal Care and approved by the institutional animal care committee (Comité de Protection des Animaux du Centre Hospitalier Universitaire de Québec, CPA-CHUQ).

PolyA⁺ RNA isolation and northern analysis

Total RNA from wild-type, *Hoxa5*^{-/-} and *Hoxa6*^{-/-} E12.5 embryos was isolated according to the TRIzol RNA extraction protocol (Invitrogen). For polyA⁺ RNA, the extraction was followed by two-step chromatography on oligo(dT)-cellulose column [58]. Seven mg of each polyA⁺ RNA preparation were used for northern analysis. Several genomic fragments covering the *HoxA* locus from *Hoxa7* to *Hoxa5* were used for antisense riboprobe synthesis (Fig. 2A). Hybridization to a GAPDH probe for quality control and quantitation was performed in parallel.

5'-Rapid Amplification of cDNA ends (RACE), 3'-RACE and RT-PCR analyses

The 5' RACE protocol was essentially based on that described in [59]. The first-strand cDNA synthesis was performed with one mg of total RNA from E12.5 wild-type mouse embryos annealed to antisense Hox primer 1 (5'-GCGACCCTGCTATTGCCAGACA-3'), primer 2 (5'CTCCGGTCGGTGCCTTCC TCAT-3') or primer 3 (5'-CTGCGGGAGAAGCAGGCTGGAAT-3'; Fig. S1A). A polyA tail was added to the cDNAs. The dA-tailed cDNAs were first amplified using the nested *Hox* primer 1' (5'-AACACAGCAGCCCCTGCACGGAA-3'), primer 2' (5'-CTGCACGCTGCCGTCAGGTTGT-3') or primer 3' (5'-GGCACCAGGGGGCAAAGCCAATA-3') with an anchor primer complementary to the added oligo(dA) tail (5'-CCAGTGAGCAGAGTGACGAGGACTCGAGCTCAAGCTTTTTTTTTTTTTTTTTTTT-3') and an adapter primer included into the anchor primer (5'-CCAGTGAGCAGAGTGACGAGGAC-3'). A second round of amplification was set up with the primary PCR products and a different set of nested primers: primer 1'' (5'-CCCTCTTCCAGGGCTCAGGAA-3'), primer 2'' (5'-AAATGCGGCCCGCCTGCTGCTCGGGAGAAAAGTG-3') or primer 3'' (5'-AAATGCGGCCGCGTCCCTGCACTGGGTCTAC-3') with the anchoring primer (5'-GACGAGGACTCGAGCTCAAGC-3'). The secondary PCR products were subcloned prior to sequencing.

The 3' RACE System for Rapid Amplification of cDNA Ends (GIBCO BRL) was used to identify the 3' extremity of the *Hoxa5* transcripts. One mg of polyA⁺ RNA from E12.5 wild-type mouse embryos was used for the first-strand cDNA synthesis with a polyT adapter primer (5'-CCATCGATGTCGACTCGAGTTTTTTTTTTTTTTTTTTTTT-

3') according to the protocol provided by the manufacturer. PCR amplification was then performed using a *Hoxa5* specific primer (5'-CTCCCCTTGTGTTCTTCTG-3') and a nested adapter primer (5'-CCATCGATGTCGACTCGAGT-3') on an aliquot of the first-strand synthesis reaction. The resulting amplified products were subcloned and sequenced.

For the RT-PCR reactions, one mg of total RNA from E12.5 wildtype mouse embryos was used for the first-strand cDNA synthesis after annealing to an oligo(dT) primer or to the specific *Hox* primers 2 and 3 (described above). The following PCR reactions were then performed using various combinations of sense and antisense primers that all contain a *NotI* site. The sense primers are: primer 4 (5'-ATATGCGGCCGCCTCCGCTCCATCCTGCGTGCTT-3'), primer 5 (5'-AAATGCGGCCGCA TCACAGTCCTGCAGAGGGGC-3'), and primer 6 (5'-AAATGCGGCCGCCACAAA CGACCGCGAGCCACA-3'). The antisense primers are: primer 7 (5'-AAATGCGGCC GCCCCGGCGAGGATACAGAGGAT-3') and primer 8 (5'-AAATGCGGCCGCACAG AGAGCTGCCCCGGCTACT-3'). The RT-PCR products were then digested with *NotI*, subcloned and sequenced.

Construction of *Hox/lacZ* transgenes and production of transgenic embryos

Genomic fragments encompassing the putative D1 and D2 promoter regions were subcloned in front of the IRES- β geo cassette obtained from the pSA-IRES β geolox2PGKDTA plasmid designed by Drs. Philippe Soriano and Valera Vasioukhin. We used a 4.19 kb *EcoRI* fragment extending from positions 27.96 kb to 23.77 kb (position +1 corresponding to the transcription start site of *Hoxa5* exon 1) for the D1 promoter and a 3.88 kb *BglII-NruI* fragment extending from positions 211.46 kb to 27.58 kb for the D2 promoter. The 2.1 kb mesodermal enhancer (MES) was inserted upstream the promoter region in both constructs [32]. The constructs were made in pBluescript SKII+ (Stratagene) and purified following a cesium chloride centrifugation.

The *Hox/lacZ* sequences were isolated using a *SalI-NotI* digestion to remove vector sequences and they were purified on agarose gel. They were injected into the pronuclei of fertilized eggs derived from (C57BL/6 x CBA) F1 hybrid intercrosses following standard procedures [60]. Transgenic founder embryos were recovered from foster mothers at

E12.5, genotyped by Southern analysis of yolk sac DNA using a *lacZ* specific probe to verify the integrity of the microinjected construct, and analyzed for *lacZ* expression by β -galactosidase staining as previously described [32].

RNA in situ hybridization and immunofluorescence analyses

The whole-mount in situ hybridization protocol was based on the one described in [61], while radioactive in situ hybridization of paraffin sections was performed according to the protocol in [62]. The following murine genomic sequences were used as templates for synthesizing either digoxigenin or [^{35}S] UTP-labeled riboprobes: a 830 bp *Bgl*II-*Hind*III fragment containing the 3'-untranslated region of the second exon of the *Hoxa5* gene (probe a), a 606 bp *Bgl*II-*Xho*I fragment present in the intergenic region between *Hoxa5* and *Hoxa6* genes (probe b), a 675 bp *Hind*III-*Eco*RI sequence located just downstream the *Hoxa7* gene (probe c), and a 356 bp *Kpn*I fragment present in the *Hoxa6* intron (probe d). The *in situ* experiments were performed on at least three specimens of each genotype.

Immunofluorescence staining with the rabbit anti-HOXA5 antibody and counterstaining with 4',6-diamidino-2-phenylindole (DAPI; Molecular Probes) was performed as described in [46].

Expression vectors, transfection, western and RT-PCR analyses

The sequence of the 1.8 kb cDNA was subcloned into the pcDNA3 expression plasmid (Invitrogen). A DNA fragment, from the p3XFLAG-MYC-CMV-24 expression vector (Sigma) and containing a MYC tag followed by the polyadenylation sequence of the human growth hormone gene, was inserted just before the stop codon of the HOXA5 protein using an overlapping PCR strategy with synthetic oligonucleotide primers covering the appropriate sequences [63]. A 5'-extended version of the 1.8 kb cDNA expression vector including the distal ATG codon in-frame with the HOXA5 open reading frame (ORF) was also designed. It contains upstream genomic sequences up to the *Eco*RV site at position 2555 bp. This plasmid was made with and without the MYC tag. We also produced a pcDNA3 expression vector carrying the HOXA5-MYC version for the 5 kb-*Hoxa5* cDNA as well as for the 5 kb-*Hoxa6/a5* and for the 9.5 kb cDNAs. For these last two

plasmids, a FLAG tag, from the p3XFLAG-MYCCMV-24 vector, was added just before the HOXA6 stop codon. The TnT7 Quick coupled transcription-translation system (Promega) was used to produce [³⁵S] methionine-labeled proteins according to the manufacturer's protocol. The translation products were analyzed by electrophoresis on a 12% sodium dodecyl sulfate-polyacrylamide gel (SDS-PAGE) and revealed by autoradiography.

HEK293 cells were transiently transfected by the calcium phosphate method in 60 mm petri dishes with 10 mg/dish of the Hox expression vectors [64]. The pEGFP-C2 expression vector was used to assess transfection efficiency (10 mg/dish; Clontech) and a pMEK1-MYC-FLAG plasmid (10 mg/dish; provided by Dr. Jean Charron) was used as a positive control for the immunodetection of the MYC and FLAG tags. Transfection of each plasmid was done in duplicate. Chloroquine was added at a final concentration of 50 mM. Three hours after transfection, the cells were shocked for 30 seconds at 37 °C with 15% glycerol in HEPES-buffered saline. Forty-eight hours after transfection, protein extracts were obtained after cell lysis in 300 µl of icecold lysis buffer (20 mM Tris-HCl pH 8.0, 1% NP-40, 10 mM EGTA, 5 mM MgCl₂, 20 mM glycerol 2-phosphate, 25 mM NaF, 1 mM Na₃VO₄ and a proteinase inhibitor cocktail (Complete Mini EDTA-free; Roche Diagnostics)). After 15 minutes on ice, the extracts were centrifuged at 15,000 g at 4°C. Protein content of the supernatant was quantified using a Lowrybased assay (DC Protein Assay, Bio-Rad), and 20 µg of total protein lysate was resolved on a denaturing 12% SDS-PAGE, electrotransferred onto nitrocellulose (PALL) and probed overnight at 4°C with either a MYC-tag rabbit monoclonal antibody at a dilution of 1/1,000 (Cell Signaling Technology) or an anti- FLAG mouse monoclonal antibody at a dilution of 1/10,000 (Sigma) according to manufacturer's instructions. Membranes were also incubated with a mouse monoclonal anti-GAPDH antibody at a dilution of 1/20,000 (Fitzgerald Industries International) for loading control. Membranes were then incubated with the appropriate secondary horseradish peroxidaseconjugated antibody (a donkey anti-rabbit IgG at a dilution of 1/100,000 or a donkey anti-mouse at a dilution of 1/80,000; Jackson ImmunoResearch Laboratories). Proteins were revealed by chemiluminescence using the Western Lighting Plus-ECL system (PerkinElmer) according to manufacturer's protocol.

Total RNA from transfected HEK293 cells was isolated according to the TRIzol RNA extraction protocol (Invitrogen) and cDNA was synthesized with Superscript II Reverse Transcriptase (Invitrogen) using random primers. Reverse-transcription-PCR (RT-PCR) of the *Hoxa5-myc* portion was used to validate the expression of each transfected cDNA. PCR was performed for 25 cycles with an annealing temperature of 60°C. A 305 bp fragment was amplified with the *Hoxa5* forward primer 5'-CCCAG ATCTACCCCTGGATG-3' and the MYC-tag reverse primer 5'-GATGAGTTT TTGTTTCGGGGC-3'.

Skeletal analysis

Whole-mount skeletons were prepared from *Hoxa5*; *Hoxa6* compound newborn pups with Alcian blue for staining the cartilage and Alizarin red for staining the bone, as described in [27]. Skeletons were observed, and left and right sides of each vertebra were scored independently for bilateral markers.

ACKNOWLEDGMENTS

We thank Dr. Jean Charron for insightful input on the manuscript and for the pMEK1-MYC-FLAG vector, Dr. Mario Capecchi for the *Hoxa6* mutant mouse line, Drs. Philippe Soriano and Valera Vasioukhin for the pSA-IRESbgeolox2PGKDTA plasmid, Éric Paquet for help with bioinformatics, and Marcelle Carter, Lisa Danielczak and Nicolas Lafond for skilled technical assistance.

AUTHOR CONTRIBUTIONS

Conceived and designed the experiments: LJ. Performed the experiments: YC ML JM JA MJ FABS ST OB FG CL. Analyzed the data: YC JM JA LJ. Contributed reagents/materials/analysis tools: CKT LJ. Wrote the paper: LJ. Contributed to manuscript redaction: JA.

REFERENCES

1. Grier DG, Thompson A, Kwasniewska A, McGonigle GJ, Halliday HL, et al. (2005) The pathophysiology of HOX genes and their role in cancer. *J Pathol* 205: 154–171.
2. Krumlauf R (1994) Hox genes in vertebrate development. *Cell* 78: 191–201.
3. Bradshaw MS, Shashikant CS, Belting H-G, Bollekens JA, Ruddle FH (1996) A long-range regulatory element of Hoxc8 identified by using the pCluster vector. *Proc Natl Acad Sci USA* 93: 2426–2430.
4. Folberg A, Kovacs EN, Featherstone MS (1997) Characterization and retinoic acid responsiveness of the murine Hoxd4 transcription unit. *J Biol Chem* 272: 29151–29157.
5. Sham MH, Hunt P, Nonchev S, Papalopulu N, Graham A, et al. (1992) Analysis of the murine Hox-2.7 gene: conserved alternative transcripts with differential distributions in the nervous system and the potential for shared regulatory regions. *EMBO J* 11: 1825–1836.
6. Gould A, Morrison A, Sproat G, White RAH, Krumlauf R (1997) Positive crossregulation and enhancer sharing: two mechanisms for specifying overlapping Hox expression patterns. *Genes Dev* 11: 900–913.
7. Kmita M, Duboule D (2003) Organizing axes in time and space; 25 years of colinear tinkering. *Science* 301: 331–333.
8. Sharpe J, Nonchev S, Gould A, Whiting J, Krumlauf (1998) Selectivity, sharing and competitive interactions in the regulation of Hoxb genes. *EMBO J* 17: 1788–1798.
9. Duboule D (2007) The rise and fall of Hox gene clusters. *Development* 134: 2549–2560.
10. Kmita M, Fraudeau N, He´rault Y, Duboule D (2002) Serial deletions and duplications suggest a mechanism for the collinearity of Hoxd genes in limbs. *Nature* 420: 145–150.
11. Kondo T, Duboule D (1999) Breaking colinearity in the mouse HoxD complex. *Cell* 97: 407–417.
12. Spitz F, Gonzalez F, Duboule D (2003) A global control region defines a chromosomal regulatory landscape containing the HoxD cluster. *Cell* 113: 405–417.
13. Chambeyron S, Bickmore WA (2004) Chromatin decondensation and nuclear reorganization of the HoxB locus upon induction of transcription. *Genes Dev* 18: 1119–1130.
14. Morey C, Da Silva NR, Perry P, Bickmore WA (2007) Nuclear reorganisation and chromatin decondensation are conserved, but distinct, mechanisms linked to Hox gene activation. *Development* 134: 909–919.
15. Brend T, Gilthorpe J, Summerbell D, Rigby PW (2003) Multiple levels of transcriptional and post-transcriptional regulation are required to define the domain of Hoxb4 expression. *Development* 130: 2717–2728.
16. Naguibneva I, Ameyar-Zazoua M, Poleskaya A, Ait-Si-Ali S, Groisman R, et al. (2005) The microRNA miR-181 targets the homeobox protein Hox-A11 during mammalian myoblast differentiation. *Nat Cell Biol* 8: 278–284.
17. Yekta S, Shih I-h, Bartel DP (2004) MicroRNA-directed cleavage of HOXB8 mRNA. *Science* 304: 594–596.

18. Hsieh-Li HM, Witte DP, Weinstein M, Brandford W, Li H, et al. (1995) Hoxa11 structure, extensive antisense transcription, and function in male and female fertility. *Development* 121: 1373–1385.
19. Mainguy G, Koster J, Woltering J, Jansen H, Durston A (2007) Extensive polycistronic and antisense transcription in the mammalian Hox clusters. *PLoS ONE* 2: e356.
20. Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, et al. (2007) Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* 129: 1311–1323.
21. Sasaki YTF, Sano M, Kin T, Asai K, Hirose T (2007) Coordinated expression of ncRNAs and HOX mRNAs in the human HOXA locus. *BBRC* 357: 724–730.
22. Sessa L, Breiling A, Lavorgna G, Silvestri L, Casari G, et al. (2007) Noncoding RNA synthesis and loss of Polycomb group repression accompanies the colinear activation of the human HOX cluster. *RNA* 13: 223–239.
23. Jeannotte L, Lemieux M, Charron J, Poirier F, Robertson EJ (1993) Specification of axial identity in the mouse: role of the Hoxa-5 (Hox1.3) gene. *Genes Dev* 7: 2085–2096.
24. Okada Y, Jiang Q, Lemieux M, Jeannotte L, Su L, et al. (2006) Leukaemic transformation by CALM-AF10 involves upregulation of Hoxa5 by hDOTIL. *Nat Cell Biol* 8: 1017–1024.
25. Gendronneau G, Lemieux M, Morneau M, Paradis J, Te'tu B, et al. (2010) Influence of Hoxa5 on p53 tumorigenic outcome in mice. *Am J Pathol* 176: 995–1005.
26. Aubin J, Lemieux M, Tremblay M, Be'nard J, Jeannotte L (1997) Early postnatal lethality in Hoxa-5 mutant mice is attributable to respiratory tract defects. *Dev Biol* 192: 432–445.
27. Aubin J, Lemieux M, Tremblay M, Behringer RR, Jeannotte L (1998) Transcriptional interferences at the Hoxa4/Hoxa5 locus: Importance of correct Hoxa5 expression for the proper specification of the axial skeleton. *Dev Dyn* 212: 141–156.
28. Aubin J, Chailier P, Me'nard D, Jeannotte L (1999) Loss of Hoxa5 gene function in mice perturbs intestinal maturation. *Am J Physiol* 277: C965–C973.
29. Aubin J, De'ry U, Lemieux M, Chailier P, Jeannotte L (2002) Stomach regional specification requires Hoxa5-driven mesenchymal-epithelial signaling. *Development* 129: 4075–4087.
30. Dony C, Gruss P (1987) Specific expression of the Hox1.3 homeo box gene in murine embryonic structures originating from or induced by the mesoderm. *EMBO J* 6: 2965–2975.
31. Gaunt SJ, Coletta PL, Pravtcheva D, Sharpe PT (1990) Mouse Hox-3.4: homeobox sequence and embryonic expression patterns compared to other members of the Hox gene network. *Development* 109: 329–339.
32. Larochelle C, Tremblay M, Bernier D, Aubin J, Jeannotte L (1999) Multiple cisacting regulatory regions are required for the restricted spatio-temporal Hox5 gene expression. *Dev Dyn* 214: 127–140.
33. Tabariès S, Lemieux M, Aubin J, Jeannotte L (2007) Comparative analysis of Hoxa5 allelic series. *Genesis* 45: 218–228.
34. Garin E, Lemieux M, Coulombe Y, Robinson GW, Jeannotte L (2006) Stromal Hoxa5 function controls the growth and differentiation of mammary alveolar epithelium. *Dev Dyn* 235: 1858–1871.

35. Meunier D, Aubin J, Jeannotte L (2003) Perturbed thyroid morphology and transient hypothyroidism symptoms in *Hoxa5* mutant mice. *Dev Dyn* 227: 367–378.
36. Odenwald WF, Taylor CF, Palmer-Hill FJ, Friedrich Jr. V, Tani M, et al. (1987) Expression of a homeo domain protein in noncontact-inhibited cultured cells and postmitotic neurons. *Genes Dev* 1: 482–496.
37. Zakany J, Tuggle CK, Patel M, Nguyen-Huu MC (1988) Spatial regulation of homeobox gene fusions in the embryonic central nervous system of transgenic mice. *Neuron* 1: 679–691.
38. Püschel A, Balling WR, Gruss P (1990) Position-specific activity of the *Hox1.1* promoter in transgenic mice. *Development* 108: 435–442.
39. Moreau J, Jeannotte L (2002) Sequence analysis of a *Hoxa4-Hoxa5* intergenic region including shared regulatory elements. *DNA Seq* 13: 203–209.
40. Nowling T, Zhou W, Krieger KE, Larochelle C, Nguyen-Huu MC, et al. (1999) *Hoxa5* gene regulation: A gradient of binding activity to a brachial spinal cord element. *Dev Biol* 208: 134–146.
41. Tabariès S, Lapointe J, Besch T, Carter M, Woollard J, et al. (2005) Cdx protein interaction with *Hoxa5* regulatory sequences contributes to *Hoxa5* regional expression along the axial skeleton. *Mol Cell Biol* 25: 1389–1401.
42. Packer AI, Crotty DA, Elwell VA, Wolgemuth DJ (1998) Expression of the murine *Hoxa4* gene requires both autoregulation and a conserved retinoic acid response element. *Development* 125: 1991–1998.
43. Jeannotte L, Ruiz JC, Robertson EJ (1991) Low level of *Hox1.3* gene expression does not preclude the use of promoterless vectors to generate a targeted gene disruption. *Mol Cell Biol* 11: 5578–5585.
44. Kostic D, Capecchi MR (1994) Targeted disruptions of the murine *Hoxa-4* and *Hoxa-6* genes result in homeotic transformations of components of the vertebral column. *Mech Dev* 46: 231–247.
45. Karolchik D, Hinrichs AS, Kent WJ (2009) The UCSC genome browser. *Curr Protoc Bioinformatics*. Dec., Chapter 1: Unit1.4.
46. Joksimovic M, Jeannotte L, Tuggle CK (2005) Dynamic expression of murine *HOXA5* protein in the central nervous system. *Gene Expr Patterns* 5: 792–800.
47. Popovic R, Erfurth F, Zeleznik-Le N (2008) Transcriptional complexity at the *HOXA9* locus. *Blood cells, Molec Diseases* 40: 156–159.
48. Dintilhac A, Bihan R, Guerrier D, Deschamps S, Pellerin I (2004) A conserved non-homeodomain *Hoxa9* isoform interacting with CBP is co-expressed with the typical *Hoxa9* protein during embryogenesis. *Gene Expr Patterns* 4: 215–222.
49. Bijl J, Sauvageau M, Thompson A, Sauvageau G (2005) High incidence of proviral integrations in the *Hoxa* locus in a new model of E2a-PBX1-induced Bcell leukemia. *Genes Dev* 19: 224–233.
50. Kozak M (2005) Regulation of translation via mRNA structure in prokaryotes and eukaryotes *Gene* 361: 13–37.
51. Simeone A, Pannese M, Acampora D, D’Esposito M, Boncinelli E (1988) At least three human homeoboxes on chromosome 12 belong to the same transcription unit. *Nucl Acids Res* 16: 5379–5390.
52. Santini S, Boore JL, Meyer A (2003) Evolutionary conservation of regulatory elements in vertebrate *Hox* clusters. *Genome Res* 13: 1111–1122.

53. Brock HW, Hodgson JW, Petruk S, Mazo A (2009) Regulatory noncoding RNAs at Hox loci. *Biochem Cell Biol* 87: 27–34.
54. Pontig CP, Oliver PL, Reik W (2009) Evolution and functions of long noncoding RNAs. *Cell* 136: 629–641.
55. Sharp PA (2009) The centrality of RNA. *Cell* 136: 577–580.
56. Gingeras TR (2009) Implications of chimeric non-co-linear transcripts. *Nature* 461: 206–211.
57. Kapranov P, Willingham AT, Gingeras TR (2007) Genome-wide transcription and the implications for genomic organization. *Nat Rev Genet* 8: 413–423.
58. Aviv H, Leder P (1972) Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid-cellulose. *Proc Natl Acad Sci US* 69: 1408–1412.
59. Chen WV, Soriano P (2003) Gene trap mutagenesis in embryonic stem cells. *Meth Enzymol* 365: 367–386.
60. Hogan B, Beddington RS, Constantini F, Lacy E (1994) *Manipulating the mouse embryo: a laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
61. Wilkinson DG, Nieto MA (1993) Detection of messenger RNA by in situ hybridization to tissue sections and whole mounts. *Meth Enzymo* 225: 361–372.
62. Jaffe L, Jeannotte L, Bikoff EK, Robertson EJ (1990) Analysis of b2- microglobulin gene expression in the developing mouse embryo and placenta. *J Immunol* 145: 3474–3482.
63. Wang W, Malcolm BA (1999) Two-stage PCR protocol allowing introductions of multiple mutation deletions and insertions using quick-change site-directed mutagenesis. *Biotechniques* 26: 680–682.
64. Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY .

Figure S5.1 Molecular characterization of the *Hoxa5* alternate transcripts by 5'-RACE and RT-PCR.

(A) The transcription initiation site of the larger transcripts was determined by 5'-RACE. Genomic organization of the *Hoxa5*, *Hoxa6* and *Hoxa7* genes along the cluster is shown. Black, grey and open boxes indicate homeobox, translated, and transcribed sequences, respectively. The primers used are indicated (arrows). With primer 1, we obtained several clones showing that the 9.5 and 11.0 kb transcripts initiate in the *Hoxa6-Hoxa7* intergenic region at position -8905 bp. With primers 2 and 3, we obtained 5'-RACE products that revealed the presence of a 4.6 kb intron in the 5.0 kb transcript. The initiation site of this transcript also coincides with that of the 9.5 and 11.0 kb transcripts. A second population of clones was obtained with a transcription start site at position -4409 bp, which corresponds to the putative first base of *Hoxa6* exon 1. (B) By using various primer combinations in RT-PCR experiments, we established the molecular structure of the different transcripts. We also demonstrated that the 5.0 kb band detected by northern analyses include minor splicing variants and two major RNA species, one initiating at -8905 bp and a second starting at -4409 bp. The latter corresponds to the putative *Hoxa6* transcript. This transcript uses the polyA site of the *Hoxa5* gene. This result correlates with the absence of a functional polyadenylation sequence at the *Hoxa6* locus as shown by northern analysis. We also showed that splicing of the *Hoxa5* intron is not always complete and a low percentage of the *Hoxa5* transcripts contained intron sequences. A, *AccI*; B, *BglII*; Ba, *BamHI*; H, *HindIII*; K, *KpnI*; RI, *EcoRI*; S, *SacI*; St, *StuI*; Xh, *XhoI*.
Found at: doi:10.1371/journal.pone.0010600.s001 (0.26 MB PPT)

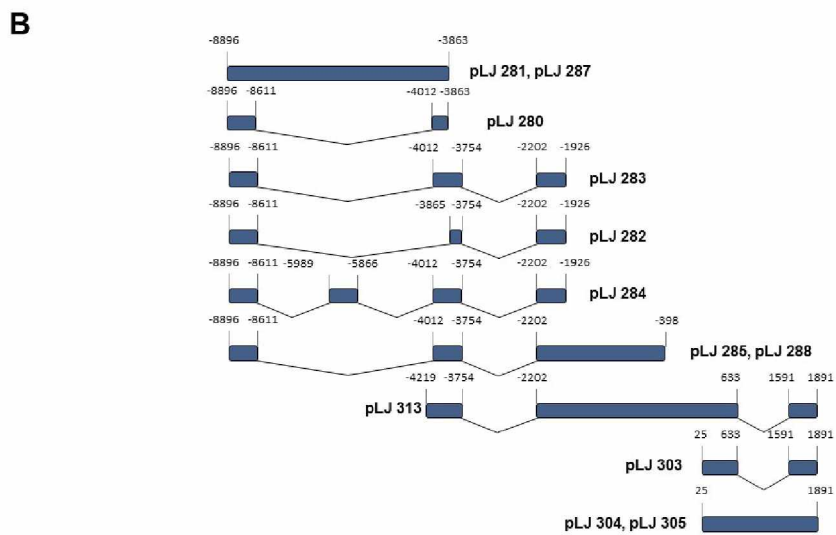
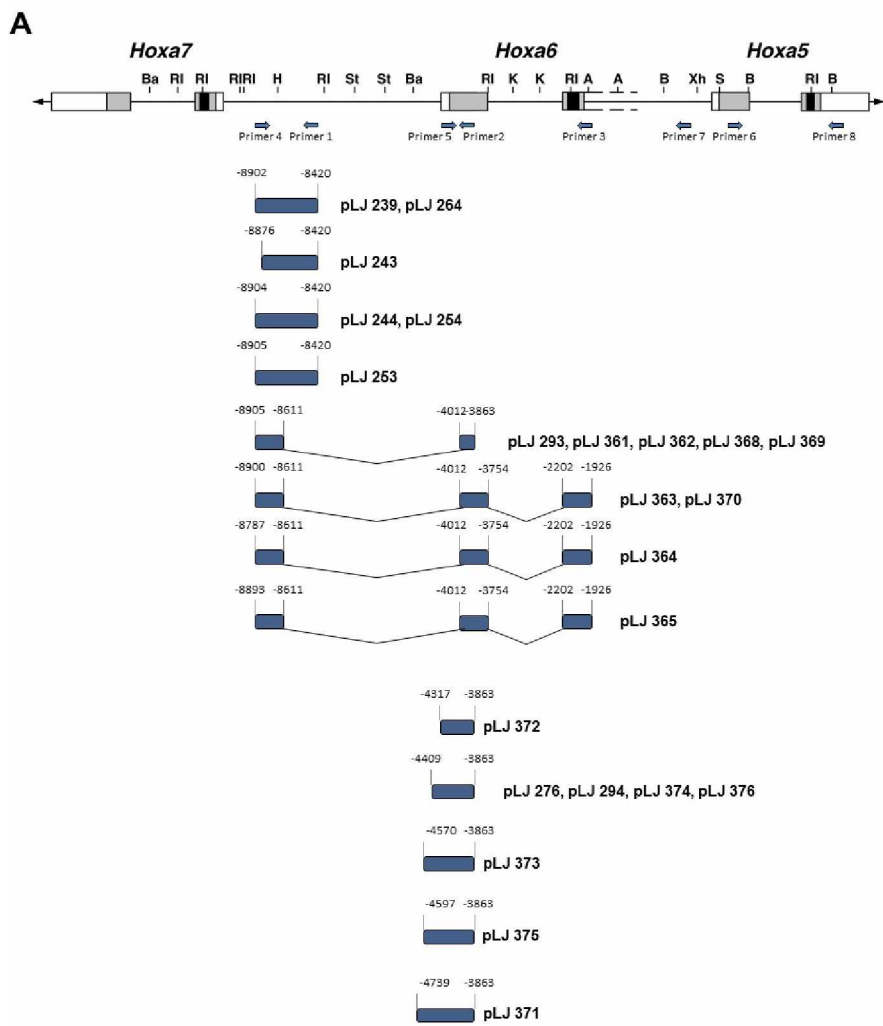
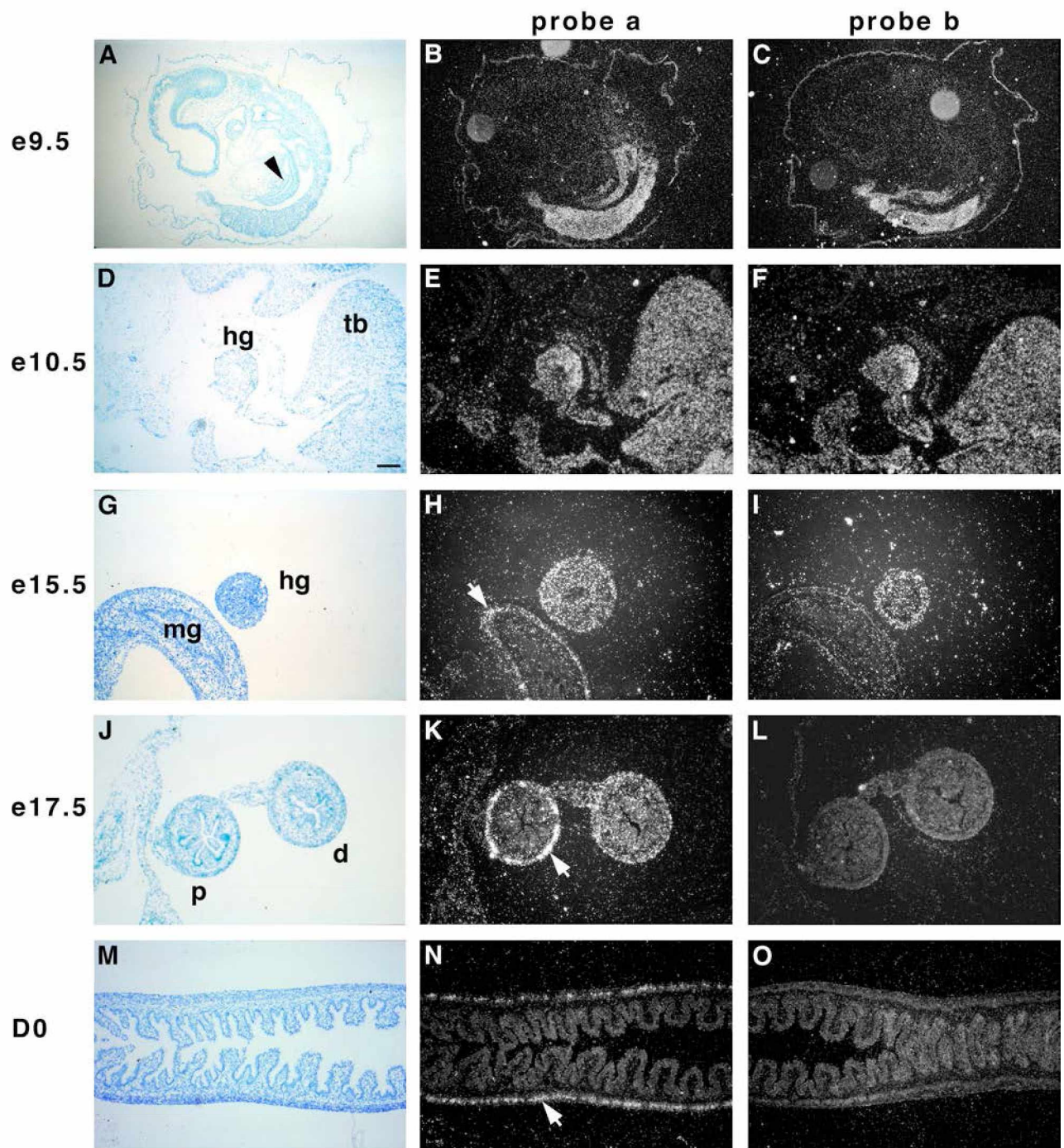


Figure S5.2 *Hoxa5* expression in the developing hindgut.

Sections of E9.5 (A-C) and E10.5 (D-F) mouse embryos, and E15.5 (G-I), e17.5 (J-L) and DO (birth; M-O) hindgut tissues were hybridized with either probe a (B, E, H, K, N) or probe b (C, F, I, L, O). Bright-field views are shown on the left panels. (A-B) At E9.5, probe a detects *Hoxa5* transcripts along the gut up to the caudal foregut (arrowhead). (C) Expression with probe b is restricted to a more posterior region. (D-I) From E10.5 to E15.5, both probes reveal signal in the mesenchyme of the hindgut, while probe a detects *Hoxa5* transcripts in the myenteric plexi of the midgut (white arrow). (J-K) At E17.5, signal with probe a is confined to myenteric plexi in the proximal part of the hindgut (white arrow) while it still displays a diffuse mesenchymal expression in the distal hindgut. (M-N) Plexi of the enteric nervous system remain positive for probe a after birth as shown for DO. (L, O) No expression is observed with probe b from E17.5 onwards. d, distal hindgut; hg, hindgut; mg, midgut; p, proximal hindgut; tb, tailbud. Scale bar, 100 μ m. Found at: doi:10.1371/journal.pone.0010600.s002 (6.19 MB TIF)



CHAPITRE VI

Partial functional redundancy between *Hoxa5* and *Hoxb5* paralog genes during lung morphogenesis.

Olivier Boucherat,^{1*} Séverine Montaron,^{1*} Félix-Antoine Bérubé-Simard,¹ Josée Aubin,¹
Polyxeni Philippidou,² Deneen M. Wellik,³ Jeremy S. Dasen,² and Lucie Jeannotte¹

1 Centre de recherche en cancérologie de l'Université Laval, Centre Hospitalier
Universitaire de Québec, L'Hôtel-Dieu de Québec, Québec, Canada

2 Howard Hughes Medical Institute, Smilow Neuroscience Program, Department of
Physiology and Neuroscience, New York University, New York, New York

3 University of Michigan Medical Center, Ann Arbor, Michigan

Article publié en 2013 dans : *Am J Physiol Lung Cell Mol Physiol* **304**: L817–L830,

AVANT-PROPOS

Les gènes *Hox* codent pour des facteurs de transcription jouant des rôles essentiels dans l'établissement du squelette axial, abaxial, appendiculaire, de même que dans de nombreux processus d'organogenèse. Bien que plusieurs gènes *Hox* soient exprimés dans le poumon en développement, seule la mutation du gène *Hoxa5* génère des malformations sévères des voies respiratoires résultant en un pourcentage élevé de mortalité périnatale, ce qui suggère une forte redondance fonctionnelle entre les autres gènes *Hox*.

Puisque les gènes *Hoxa5* et *Hoxb5* sont tous deux exprimés au niveau du mésenchyme pulmonaire, nous avons entrepris de discriminer les différentes fonctions propres à chacun de ces paralogues de celles partagées entre les deux gènes dans l'ontogénie du système respiratoire. Tout d'abord, même si le taux de mortalité périnatale observé chez les individus *Hoxa5*^{-/-} augmente suite à l'inactivation des allèles du gène *Hoxb5*, il semble y avoir une tendance à l'augmentation de la quantité du transcrit *Hoxa5* chez les souris *Hoxb5*^{-/-}, suggérant une certaine compensation du gène *Hoxa5* pour maintenir ses fonctions partagées avec *Hoxb5*. Ensuite, le ratio poids poumon/poids embryon, de même que le nombre d'acini pulmonaires diminuent en fonction du nombre d'allèles *Hox5* mutés, alors que la diminution de prolifération semble davantage associée aux poumons mutants *Hoxa5*^{-/-}. De plus, les gènes *Hoxa5* et *Hoxb5* sont tous les deux impliqués dans la différenciation cellulaire, avec un rôle spécifique au gène *Hoxa5* dans l'établissement des cellules de Clara et un rôle prépondérant pour le gène *Hoxa5* dans l'établissement des pneumocytes de type 1 ainsi que des cellules endothéliales. Chez les souris adultes, même si l'augmentation du nombre de cellules productrice de mucus associée à l'activité de la voie NOTCH est également observée chez les souris *Hoxb5*^{-/-}, ce phénotype est beaucoup plus important chez les mutants *Hoxa5*^{-/-}.

En ce qui me concerne, suite à la validation du rôle spécifique du gène *Hoxa5* dans le développement de la trachée et à l'observation que l'expression gène *Fstll1*, nécessaire à la formation du cartilage trachéal, est diminué dans les mutants *Hoxa5*^{-/-}, j'ai vérifié si la protéine HOXA5 lie directement les séquences du gène *Fstll1*. Pour ce faire, j'ai fait ressortir par analyse bioinformatique des régions riches en sites de liaison pour les

protéines HOX en amont du promoteur du gène *Fstl1* (Fig. 6.4J). J'ai ensuite démontré, par immunoprécipitation de la chromatine (ChIP) de trachées d'embryons de souris E18.5, que la protéine HOXA5 se lie spécifiquement à ces différentes régions *in vivo* (Fig. 6.4K).

Enfin, le problème d'innervation du diaphragme par le nerf phrénique semble spécifique à l'inactivation du gène *Hoxa5* tel qu'observé par immunofluorescence.

ABSTRACT

Hox genes encode transcription factors governing complex developmental processes in several organs. A subset of *Hox* genes are expressed in the developing lung. Except for *Hoxa5*, the lack of overt lung phenotype in single mutants suggests that *Hox* genes may not play a predominant role in lung ontogeny or that functional redundancy may mask anomalies. In the *Hox5* paralog group, both *Hoxa5* and *Hoxb5* genes are expressed in the lung mesenchyme whereas *Hoxa5* is also expressed in the tracheal mesenchyme. Herein, we generated *Hoxa5*;*Hoxb5* compound mutant mice to evaluate the relative contribution of each gene to lung development. *Hoxa5*;*Hoxb5* mutants carrying the four mutated alleles displayed an aggravated lung phenotype, resulting in the death of the mutant pups at birth. Characterization of the phenotype highlighted the role of *Hoxb5* in lung formation, the latter being involved in branching morphogenesis, goblet cell specification, and postnatal air space structure, revealing partial functional redundancy with *Hoxa5*. However, the *Hoxb5* lung phenotypes were less severe than those seen in *Hoxa5* mutants, likely because of *Hoxa5* compensation. New specific roles for *Hoxa5* were also unveiled, demonstrating the extensive contribution of *Hoxa5* to the developing respiratory system. The exclusive expression of *Hoxa5* in the trachea and the phrenic motor column likely underlies the *Hoxa5*-specific trachea and diaphragm phenotypes. Altogether, our observations establish that the *Hoxa5* and *Hoxb5* paralog genes shared some functions during lung morphogenesis, *Hoxa5* playing a predominant role.

INTRODUCTION

To permit respiratory gas exchanges between blood and environmental air, the lung must have an efficient diffusible interface that results from finely regulated developmental processes. In the mouse, lung ontogeny initiates around embryonic day (E) 9 with the outpocketing of the ventral foregut endoderm into the splanchnic mesenchyme to form the laryngotracheal groove. The lung bud then elongates and undergoes a stereotypic pattern of branching leading to the formation of the respiratory tree. Starting at E16.5, the prospective gas-exchange region develops and capillaries get apposed to the flattened respiratory epithelium to form the first blood-air barriers. Lung development culminates postnatally from day (D) 5 to D30 with the formation of alveoli, the final gasexchange units. Throughout lung formation, undifferentiated epithelial cells present in the primary buds proliferate and differentiate into specialized cell types regionally located along the proximodistal axis of the respiratory tract (38, 55).

Lung morphogenesis is influenced by physical forces. For instance, impairment of fetal breathing movements due to a nonfunctional diaphragm can cause lung hypoplasia (25, 27, 32). Moreover, lung development is under the concerted action of multiple molecules organized in distinct networks that control the patterning of the respiratory tract (38). Reciprocal interactions between lung mesenchyme and its flanking epithelium also govern lung morphogenesis and epithelial cell fate determination (48). Despite accumulated evidence showing that mesenchymal cells can instruct the epithelium, the nature of the mesenchymal factors involved has yet to be resolved.

Hox genes encode transcription factors specifying the regionalization of the body plan and regulating morphogenesis during animal development. In human and mouse, 39 *Hox* genes are organized in four clusters located on different chromosomes. The 3' to 5' position of each *Hox* gene within a cluster corresponds to its spatiotemporal expression domain along the anterior-posterior axis of the embryo. Different members of the *Hox* complexes are thus expressed in overlapping domains along the developing body, suggesting that specific combinations of HOX proteins provide a unique address to a particular region. On the basis of sequence homologies and position within clusters, *Hox*

genes are also classified into 13 paralog groups (34). The similarities in protein structure and expression pattern among genes from the same paralog group have led to the hypothesis that *Hox* paralogs perform partially redundant and/or overlapping functions. Indeed, compound mutant mice for *Hox* paralogs often exhibit a more severe phenotype than mutant mice for a single *Hox* gene (10, 35, 44, 45, 56). Furthermore, knock-in substitutions of *Hox* genes by their paralogs have demonstrated that they can fulfill similar roles (22, 51).

In the lung, the proximodistal distribution of the different structures (trachea, bronchi, bronchioli, alveoli) and cell types suggests that the respiratory tract can be specified by *Hox* regulated mechanisms similar to those governing anteriorposterior axial skeleton formation. *Hox* genes, predominantly from paralogous groups 1 to 8, are mainly expressed in lung mesenchyme with a distinct spatiotemporal profile, supporting a role in the regional specification of the respiratory tract (5, 26). However, except for *Hoxa5*, the lack of overt lung phenotype in single mutants suggests that *Hox* genes may not play a predominant role in lung ontogeny or that functional redundancy may mask anomalies.

Hoxa5, *Hoxb5*, and *Hoxc5* are members of the *Hox5* paralog group. *Hoxb5* and *Hoxc5* mutant mice are viable and no organ defect has been described (8, 43). In contrast, the loss of *Hoxa5* function results in a panoply of phenotypes indicative of the broad range of *Hoxa5* actions throughout life, including tracheal and lung dysmorphogenesis responsible for the high neonatal mortality rate of *Hoxa5*^{-/-} pups (1, 2, 4, 17, 19, 29, 36). Surviving *Hoxa5*^{-/-} mice display an emphysema-like phenotype characterized by air space enlargement and goblet cell metaplasia (33).

Even though no lung phenotype was reported in *Hoxb5* mutant mice, *Hoxb5* lung expression, in vitro studies, and expression data in human lung diseases suggest a role for *Hoxb5* in lung development (6, 43, 52, 53). The severity of the *Hoxa5* lung phenotype indicates that *Hoxa5* function is less subject to rescue by other genes. However, it is possible that the other *Hox5* paralogs exert functions in lung development hidden by *Hoxa5* compensation. A threshold level of HOX5 proteins may also be required for specific aspects of development, and mutation of several paralogs may be needed to reveal defects

otherwise not detectable (10, 35, 45). Herein, genetic interactions between *Hoxa5* and *Hoxb5* paralog genes during lung morphogenesis were investigated. Our in vivo data confirm the importance of *Hoxa5* and uncovers the role of *Hoxb5* in lung morphogenesis. *Hoxa5* and *Hoxb5* genes share functions in the developing lung, *Hoxa5* playing a predominant role.

MATERIALS AND METHODS

Mice, genotyping, and tissue collection.

The *Hoxa5* and *Hoxb5* mutant mouse lines were maintained in the 129/Sv inbred background. *Hoxa5*^{-/-} mice were mated with *Hoxb5*^{-/-} mice to obtain double heterozygous animals (*Hoxa5*^{-/-};*Hoxb5*^{-/-}). The latter were intercrossed to generate mice of all possible allelic combinations. Age of the embryos was estimated by considering the morning of the day of the vaginal plug as E0.5. Experimental animals were genotyped by Southern blot analysis (29, 43).

Lungs were collected from wild-type (WT) and mutant embryos at E13.5, E15.5, and E18.5 and from adults at D65±5. The wet lung and body weights were obtained by direct measurement to determine the ratio. Excised lungs were fixed in 4% cold paraformaldehyde, paraffin embedded, and sectioned at 4 µm (28). For adult specimens, lungs were instilled with 4% cold paraformaldehyde before processing. Diaphragms were dissected from E18.5 embryos and fixed in 1% paraformaldehyde before processing for immunofluorescence (IF) staining. For RNA extraction, lungs were snap frozen in N2. Animal experimentations were performed according to the guidelines of the Canadian Council on Animal Care and approved by the institutional animal care committee.

Histology, IHC, and IF analyses.

Lung sections were stained according to standard histological procedures: hematoxylin and eosin for morphology, Alcian blue for detection of mucus-producing goblet cells, and Weigert with tartrazine counterstaining to visualize elastic fibers. Immunohistochemistry (IHC) experiments were performed as described (19). Slides were

counterstained with either Alcian blue and nuclear fast red or methyl green. For phosphohistone H3 (pHH3) immunostaining, ratio of positive-stained cells to total cell number was evaluated for a minimum of six random areas with a total of at least 2,200 cells per specimen. Whole mounts of diaphragms were IF stained to label axons and nerve terminals (24). The primary antibodies used were a rabbit monoclonal antibody against pHH3 (1/200 dilution; Cell Signaling), a rabbit antibody against cleaved caspase-3 (1/200; Cell Signaling), a goat antibody against CC10 (1/500; gift from Dr. G. Singh), a Syrian hamster monoclonal antibody against podoplanin (T1 α ; 1/1,000; DSHB), a goat antibody against FOXA2 (1/100; Santa Cruz Biotechnology), a rabbit antibody against pro- SP-C (1/500; Millipore), a rat monoclonal against PECAM-1 (1/75; BD Biosciences Pharmingen), a rabbit antibody against neurofilament (1/1,000; Millipore), a rabbit antibody against synaptophysin (1/5; Invitrogen), a rabbit antibody against N1ICD (1/250; Abcam), and a rabbit antibody against HEY2 (1/250; Millipore). The biotinylated secondary antibodies used were a goat anti-rabbit antibody (1/300; Vector Laboratories), a swine anti-goat antibody (1/300; Cedarlane), a goat anti-Syrian hamster antibody (1/300; Cedarlane), and a goat anti-rat antibody (1/500; Cedarlane). The fluorescent secondary antibodies FITC-conjugated donkey anti-goat IgG (1/250; Cedarlane) and Alexa Fluor 488-conjugated donkey anti-rabbit IgG (1/1,000; Invitrogen) were also used. Finally, the Alexa Fluor 555-conjugated α -bungorotoxin (1/1,000; Invitrogen) was used.

Morphometry.

Radial alveolar count, a measure of the complexity of the respiratory acini, was determined by counting the number of air spaces lying on a line drawn perpendicularly from the center of a terminal or respiratory bronchiole to the closest edge of the pleural tissue (15). For tracheal luminal measurement, images of histological sections of WT and mutant specimens were visualized with a Leica SCN400 slide scanner (Leica Microsystems, Concord, ON, Canada). The apical epithelial surface of each image was traced and quantified by using the Adobe Photoshop CS3 software.

In situ hybridization analyses.

Radioactive in situ hybridization experiments were performed on frontal sections of E13.5 embryos as described (28). The following murine sequences were used as specific templates for synthesizing [³⁵S]UTP-labeled riboprobes: a 850-bp *Bgl*II-*Hind*III genomic fragment containing the 3'-untranslated region of *Hoxa5* exon 2 (11), a 430-bp *Bam*HI-*Hind*III cDNA fragment covering part of *Hoxb5* exon 1 (31), and a 200-bp *Eco*RI-*Xba*I genomic fragment corresponding to *Hoxc5* exon 2 (18). After exposure, slides were counterstained with toluidine blue. Experiments were performed on two to five specimens per genotype tested.

qRT-PCR experiments.

Total RNA was isolated from the trachea/primary bronchi and the lungs of E18.5 embryos and from the entire respiratory tract of E15.5 embryos according to the TRIzol reagent procedure (Invitrogen, Carlsbad, CA). cDNA was synthesized with the Superscript II Reverse Transcriptase (Invitrogen) using random primers. Quantitative (q)PCR was performed with Power SYBR Green PCR Master Mix (Applied Biosystems, Foster City, CA) and a thermal cycler ABI PRISM 7000. Samples were analyzed in triplicate. Five to eight specimens were used for each genotype tested. Primer sequences used were as follows: *Fstl1* forward ATGGCGACTCTCACCTGGAC, reverse CAATGAGGGCGTCAACACAG; *Hoxa5* forward CCCAGATCTACCCCTGGATG, reverse GGCATGAGCTATTTTCGATCCT; *Hoxb5* forward TATCCCCTGGA TGAGGAAG, reverse GGGTCAGGTAGCGATTGAAG; *T1α* forward TGGCAAGG CACCTCTGGTA, reverse GGTGGACAGTTCCTCTAAGGGA; *Vegfa* forward TGC ACTGGACCCTGGCTTTAC, reverse CGGCAGTAGCTTCGCTGGTAG. The *Rpl19* gene was used as control with the primers forward GATCATCC GCAAGCCTGTGA and reverse GCATCCGAGCATTGGCAGTA.

ChIP assays.

Trachea from E18.5 WT embryos was isolated and mechanically disrupted prior to be cross-linked with 1% formaldehyde in PBS for 15 min at room temperature. Cross-linking was stopped by adding glycine to a final concentration of 0.125 M. Extracts were

then disrupted into crude lysates in 2 ml of swelling buffer (5 mM PIPES pH 8.0, 85 mM KCl, 1% Nonidet P-40, and protease inhibitors), equilibrated 20 min on ice, and centrifuged for 5 min at 3,000 rpm at 4°C. The pellets were eluted in 1 ml of chromatin immunoprecipitation (ChIP) lysis buffer (50 mM Tris·HCl pH 8.0, 10 mM EDTA, 1% SDS) with protease inhibitors (Complete Mini-EDTA-free; Roche Diagnostics, Mannheim, Germany). Sonication was performed by using a Bioruptor (Diagenode, Liège, Belgium) for 10 cycles of a 30-s pulse followed by a 30-s pause at the highest setting to obtain an average DNA size of 300–600 bp. Fragmented chromatin (100 µg) was incubated overnight at 4°C with Dynabeads linked to protein G (Invitrogen) and 2 µg of either rabbit anti-HOXA5 antibody (Aviva Systems Biology, San Diego, CA), rabbit antihistone H3 (ab1791; Abcam, Cambridge, MA), or rabbit IgG (Santa Cruz Biotechnology, Santa Cruz, CA). Postimmunoprecipitation washes, chromatin elution, cross-link reversion, and purification of the DNA fragments were done prior to qPCR amplification as described (41). qPCR-ChIP analyses were performed with the following primers: *FstII-1* forward AATACTGCCACTGGCCACTAACCA, reverse CTTTGGTCCCATGCCTTTGGGTT T; *FstII-2* forward AACACAGGTCCACCAGCTTGTATG, reverse GACATGTGATAGTCAATACGGTGC; *FstII-3* forward ACAGACTGTAGCT CTCAGGAAGCA, reverse TGGGACAAGGGCTTGATTTAGGGA; *FstII-4* forward ACTCTGAATCAAGGAGAATTATGTCCTA, reverse CCAGAGCTTGCAACATGC CGTATT; *ctl* locus forward TGGAGTGCTTACAAGGAAGGCAGA, reverse ACTAACCGCCATGTTCTTTTCGCTC. The values for the samples immunoprecipitated by anti-HOXA5, anti-histone H3, or control IgG were recorded as the percentage relative to input. ChIP results were confirmed by three independent experiments and qPCR was performed in triplicate for each sample.

Statistical analyses.

Student's *t*-test was performed to compare the lung weight-to-body weight ratio, the number of pHH3-positive cells, the number of acini, the CC10 intensity staining, the number of tracheal rings, the luminal surface of the trachea, the radial alveolar counts as well as for gene expression experiments. A significance level inferior to 5% ($P < 0.05$) was considered statistically significant.

RESULTS

Expression of *Hox5* paralog genes in the developing lung.

To define whether the three *Hox5* paralog genes share a similar expression profile along the respiratory tract, we performed comparative in situ hybridization experiments on lungs from E13.5 embryos using specific riboprobe for each gene. *Hoxa5* and *Hoxb5* transcripts were readily seen throughout the mesenchymal component of the distal lung, whereas only *Hoxa5* expression was observed in the mesenchyme surrounding the epithelium of the trachea and the primary bronchi (Fig. 6.1, A and B). No evident *Hoxc5* expression was detected along the respiratory tract (Fig. 6.1C). Thus only *Hoxa5* and *Hoxb5* genes displayed specific and overlapping expression domains in the developing lung. Therefore, we focused our attention on these two *Hox5* paralog genes.

The loss of *Hoxb5* alleles reduces the viability of *Hoxa5*^{-/-} pups.

Given that *Hox5* paralog genes demonstrate functional redundancy in axial patterning, we assessed whether *Hoxb5* plays a role in lung development that can be masked by *Hoxa5* compensation (35). In situ hybridization and qRT-PCR experiments revealed an increased *Hoxa5* expression in *Hoxb5*^{-/-} specimens that supported this notion (Fig. 6.1, F and H). In contrast, *Hoxb5* expression remained unchanged in lungs from *Hoxa5*^{-/-} embryos (Fig. 6.1, E and I). Even though mutant transcripts are produced from the *Hoxa5* and *Hoxb5* mutant alleles, both *Hoxa5* and *Hoxb5* mutations generate a null allele with no production of protein (11, 43). We observed a loss of *Hoxa5* expression in lungs from *Hoxa5*^{-/-} embryos and a reduced *Hoxb5* expression in *Hoxb5*^{-/-} specimens that might reflect the interference of the *neo* selection cassette on *Hox* transcription as reported (Fig. 6.1, D and G-I; Ref. 3).

We generated *Hoxa5*;*Hoxb5* compound mutant mice. At weaning age, only 2% of *Hoxa5*^{-/-} mice were recovered instead of the expected 6.25%, reflecting the high rate (66%) of lethality associated with the *Hoxa5* mutation (Table 6.1; Ref. 29). The survival of animals decreased with the number of *Hoxb5* mutant alleles: less than 10% of the

Fig. 6.1. Expression analyses of *Hox5* paralog genes in the developing lung.

In situ hybridization was performed on frontal sections of embryonic day (E)13.5 wild-type (WT) (A–C), *Hoxa5*^{-/-} (D, E), and *Hoxb5*^{-/-} (F, G) mouse embryos with *Hoxa5* (A, D, F), *Hoxb5* (B, E, G), and *Hoxc5* (C) specific riboprobes. Representative bright-field (*insets*) and dark-field images are shown. A strong *Hoxa5* signal was detected in the mesenchyme surrounding the entire respiratory tract from the trachea to the distal lung. *Hoxb5* transcripts were restricted to the distal lung mesenchyme, whereas no *Hoxc5* expression was observed in the lung. m, Mesenchyme; nt, neural tube; tr, trachea. Scale bar: 500 μ m. H and I: quantitative (q)RT-PCR analysis of *Hoxa5* (H) and *Hoxb5* (I) expression levels in the entire respiratory tract (from trachea to lungs) of E15.5 embryos. *Hoxa5* expression was upregulated in *Hoxb5*^{-/-} specimens suggesting *Hoxa5* compensatory effects. No change in *Hoxb5* expression was observed in *Hoxa5* mutants. The decreased *Hoxa5* expression in lungs from *Hoxa5*^{-/-} embryos and the reduced *Hoxb5* expression in *Hoxb5*^{-/-} specimens might reflect the interference of the *neo* selection cassette on *Hox* transcription. Values are expressed as means \pm SE. *** $P < 0.001$.

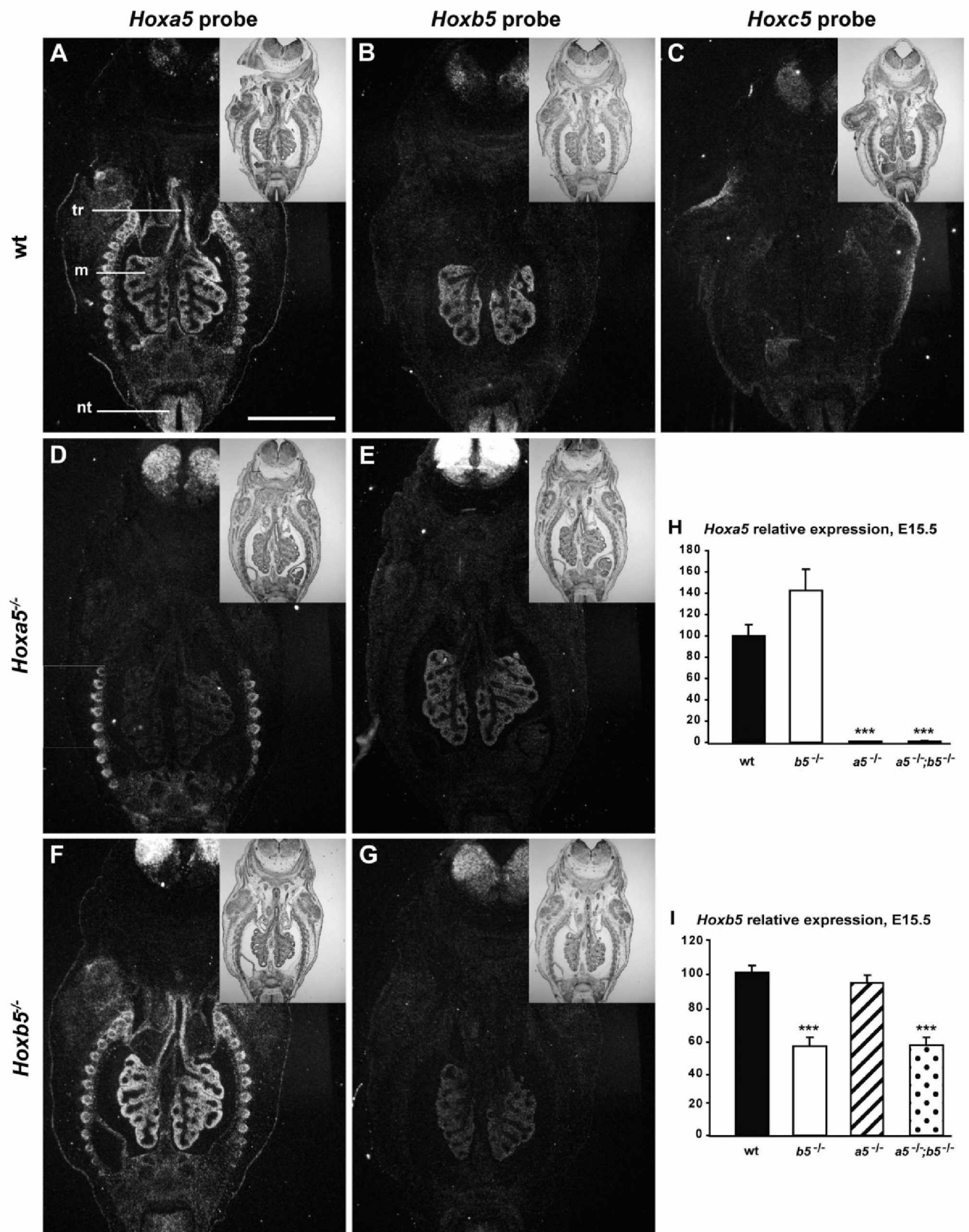


Table 6.1. Ratios of genotypes of litters from intercrosses of *Hoxa5*^{+/-}; *Hoxb5*^{+/-} mutant mice

Age	No of litters	No of animal	Genotypes*								
			<i>a5</i> ^{+/+} ; <i>b5</i> ^{+/+}	<i>a5</i> ^{+/-} ; <i>b5</i> ^{+/-}	<i>a5</i> ^{+/+} ; <i>b5</i> ^{-/-}	<i>a5</i> ^{+/-} ; <i>b5</i> ^{+/+}	<i>a5</i> ^{+/-} ; <i>b5</i> ^{+/-}	<i>a5</i> ^{+/-} ; <i>b5</i> ^{-/-}	<i>a5</i> ^{-/-} ; <i>b5</i> ^{+/+}	<i>a5</i> ^{-/-} ; <i>b5</i> ^{+/-}	<i>a5</i> ^{-/-} ; <i>b5</i> ^{-/-}
E10.5	24	138	7	15	11	19	30	27	8	12	9
% obtained			5%	11%	8%	14%	22%	20%	6%	9%	7%
E15.5	20	125	9	20	7	15	35	11	7	16	5
% obtained			7%	16%	6%	12%	28%	9%	6%	13%	4%
E18.5	63	325	18	36	33	45	68	52	16	37	20
% obtained			6%	11%	10%	14%	21%	16%	5%	11%	6%
D0	41	181 (62)	10	25	17	31 (2)	65 (4)	25	5 (7)	3 (21)	0 (28)
% obtained			4%	10%	7%	12.7% (0.8%)	26.7% (1.7%)	10%	2.1% (2.9%)	1.2% (8.6%)	0% (11.5%)
D21-D28	54	242	14	39	23	37	85	35	6	3	0
% obtained			6%	16%	10%	15%	35%	14%	2%	1%	0%
% Expected		12.5	6.25	12.5	6.25	12.5	25	12.5	6.25	12.5	6.25

*Numbers in parentheses correspond to pups found dead at birth. E, embryonic day; D, day.

expected number of *Hoxa5*^{-/-};*Hoxb5*^{+/-} mice and no double mutants *Hoxa5*^{-/-};*Hoxb5*^{-/-} were recovered at weaning. Timed matings showed that, during gestation, specimens of all nine possible genotypes were obtained at the expected Mendelian ratio. However, a close monitoring of the litters at birth revealed a high rate of mortality of the compound mutants deficient for both *Hoxa5* alleles and demonstrated that the loss of *Hoxb5* function worsened the neonatal lethality phenotype of the *Hoxa5* mutation.

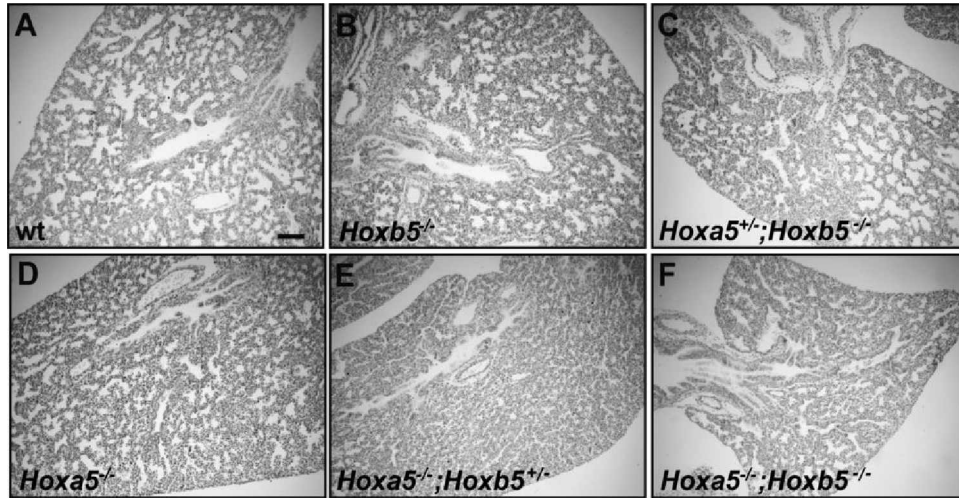
An exacerbated lung hypoplasia underlies the increased neonatal lethality of *Hoxa5*;*Hoxb5* compound mutants.

To identify the cause(s) of the higher rate of death of the *Hoxa5*;*Hoxb5* mutants, we verified whether the lung defects associated with the *Hoxa5* mutation were aggravated by the reduction of *Hoxb5* dosage. At E18.5, the number of lobes and their orientation were preserved in compound mutants indicating that the first bronchial ramifications occurred correctly. Comparative lung morphology of E18.5 embryos revealed that WT, *Hoxb5*^{-/-}, and *Hoxa5*^{+/-};*Hoxb5*^{-/-} specimens presented a normal structure with dilated peripheral lung saccules and thin mesenchyme (Fig. 6.2, A–C). However, lungs from *Hoxa5*^{-/-} embryos exhibited narrower air spaces and thicker mesenchyme as reported (4). The defects appeared more pronounced in *Hoxa5*^{-/-};*Hoxb5*^{+/-} and *Hoxa5*^{-/-};*Hoxb5*^{-/-} embryos (Fig. 6.2, D–F).

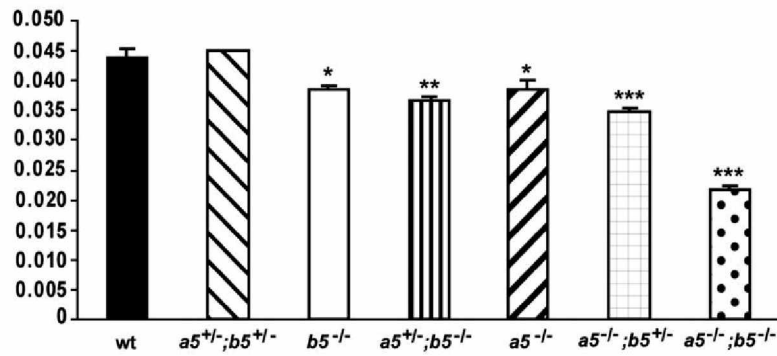
At E18.5, the lung weight-to-body weight ratio was significantly diminished in *Hoxa5*^{-/-}, *Hoxb5*^{-/-}, *Hoxa5*^{-/-};*Hoxb5*^{-/-}, *Hoxa5*^{+/-};*Hoxb5*^{-/-}, and *Hoxa5*^{-/-};*Hoxb5*^{-/-} embryos (Fig. 2G). Lungs from *Hoxa5*^{-/-};*Hoxb5*^{-/-} embryos were considerably smaller and the lung weight-to-body weight ratio was ~50% of that of WT embryos. To identify the factors contributing to pulmonary hypoplasia, we measured cell proliferation (Fig. 6.2H). At E18.5, a statistically reduced number of pHH3-stained lung cells was obtained for the *Hoxa5*^{-/-}, *Hoxa5*^{-/-};*Hoxb5*^{+/-}, and *Hoxa5*^{-/-};*Hoxb5*^{-/-} genotypes compared with controls, indicating that decreased lung proliferation was specifically associated with the *Hoxa5* mutation. Comparable results were obtained at E15.5 (not shown). Immunostaining for the activated form of caspase-3 did not show any significant variations in the extent of apoptosis in E18.5 embryonic lungs between all genotypes (not shown).

Fig. 6.2. Comparative analyses of lungs from *Hoxa5*;*Hoxb5* mutant embryos.

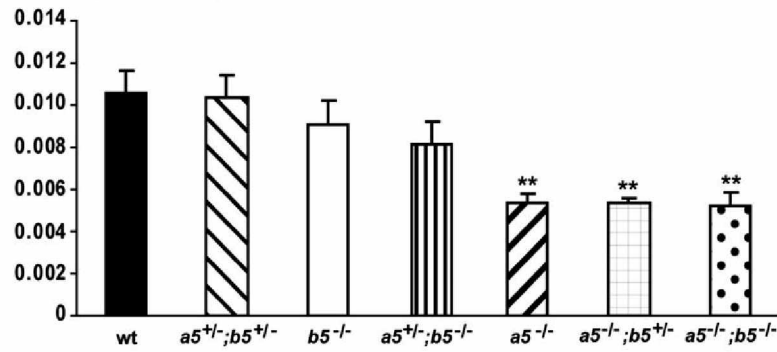
A–F: comparative lung histology of E18.5 *Hoxa5*; *Hoxb5* embryos. WT, *Hoxb5*^{-/-}, and *Hoxa5*^{+/-};*Hoxb5*^{-/-} specimens presented a normal lung structure with dilated peripheral saccules and thin mesenchyme, whereas lungs from *Hoxa5*^{-/-} embryos exhibited narrower air spaces and thicker mesenchyme. The phenotype worsened with the number of *Hoxb5* mutant alleles, and lungs from *Hoxa5*^{-/-};*Hoxb5*^{-/-} embryos were smallest. Scale bar: 100 μm. *G*: at E18.5, the lung weight-to-body weight ratio was statistically lower for *Hoxa5*^{-/-}, *Hoxb5*^{-/-}, *Hoxa5*^{+/-};*Hoxb5*^{-/-}, *Hoxa5*^{-/-};*Hoxb5*^{+/-}, and *Hoxa5*^{-/-};*Hoxb5*^{-/-} embryos compared with WT. *H*: lung hypoplasia in E18.5 *Hoxa5*^{-/-}, *Hoxa5*^{-/-};*Hoxb5*^{+/-} and *Hoxa5*^{-/-};*Hoxb5*^{-/-} embryos correlated with a statistically reduced proliferation in the lung. *I*: lung branching at E15.5, as estimated by the number of acini per field, was statistically less extensive in *Hoxa5*^{-/-}, *Hoxb5*^{-/-} and *Hoxa5*^{-/-};*Hoxb5*^{-/-} embryos compared with WT. Values are expressed as means ± SD. **P* < 0.05, ***P* < 0.01, ****P* < 0.001.



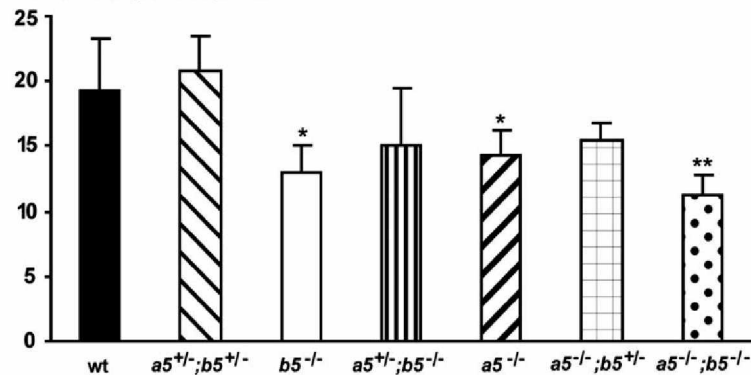
G Lung weight / body weight ratio, E18.5



H % of pHH3+ cells, E18.5



I # of acini per field, E15.5



As reduced lung branching may participate to lung hypoplasia, the number of acini was evaluated during the pseudoglandular stage at E15.5 (9). At this stage, the ramifying bronchial tree gives rise to tubular branches ending in terminal acini. As previously observed, branching was less extensive in *Hoxa5*^{-/-} embryos with a mean reduction of 27% (Fig. 6.2I; Ref. 4). *Hoxb5*^{-/-} mutants also presented a reduced number of acini indicating that both genes are involved in the branching process. In *Hoxa5*^{-/-}; *Hoxb5*^{-/-} specimens, the phenotype was amplified. Thus insufficient proliferation and decreased branching likely participate to the important lung hypoplasia that may explain the increased lethality of *Hoxa5*^{-/-}; *Hoxb5*^{-/-} pups at birth.

***Hoxa5* and *Hoxb5* genes are involved in lung epithelial cell differentiation..**

The role of *Hoxb5* in lung branching raised the possibility that *Hoxb5* can also be involved in other aspects of lung development. To assess the impact of the *Hoxa5*; *Hoxb5* compound mutations on the integrity of the airway epithelium, we examined the expression of specific cell markers in lungs from E18.5 embryos. Decreased expression of the club cell (Clara cell) marker CC10 (encoded by the *Scgb1a1* gene) was observed in lung airways of embryos carrying two *Hoxa5* mutant alleles, as shown by IF staining and CC10 intensity measurement by using ImageJ as well as by qRT-PCR (Fig. 6.3, A–H). The decreased CC10 immunostaining in *Hoxa5*^{-/-}; *Hoxb5*^{-/-} specimens did not correlate with the *Scgb1a1* high expression obtained by qRT-PCR. One likely explanation could be that the important hypoplasia of *Hoxa5*^{-/-}; *Hoxb5*^{-/-} lung specimens causes an increase in the proportion of airway epithelial cells vs. the total lung tissue obtained for RNA extraction.

We looked at the expression of podoplanin (T1 α), a marker of type I pneumocytes, which are flat epithelial cells lining the air spaces and closely interacting with the underlying vascular endothelium for gas exchanges (42). T1 α expression was comparable between WT, *Hoxb5*^{-/-}, and *Hoxa5*^{+/-}; *Hoxb5*^{-/-} embryos (Fig. 6.3, I–K). In contrast, it was dramatically decreased in lungs from *Hoxa5*^{-/-} embryos, a phenotype accentuated in *Hoxa5*^{-/-}; *Hoxb5*^{-/-} mutants as confirmed by qRT-PCR analysis (Fig. 6.3, L–O). Thus *Hoxa5* plays a predominant role in the specification of type I pneumocytes, and *Hoxb5* also seems to collaborate.

To investigate whether lung vascular development was also perturbed, we examined endothelial cells using the platelet endothelial cell adhesion molecule 1 (PECAM-1) marker. A growing microvasculature network was observed in the well-developed saccules of WT, *Hoxb5*^{-/-} and *Hoxa5*^{+/-}; *Hoxb5*^{-/-} specimens with staining in close proximity to the lumen (Fig. 6.3, P–R). In contrast, a faint PECAM-1 staining embedded in a thick mesenchyme was detected in *Hoxa5*; *Hoxb5* compound mutants carrying two *Hoxa5* mutant alleles, indicating a relatively undeveloped pulmonary capillary bed (Fig. 6.3, S–U). The VEGF pathway is involved in driving the epithelial to endothelial cross talk during lung morphogenesis (14). qRT-PCR expression analysis revealed significantly reduced *Vegfa* expression levels in *Hoxa5*^{-/-} and *Hoxa5*^{+/-}; *Hoxb5*^{-/-} specimens (Fig. 6.3V). This suggests that *Vegfa* downregulation might contribute to the abnormal lung microvasculature and supports the notion that this phenotype was mainly associated with the *Hoxa5* mutation.

Tracheal malformations are specific to *Hoxa5* mutant embryos.

Hoxa5^{-/-} mice exhibit tracheal malformations characterized by disorganized banding pattern and narrowing of the trachea (4). Even though *Hoxa5* is the only *Hox5* member to be expressed in the upper respiratory airways, a possible worsening of the phenotype was investigated in E18.5 *Hoxa5*; *Hoxb5* embryos. Macroscopic examination of whole trachea and analysis of transverse sections showed a reduced number of cartilaginous rings and a diminished luminal surface in *Hoxa5*; *Hoxb5* mutants carrying two *Hoxa5* mutant alleles, independently of the *Hoxb5* mutation (Fig. 6.4, A–H). Thus correct tracheal formation is specifically under the control of the *Hoxa5* gene.

***Hoxa5* controls *Fstll* expression during trachea formation.**

Mutation in the follistatin-like 1 (*Fstll*) gene, which encodes a bone morphogenetic protein (BMP) antagonist, causes a malformed trachea with a reduced number of rings and a disorganized epithelial layer reminiscent of the *Hoxa5* mutant phenotype (20). We monitored *Fstll* expression in the upper airways (trachea and primary bronchi) of E18.5 *Hoxa5*; *Hoxb5* embryos by qRT-PCR. Consistent with our morphological observations,

Fig. 6.3. Characterization of lung epithelium and microvasculature of E18.5 *Hoxa5*;*Hoxb5* mutant embryos.

A–F: as assessed by immunofluorescence (IF), CC10 expression was decreased in lung airways of embryos carrying 2 *Hoxa5* mutant alleles indicating the predominant role of *Hoxa5* in club cell specification. G: mean CC10 intensity within the airway epithelium was measured by use of ImageJ. Values represent the average fluorescence intensity of the bronchial epithelial areas (surrounded by a dotted line). At least 4 pictures per specimen from 3–5 specimens were measured for each genotype. Values are expressed as means \pm SE. I–N: immunostaining with podoplanin ($T1\alpha$), a marker of type I pneumocytes, was considerably decreased in lungs from *Hoxa5*^{-/-}, *Hoxa5*^{-/-};*Hoxb5*^{+/-} and *Hoxa5*^{-/-};*Hoxb5*^{-/-} embryos. P–U: PECAM-1 immunostaining revealed an undeveloped vascular network embedded in the thick lung mesenchyme in *Hoxa5*^{-/-}, *Hoxa5*^{-/-};*Hoxb5*^{+/-}, and *Hoxa5*^{-/-};*Hoxb5*^{-/-} embryos. Scale bars: 100 μ m (I–N), 50 μ m (A–F, P–U). H, O, V: qRT-PCR analysis for *Scgb1a1*, $T1\alpha$, and *Vegfa* expression in lungs from WT, *Hoxa5*^{-/-}, *Hoxb5*^{-/-}, and *Hoxa5*^{-/-};*Hoxb5*^{-/-} E18.5 embryos. Comparison was made against WT. O: $T1\alpha$ expression was also significantly reduced in *Hoxa5*^{-/-};*Hoxb5*^{-/-} specimens compared with *Hoxa5*^{-/-} samples. Values are expressed as means \pm SE. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

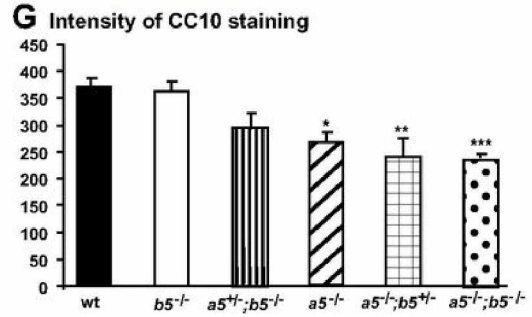
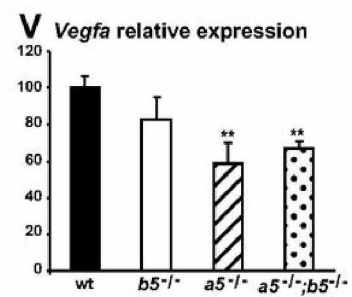
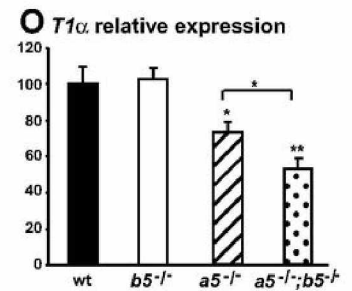
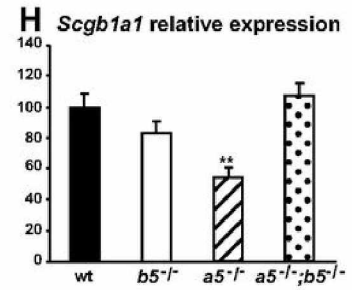
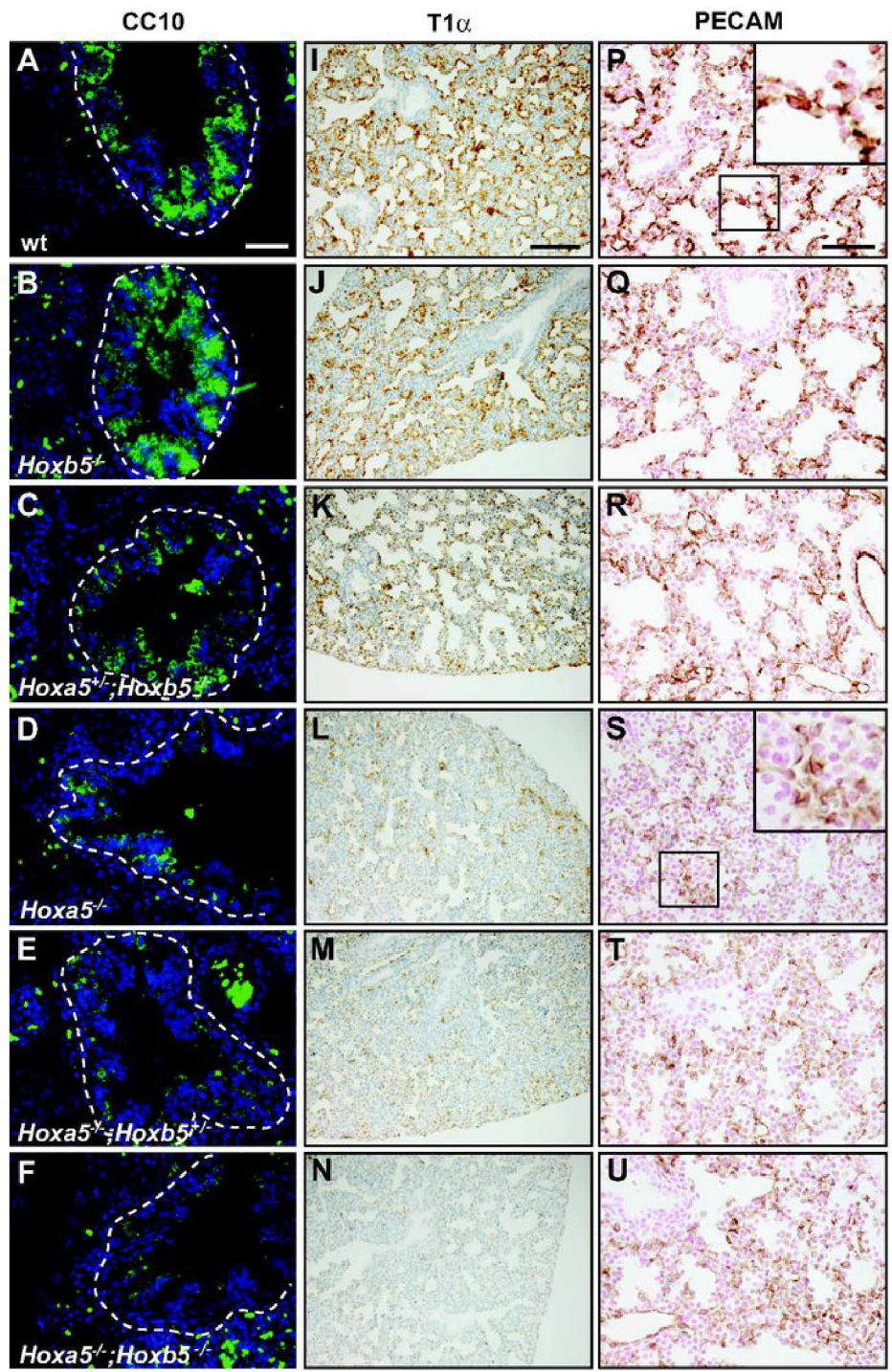
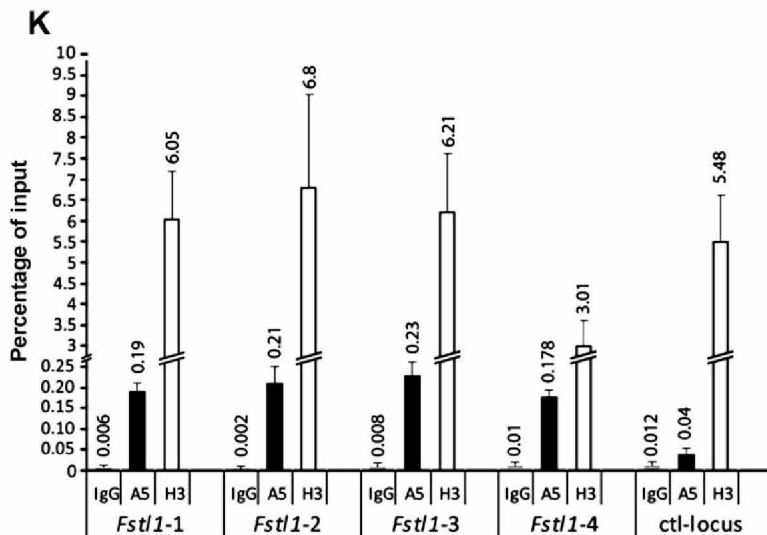
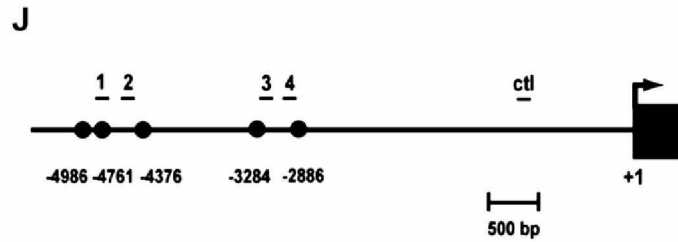
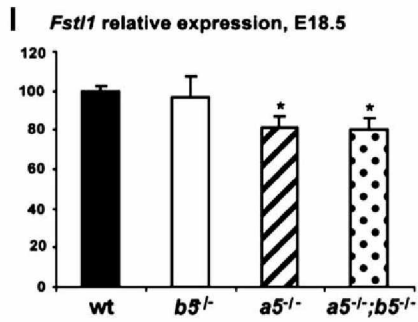
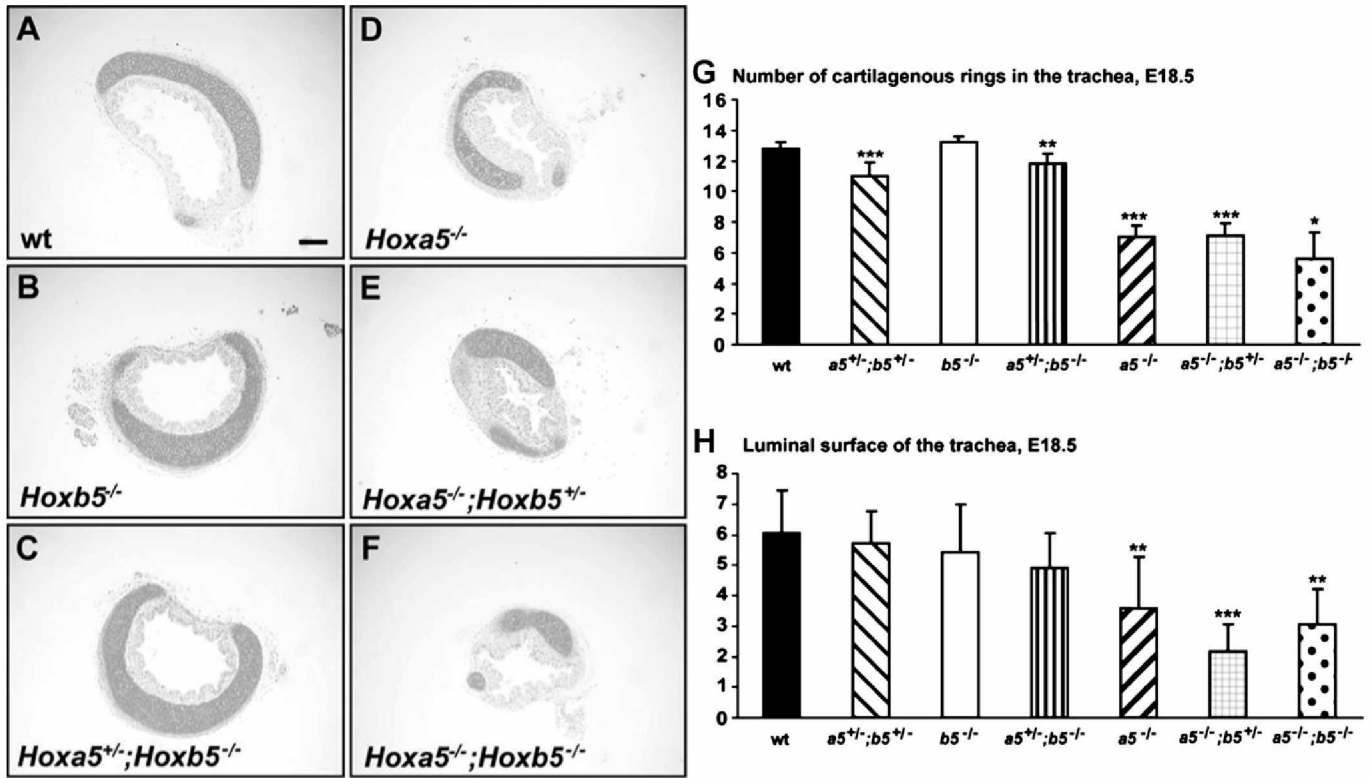


Fig. 6.4. *Hoxa5* plays a predominant role in tracheal formation and it regulates *Fstll* expression in the developing trachea.

A–F: transverse sections of trachea from E18.5 *Hoxa5*;*Hoxb5* embryos were Alcian blue stained for morphology analysis. The trachea from *Hoxa5*^{-/-}, *Hoxa5*^{-/-};*Hoxb5*^{+/-}, and *Hoxa5*^{-/-};*Hoxb5*^{-/-} embryos was narrower and had a diminished luminal surface. Scale bar: 100 μm. G: reduced number of cartilage rings was observed in *Hoxa5*^{-/-} specimens independently of the *Hoxb5* mutation. H: the surface of the lumen was quantified and the numbers correspond to arbitrary units. Values are expressed as means ± SD. **P* < 0.05, ***P* < 0.01, ****P* < 0.001. I: qRT-PCR analysis of *Fstll* expression revealed diminished levels in trachea/primary bronchi from *Hoxa5*^{-/-} and *Hoxa5*^{-/-};*Hoxb5*^{-/-} embryos compared with WT. Values are expressed as means ± SE. **P* < 0.05. J: schematic representation and position relative to the transcription start site of the 5 putative HOX binding sites in the *Fstll* promoter. The position of the qPCR fragments is indicated. K: ChIP analysis of endogenous *Fstll* regulatory sequences in E18.5 trachea. Chromatin was immunoprecipitated with rabbit IgG, anti-HOXA5 and anti-histone H3 antibodies. Recruitment of HOXA5 and H3 on regulatory sequences of the *Fstll* locus was evaluated by qPCR and is indicated as the percentage of input. HOXA5 binds to the *Fstll* genomic region spanning the HOX binding sites whereas no binding was observed with the *Fstll* DNA negative control. The data are means ± SD of 3 independent experiments.



decreased *Fstll* expression was observed in *Hoxa5*^{-/-} and *Hoxa5*^{-/-};*Hoxb5*^{-/-} specimens but not in *Hoxb5*^{-/-} ones (Fig. 6.4I).

Fstll is expressed in the mesenchyme surrounding the airways like *Hoxa5* (49). Sequence analysis of the mouse *Fstll* gene with the Transcription Element Search System (TESS) software revealed five HOX binding sites into a 2.1-kb region located between positions -5.0 and -2.9 kb upstream the *Fstll* transcription start site (Fig. 6.4J). To determine whether *Fstll* could be a potential HOXA5 transcriptional target, we performed ChIP assays on cross-linked chromatin isolated from trachea of E18.5 WT mouse embryos. DNA from the immunoprecipitate was subjected to qPCR analyses with primers covering the 2.1-kb region. An *Fstll* DNA genomic fragment containing no HOX binding site was used as a negative control for the locus. As shown in Fig. 6.4K, HOXA5 was recruited to the *Fstll* genomic region spanning the HOX binding sites whereas no binding was observed with the *Fstll* DNA negative control. Experiments done with an HOXB5 antibody did not reveal any association with the *Fstll* sequences, consistent with the lack of *Hoxb5* expression in the trachea (not shown). Altogether, these data suggest that HOXA5 may act on trachea formation via in part the regulation of the transcriptional target *Fstll*.

Aberrant diaphragm innervation in *Hoxa5*^{-/-} embryos.

Insufficiency or lack of fetal breathing movements can participate to lung hypoplasia (25, 27, 32). *Hoxa5* is expressed in motor neurons from the phrenic motor column, the latter controlling the contraction of the diaphragm (40). Motor axons from the phrenic nerve exit the cervical spinal cord to contact the diaphragmatic primordium in E12 mouse embryos. Nerve and muscle then develop together with branching and projection of the motor axons in the dorsal and ventral portions of the diaphragm (21). To assess whether the *Hoxa5* and *Hoxb5* mutations interfere with normal diaphragm innervation and neuromuscular synapse formation, we examined the motor axons in the diaphragm of E18.5 embryos by whole-mount IF staining with antibodies against neurofilament, synaptophysin (for presynaptic nerve terminals), and with α -bungarotoxin (for postsynaptic acetylcholine receptor). In WT, *Hoxb5*^{-/-}, *Hoxa5*^{+/-};*Hoxb5*^{+/-} and

Hoxa5^{+/-};*Hoxb5*^{-/-} specimens, phrenic nerves reached the diaphragm, split and projected correctly. Synapses also formed between the phrenic nerves and the diaphragm as shown by the costaining of synaptophysin and α -bungarotoxin (Fig. 6.5, A–D). The neuromuscular junctions were confined to the endplate band, a central region of the muscle fiber in which intramuscular nerves are located. No obvious phenotype was associated with the loss of *Hoxb5* function. In contrast, *Hoxa5*;*Hoxb5* compound mutants carrying two *Hoxa5* mutant alleles presented fewer intramuscular branches and synaptic contacts (Fig. 6.5, E–G). There was no clear evidence that the phenotype worsened with the presence of *Hoxb5* mutant alleles, suggesting that the loss of *Hoxa5* function results in improper innervation of the diaphragm that may impact negatively on fetal breathing movements.

Functional redundancy between *Hoxa5* and *Hoxb5* genes in postnatal lung development.

Hoxa5^{-/-} surviving mice display lung air space enlargement and goblet cell metaplasia (33). Lung morphology of *Hoxa5*;*Hoxb5* mutants was first investigated using a quantitative morphometric method based on radial alveolar counts. We observed decreased septa number, indicative of disruption of the alveolar growth, in *Hoxa5*^{-/-} and *Hoxb5*^{-/-} mutants (Fig. 6.6A). The phenotype worsened with the number of *Hox5* mutant alleles, but again *Hoxa5* appeared to play a main role in pulmonary alveogenesis.

In *Hoxa5*^{-/-} mice, goblet cell metaplasia originates from club to goblet cell transdifferentiation and is associated with decreased CC10 expression (7). We assessed whether the phenotype was exacerbated in *Hoxa5*;*Hoxb5* compound mutants by evaluating club cell integrity and goblet cell metaplasia on sections of upper airways from adult mice stained, respectively, with CC10 antibody and Alcian blue for mucus detection. In WT specimens, as expected, an intense CC10 staining was detected along the bronchial airway epithelium while Alcian blue-positive cells were rarely seen (Fig. 6.6, B and G). In *Hoxb5*^{-/-} mice, CC10 staining was strong but few goblet cell metaplasia areas were detected in upper airways (Fig. 6.6, C–H). A weaker CC10 labeling was observed in *Hoxa5*^{+/-};*Hoxb5*^{-/-}, *Hoxa5*^{-/-} and *Hoxa5*^{-/-};*Hoxb5*^{+/-} specimens concomitant with the presence of Alcian blue positive cells (Fig. 6, D–F, I–K). The phenotype worsened with the number of *Hox5*

Fig. 6.5. Abnormal diaphragm innervation pattern in *Hoxa5;Hoxb5* embryos carrying *Hoxa5* mutant alleles.

A–G: analysis of diaphragm innervation patterns at E18.5 by whole-mount IF staining for neurofilament (NF), synaptophysin (SYN), and α -bungarotoxin (AChR). The dorsal-ventral (D-V) and left-right (L-R) axes are indicated on the view of the entire diaphragm (A). A–D: diaphragms from WT, *Hoxb5*^{-/-}, *Hoxa5*^{+/-};*Hoxb5*^{+/-}, and *Hoxa5*^{+/-};*Hoxb5*^{-/-} embryos appeared normal with the phrenic nerves reaching the diaphragm correctly and synapses forming between the phrenic nerves and the diaphragm. E–G: diaphragms from *Hoxa5;Hoxb5* compound mutants carrying 2 *Hoxa5* mutant alleles presented fewer intramuscular branches and synaptic contacts. Scale bars: 1 mm.

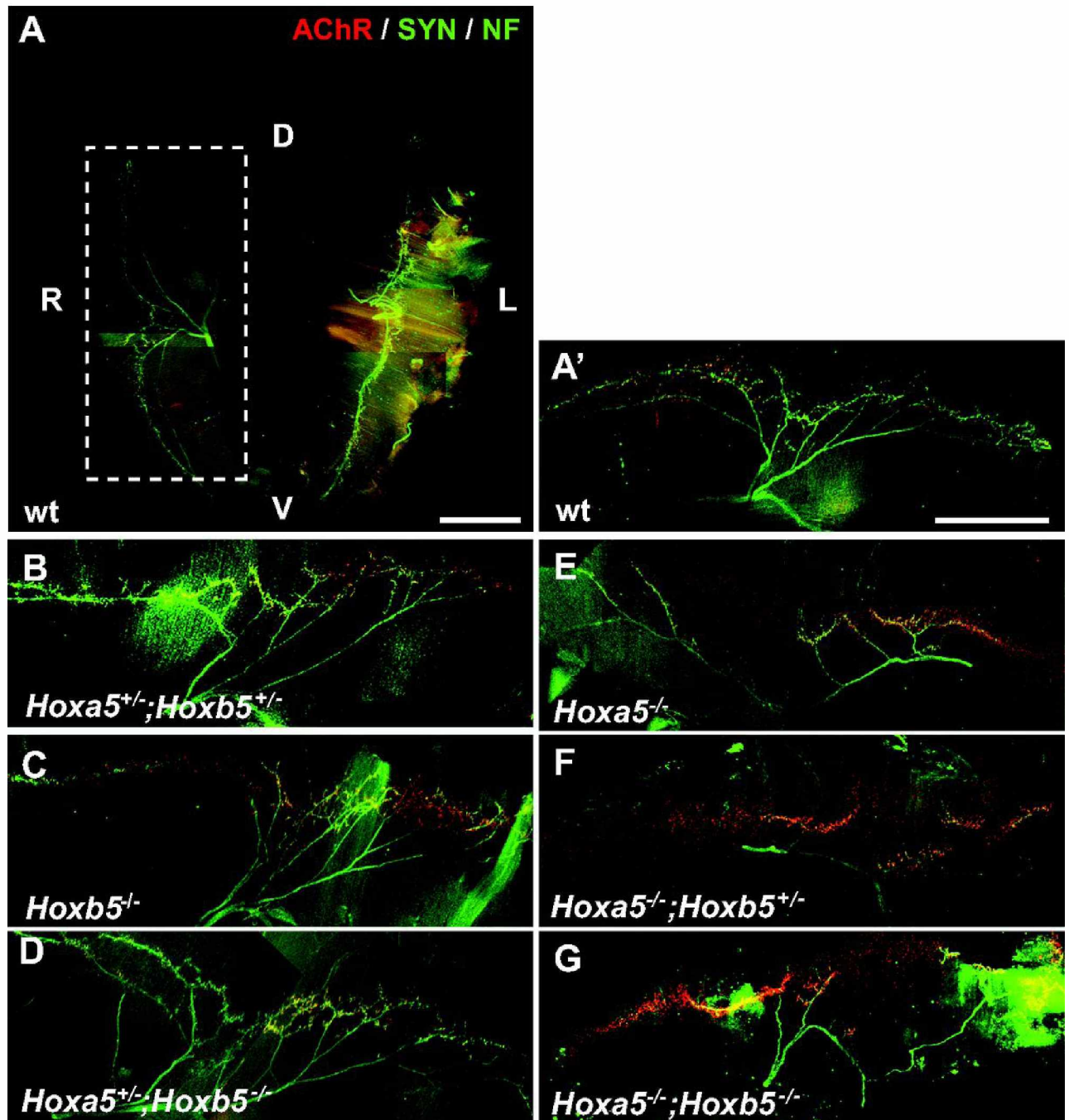
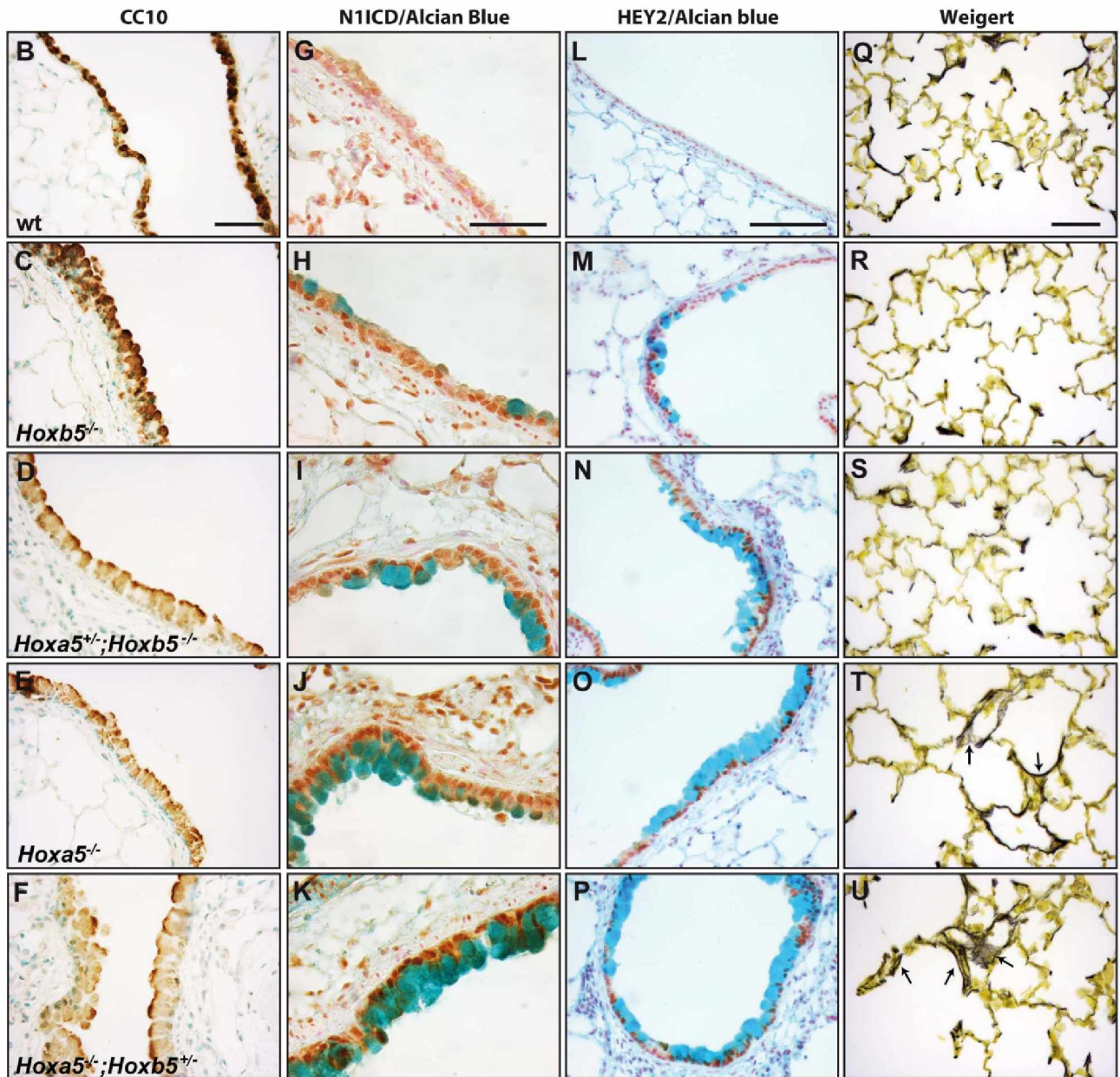
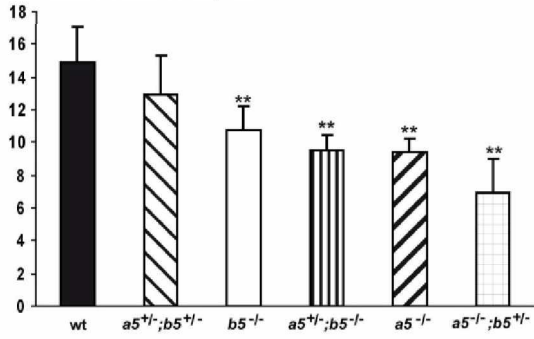


Fig. 6.6. Comparative analyses of lungs from *Hoxa5;Hoxb5* adult mutant mice.

A: comparison of the radial alveolar counts revealed a decreased septa number in the lungs from *Hoxb5*^{-/-}, *Hoxa5*^{+/-};*Hoxb5*^{-/-}, *Hoxa5*^{-/-} and *Hoxa5*^{-/-};*Hoxb5*^{+/-} mutants. Values are expressed as means ±SD. ****P < 0.01.** B–F: detection of club cells by CC10-immunostaining in lungs from *Hoxa5;Hoxb5* adult mice. Decreased CC10-staining was seen in *Hoxb5*^{-/-}, *Hoxa5*^{+/-};*Hoxb5*^{-/-}, *Hoxa5*^{-/-}, and *Hoxa5*^{-/-};*Hoxb5*^{+/-} specimens. G–P: increased Notch signaling in goblet cell metaplasia areas was revealed by immunostaining for N1ICD (G–K) and HEY2 (L–P) in *Hoxb5*^{-/-}, *Hoxa5*^{+/-};*Hoxb5*^{-/-}, *Hoxa5*^{-/-}, and *Hoxa5*^{-/-};*Hoxb5*^{+/-} specimens. Q–U: Weigert staining was used for visualization of elastic fibers. In WT, *Hoxb5*^{-/-} and *Hoxa5*^{+/-};*Hoxb5*^{-/-} specimens, elastic fibers appeared well-distributed along the alveoli and at the tips of septa. In contrast, lungs from *Hoxa5*^{-/-} and *Hoxa5*^{-/-};*Hoxb5*^{+/-} mutant mice displayed disorganized and fragmented elastin aggregates within the saccular walls (arrows). Scale bars: 50 μm.

A. Radial alveolar count, D65



mutant alleles and *Hoxa5* appeared to play a key function in the specification of airway secretory cell lineages.

The molecular mechanisms underlying goblet cell metaplasia in *Hoxa5*^{-/-} mice are independent of FOXA2, a transcription factor known to govern lung epithelial cell differentiation (7). Similarly, goblet cell metaplasia in lungs from *Hoxb5*^{-/-} mice and *Hoxa5*;*Hoxb5* compound mutants did not impact on FOXA2 expression when tested by immunostaining (not shown).

Notch signaling plays a key role in controlling the delicate balance between airway epithelial cell fates (23, 37, 50). As well, specific activation of the Notch pathway in airway epithelium of *Hoxa5*^{-/-} mice was observed with HEY2 as a potential Notch transcriptional effector of goblet cell metaplasia (7). We evaluated the functional status of the Notch signaling pathway in airways from *Hoxa5*;*Hoxb5* mutants by performing IHC experiments for the active intracellular domain of NOTCH1 receptor (N1ICD) and HEY2. Expression of both proteins was markedly increased at sites of goblet cell metaplasia in *Hoxb5*^{-/-} and compound mutants (Fig. 6.6, G–P). Thus the loss of *Hoxa5* function and to a lower extent the *Hoxb5* mutation induce club to goblet cell transdifferentiation, a FOXA2-independent process accompanied by an increased activity of Notch signaling.

Correct formation and deposition of elastic fibers are essential for lung alveogenesis (57). Using Weigert staining to visualize the spatial distribution of elastic fibers, we observed that WT, *Hoxb5*^{-/-}, and *Hoxa5*^{+/-};*Hoxb5*^{-/-} mutant mice displayed focal elastic fiber deposits localized primarily at the tips of the septa, indicating no major impact of the *Hoxb5* mutation on elastic fibers production (Fig. 6.6, Q–S). In contrast, in compound mutants carrying two *Hoxa5* mutant alleles, aggregates of elastic fibers occurred within the saccular walls rather than being focused at the tips of secondary septa. They also appeared disorganized and fragmented (Fig. 6.6, T and U). Altogether, the results suggest that both *Hoxa5* and *Hoxb5* are implicated in correct lung morphology and goblet cell specification with *Hoxa5* having a predominant role.

DISCUSSION

***Hoxb5* is involved in lung morphogenesis.**

Despite the clear presence of *Hox* gene expression in the developing lung, *Hox* function during lung morphogenesis remains largely unknown since only *Hoxa5*^{-/-} mice exhibit primary lung defects leading to respiratory distress and to a high rate of lethality at birth (4, 26). In the *Hox5* paralog group, *Hoxa5* and *Hoxb5*, but not *Hoxc5*, are expressed along the developing respiratory tract, *Hoxa5* having the broader expression domain. However, we cannot rule out the possibility of a later onset for *Hoxc5* expression and some impact on lung development. During the characterization of the *Hoxa5*;*Hoxb5* compound mutant mice, we have uncovered the role of *Hoxb5* during lung development, unveiled new roles for *Hoxa5* in this process and demonstrated the partial functional overlap between *Hoxa5* and *Hoxb5* genes.

Like *Hoxa5*, *Hoxb5* is involved in branching morphogenesis as revealed by the decreased number of lung epithelial tubules in *Hoxb5*^{-/-} embryos. This finding agrees with those of Volpe et al. (54), who reported that *Hoxb5* siRNA-treated whole lung cultures showed decreased branching through the downregulation of tenascin-C, an extracellular matrix component required for branching morphogenesis. Decreased expression of tenascin-C following *Hoxb5* inhibition was also shown in cultured mouse lung fibroblasts (46).

As reported in *Hoxa5*^{-/-} mice, an increased number of goblet cells and air space enlargement were observed in *Hoxb5*^{-/-} adult mice. Similar molecular mechanisms appeared to drive goblet cell metaplasia in *Hoxb5*^{-/-} mice with no change in FOXA2 expression, and an increased Notch signaling activity in areas of goblet cell metaplasia, suggesting that both proteins act via identical means during the process of secretory epithelial cell fate in upper airways (7).

Thus the loss of *Hoxb5* function impacts on lung development. The more *Hoxb5* alleles are knocked out in the *Hoxa5* mutant environment, the worse the lung phenotype

is. *Hoxb5*^{-/-} lung phenotypes are shared by *Hoxa5*^{-/-} mice, indicating some functional redundancy. However, the *Hoxb5* lung phenotypic traits are less severe than those encountered in *Hoxa5*^{-/-} mice, suggesting that *Hoxb5* is less prevalent in lung morphogenesis. This may be explained by *Hoxa5* compensation in *Hoxb5* mutants, as pointed out by the increased *Hoxa5* expression in *Hoxb5*^{-/-} lungs.

Extensive contribution of *Hoxa5* to the developing respiratory system.

Pulmonary hypoplasia was observed in both *Hoxa5* and *Hoxb5* mutant mice and the phenotype was dramatically exacerbated in *Hoxa5*^{-/-};*Hoxb5*^{-/-} embryos. Pulmonary hypoplasia is described as wet lung weight to body weight ratio inferior to 67% of normal at the time of birth (47). In E18.5 *Hoxa5*^{-/-};*Hoxb5*^{-/-} embryos, the ratio was ~50% of that of controls. Thus pulmonary hypoplasia likely contributes to *Hoxa5*^{-/-};*Hoxb5*^{-/-} newborn death. In *Hoxa5*; *Hoxb5* mutants, pulmonary hypoplasia can be explained by reduced branching due to the *Hoxa5* and *Hoxb5* mutations and decreased proliferation in specimens carrying two *Hoxa5* mutant alleles, demonstrating the specific role of *Hoxa5* in governing lung growth.

Mechanical forces can influence fetal lung development through pulmonary distension (30). Deficient fetal breathing movements produced by rhythmic contraction of the respiratory muscles is a major cause of pulmonary hypoplasia and lung immaturity (25, 32). Since *Hox* genes are involved in motor neuron identity and connectivity, requirement for *Hoxa5* and *Hoxb5* in diaphragm innervation was investigated (12, 13). We observed that branching and projection of the phrenic nerve were altered and that diaphragm innervation was abnormal in specimens carrying two *Hoxa5* mutant alleles, suggesting a specific role for *Hoxa5* in the control of axon guidance to diaphragm. The lack of such phenotype in *Hoxb5* mutants correlates with the lack of *Hoxb5* expression in the phrenic motor column (40). Moreover, even though the innervation of the diaphragm was altered, the diaphragm muscle appeared intact across the abdominal wall and no visceral herniation was detected. Since spinal cord or phrenic nerve transection abolishes fetal breathing resulting in secondary lung hypoplasia, the abnormal innervation of the diaphragm specifically detected in *Hoxa5*^{-/-} can contribute to the pulmonary phenotype (16, 58).

The transition from a fluid-filled lung to a functional gas-exchanging organ at birth requires the expansion of the alveolar space and a thinning of the blood-gas barrier. The distal epithelium is comprised of two specialized cell populations: type II pneumocytes, which produce surfactant for the maintenance of alveolar surface tension, and type I pneumocytes, closely associated with vascular endothelial cells for optimal gas exchanges. Interestingly, staining for T1 α and PECAM was markedly reduced in *Hoxa5*^{-/-} mice but no difference was detected in *Hoxb5*^{-/-} mice, suggesting that *Hoxb5* is not likely to have a significant role in alveolar epithelial cell and endothelial cell differentiation. In support of this, the endothelial differentiation defect was not exacerbated in *Hoxa5/Hoxb5* double mutants. Thus the lack of functional type I cells in *Hoxa5*^{-/-} mutants combined with the defective capillary growth might contribute to the reduction of the gas exchange surface area to a level close to that required for survival of newborns.

In situ hybridization experiments clearly revealed that among the *Hox5* paralogous genes, only *Hoxa5* is expressed in the mesenchyme surrounding the developing trachea. Consistent with this, the tracheal defects seen in *Hoxa5* mutant mice were not observed in *Hoxb5*^{-/-} mice and not amplified in *Hoxa5;Hoxb5* double mutants. Similarities in tracheal cartilage defects between *Hoxa5*^{-/-} and *Fstll*^{-/-} mice adding to the fact that both are confined to the mesenchymal component of the developing lung suggested that *Fstll* might be a transcriptional target of HOXA5 (20). Expression of the BMP antagonist *Fstll* was downregulated in *Hoxa5*^{-/-} trachea, and ChIP experiments support the hypothesis that HOXA5 may control trachea formation through regulation of the BMP signaling pathway.

Functional predominance of *Hoxa5* in lung development.

Hoxb5^{-/-} mice survived while *Hoxa5*^{-/-} mice displayed reduced viability at birth. However, the loss of both *Hoxa5* and *Hoxb5* gene functions caused the death at birth of all *Hoxa5*^{-/-};*Hoxb5*^{-/-} pups underlying an aggravated pulmonary phenotype compared with that previously described in *Hoxa5*^{-/-} mice. Characterization of compound mutants from *Hox* paralog groups has led to the discovery of new functions for *Hox* genes. For instance, *Hox11* single mutants do not have kidney abnormalities, whereas *Hoxa11;Hoxd11* double or *Hoxa11;Hoxc11;Hoxd11* triple mutants present kidney hypoplasia and kidney agenesis,

respectively, revealing the convergent role of all *Hox11* paralog genes in kidney organogenesis (39, 56). Our findings on *Hox5* gene action in lung development somewhat differ from those reported for the role of *Hox11* genes in kidney formation since both *Hoxa5* and *Hoxb5* single mutants present phenotypes. However, the severity of the lung phenotype increases with the number of mutant alleles, the deletion of the *Hoxa5* allele being more deleterious.

Hoxa5 can compensate for the *Hoxb5* lung phenotype but the opposite is not true. This nonreciprocal compensation and the predominant *Hoxa5* role may have several explanations not mutually exclusive. The mouse HOXA5 protein is 60% identical and 70% similar to HOXB5, with the homeobox and the hexapeptide sequences being 100 and 86% identical, respectively. Possibilities for the distinctive roles of *Hoxa5* and *Hoxb5* genes include qualitative and quantitative differences in expression as well as variations in the regulation of downstream target genes due to variable affinities in DNA binding specificity or to difference in interactions with specific cofactors.

Taken together, our results indicate that *Hoxa5*; *Hoxb5* compound mutants exhibit synergistic lung abnormalities underscoring that both genes coordinate lung development and size. By examining the lung phenotype of *Hoxa5*; *Hoxb5* specimens, we have unveiled *Hoxb5* functions but most importantly confirmed the hierarchical importance of *Hoxa5*. In addition to shared functions with *Hoxb5*, *Hoxa5* plays specific roles in lung proliferation, specification of type I pneumocytes, lung microvascular development, elastic fiber deposition, trachea formation, and diaphragm innervation. Although the specific expression of *Hoxa5* in the trachea and the phrenic motor column likely underlies the trachea and the diaphragm phenotypes, an explanation for the predominant role of *Hoxa5* in the lung remains to be found.

ACKNOWLEDGMENTS

We thank Dr. J. Charron for helpful comments on the manuscript, Drs. A. Boulet and M. Capecchi for collaboration in obtaining the *Hoxb5* mutant mouse line, Drs. B. Hogan and

P. Sharpe for in situ probes, Dr. G. Singh for antibody, Dr. J. Côté for advices for ChIP experiments, and M. Lemieux for technical assistance.

GRANTS

This work was supported by a grant from the Canadian Institutes of Health Research (MOP-15139, to L. Jeannotte), by a National Institute of Neurological Disorders and Stroke grant (R01-NS062822; to J. S. Dasen), and by a postdoctoral fellowship from the Fonds de la Recherche en Santé du Québec-Institut National de la Santé et de la Recherche Médicale (to O. Boucherat).

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS

O.B. and L.J. conception and design of research; O.B., S.M., F.-A.B.-S., J.A., P.P., and D.M.W. performed experiments; O.B., S.M., F.-A.B.-S., J.A., P.P., and L.J. analyzed data; O.B. and L.J. interpreted results of experiments; O.B. prepared figures; O.B. drafted manuscript; J.S.D. and L.J. edited and revised manuscript; L.J. approved final version of manuscript.

REFERENCES

1. **Aubin J, Chailer P, Ménard D, Jeannotte L.** Loss of Hoxa5 gene function in mice perturbs intestinal maturation. *Am J Physiol Cell Physiol* 277: C965–C973, 1999.
2. **Aubin J, Déry U, Lemieux M, Chailer P, Jeannotte L.** Stomach regional specification requires Hoxa5-driven mesenchymal-epithelial signaling. *Development* 129: 4075–4087, 2002.
3. **Aubin J, Lemieux M, Tremblay M, Behringer RR, Jeannotte L.** Transcriptional interferences at the Hoxa4/Hoxa5 locus: importance of correct Hoxa5 expression for the proper specification of the axial skeleton. *Dev Dyn* 212: 141–156, 1998.

4. **Aubin J, Lemieux M, Tremblay M, Bérard J, Jeannotte L.** Early postnatal lethality in *Hoxa-5* mutant mice is attributable to respiratory tract defects. *Dev Biol* 192: 432–445, 1997.
5. **Bogue CW, Gross I, Vasavada H, Dynia DW, Wilson CM, Jacobs HC.** Identification of *Hox* genes in newborn lung and effects of gestational age and retinoic acid on their expression. *Am J Physiol Lung Cell Mol Physiol* 266: L448–L454, 1994.
6. **Bogue CW, Lou LJ, Vasavada H, Wilson CM, Jacobs HC.** Expression of *Hoxb* genes in the developing mouse foregut and lung. *Am J Respir Cell Mol Biol* 15: 163–171, 1996.
7. **Boucherat O, Chakir J, Jeannotte L.** The loss of *Hoxa5* function in mice promotes NOTCH-dependent goblet cell metaplasia in lung airways. *Biol Open* 1: 677–691, 2012.
8. **Boulet AM, Capecchi MR.** Targeted disruption of *hoxc-4* causes esophageal defects and vertebral transformations. *Dev Biol* 177: 232–249, 1996.
9. **Colvin JS, White AC, Pratt SJ, Ornitz DM.** Lung hypoplasia and neonatal death in *Fgf9*-null mice identify this gene as an essential regulator of lung mesenchyme. *Development* 128: 2095–2106, 2001.
10. **Condie BG, Capecchi MR.** Mice with targeted disruptions in the paralogous genes *hoxa-3* and *hoxd-3* reveal synergistic interactions. *Nature* 370: 304–307, 1994.
11. **Coulombe Y, Lemieux M, Moreau J, Aubin J, Joksimovic M, Bérubé-Simard FA, Tabariès S, Boucherat O, Guillou F, Larochelle C, Tuggle CK, Jeannotte L.** Multiple promoters and alternative splicing: *Hoxa5* transcriptional complexity in the mouse embryo. *PLoS One* 5: e10600, 2010.
12. **Dasen JS, Jessell TM.** *Hox* networks and the origins of motor neuron diversity. *Curr Top Dev Biol* 88: 169–200, 2009.
13. **Dasen JS, Tice BC, Brenner-Morton S, Jessell TM.** A *Hox* regulatory network establishes motor neuron pool identity and target-muscle connectivity. *Cell* 123: 477–491, 2005.
14. **Del Moral PM, Sala FG, Tefft D, Shi W, Keshet E, Bellusci S, Warburton D.** VEGF-A signaling through *Flk-1* is a critical facilitator of early embryonic lung epithelial to endothelial crosstalk and branching morphogenesis. *Dev Biol* 290: 177–188, 2006.
15. **Emery JL, Mithal A.** The number of alveoli in the terminal respiratory unit of man during late intrauterine life and childhood. *Arch Dis Child* 35: 544–547, 1960.
16. **Fewell JE, Lee CC, Kitterman JA.** Effects of phrenic nerve section on the respiratory system of fetal lambs. *J Appl Physiol* 51: 293–297, 1981.
17. **Garin E, Lemieux M, Coulombe Y, Robinson GW, Jeannotte L.** Stromal *Hoxa5* function controls the growth and differentiation of mammary alveolar epithelium. *Dev Dyn* 235: 1858–1871, 2006.
18. **Gaunt SJ, Coletta PL, Pravtcheva D, Sharpe PT.** Mouse *Hox-3.4*: homeobox sequence and embryonic expression patterns compared with other members of the *Hox* gene network. *Development* 109: 329–339, 1990.
19. **Gendronneau G, Boucherat O, Aubin J, Lemieux M, Jeannotte L.** The loss of *Hoxa5* function causes estrous acyclicity and ovarian epithelial inclusion cysts. *Endocrinology* 153: 1484–1497, 2012.

20. **Geng Y, Dong Y, Yu M, Zhang L, Yan X, Sun J, Qiao L, Geng H, Nakajima M, Furuichi T, Ikegawa S, Gao X, Chen YG, Jiang D, Ning W.** Follistatin-like 1 (Fstl1) is a bone morphogenetic protein (BMP) 4 signaling antagonist in controlling mouse lung development. *Proc Natl Acad Sci USA* 108: 7058–7063, 2011.
21. **Greer JJ, Allan DW, Martin-Caraballo M, Lemke RP.** An overview of phrenic nerve and diaphragm muscle development in the perinatal rat. *J Appl Physiol* 86: 779–786, 1999.
22. **Greer JM, Puetz J, Thomas KR, Capecchi MR.** Maintenance of functional equivalence during paralogous Hox gene evolution. *Nature* 403: 661–665, 2000.
23. **Guseh JS, Bores SA, Stanger BZ, Zhou Q, Anderson WJ, Melton DA, Rajagopal J.** Notch signaling promotes airway mucous metaplasia and inhibits alveolar development. *Development* 136: 1751–1759, 2009.
24. **Hallock PT, Xu CF, Park TJ, Neubert TA, Curran T, Burden SJ.** Dok-7 regulates neuromuscular synapse formation by recruiting Crk and Crk-L. *Genes Dev* 24: 2451–2461, 2010.
25. **Harding R, Hooper SB, Han VK.** Abolition of fetal breathing movements by spinal cord transection leads to reductions in fetal lung liquid volume, lung growth, and IGF-II gene expression. *Pediatr Res* 34: 148–153, 1993.
26. **Herriges JC, Yi L, Hines EA, Harvey JF, Xu G, Gray PA, Ma Q, Sun X.** Genome-scale study of transcription factor expression in the branching mouse lung. *Dev Dyn* 241: 1432–1453, 2012.
27. **Inanlou MR, Baguma-Nibasheka M, Kablar B.** The role of fetal breathing-like movements in lung organogenesis. *Histol Histopathol* 20: 1261–1266, 2005.
28. **Jaffe L, Jeannotte L, Bikoff EK, Robertson EJ.** Analysis of beta 2-microglobulin gene expression in the developing mouse embryo and placenta. *J Immunol* 145: 3474–3482, 1990.
29. **Jeannotte L, Lemieux M, Charron J, Poirier F, Robertson EJ.** Specification of axial identity in the mouse: role of the Hoxa-5 (Hox1.3) gene. *Genes Dev* 7: 2085–2096, 1993.
30. **Kitterman JA.** The effects of mechanical forces on fetal lung growth. *Clin Perinatol* 23: 727–740, 1996.
31. **Krumlauf R, Holland PW, McVey JH, Hogan BL.** Developmental and spatial patterns of expression of the mouse homeobox gene, Hox 2.1. *Development* 99: 603–617, 1987.
32. **Liggins GC, Vilos GA, Campos GA, Kitterman JA, Lee CH.** The effect of spinal cord transection on lung development in fetal sheep. *J Dev Physiol* 3: 267–274, 1981.
33. **Mandeville I, Aubin J, LeBlanc M, Lalancette-Hébert M, Janelle MF, Tremblay GM, Jeannotte L.** Impact of the loss of Hoxa5 function on lung alveogenesis. *Am J Pathol* 169: 1312–1327, 2006.
34. **McGinnis W, Krumlauf R.** Homeobox genes and axial patterning. *Cell* 68: 283–302, 1992.
35. **McIntyre DC, Rakshit S, Yallowitz AR, Loken L, Jeannotte L, Capecchi MR, Wellik DM.** Hox patterning of the vertebrate rib cage. *Development* 134: 2981–2989, 2007.

36. **Meunier D, Aubin J, Jeannotte L.** Perturbed thyroid morphology and transient hypothyroidism symptoms in *Hoxa5* mutant mice. *Dev Dyn* 227: 367–378, 2003.
37. **Morimoto M, Liu Z, Cheng HT, Winters N, Bader D, Kopan R.** Canonical Notch signaling in the developing lung is required for determination of arterial smooth muscle cells and selection of Clara versus ciliated cell fate. *J Cell Sci* 123: 213–224, 2010.
38. **Morrissey EE, Hogan BL.** Preparing for the first breath: genetic and cellular mechanisms in lung development. *Dev Cell* 18: 8–23, 2010.
39. **Patterson LT, Pembaur M, Potter SS.** *Hoxa11* and *Hoxd11* regulate branching morphogenesis of the ureteric bud in the developing kidney. *Development* 128: 2153–2161, 2001.
40. **Philippidou P, Walsh C, Aubin J, Jeannotte L, Dasen JS.** Sustained *Hox5* gene activity is required for respiratory motor neuron development. *Nat Neurosci* 15: 1636–1644, 2012.
41. **Potvin E, Beuret L, Cadrin-Girard JF, Carter M, Roy S, Tremblay M, Charron J.** Cooperative action of multiple cis-acting elements is required for N-myc expression in branchial arches: specific contribution of GATA3. *Mol Cell Biol* 30: 5348–5363, 2010.
42. **Ramirez MI, Millien G, Hinds A, Cao Y, Seldin DC, Williams MC.** T1alpha, a lung type I cell differentiation gene, is required for normal lung cell proliferation and alveolus formation at birth. *Dev Biol* 256: 61–72, 2003.
43. **Rancourt DE, Tsuzuki T, Capecchi MR.** Genetic interaction between *hoxb-5* and *hoxb-6* is revealed by nonallelic noncomplementation. *Genes Dev* 9: 108–122, 1995.
44. **Rijli FM, Chambon P.** Genetic interactions of *Hox* genes in limb development: learning from compound mutants. *Curr Opin Genet Dev* 7: 481–487, 1997.
45. **Rossel M, Capecchi MR.** Mice mutant for both *Hoxa1* and *Hoxb1* show extensive remodeling of the hindbrain and defects in craniofacial development. *Development* 126: 5027–5040, 1999.
46. **Roth-Kleiner M, Hirsch E, Schittny JC.** Fetal lungs of tenascin-C deficient mice grow well, but branch poorly in organ culture. *Am J Respir Cell Mol Biol* 30: 360–366, 2004.
47. **Seegmiller RE, Cooper CA, Houghton MJ, Carey JC.** Pulmonary hypoplasia in chondrodystrophic mice. *Teratology* 33: 339–347, 1986.
48. **Shannon JM, Nielsen LD, Gebb SA, Randell SH.** Mesenchyme specifies epithelial differentiation in reciprocal recombinants of embryonic lung and trachea. *Dev Dyn* 212: 482–494, 1998.
49. **Sylva M, Li VSW, Buffing AAA, van Es JH, van den Born M, van der Velden S, Gunst Q, Koolstra JH, Moorman AFM, Clevers H, van den Hoff MJB.** The BMP antagonist follistatin-like 1 is required for skeletal and lung organogenesis. *PLoS ONE* 6: e22616, 2011.
50. **Tsao PN, Vasconcelos M, Izvolsky KI, Qian J, Lu J, Cardoso WV.** Notch signaling controls the balance of ciliated and secretory cell fates in developing airways. *Development* 136: 2297–2307, 2009.
51. **Tvrđik P, Capecchi MR.** Reversal of *Hox1* gene subfunctionalization in the mouse. *Dev Cell* 11: 239–250, 2006.

52. **Volpe MV, Archavachotikul K, Bhan I, Lessin MS, Nielsen HC.** Association of bronchopulmonary sequestration with expression of the homeobox protein Hoxb-5. *J Pediatr Surg* 35: 1817–1819, 2000.
53. **Volpe MV, Pham L, Lessin M, Ralston SJ, Bhan I, Cutz E, Nielsen HC.** Expression of Hoxb-5 during human lung development and in congenital lung malformations. *Birth Defects Res A Clin Mol Teratol* 67: 550–556, 2003.
54. **Volpe MV, Ramadurai SM, Pham LD, Nielsen HC.** Hoxb-5 down regulation alters Tenascin-C, FGF10 and Hoxb gene expression patterns in pseudoglandular period fetal mouse lung. *Front Biosci* 12: 860–873, 2007.
55. **Warburton D, El-Hashash A, Carraro G, Tiozzo C, Sala F, Rogers O, De Langhe S, Kemp PJ, Riccardi D, Torday J, Bellusci S, Shi W, Lubkin SR, Jesudason E.** Lung organogenesis. *Curr Top Dev Biol* 90: 73–158, 2010.
56. **Wellik DM, Hawkes PJ, Capecchi MR.** Hox11 paralogous genes are essential for metanephric kidney induction. *Genes Dev* 16: 1423–1432, 2002.
57. **Wendel DP, Taylor DG, Albertine KH, Keating MT, Li DY.** Impaired distal airway development in mice lacking elastin. *Am J Respir Cell Mol Biol* 23: 320–326, 2000.
58. **Wigglesworth JS, Desai R.** Effect on lung growth of cervical cord section in the rabbit fetus. *Early Hum Dev* 3: 51–65, 1979.

Chapitre VII

SECTION 7.1

ADDENDUM AU CHAPITRE III (*Hoxa5/Cre* transgenic mice: novel tools for regional deletion along the anterior-posterior axis)

AVANT-PROPOS

Malgré le fait que les résultats concernant la caractérisation des deux différentes lignées de souris transgéniques exprimant la recombinase Cre sous le contrôle des séquences régulatrices du gène *Hoxa5* soient déjà publiés (Bérubé-Simard et Jeannotte, 2014; Chapitre III), des expériences additionnelles ont été réalisées afin d'approfondir la validation de nos deux outils.

Premièrement, le géotypage par buvardage Southern présenté au chapitre III (Fig. 3.2H) soulève au moins deux points laissant place à amélioration: la proximité des bandes représentant les allèles *Hoxa5*⁺ et *Hoxa5*^{floxed} (respectivement 3.7-kb et 3.9-kb) ainsi que la faible intensité de la bande représentant l'allèle *Hoxa5*^Δ, lorsque les échantillons d'ADN génomique sont digérés par l'enzyme *EcoRI* et hybridés avec une sonde *Bg/III-XhoI* (pLJ36). Pour remédier à ces problèmes, j'ai effectué différentes digestions sur des extraits d'ADN génomique issus de tissus où la recombinase Cre est davantage exprimée, à savoir les carcasses d'embryons éviscérés à E12.5 en ce qui concerne la lignée 7.1-kb *Hoxa5/Cre*, alors que j'ai opté plutôt pour des poumons d'embryons E18.5 dans le cas de la lignée 14.5-kb *Hoxa5/Cre* (Figures 7.1.1-7.1.4). Ces résultats seront discutés à la section 8.2. Deuxièmement, j'ai testé l'efficacité de la lignée 14.5-kb *Hoxa5/Cre* à déléter l'allèle *Yy1*^{flox} et comparer les phénotypes obtenus à ceux des souris *Yy1*^{flox/flox}, *Dermo1*^{+Cre} (Chapitre II), des souris *Hoxa5*^{flox/flox}, 14.5-kb *Hoxa5/Cre* (Chapitre III) et des souris *Hoxa5*^{-/-} (Figure 7.1.5). Ces résultats seront discutés à la section 8.1.1.

Figure 7.1.1. Génotypage par Southern blot pour contrôler l'efficacité de la délétion conditionnelle du gène *Hoxa5* à l'aide de deux lignées *Hoxa5/Cre*

(A) Cartes de restriction partielles des différents allèles produits suite à la délétion spécifique de la recombinaise Cre. Les boîtes noires correspondent aux homéoboîtes, les grises, aux régions traduites et les blanches, aux régions transcrites. La flèche représente le site d'initiation de la transcription (TSS) du gène *Hoxa5*. Le fragments d'ADN, sur lequel la sonde utilisée pour les analyses de type Southern blot s'hybrident (pLJ36), est représenté sur la carte de l'allèle sauvage (wt). Les fragments d'ADN attendus suite à une digestion *EcoRI* sont indiqués sous les cartes des différents allèles. Ap, *ApaI*; B, *BglII*; H, *HindIII*; Nh, *NheI*; RI, *EcoRI*; Sc, *SacI*; Xh, *XhoI* (B) Cinq échantillons d'ADN génomique différents (un échantillon *Hoxa5*^{fllox/+}, deux *Hoxa5*^{fllox/fllox}; 7.1-kb *Hoxa5/Cre* et deux *Hoxa5*^{fllox/fllox}; 14.5-kb *Hoxa5/Cre*) ont été utilisés. L'analyse de type Southern blot permet de discriminer les différents allèles *Hoxa5* suite à une digestion *EcoRI*: *Hoxa5*⁺ (3793-pb), *Hoxa5*^{fllox} (3910-pb), *Hoxa5*^Δ (7679-pb) et *Hoxa5/Cre* (2232-pb). Les astérisques représentent des fragments d'ADN non attendus (digestion incomplète, intégration incomplète et intégration en concatémère du transgène).

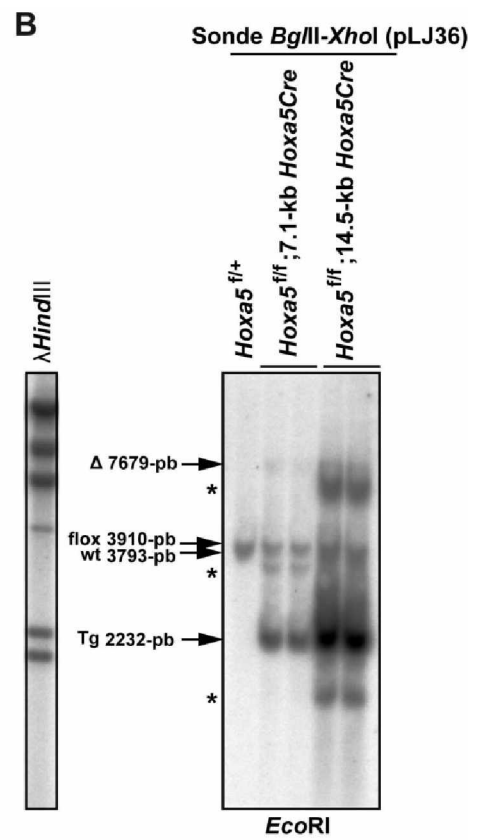
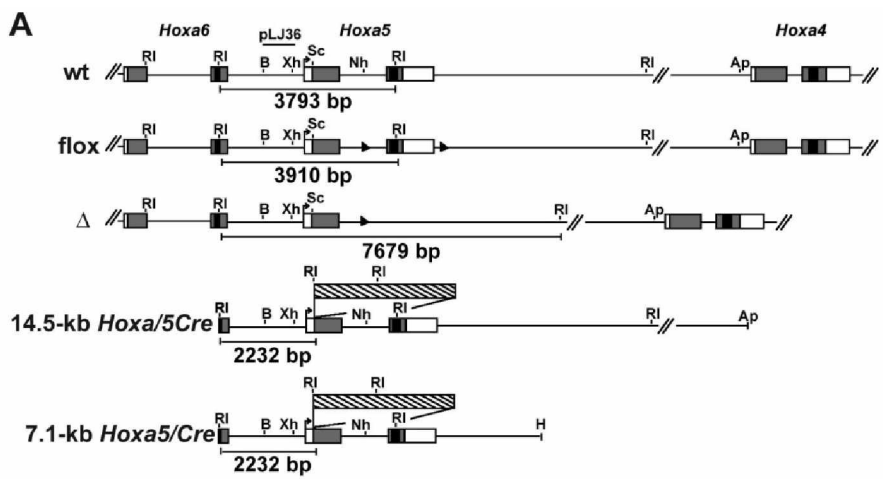


Figure 7.1.2. Génotypage par Southern blot pour contrôler l'efficacité de la délétion conditionnelle du gène *Hoxa5* à l'aide de deux lignées *Hoxa5/Cre*

(A) Cartes de restriction partielles des différents allèles produits suite à la délétion spécifique de la recombinaise Cre. Les boîtes noires correspondent aux homéoboîtes, les grises, aux régions traduites et les blanches, aux régions transcrites. La flèche représente le site d'initiation de la transcription (TSS) du gène *Hoxa5*. Le fragments d'ADN, sur lequel la sonde utilisée pour les analyses de type Southern blot s'hybrident (pLJ36), est représenté sur la carte de l'allèle sauvage (wt). Les fragments d'ADN attendus suite à une digestion *EcoRI* sont indiqués sous les cartes des différents allèles. Ap, *ApaI*; H, *HindIII*; Nh, *NheI*; RI, *EcoRI*; Sc, *SacI*; Xh, *XhoI* (B) Cinq échantillons d'ADN génomique différents (un échantillon *Hoxa5*^{fllox/+}, deux *Hoxa5*^{fllox/fllox}, 7.1-kb *Hoxa5/Cre* et deux *Hoxa5*^{fllox/fllox}, 14.5-kb *Hoxa5/Cre*) ont été utilisés. L'analyse de type Southern blot permet de discriminer les différents allèles *Hoxa5* suite à une digestion *ApaI*: *Hoxa5*⁺ (2534-pb), *Hoxa5*^{fllox} (2117-pb), *Hoxa5*^Δ (2117-pb) et *Hoxa5/Cre* (1137-pb). Les astérisques représentent des fragments d'ADN non attendus (digestion incomplète, intégration incomplète et intégration en concatémère du transgène).

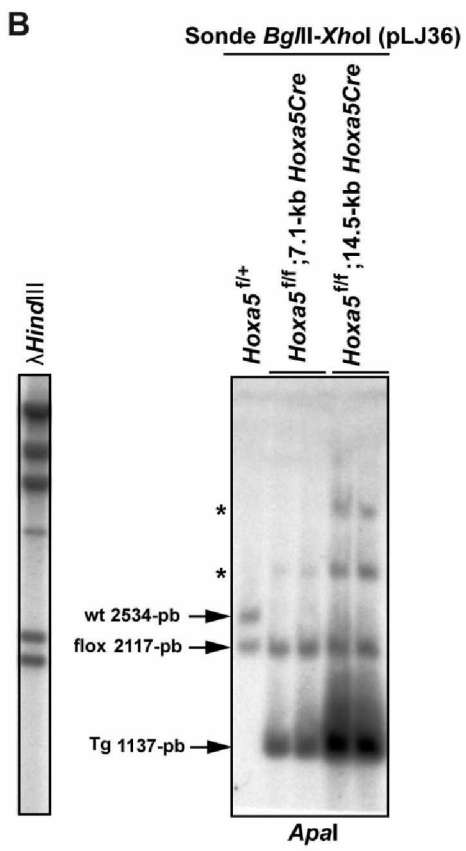
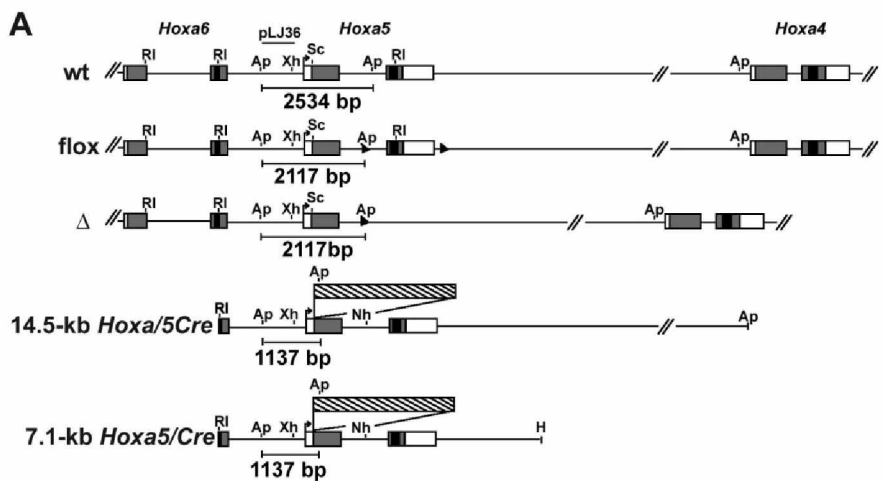


Figure 7.1.3. Génotypage par Southern blot pour contrôler l'efficacité de la délétion conditionnelle du gène *Hoxa5* à l'aide de deux lignées *Hoxa5/Cre*

(A) Cartes de restriction partielles des différents allèles produits suite à la délétion spécifique de la recombinaise Cre. Les boîtes noires correspondent aux homéoboîtes, les grises, aux régions traduites et les blanches, aux régions transcrites. La flèche représente le site d'initiation de la transcription (TSS) du gène *Hoxa5*. Le fragments d'ADN, sur lequel la sonde utilisée pour les analyses de type Southern blot s'hybrident (pLJ35), est représenté sur la carte de l'allèle sauvage (wt). Les fragments d'ADN attendus suite à une digestion *EcoRI* sont indiqués sous les cartes des différents allèles. Ap, *ApaI*; B, *BglII*; H, *HindIII*; Nh, *NheI*; RI, *EcoRI*; Xb, *XbaI*; Xh, *XhoI* (B) Cinq échantillons d'ADN génomique différents (un échantillon *Hoxa5*^{flox/+}, deux *Hoxa5*^{flox/flox}, 7.1-kb *Hoxa5/Cre* et deux *Hoxa5*^{flox/flox}, 14.5-kb *Hoxa5/Cre*) ont été utilisés. L'analyse de type Southern blot permet de discriminer les différents allèles *Hoxa5* suite à une digestion *XbaI*: *Hoxa5*⁺ (8590-pb), *Hoxa5*^{flox} (1408-pb), *Hoxa5*^Δ (6953-pb) 7.1-kb *Hoxa5/Cre* (5096-pb) et 14.5-kb *Hoxa5/Cre* (8618-pb). Les astérisques représentent des fragments d'ADN non attendus (digestion incomplète, intégration incomplète et intégration en concatémère du transgène).

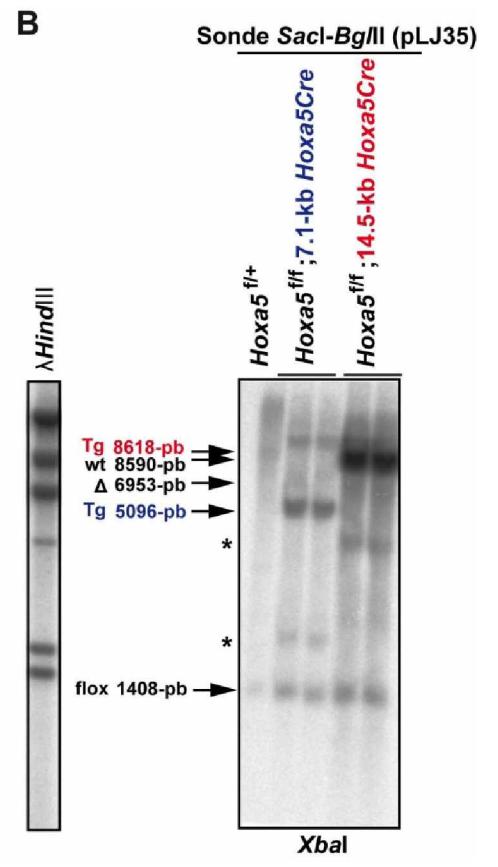
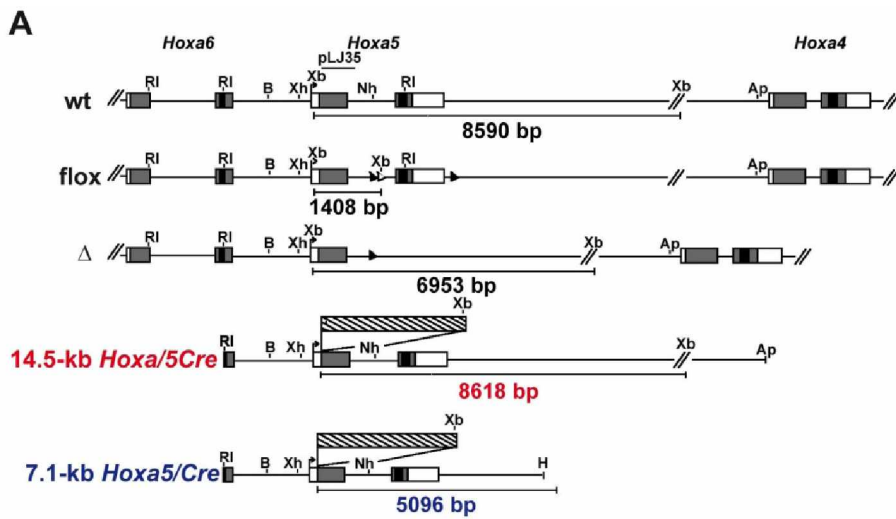


Figure 7.1.4. Génotypage par Southern blot pour contrôler l'efficacité de la délétion conditionnelle du gène *Hoxa5* à l'aide de deux lignées *Hoxa5/Cre*

(A) Cartes de restriction partielles des différents allèles produits suite à la délétion spécifique de la recombinaise Cre. Les boîtes noires correspondent aux homéoboîtes, les grises, aux régions traduites et les blanches, aux régions transcrites. La flèche représente le site d'initiation de la transcription (TSS) du gène *Hoxa5*. Le fragments d'ADN, sur lequel la sonde utilisée pour les analyses de type Southern blot s'hybrident (pLJ35), est représenté sur la carte de l'allèle sauvage (wt). Les fragments d'ADN attendus suite à une digestion *EcoRI* sont indiqués sous les cartes des différents allèles. Ap, *ApaI*; H, *HindIII*; Nh, *NheI*; RI, *EcoRI*; Pv, *PvuII*; Xb, *XbaI*; Xh, *XhoI* (B) Cinq échantillons d'ADN génomique différents (un échantillon *Hoxa5*^{flox/+}, deux *Hoxa5*^{flox/flox}, 7.1-kb *Hoxa5/Cre* et deux *Hoxa5*^{flox/flox}; 14.5-kb *Hoxa5/Cre*) ont été utilisés. L'analyse de type Southern blot permet de discriminer les différents allèles *Hoxa5* suite à une digestion *PvuII*: *Hoxa5*⁺ (6493-pb), *Hoxa5*^{flox} (6655-pb), *Hoxa5*^Δ (4856-pb) et *Hoxa5/Cre* (9195-pb). Les astérisques représentent des fragments d'ADN non attendus (digestion incomplète, intégration incomplète et intégration en concatémère du transgène).

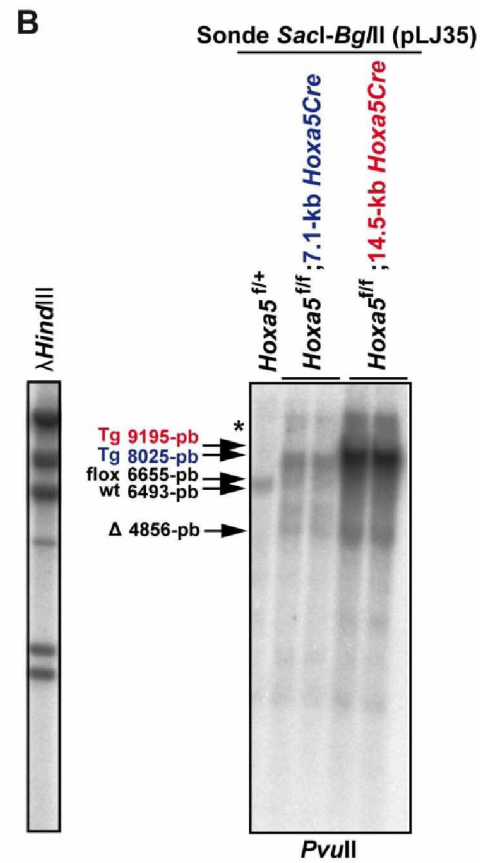
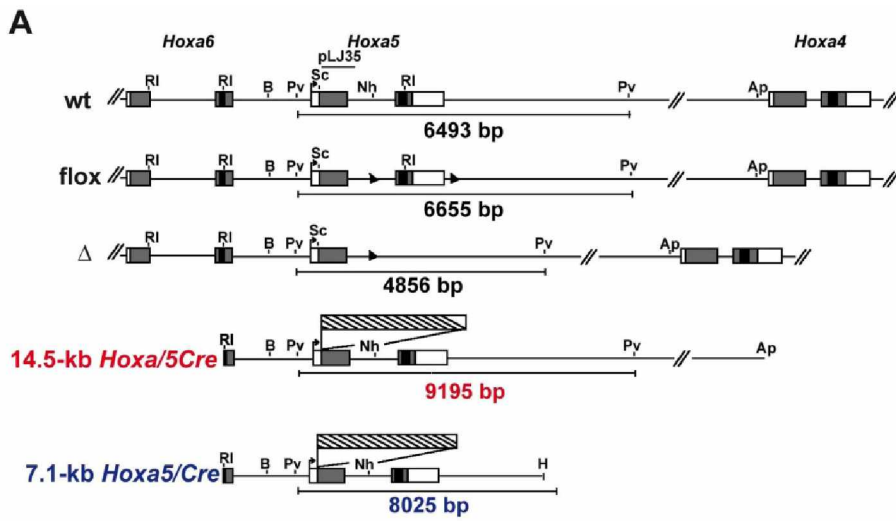


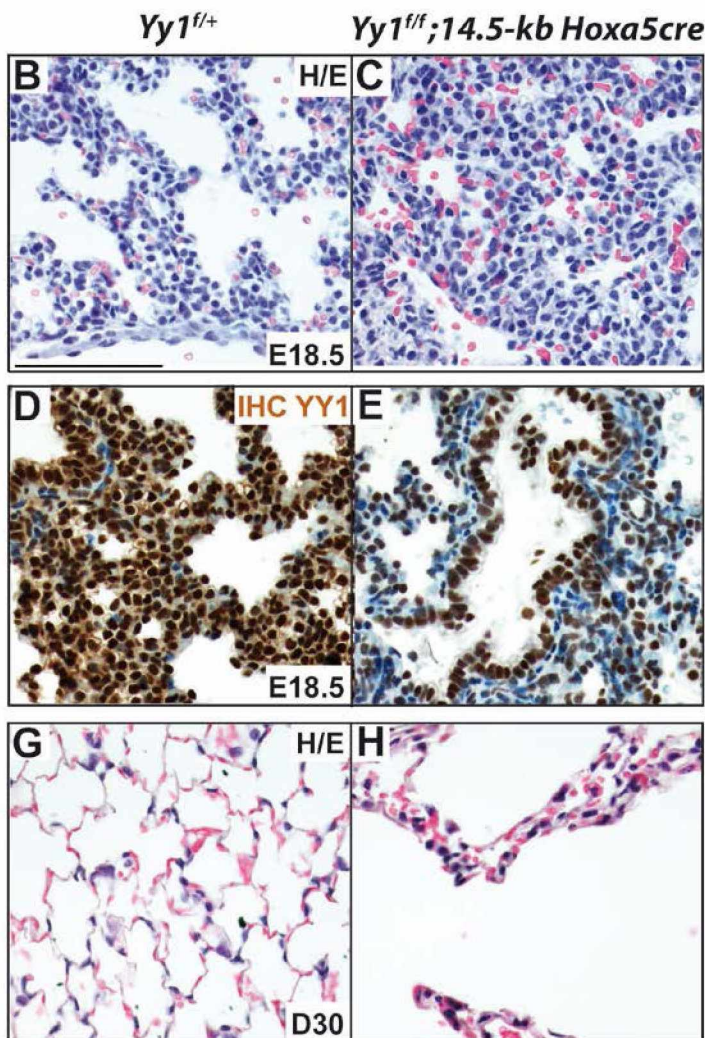
Figure 7.1.5. Impact de la délétion conditionnelle du gène *Yy1* suite à l'action de la 14.5-kb *Hoxa5/Cre*

(A) Ratios des génotypes des portées obtenues suite aux croisements entre des souris *Yy1*^{flox/+}; 14.5-kb *Hoxa5/Cre* et *Yy1*^{flox/flox} (B-C) Histologie comparative de poumons *Yy1*^{flox/+} et *Yy1*^{flox/flox}; 14.5-kb *Hoxa5/Cre* au jour embryonnaire E18.5. Les individus *Yy1*^{flox/+} présentent une structure pulmonaire normale, alors que les poumons des individus *Yy1*^{flox/flox}; 14.5-kb *Hoxa5/Cre* sont affaissés. Échelle: 200µm. (D-E) Immunodétection de la protéine YY1 montrant une expression ubiquitaire au niveau de l'épithélium et du mésenchyme pulmonaire d'embryons de souris *Yy1*^{flox/+} au jour E18.5, mais une diminution importante du marquage mésenchymal chez les individus *Yy1*^{flox/flox}; 14.5-kb *Hoxa5/Cre*. (F) Analyse qRT-PCR pour l'expression des gènes *Yy1*, *Hoxa5* et *Hoxa4* dans des poumons d'embryons *Yy1*^{+/+}; 14.5-kb *Hoxa5/Cre* et *Yy1*^{flox/flox}; 14.5-kb *Hoxa5/Cre* au jour E18.5. Le niveau d'expression des gènes *Yy1* et *Hoxa5* est significativement diminué. Les résultats sont présentés en valeurs moyennes ± SEM. **p*<0.05, ***p*<0.01, ****p*<0.001. (G-H) Histologie comparative de poumons *Yy1*^{flox/+} et *Yy1*^{flox/flox}; 14.5-kb *Hoxa5/Cre* à J30. Les individus *Yy1*^{flox/+} présentent une structure pulmonaire normale, alors que les poumons des individus *Yy1*^{flox/flox}; 14.5-kb *Hoxa5/Cre* montrent des espaces aériens anormalement élargis.

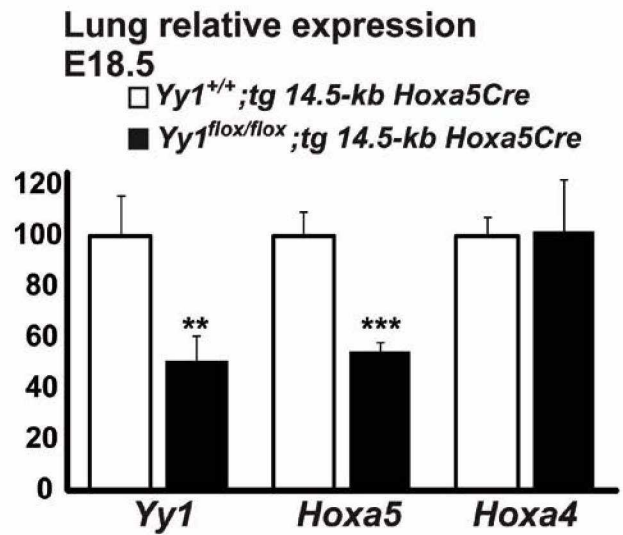
A

Age	# of litters	# of pups	<i>Yy1^{flox/flox}</i> X <i>Yy1^{flox/+}; 14.5-kb Hoxa5Cre</i>			
			Cre-		Cre+	
			<i>Yy1^{flox/+}</i>	<i>Yy1^{flox/flox}</i>	<i>Yy1^{flox/+}</i>	<i>Yy1^{flox/flox}</i>
E18.5	13	92	28(30%)	22(24%)	17(19%)	25(27%)
D30	19	102	24(23%)	25(25%)	22(21%)	13(13%)-18*(18%)
% expected			25	25	25	25

* Found dead at birth



F



SECTION 7.2

Rôle de l'expression du gène *Yy1* au niveau de l'épithélium des voies respiratoires

AVANT-PROPOS

Tel que mentionné au chapitre II, la délétion du gène *Yy1* au niveau du mésenchyme pulmonaire, grâce à l'action de la *Dermo1/Cre*, entraîne une mortalité périnatale associée à une détresse respiratoire. De plus, les individus mutants au jour embryonnaire E18.5 présentent des poumons anormalement compacts qui ressemblent aux poumons des souris *Hoxa5*^{-/-}. D'ailleurs, ces observations corrélerent avec d'autres données présentées au chapitre II supportant une diminution d'expression du gène *Hoxa5* (transgénèse, qPCR) dans un contexte où la protéine YY1 n'est pas en mesure de lier l'élément régulateur Poumon/Intestin situé dans la séquence intergénique *Hoxa4-Hoxa5*. Comme il a été démontré que le gène *Yy1* remplit des fonctions essentielles à la progression du cycle cellulaire ainsi qu'au développement embryonnaire tardif, et ce de façon dépendante à la quantité totale de la protéine YY1 (Donohoe *et al.*, 1999; Affar *et al.*, 2006), j'ai voulu m'assurer que le phénotype observé chez les individus *Yy1*^{flox/flox}, *Dermo1*^{+Cre} soit bien spécifique à sa perte au niveau du mésenchyme pulmonaire et non à l'atteinte d'un seuil critique global au sein du poumon.

Pour se faire, j'ai initié la délétion du gène *Yy1* au niveau de l'épithélium pulmonaire en tirant avantage de la lignée *Shh*^{+Cre}. Bien que j'ai également observé une mortalité périnatale associée à une détresse respiratoire chez les individus mutants, le phénotype pulmonaire observé chez les embryons E18.5 *Yy1*^{flox/flox}, *Shh*^{+Cre} est beaucoup plus sévère que celui des souris *Yy1*^{flox/flox}, *Dermo1*^{+Cre}, soit la présence de deux sacs vides en guise de poumon (Figure 7.2.1). Par conséquent, non seulement le phénotype pulmonaire observé au chapitre II est spécifique à la délétion mésenchymale du gène *Yy1*, mais tout indique que le gène *Yy1* joue également un rôle essentiel au niveau de l'épithélium pulmonaire. Puisque mes collègues Kim Landry-Truchon et Olivier Boucherat poursuivent le projet et suggèrent un lien entre les gènes *Yy1* et *Shh*, j'ai démontré par immunoprécipitation de la

chromatine (ChIP) que les protéines YY1 et BAF53A, un membre du complexe INO80 qui est connu pour interagir avec YY1 dans un contexte d'activation transcriptionnelle de gènes cibles (Cai *et al.*, 2007), lient spécifiquement des sites de liaisons localisés dans la région promotrice du gène *Shh* chez la souris (Figure 7.2.2). Ces résultats seront discutés à la section 8.1.1 et feront partie d'un article actuellement en préparation (Boucherat O, Landry-Truchon K, Bérubé-Simard FA, Houde N, Lezmi, Delacourt C, Jeannotte L. Epithelial inactivation of *Yy1* results in defective lung branching morphogenesis and cyst formation reminiscent of type I pleuropulmonary blastoma).

Figure 7.2.1. Analyse du phénotype pulmonaire des souris $YyI^{\text{flox/flox}}$; $Shh^{+/Cre}$

(A) Ratios des génotypes des portées obtenues suite aux croisements entre des souris $YyI^{\text{flox/+}}$; $Shh^{+/Cre}$ et $YyI^{\text{flox/flox}}$ (B-D) Analyse macroscopique de poumons entiers d'embryons $YyI^{\text{flox/+}}$ et $YyI^{\text{flox/flox}}$; $Shh^{+/Cre}$ au jour embryonnaire E18.5. Les individus $YyI^{\text{flox/+}}$ présentent un système respiratoire d'apparence normale, alors que les trachées des individus $YyI^{\text{flox/flox}}$; $Shh^{+/Cre}$ semblent plus longues, plus étroites et leurs anneaux de cartilages semblent désorganisés. Quant à leur poumon, il semble plus petit et présente des sacs vides en guise de lobes. Échelle: 4mm (E-F) Histologie comparative de poumons $YyI^{\text{flox/flox}}$ et $YyI^{\text{flox/flox}}$; $Shh^{+/Cre}$ au jour E18.5. Tel qu'observé macroscopiquement, les individus $YyI^{\text{flox/flox}}$ présentent une structure pulmonaire normale, alors que l'architecture pulmonaire des individus $YyI^{\text{flox/flox}}$; $Shh^{+/Cre}$ est dramatiquement affectée, caractérisée par la présence de kystes. Échelle: 25 μm .

A

Age	# of litters	# of pups	<i>Yy1</i> ^{flox/flox} X <i>Yy1</i> ^{flox/+} ; <i>Shh</i> ^{+/-Cre}			
			<i>Shh</i> ^{+/+}		<i>Shh</i> ^{+/-Cre}	
			<i>Yy1</i> ^{flox/+}	<i>Yy1</i> ^{flox/flox}	<i>Yy1</i> ^{flox/+}	<i>Yy1</i> ^{flox/flox}
E18.5	6	46	8(17,4%)	15(32,6%)	8(17,4%)	15(32,6%)
D30	13	79	26(33%)	21(27%)	27(34%)	0(0%) 5*(6%)
% expected			25	25	25	25

* Found dead at birth

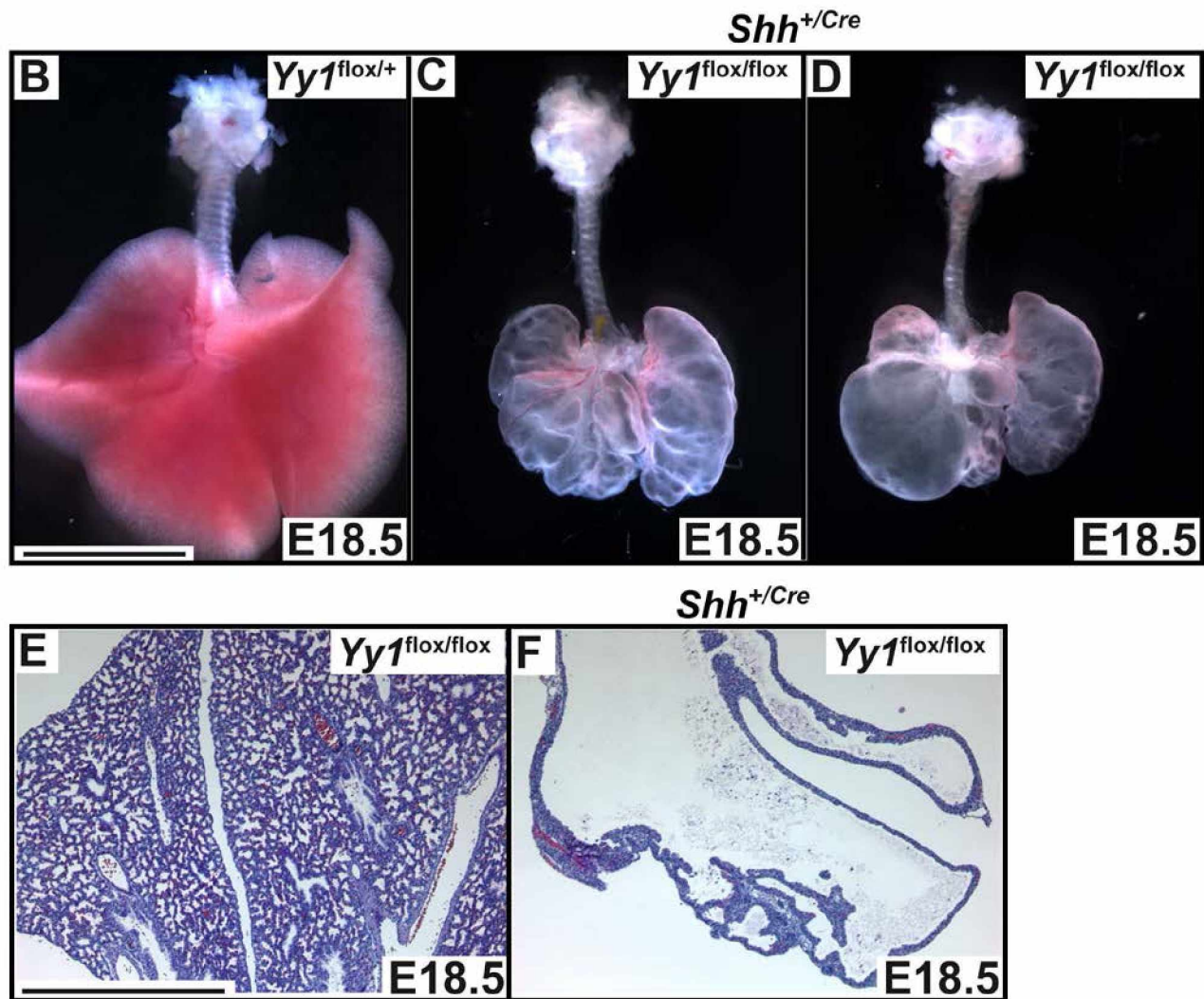
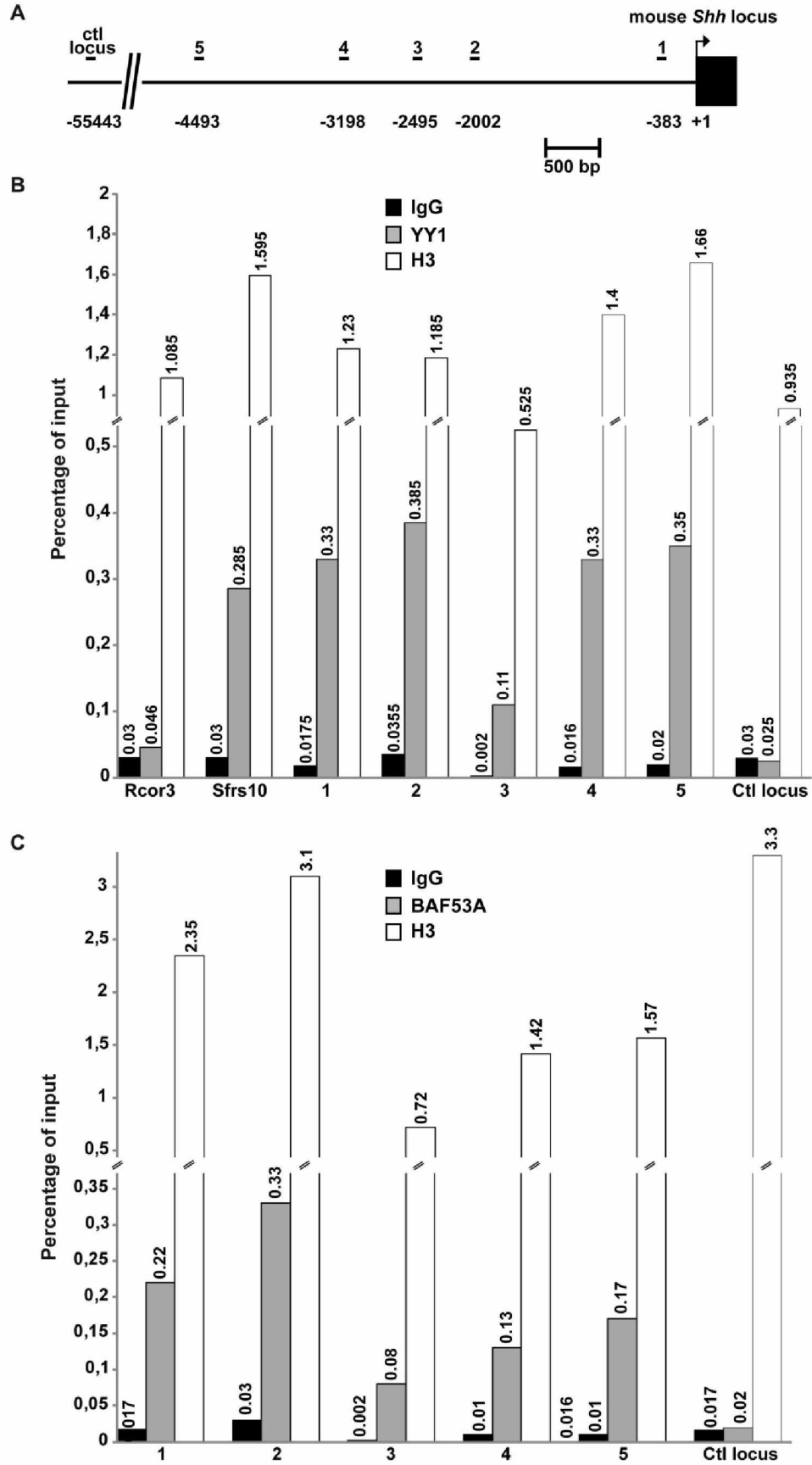


Figure 7.2.2. Détection de la liaison des protéines YY1 et BAF53A au niveau de la région promotrice du gène *Shh* chez la souris.

(A) Représentation schématique de la région promotrice du gène *Shh* avec les positions relatives à son site d'initiation de la transcription (TSS). Les positions correspondant aux 5 régions enrichies en sites de liaison YY1, de même que celle du contrôle locus négatif, sont indiquées. (B-C) Analyse CHIP sur 5000 pb à l'extrémité 5' du promoteur du gène *Shh* dans des poumons d'embryons de souris au jour E14.5. (B) La chromatine a été immunoprécipitée avec des anticorps anti-IgG de lapin (contrôle négatif), anti-YY1 et anti-histone H3 (contrôle positif de l'intégrité de la chromatine). Les différents pourcentages d'immunoprécipitation ont été évalués par qPCR. Ces résultats sont la moyenne \pm SEM de deux expériences indépendantes. (C) La chromatine a été immunoprécipitée avec des anticorps anti-IgG de lapin (contrôle négatif), anti-BAF53A et anti-histone H3 (contrôle positif de l'intégrité de la chromatine). Les différents pourcentages d'immunoprécipitation ont été évalués par qPCR. Ces résultats proviennent d'une seule expérience.



SECTION 7.3

Rôle critique de la signalisation ERK/MAPK dans le développement du système respiratoire

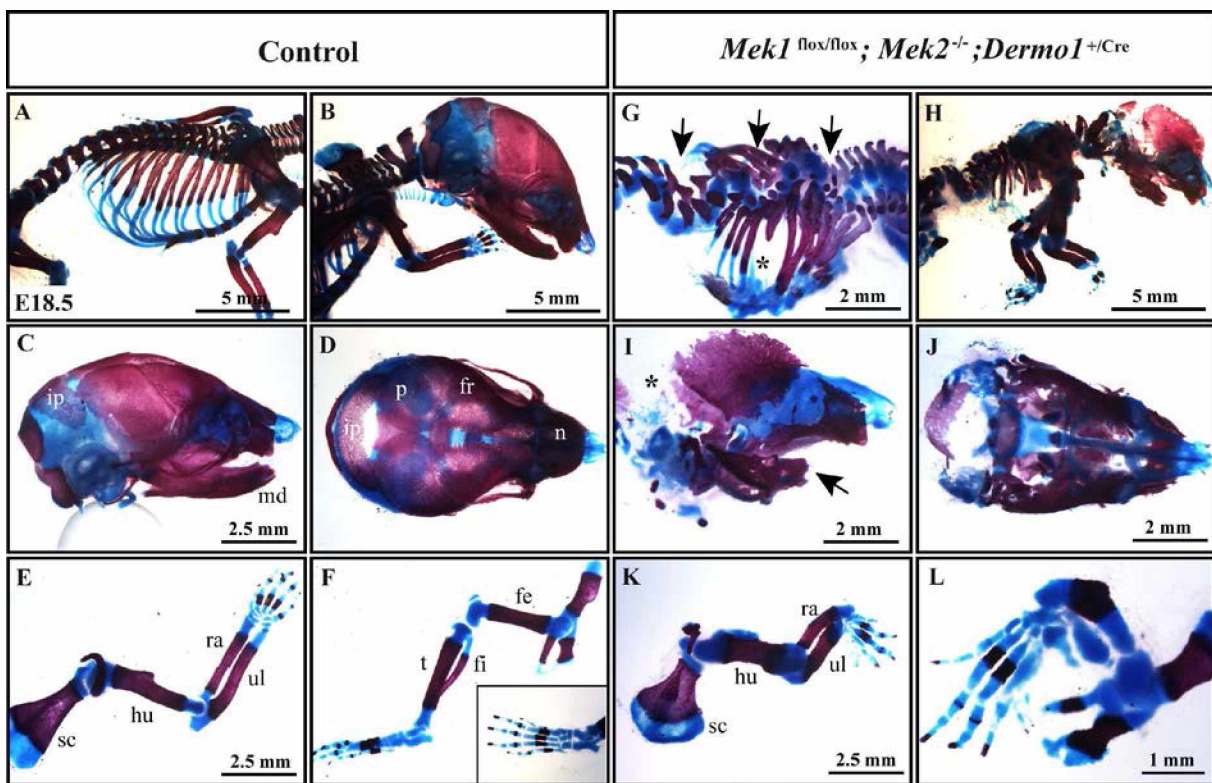
PRÉSENTATION

Bien que des expériences *in vivo* et *in vitro* aient démontré l'importance de certains facteurs de croissance impliqués dans l'activation des sentiers de signalisation MAPK lors du développement pulmonaire, notamment FGF1, -2, -7, -9, -10, -18, le PDGF et l'EGF, (Warburton et al., 2010), la contribution du signal ERK/MAPK dans la morphogenèse pulmonaire n'a jamais été clairement démontrée. Afin de vérifier cette hypothèse, l'ablation de la fonction du gène *Mek1*, dans un environnement mutant pour le gène *Mek2*, a été réalisée spécifiquement au niveau du mésenchyme en utilisant la lignée de souris *Dermo1^{+Cre}*, de même qu'au niveau de l'épithélium grâce aux souris *Shh^{+Cre}*. La délétion mésenchymale résulte en un retard de croissance intra-utérine, un omphalocèle, un problème de fermeture des paupières, des anomalies squelettiques, une hypoplasie pulmonaire et une mortalité périnatale. Quant à la délétion épithéliale, elle génère également un phénotype de mortalité périnatale. Toutefois, on observe plutôt une agénie pulmonaire bilatérale ainsi que l'absence de cellules épithéliales différenciées au niveau de la trachée des embryons *Mek1^{flox/flox};Mek2^{-/-};Shh^{+Cre}*.

En ce qui me concerne, j'ai réalisé la caractérisation des squelettes des embryons *Mek1^{flox/flox};Mek2^{-/-};Dermo1^{+Cre}* au stade E18.5 (Fig. S2; Figure 7.3.1). Ils présentent plusieurs difformités, dont des anomalies sévères au niveau de la cage thoracique et des membres, l'absence de cartilage dans la région occipitale, la polydactylie et la malformation des structures maxillaires. Ces résultats, qui suggèrent que les protéines MEK sont essentielles au mésenchyme en développement lors de la spécification des cellules ostéochondro-progénitrices et de la formation de la paroi corporelle, seront présentés dans le journal *Development* (Boucherat O, Nadeau V, Bérubé-Simard FA, Charron J*, Jeannotte L*. Critical requirement of ERK/MAPK signaling in respiratory tract development) Article accepté en juin 2014.

Figure 7.3.1. Dismorphogénèse cranio-faciale et anomalies squelettiques observées chez les embryons *Mek1^{flox/flox};Mek2^{-/-};Dermo1^{+/-Cre}*

(A, B, G, H) Préparations de squelettes entiers d'embryons de souris au stade E18.5 colorés au bleu d'alcan (cartilage) et au rouge alizarine (os). Les embryons *Mek1^{flox/flox};Mek2^{-/-};Dermo1^{+/-Cre}* présentent une cage thoracique plus petite et anormale, caractérisée par des difformités cyphotiques de la colonne vertébrale thoracique (flèches) ainsi que la troncation/absence de côtes (astérisque). (C, D, I, J) Vues latérales (C, I) et vues de dessus (D, J) du crâne montrant plusieurs anomalies incluant une ossification incomplète (astérisque) et un rétrécissement du mandibule (flèche) chez les embryons *Mek1^{flox/flox};Mek2^{-/-};Dermo1^{+/-Cre}*. (E, F, K, L) Rétrécissement des éléments squelettiques des membres supérieurs (E, K) et inférieurs (F, L) observé chez les embryons *Mek1^{flox/flox};Mek2^{-/-};Dermo1^{+/-Cre}*. Os intrapariétal (ip), pariétal (p), frontal (fr) et nasal (n). sc, scapula; hu, humérus; ra, radius; ul, ulna; t, tibia; fi, fibula; fe, fémur. Barres d'échelle, 5 mm (A, B, H), 2.5 mm (C-F, K), 2 mm (G, I, J) and 1 mm (L).



CHAPITRE VIII: DISCUSSION

Le but de mon projet de thèse était d'approfondir les mécanismes contrôlant la régulation transcriptionnelle du gène *Hoxa5* chez la souris afin d'acquérir de nouveaux outils pour mieux comprendre les pathologies où le gène *Hoxa5* est régulé de façon aberrante, comme les maladies pulmonaires obstructives chroniques (MPOC) (Boucherat *et al.*, 2012) et certains types de cancers (lymphome, carcinosarcome mammaire, MIN, leucémie lymphoblastique aigüe) (Okada *et al.*, 2006; Gendronneau *et al.*, 2010). Pour atteindre ce but, je m'étais fixé comme objectif principal de caractériser les éléments régulateurs "poumon/intestin" (Moreau et Jeannotte, 2002; Chapitre II), ainsi que le "mesodermal enhancer sequence" (MES) (Tabariès *et al.*, 2005; Chapitre IV), schématisés à la Figure 1.14. Parallèlement, la caractérisation des séquences régulatrices m'a amené à générer des lignées de souris transgéniques exprimant la recombinaison Cre sous le contrôle des différents éléments de contrôle du gène *Hoxa5* (Chapitre III). Enfin, j'ai participé à d'autres résultats présentés aux Chapitres V à VII qui seront également discutés.

8.1 L'élément poumon/intestin contient plusieurs séquences régulatrices importantes

De tous les phénotypes découlant de la perte de fonction du gène *Hoxa5*, le phénotype de mortalité périnatale associée à la dysmorphogénèse des voies respiratoires est de loin le plus important (Aubin *et al.*, 1997; Mandeville *et al.*, 2006; Boucherat *et al.*, 2012, 2013). Une analyse par délétions successives dans un contexte de transgénèse *Hoxa5/lacZ* a permis d'identifier une séquence d'ADN de 1.5 kb (élément poumon/intestin), située dans la séquence intergénique *Hoxa4-Hoxa5* entre les positions +9.4 kb et +10.8 kb, pouvant récapituler en bonne partie le profil d'expression endogène du gène *Hoxa5* au niveau des poumons, de l'estomac et de l'intestin (Fig. 2.3H; Moreau et Jeannotte, 2002).

Une caractérisation plus approfondie de l'élément poumon/intestin m'a permis d'identifier deux séquences régulatrices, dont un sous-fragment *XbaI-BssHII* de 259 pb important pour diriger l'expression du gène *Hoxa5* au niveau des systèmes respiratoire et

digestif (Fig. 2.3I). Comme ce fragment renferme un élément régulateur positif sensible à l'acide rétinoïque et dirigeant l'expression du gène *Hoxa4* au niveau du poumon, de l'intestin et du métanéphros (Packer et al., 1998), nous avons vérifié si ce site RARE était aussi important pour l'expression du gène *Hoxa5*. Notre étude par transgénèse a démontré que ce site RARE ne semble pas être partagé entre ces deux gènes en ce qui concerne leur expression respective au niveau des poumons et de l'intestin, mais serait plutôt spécifique à l'expression du gène *Hoxa4*. Par contre, la mutation du site RARE affecte la limite antérieure d'expression du transgène *Hoxa5/lacZ* au niveau du tube neural (Fig. 2.2E, 2.6C), ce qui suggère que cet élément régulateur positif pourrait être partagé entre les gènes *Hoxa4* et *Hoxa5* pour diriger l'expression au niveau de cette structure. De plus, ce site RARE semble posséder des propriétés activatrices puisque le fragment *XbaI-XbaI* de 1.5 kb dirige l'expression de la β -galactosidase de façon similaire au niveau de la partie antérieure du tube neural lorsqu'il est jumelé au transgène *hsp68/lacZ* (Fig. 2.2G-H) alors que le fragment *BssHIII-SacI* de 751 pb semble avoir un profil d'expression non-spécifique dans cette structure (Fig. 2.4D). Ce site RARE est conservé au niveau de la séquence intergénique *Hoxb4-Hoxb5* et contrôle l'établissement de la limite antérieure d'expression des gènes *Hoxb* situés plus en 5' du complexe, au niveau du tube neural (Oosterveen *et al.*, 2003). Par conséquent, il serait intéressant de vérifier si le site RARE situé dans la séquence intergénique *Hoxa4-Hoxa5* agit sur l'expression d'autres gènes *Hoxa* plus postérieurs. Nous pourrions également générer une construction transgénique contenant les promoteurs des gènes *Hoxa4* et *Hoxa5*, dirigeant respectivement l'expression de deux gènes rapporteurs différents (*Hoxa4/cherry*; *Hoxa5/GFP*), et vérifier si le site RARE est partagé de façon équivalente entre les deux promoteurs ou s'il y possède une certaine affinité pour l'un ou l'autre des promoteurs lorsqu'il dirige l'expression au niveau du tube neural.

Les résultats obtenus par EMSA ont également corroboré nos observations recueillies par transgénèse concernant l'absence de rôle important pour le site RARE présent au niveau du fragment *XbaI-BssHIII* de 259 pb dans le développement des systèmes respiratoire et digestif chez la souris au stade E13.5 (Fig. 2.5B, puits 1-6). Toutefois, nous ne pouvons pas exclure la possibilité que ce site RARE soit occupé par des RARs et/ou RXR à des stades plus précoces lors du développement pulmonaire, d'autant plus que

l'ARNm *Raldh2* codant l'enzyme principal synthétisant le RA ne semble plus être exprimé au niveau pulmonaire au stade E13.5 (Niederreither *et al.*, 1997). Par conséquent, l'efficacité du site RARE pourrait être évaluée par EMSA avec des extraits protéiques de poumons issus d'embryons de souris à des stades développementaux plus précoces. De plus, l'utilisation d'extrait de cellules entières a peut-être favorisé l'identification d'une protéine ubiquitaire abondante, comme YY1 par exemple (Fig. 2.5B-D), au détriment d'un facteur de transcription spécifique au mésenchyme. Par conséquent, l'utilisation d'extraits protéiques nucléaires provenant d'un enrichissement de cellules mésenchymales pulmonaires serait intéressante pour confirmer que YY1 est bien le facteur de transcription clé liant cette séquence. Malgré tout, le rôle spécifique du facteur de transcription YY1 dans le contrôle de l'expression du gène *Hoxa5* a été confirmé de plusieurs façons et son implication biologique au niveau du développement des systèmes respiratoire et digestif sera abordée à la section 8.1.1.

L'autre séquence régulatrice que j'ai identifiée au sein de l'élément poumon/intestin est un fragment *NcoI-SacI* de 163 pb essentiel à l'expression du gène *Hoxa5* dans les systèmes respiratoire et digestif chez les souris au stade E13.5 (Fig. 2.4J-K). Ce fragment semble posséder des propriétés activatrices puisqu'il est en mesure de diriger l'expression du gène *lacZ* au niveau des poumons, de l'estomac et de l'intestin lorsqu'il est accolé au transgène *hsp68/lacZ* (Fig. 2.4L). Il serait également intéressant de vérifier l'influence de ce fragment sur l'expression du gène *Hoxa4* au niveau des organes. Nos essais de retard sur gel ont également identifié le facteur de transcription YY1 comme étant en mesure de lier spécifiquement le fragment *NcoI-SacI* de 167 pb.

Bien que j'ai identifié deux régions d'ADN à l'intérieur du fragment *XbaI-XbaI* de 1.5 kb qui agissent conjointement pour diriger l'expression du transgène *Hoxa5/lacZ* au niveau des systèmes respiratoire et digestif, il est possible que d'autres séquences d'ADN soient également importantes pour réguler adéquatement l'expression du gène *Hoxa5* dans ces organes. Par exemple, l'incapacité du fragment *XbaI-XbaI* de 1.5 kb à diriger l'expression du gène *lacZ* lorsque qu'il est placé en amont du transgène *hsp68/lacZ* (Fig. 2.2M-N), comparativement au fragment *BssHIII-SacI* de 751 pb (Fig. 2.4I), suggère la présence d'une séquence capable d'inhiber la propriété activatrice dirigeant l'expression

dans les systèmes respiratoire et digestif. Le fragment *SacI-XbaI* de 455 pb pourrait être un candidat potentiel puisqu'il semble interférer avec les éléments BSC et MES lorsqu'il est à proximité (Fig. 2.3B, D). D'ailleurs, ce fragment de 455 pb contient une séquence d'ADN reconnue par le miRNA *miR-130a* et qui est responsable de réguler négativement l'ARNm du gène *Hoxa5* (Chen *et al.*, 2008). De plus, puisque l'on retrouve plusieurs sites de liaison à la protéine PEA3 au sein du fragment *SacI-XbaI* de 455 pb et que la surexpression de ce facteur est associée au MPOC, notamment en activant directement les gènes *Cox2* et *Mmp9* (Ratovitski, 2010; Xu *et al.*, 2012), la régulation négative de l'expression du gène *Hoxa5* par PEA3 pourrait également participer aux anomalies pulmonaires observées chez les patients souffrant de MPOC. Un autre exemple implique le fragment *BssHII-NcoI* de 588 pb, lequel est toujours en mesure de diriger une certaine expression du transgène *Hoxa5/lacZ* au niveau de l'estomac, plus particulièrement au niveau du pylore, dans 57% des embryons colorés au X-gal (Fig. 2.4B, J). Même si nous ne pouvons exclure la possibilité que l'expression du transgène *Hoxa5/lacZ* observée pour la construction 14 (Fig 2.4B, J) soit dirigée par l'élément MES (Larochelle *et al.*, 1999; Bérubé-Simard et Jeannotte, 2014), des résultats similaires ont été observés pour un fragment *BssHII-MfeI* de 318 pb (Figure 8.1A-D). De plus, un fragment *BsmI-MfeI* de 216 pb est en mesure de lier spécifiquement un extrait protéique d'organes de souris au stade E13.5 par EMSA, comparativement au fragment *BssHII-BsmI* de 102 pb (Figure 8.1E), ce qui suggère la présence d'une séquence régulatrice responsable de raffiner l'expression du gène *Hoxa5* au niveau de l'estomac à l'intérieur du fragment *BsmI-MfeI* de 216 pb. Une caractérisation plus approfondie de ce fragment par transgénèse et par EMSA pourrait être envisagée.

8.1.1 Le facteur de transcription YY1 régule positivement l'expression du gène *Hoxa5* dans les poumons

Plusieurs expériences ont été réalisées afin de démontrer que le facteur de transcription YY1 lie bien les régions d'ADN C3 et G3 *in vivo*, respectivement situées dans les fragments *XbaI-BssHII* de 259 pb et *NcoI-SacI* de 163 pb (Figs 2.5 et 2.6), et de comprendre son rôle dans la régulation transcriptionnelle du gène *Hoxa5* lors du développement des systèmes respiratoire et digestif.

Mes expériences de CHIP ont démontré que la protéine YY1 est bien présente au niveau des deux régions identifiées préalablement par EMSA, à la fois dans des extraits de chromatine de poumons d'embryons au stade E13.5 (Fig. 2.7) et dans ceux d'estomacs (Figure 8.2). Cependant, cette expérience ne permet pas de localiser précisément les sites sur lesquels les protéines se lient puisque l'amplification par PCR s'effectue sur des fragments matrices de chromatine immunoprécipitée d'une longueur de 300-600 pb.

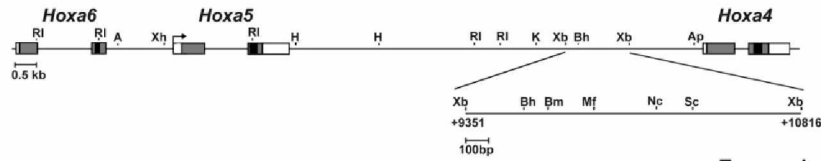
Les questions soulevées par les expériences d'EMSA et de CHIP ont toutefois été répondues suite à la mutation des différents sites de liaison à la protéine YY1, situés à l'intérieur des fragments C3 et G3, dans un contexte de transgénèse. En effet, on observe la perte de l'expression de la β -galactosidase, au niveau des systèmes respiratoire et digestif d'embryons de souris au stade E13.5, lorsqu'on mute les sites de liaison YY1 identifiés préalablement par EMSA et ce, que ce soit dans un contexte impliquant le promoteur du gène *Hoxa5* ou *hsp68* (Fig. 2.8 K-N). Bien que ces résultats suggèrent que le facteur de transcription YY1 régule positivement l'expression du gène *Hoxa5* au niveau du mésenchyme des poumons, de l'estomac et de l'intestin, la possibilité que la mutation des sites de liaison YY1 empêche un autre facteur de transcription d'effectuer son rôle activateur sur l'expression du gène *Hoxa5* ne pouvait pas être exclue tant que nous n'avions pas inactivé le gène *Yy1*. En d'autres termes, se pourrait-il que la protéine YY1 régule négativement l'expression du gène *Hoxa5* en entrant en compétition avec un autre facteur pour la liaison à l'élément poumon/intestin, d'autant plus que ce genre de compétition entre les protéines YY1 et NFY a été décrite pour un élément régulateur responsable de diriger l'expression de *Hoxb4* au niveau du tube neural et du mésoderme prévertébral le long de l'axe AP (Gilthorpe *et al.*, 2002).

La délétion conditionnelle du gène *Yy1* au niveau du mésenchyme nous a permis de valider son rôle de régulateur positif sur l'expression du gène *Hoxa5* au niveau du système respiratoire (Fig. 2.9). En effet, les poumons des souris *Yy1*^{fllox/fllox}; *Dermo1*^{+Cre} sont anormalement compacts au stade E18.5, ce qui ressemble fortement au phénotype des souris *Hoxa5*^{-/-}. D'ailleurs, des gènes tels que *Scgb1a1* et *T1a*, exprimés respectivement au niveau des cellules de Clara et des pneumocytes de type I, sont diminué à la fois dans les poumons des souris *Yy1*^{fllox/fllox}; *Dermo1*^{+Cre} et *Hoxa5*^{-/-}, ce qui renforce le rôle du facteur

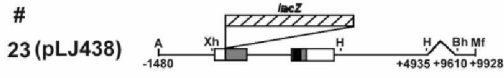
Figure 8.1. Caractérisation du fragment d'ADN *Bss*HIII-*Nco*I de 588-pb chez des embryons transgéniques F0 au stade E13.5

(A) Représentation schématique du fragment d'ADN *Xba*I-*Xba*I de 1.5-kb à l'intérieur de la région génomique *Hoxa4-Hoxa6* (B) Diagramme de la construction *Hoxa5/lacZ* utilisée pour générer les embryons transgéniques F0 au stade E13.5 de même que le sommaire de l'analyse de l'expression du transgène. A, *Acc*I; Ap, *Apa*I; Bh, *Bss*HIII; Bm, *Bsm*I; RI, *Eco*RI; H, *Hind*III; K, *Kpn*I; Mf, *Mfe*I; Nc, *Nco*I; Sc, *Sac*I; Xb, *Xba*I; Xh, *Xho*I. (C) Carcasse représentative des embryons transgéniques E13.5 (D) et leurs organes associés colorés pour leur activité β -galactosidase montrant l'effet de la délétion sur le patron d'expression. i, intestine; l, lung; nt, neural tube; pv, prevertebrae; s, stomach. (E) EMSA avec un extrait cellulaire entier de poumon/estomac/intestin d'embryons de souris au stade E13.5 et une sonde radioactive *Bss*HIII-*Mfe*I de 318-pb où une compétition par un excès (100X) du fragment d'ADN *Bsm*I-*Mfe*I de 216-pb montre que les protéines se lient spécifiquement sur ce fragment (puit 5). Les flèches noirs coresspondent aux liaisons protéiques. COMP, compétiteur; WCE, Whole Cell Extract

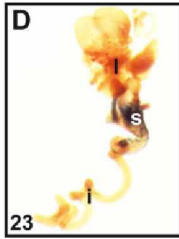
A



B



<i>lacZ</i> ⁺ / <i>#</i> Tg	Expression				
	nt	pv	l	s	i
6/7	6	6	1	4	2



E

Probe	<i>Bss</i> HI- <i>Mfe</i> (318-bp)				
WCE	-	+	+	+	+
Comp.			<i>Bss</i> HI- <i>Mfe</i> (318-bp)	<i>Bss</i> HI- <i>Bsm</i> I (102-bp)	<i>Bsm</i> I- <i>Mfe</i> I (216-bp)

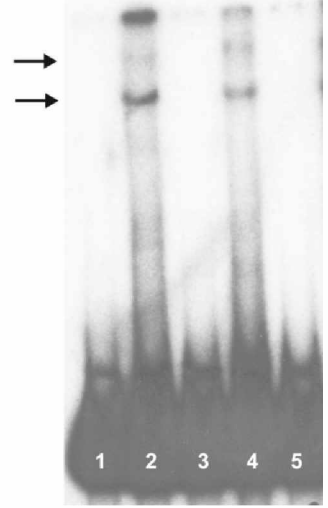
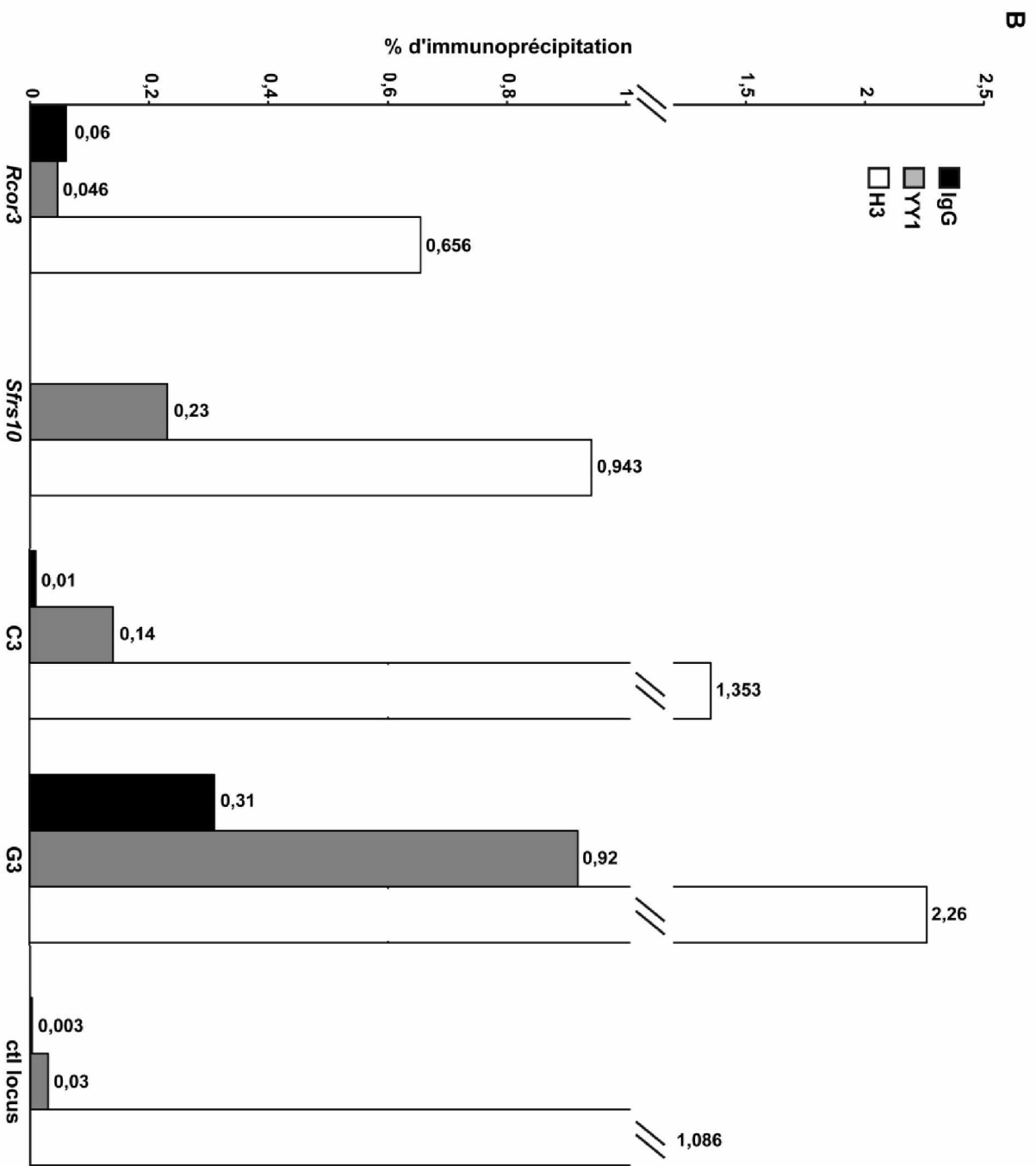
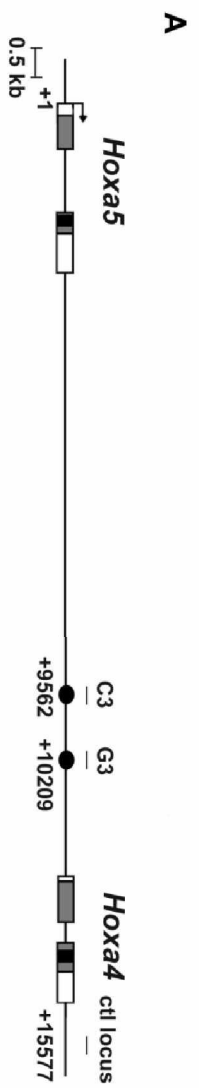


Figure 8.2. Détection de la liaison de la protéine YY1 au niveau des régions C3 et G3 par ChIP chez la souris.

(A) Représentation schématique de la région intergénique *Hoxa4-Hoxa5* avec les positions relatives au site d'initiation de la transcription (TSS) du gène *Hoxa5*. Les positions correspondant aux régions C3 et G3, de même que celle du contrôle locus négatif, sont indiquées. (B) Analyse ChIP sur la région intergénique *Hoxa4-Hoxa5* dans des estomacs d'embryons de souris au jour E13.5. La chromatine a été immunoprécipitée avec des anticorps anti-IgG de lapin (contrôle négatif), anti-YY1 et anti-histone H3 (contrôle positif de l'intégrité de la chromatine). Les différents pourcentages d'immunoprécipitation ont été évalués par qPCR. Ces résultats sont la moyenne \pm SEM de trois expériences indépendantes.



de transcription YY1 en tant que régulateur positif du gène *Hoxa5* au niveau du système respiratoire. Toutefois, aucune souris *Yy1^{flox/flox}*; *Dermo1^{+Cre}* survit à la naissance, comparativement aux souris *Hoxa5^{-/-}* où certaines sont viables, ce qui corrèle avec le fait que la protéine YY1 régule plusieurs gènes dont certains sont impliqués dans le développement pulmonaire, notamment les gènes *Scgb1a1* (Klug et Beato, 1996), *c-Myc* (Riggs *et al.*, 1993), *SM22 α* (Favot *et al.*, 2005) et potentiellement d'autres gènes *Hox* comme le gène *Hoxa4* (Fig. 2.9S). D'ailleurs, en croisant des souris double hétérozygotes *Hoxa5^{+/-}*; *Yy1 Δ ⁺* entre elles, nous avons observé une synergie entre ces deux gènes dans le développement pulmonaire (Figure 8.3A). En effet, alors que les souris simple hétérozygotes *Hoxa5^{+/-}* et *Yy1 Δ ⁺* ne montrent aucun phénotype apparent à la naissance, la perte combinée d'un allèle de chacun de ces gènes entraîne un certain affaissement des poumons accompagné d'un taux de mortalité périnatale d'environ 10%. De la même façon, la perte d'un allèle *Yy1* supplémentaire fait passer le taux de mortalité périnatal des individus *Hoxa5^{-/-}* de 60%, à 100% chez les souris *Hoxa5^{-/-}*; *Yy1 Δ ⁺*. Par conséquent, le gène *Yy1* régulerait directement l'expression du gène *Hoxa5* de façon positive, alors qu'il pourrait aussi agir de façon parallèle sur des gènes en aval du gène *Hoxa5*, comme le gène *Scgb1a1*, de même que sur d'autres gènes importants lors de l'ontogénie du système respiratoire.

De façon surprenante, le niveau d'expression des gènes *Yy1* et *Hoxa5* est relativement élevé à l'intérieur des poumons *Yy1^{flox/flox}*; *Dermo1^{+Cre}* au stade E18.5, soit respectivement $68.5 \pm 6.3\%$ et $77.3 \pm 2.7\%$ (Fig. 2.9R-S), alors que leur niveau d'expression est davantage diminué dans les poumons *Yy1^{flox/flox}*; 14.5-kb *Hoxa5/Cre*, respectivement $51 \pm 10\%$ pour le gène *Yy1* et $55 \pm 3.3\%$ pour le gène *Hoxa5* (Fig. 7.2F). Plusieurs hypothèses non mutuellement exclusives peuvent expliquer ces résultats. Tout d'abord, il est possible que l'allèle "knock-in" *Dermo1^{+Cre}* (*Twist2^{+Cre}*) soit moins efficace que le transgène 14.5-kb *Hoxa5/Cre* pour déléter les allèles *Yy1^{flox}*, donc par le fait même influence moins l'expression de son gène cible (*Hoxa5*). Comme certaines analyses *in silico* prédisent à la fois un lien fonctionnel entre les gènes *Yy1* et *Twist2* (<http://biograph.be/concept/graph/C1427070/C1421565>), de même que la présence de sites de liaison YY1 potentiels en amont du promoteur du gène *Twist2* (TRANSFAC;

chr1:93,627,036-93,632,035), il est envisageable que la délétion mésenchymale du gène *Yy1* affecte négativement l'expression de la *Dermo1/Cre*. D'ailleurs, la présence d'un marquage résiduel pour la protéine YY1 dans le mésenchyme pulmonaire des souris *Yy1^{flox/flox}; Dermo1^{+Cre}* au stade E15.5 pourrait suggérer que la délétion est incomplète (Fig. 2.9P). De la même façon, comme la protéine YY1 régule positivement l'expression du gène *Hoxa5*, la délétion du gène *Yy1* pourrait également interférer avec l'expression du transgène 14.5-kb *Hoxa5/Cre*. L'utilisation d'une autre lignée délétant dans le mésenchyme pulmonaire, comme les lignées *Tbx4^{tm1 (cre) Tmj}* et *Tcf21^{tm1 (cre) Seq}* (Luria *et al.*, 2008; Maezawa *et al.*, 2012), pourrait être envisagée pour confirmer cette hypothèse, puisqu'aucune évidence dans la littérature suggère que les gènes *Tbx4* et *Tcf21* soient régulés par la protéine YY1.

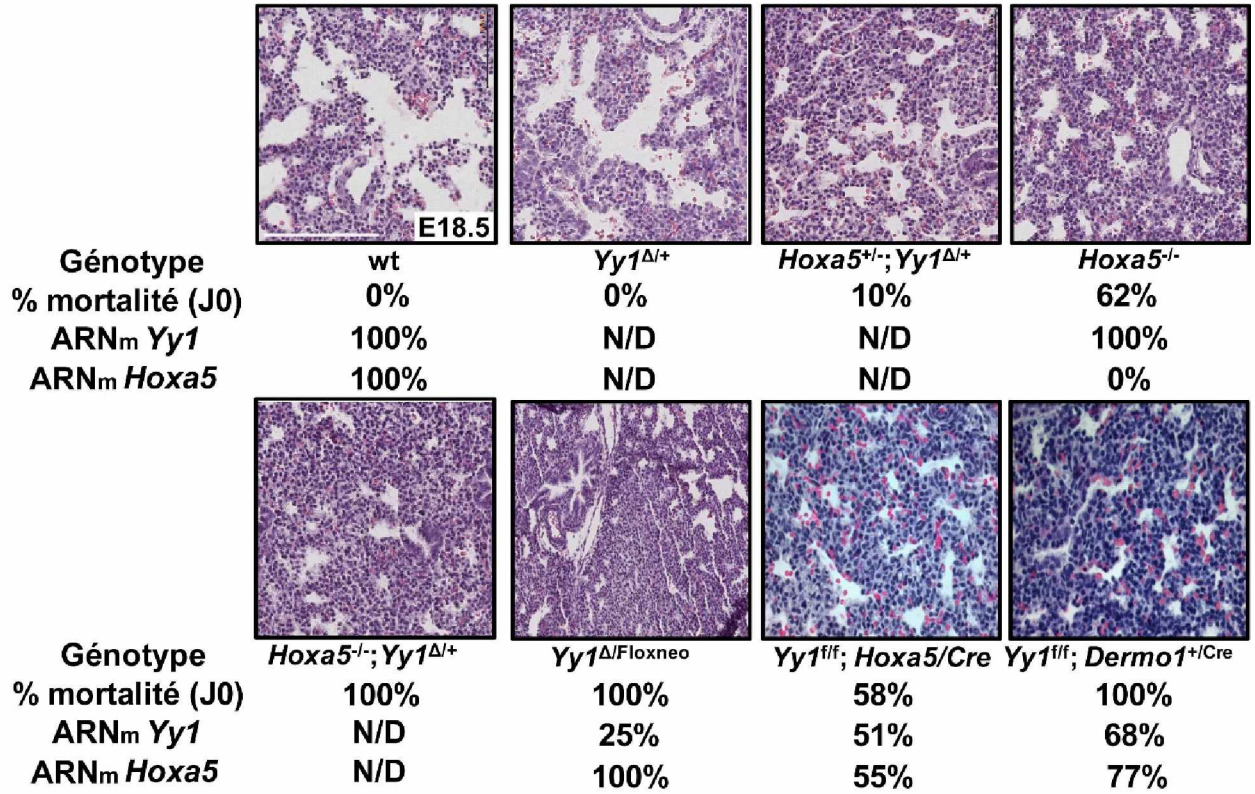
D'un autre côté, les cellules exprimant la protéine YY1 au niveau du mésenchyme pulmonaire des souris *Yy1^{flox/flox}; Dermo1^{+Cre}* pourraient également refléter un problème de transition épithélium-mésenchyme (EMT), puisque les gènes *Yy1* et *Twist2* sont essentiels à cette transition, notamment en réprimant respectivement l'expression de l'occludine et de la E-cadhérine pour ainsi favoriser les cellules mésenchymales au détriment des cellules épithéliales (Peinado *et al.*, 2007; Sade *et al.*, 2009; Trask *et al.*, 2012). De plus, le facteur de transcription YY1 fait également partie du sentier pro-EMT NF- κ B/Snail/YY1/RKIP (Zaravinos *et al.*, 2014). Par conséquent, la délétion mésenchymale complète du gène *Yy1* dans un environnement hétérozygote pour *Twist2* pourrait bloquer les cellules dans un état "épithéliale" et donc débalancer le niveau relatif d'ARNm *Yy1* résiduel dans les poumons *Yy1^{flox/flox}; Dermo1^{+Cre}*. Quant aux souris *Yy1^{flox/flox}; 14.5-kb Hoxa5/Cre*, les cellules mésenchymales pulmonaires qui n'expriment pas le gène *Hoxa5* (donc qui expriment toujours le gène *Yy1* puisque la recombinase Cre n'est pas présente) seraient aptes à gérer la EMT normalement. Une immunohistochimie dirigée contre un marqueur des cellules épithéliales (occludine, E-cadhérine, cytokératine 5/14, claudine 3/18, *etc.*) confirmerait la nature de ces cellules.

Comme le gène *Yy1* est essentiel à la progression du cycle cellulaire et au développement embryonnaire tardif, et ce de manière dépendante à la quantité de protéine YY1 produite (Donohoe *et al.*, 1999; Affar *et al.*, 2006), nous avons validé que les

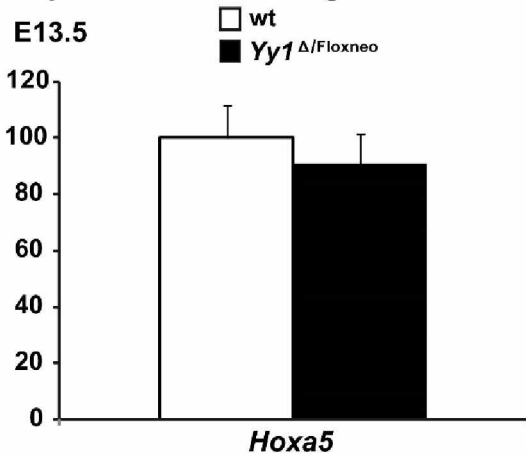
Figure 8.3. Impact de différentes délétions alléliques des gènes *Yy1* et *Hoxa5* sur le poumon durant le développement.

(A) Histologie comparative de poumons sauvages (wt), *Yy1*^{Δ/+}, *Hoxa5*^{+/-}; *Yy1*^{Δ/+}, *Hoxa5*^{-/-}, *Hoxa5*^{-/-}; *Yy1*^{Δ/+}, *Yy1*^{Δ/Floxneo}, *Yy1*^{flox/flox}, *Hoxa5/Cre* et *Yy1*^{flox/flox}; *Dermo1*^{+/-Cre} au jour E18.5. L'expression relative des gènes *Yy1* et *Hoxa5* au niveau de poumons embryonnaires au stade E18.5 a été mesurée par PCR quantitatif. N/D, Non Déterminé. Échelle: 200μm. (B-C) Analyse qRT-PCR mesurant l'expression du gène *Hoxa5* dans les poumons d'embryons *Yy1*^{Δ/Floxneo} relativement aux poumons d'embryons sauvages (B) au stade E13.5 et (C) au stade E18.5. Le niveau d'expression du gène *Hoxa5* n'est pas significativement modifié. Les résultats sont présentés en valeurs moyennes ± SEM.

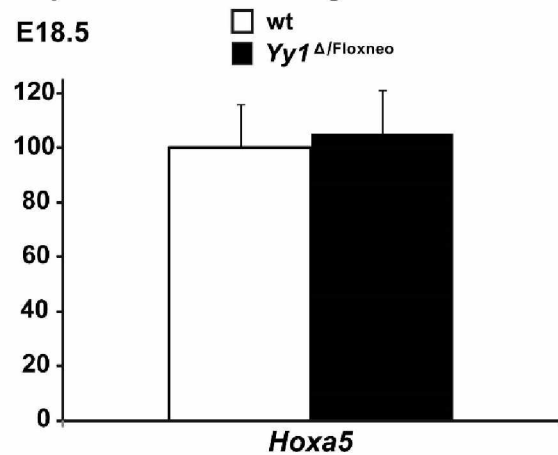
A



B Expression relative du gène *Hoxa5*



C Expression relative du gène *Hoxa5*



phénotypes observés chez les souris $Yy1^{flox/flox}$, $Dermo1^{+/Cre}$ et $Yy1^{flox/flox}$, 14,5-kb $Hoxa5/Cre$ quant à l'affaissement des poumons, étaient causés par la diminution d'expression du gène $Hoxa5$ associée à la perte d'expression mésenchymale du gène $Yy1$ et non à une diminution globale du niveau d'expression de ce gène. En délétant l'expression du gène $Yy1$ au niveau de l'épithélium pulmonaire avec l'aide de la lignée $Shh^{tm2(cre-ERT2)Cjt}$ (Harfe *et al.*, 2004), j'ai confirmé que l'expression du gène $Yy1$ au niveau de l'épithélium joue un rôle différent, mais tout aussi essentiel, que son expression mésenchymale lors du développement pulmonaire, puisque que le phénotype de "sacs pulmonaires vides" observé au stade E18.5 des souris $Yy1^{flox/flox}$, $Shh^{+/Cre}$ diffère complètement de l'affaissement pulmonaire des souris $Yy1^{flox/flox}$, $Dermo1^{+/Cre}$ (Figs 2.9D, 7.2.1F). De plus, aucune variation d'expression du gène $Hoxa5$ n'a été observée au niveau des poumons $Yy1^{flox/flox}$, $Shh^{+/Cre}$ au stade E14.5 (Boucherat *et al.*, non publié), suggérant non seulement que $Yy1$ ne régule pas négativement l'expression du gène $Hoxa5$ au niveau de l'épithélium pulmonaire, mais également que la diminution d'expression du gène $Hoxa5$ observée chez les individus $Yy1^{flox/flox}$, $Dermo1^{+/Cre}$, de même que chez les individus $Yy1^{flox/flox}$, 14.5-kb $Hoxa5/Cre$, est une conséquence directe de l'absence de protéines YY1 dans le mésenchyme pulmonaire plutôt qu'une conséquence indirecte liée au niveau global de YY1. D'ailleurs, ces résultats concordent également avec l'absence de changement du niveau d'expression du gène $Hoxa5$ dans les poumons des souris hypomorphes $Yy1^{\Delta/Floxneo}$ aux stades E13.5 et E18.5, même si celle-ci meurent à la naissance de détresse respiratoire (Affar *et al.*, 2006; Figure 8.3B). L'étude de la fonction du gène $Yy1$ au niveau de l'épithélium pulmonaire est présentement en cours dans le laboratoire.

Le profil d'expression spatio-temporel spécifique à chacune des deux lignées Cre pourrait également expliquer les différentes baisses d'expression relative des gènes $Hoxa5$ et $Hoxa4$ entre les poumons de souris $Yy1^{flox/flox}$, $Dermo1^{+/Cre}$ et $Yy1^{flox/flox}$, 14.5-kb $Hoxa5/Cre$ au stade E18.5. En effet, l'expression du gène $Hoxa4$ est initiée dans le mésoderme embryonnaire au stade E7.5 (Galliot *et al.*, 1989), alors que l'expression du gène $Hoxa5$ débute entre E8.25 et E8.5 au niveau des somites 5-7 (Larochelle *et al.*, 1999). Il est donc possible que la protéine YY1 régule positivement l'expression précoce du gène $Hoxa4$, mais qu'elle ne soit pas impliquée dans le maintien de son expression par la suite. Par conséquent, la délétion du gène $Yy1$ dès le stade E7.5 aurait une incidence sur

l'expression du gène *Hoxa4*, alors qu'une délétion plus tardive, comme celle induite par la 14.5-kb *Hoxa5/Cre*, n'en aurait pas. Afin de vérifier que les souris *Dermo1^{+ / Cre}* expriment la recombinase Cre au niveau du mésoderme embryonnaire au stade E7.5, des croisements avec des souris rapportrices *Rosa26R-lacZ* pourraient être entrepris. De plus, comme le profil d'expression du gène *Hoxa4* au niveau des poumons ne chevauche pas forcément celui du gène *Hoxa5*, les cellules des souris *Yy1^{flox/flox}*; 14.5-kb *Hoxa5/Cre* qui exprimeraient le gène *Hoxa4*, mais non *Hoxa5*, ne seraient pas affectées et leurs contributions pourraient être suffisantes pour maintenir un niveau d'expression du gène *Hoxa4* relativement normal, comparativement aux souris *Yy1^{flox/flox}*; *Dermo1^{+ / Cre}* où toutes les cellules mésenchymales sont affectées. Une co-immunofluorescence dirigée contre les protéines HOXA4 et HOXA5 à différents stades du développement pulmonaire nous renseignerait sur leur profil d'expression respectif.

Quant au niveau relativement élevé d'expression du gène *Hoxa5* dans les poumons d'embryons *Yy1^{flox/flox}*; *Dermo1^{+ / Cre}* et *Yy1^{flox/flox}*; 14.5-kb *Hoxa5/Cre* au stade E18.5, différentes hypothèses non mutuellement exclusives peuvent expliquer ces résultats. Premièrement, le rôle régulateur positif de la protéine YY1 sur l'expression du gène *Hoxa5* a été révélé initialement par la perte d'expression du transgène *Hoxa5/lacZ* au niveau des systèmes respiratoire et digestif suite à la délétion de deux fragments d'ADN (section 8.1), situés dans la séquence intergénique *Hoxa4-Hoxa5*, et à la mutation des sites de liaison YY1 circonscrits à l'intérieur de ces fragments. Comme ce facteur de transcription est plutôt reconnu pour son action négative sur l'expression des gènes *Hox* (Gilthorpe *et al.*, 2002; Kobrossy *et al.*, 2006; Luke *et al.*, 2006; Kobrossy et Featherstone, 2007), il n'est pas impossible que d'autres sites de liaison YY1 soient importants pour réguler négativement l'expression du gène *Hoxa5* et, par le fait même, que le niveau relatif d'expression du gène *Hoxa5* observé par qPCR suite à la délétion mésenchymale du gène *Yy1* reflète cette balance activation/répression. D'ailleurs, les complexes YY1/EED ainsi que YY1/BMI1, deux membres de la famille Polycomb, ont été détectés en amont du promoteur du gène *Hoxa5* dans des extraits d'embryon éviscérés au stade E12.5 (Kim *et al.*, 2006). Deuxièmement, puisque le facteur de transcription YY2 reconnaît le même site de liaison consensus que son homologue YY1 (Yant *et al.*, 1995; Kim *et al.*, 2007), il est suggéré qu'une certaine redondance fonctionnelle puisse exister entre ces deux protéines

(Klar et Bode, 2005), ce qui pourrait expliquer l'expression résiduelle relative du gène *Hoxa5* au niveau des poumons de souris *Yy1^{flox/flox}*; 14.5-kb *Hoxa5/Cre* au stade E18.5. Donc, il serait intéressant de vérifier par EMSA et par ChIP si la protéine YY2 se retrouve liée aux régions C3 et G3, ainsi que de mesurer le niveau relatif d'expression du gène *Hoxa5* dans des individus *Yy1^{flox/flox}*, *Yy2^{flox/flox}*, 14.5-kb *Hoxa5/Cre*, d'autant plus qu'aucune souris mutante pour le gène *Yy2* n'est répertoriée pour le moment. Troisièmement, comme le facteur de transcription YY1 remplit généralement ses fonctions régulatrices avec l'aide de différents cofacteurs, il est envisageable qu'une synergie entre YY1 et d'autres facteurs de transcription soit nécessaire pour que l'expression du gène *Hoxa5* au niveau du système respiratoire soit complète, ce qui expliquerait que la délétion seule du gène *Yy1* au niveau du mésenchyme pulmonaire ne soit pas suffisante pour abolir totalement l'expression du gène *Hoxa5*. Une analyse bioinformatique tirant avantage de la plate-forme PReMod a identifié un fragment de 923 pb, localisé dans la séquence intergénique *Hoxa4-Hoxa5* (entre les positions +9388 et + 10311) et englobant les régions C3 et G3 (section 8.1), qui contient certains enrichissements pour des sites de liaison conservés chez les mammifères (Figure 8.4). En couplant cette analyse avec deux autres outils bioinformatiques (TFSEARCH: Searching Transcription Factor Binding Sites et TESS-Transcription Element Search System), les facteurs de transcription de la famille OCT, de même que AP1/2 s'avèrent des candidats intéressants. En effet, ces différents facteurs de transcription sont fréquemment retrouvés en étroite collaboration avec la protéine YY1 afin de réguler finement certains gènes. Par exemple, des sites de liaison pour les facteurs de transcription OCT1, AP1 et YY1 sont présents au niveau d'un promoteur alternatif fonctionnel présent dans l'intron 1 du gène codant pour la β -Caséine (Kolb, 2003), de même qu'au niveau du promoteur du gène *Scgb1a1* (Sawaya *et al.*, 1993). De plus, les protéines YY1 et AP1 réguleraient positivement le gène suppresseur de tumeur *Hlj1*, codant pour une protéine chaperonne de la famille HSP40/DNAJ, au niveau du sein et du poumon (Wang *et al.*, 2007). Il serait donc intéressant de vérifier si les protéines OCT et AP1/2 sont en mesure de lier spécifiquement les régions C3 et G3, que ce soit par EMSA ou par ChIP. Des techniques comme la chromatographie d'affinité utilisant un anticorps dirigé contre la protéine YY1, ou une purification biotine/streptavidine à l'aide des régions C3 et G3, couplées à la spectrométrie de masse pourraient nous renseigner sur les interacteurs

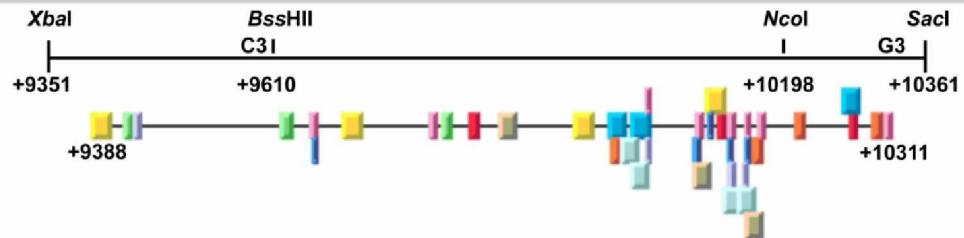
potentiels du facteur de transcription YY1 à différents stades du développement pulmonaire chez la souris. Par la suite, la mutation conditionnelle du gène *Yy1* au niveau du mésenchyme, combinée à celle du (des) cofacteur(s) identifié(s) par spectrométrie de masse, nous permettrait de vérifier si le niveau d'expression du gène *Hoxa5* serait davantage diminué au niveau des poumons d'embryons de souris au stade E18.5 que dans ceux des souris *Yy1^{flox/flox}*; 14.5-kb *Hoxa5/Cre*.

Il est également possible que la protéine YY1 potentialise l'effet d'un autre facteur de transcription sur l'expression du gène *Hoxa5* en modifiant localement la structure de la chromatine et, par le fait même, en rapprochant le facteur de transcription d'intérêt du promoteur *Hoxa5*. Les protéines membres du complexe INO80 en sont de bons exemples. En effet, le complexe de remodelage de la chromatine INO80 se lierait au facteur de transcription YY1 afin de faciliter son accès au promoteur lors de l'activation des gènes d'intérêt (Cai *et al.*, 2007). La capacité des différents membres du complexe INO80 à lier les régions C3 et G3 pourrait être testée par ChIP. Par exemple, les protéines YY1, BAF53A/ARP4 et ARP8 forment le module HSA (helicase-SANT-associated) du complexe INO80 responsable de la fonction hélicase (Chen *et al.*, 2011, 2013). D'ailleurs, j'ai montré, par ChIP, que BAF53A est en mesure de se fixer aux mêmes régions que la protéine YY1 au niveau des séquences en amont du promoteur du gène *Shh* (Figure 7.2.2). Ensuite, puisque Corto et DSP1 se fixent conjointement à un élément de maintenance de l'expression du gène *Src* chez la drosophile (Salvaing *et al.*, 2006), il serait également intéressant de tester par ChIP si des protéines membres de la famille Trithorax sont associées aux régions C3 et G3. En effet, Corto et DSP1 appartiennent à une classe de protéine appelée "activatrice des Polycomb et Trithorax" (Enhancers of trithorax and Polycomb (ETP)) dont YY1 fait également partie. Il est donc possible que YY1 remplisse chez les mammifères le même rôle que DSP1 et Corto chez la drosophile. Le facteur de transcription à doigts de zinc CTCF serait également un candidat potentiel intéressant pour interagir avec YY1, remodeler localement la structure de la chromatine et ainsi rapprocher les régions C3 et G3 du promoteur *Hoxa5*. En effet, plusieurs exemples d'interactions entre les protéines YY1 et CTCF sont recensés dans la littérature. L'exemple le mieux documenté concerne la capacité de ces deux protéines à activer conjointement la transcription du

Figure 8.4 Regroupements de sites potentiels de liaison dans la séquence intergénique *Hoxa4-Hoxa5*

Analyse bioinformatique utilisant la base de données PReMod (Predicted Regulatory Modules; <http://genomequebec.mcgill.ca/PReMod/>) pour identifier des modules enrichis en sites de liaison pour des facteurs de transcription (TRANSFAC 7.2) qui sont conservés entre les génomes de l'humain, de la souris et du rat (projet ENCODE; <https://genome.ucsc.edu/ENCODE/>). Les 10 meilleures prédictions sont affichées à gauche et un résultat leur est attribué en fonction du degré de confiance de cette prédiction. L'analyse de la séquence intergénique *Hoxa4-Hoxa5* fait ressortir un module conservé de 923 pb englobant les fragments d'ADN C3 et G3 identifiés au chapitre II.

Module mod070417 (923bp)



Matrix name	Total Unit Score
ZF5 (M00716)	30
Pax-3 (M00327)	14
Oct-1 (M00138)	13
FOX (M00809)	10
HOXA3 (M00395)	10
E2F (M00803)	19
HNF-4 direct repeat 1 (M00764)	17
cap (M00253)	17
ER (M00191)	17
YY1 (M00069)	16

lncRNA *Tsix* maintenant ainsi l'état actif du chromosome X (Donohoe *et al.*, 2007). Par ailleurs, des expériences d'immunoprécipitation de la chromatine à la grandeur du génome couplées au séquençage à haut débit (ChIP-seq) ont permis d'observer la colocalisation des protéines YY1 et CTCF à plusieurs endroits transcriptionnellement actifs sur le génome de primates et de l'humain (Dunham *et al.*, 2012; Guo *et al.*, 2012; Schwalie *et al.*, 2013). De plus, une équipe a récemment réalisé un modèle intégratif de l'architecture de la chromatine du complexe *HoxA* selon son état transcriptionnel en fonction du temps, et ce en tirant avantage de la technique 5C (Chromosome Conformation Capture Carbon Copy), de la modélisation informatique et d'analyses épigénomiques (Rousseau *et al.*, 2013). Dans un premier temps, cette étude a permis de visualiser la dynamique chromatinienne du complexe *HoxA* qui se relaxe globalement lors du début de la différenciation dans un modèle cellulaire de leucémie myéloïde humaine, et où l'architecture des gènes situés en 3' et celle des gènes en 5' du complexe *HoxA* varie indépendamment comme deux domaines topologiques distincts. Dans un deuxième temps, ils ont observé que la région centrale du complexe *HoxA* séparant le domaine antérieur du domaine postérieur, composée des gènes *Hoxa4-Hoxa9*, correspond à la région du complexe *HoxA* où l'on retrouverait à la fois le plus de sites de liaison pour la protéine CTCF et le plus de remodelage de la chromatine durant la différenciation. Par conséquent, on pourrait penser que les facteurs de transcription CTCF et YY1 agissent conjointement pour permettre la transition entre la vague précoce d'activation engendrée par le RA sur les gènes situés en 3' et la vague tardive activant l'expression des gènes situés en 5' du complexe *HoxA*. Pour vérifier cette hypothèse, nous pourrions effectuer des expériences de ChIP-reChIP, où la chromatine immunoprécipitée avec un anticorps dirigé contre la protéine X est soumise à une deuxième ronde d'immunoprécipitation en utilisant un anticorps dirigé contre la protéine Y et soupçonnée d'interagir avec la protéine X. De cette façon, il serait possible de vérifier si les protéines YY1 et CTCF lient conjointement les régions C3 et G3 à différents stades du développement pulmonaire chez la souris. Parallèlement, la présence de marques associées à l'état relaxé de la chromatine (H3K9ac, H3K4me3, H3K27me, H3K36me3, *etc.*) sur les régions C3 et G3 pourrait être testée par ChIP. Il serait également intéressant de tester par technique 5C si la dynamique des changements architecturaux de la chromatine au niveau de la région centrale du complexe *HoxA* est modifiée chez nos mutants *Yy1^{flox/flox}*,

Dermo1^{+Cre} et *Yy1*^{flox/flox}; 14.5-kb *Hoxa5/Cre*. Enfin, pour valider que la protéine YY1 rapproche le promoteur *Hoxa5* des régions C3 et G3 à différents stades du développement pulmonaire chez la souris, des essais de ChIP-loop pourraient être entrepris. Cet essai combine les techniques de 3C et de ChIP, où la chromatine, sur laquelle les protéines ont été préalablement fixées, est d'abord digérée avec une enzyme coupant de part et d'autre des deux régions d'ADN susceptibles d'interagir. Par la suite, la chromatine récoltée par immunoprécipitation est soumise à une ligation et une amplification par qPCR, laquelle n'est possible que si les deux régions normalement éloignées ont été rapprochées par une interaction ADN-protéine(s)-ADN.

En résumé, j'ai démontré à l'intérieur du Chapitre II de ma thèse i) que l'expression du gène *Hoxa5* au niveau des poumons, de l'estomac et de l'intestin est contrôlée par plusieurs séquences régulatrices, ii) que le site RARE situé dans le fragment *XbaI-XbaI* et qui est nécessaire à l'expression du gène *Hoxa4* dans les organes (Packer *et al.*, 1998), est plutôt essentiel au positionnement de la frontière antérieure d'expression du gène *Hoxa5* au niveau du SNC, et iii) que le facteur de transcription YY1 joue un rôle important comme régulateur positif de l'expression du gène *Hoxa5* dans les systèmes respiratoire et digestif. Ces données font l'objet d'un manuscrit publié dans la revue *PLOS ONE* en avril 2014.

8.2 Les lignées *Hoxa5/Cre* permettent d'étudier la fonction de gènes dans certains tissus le long de l'axe AP

Les souris mutantes pour le gène *Hoxa5* présentent une panoplie de phénotypes affectant le squelette axial, le squelette appendiculaire, de même que le développement de plusieurs organes, dont le système respiratoire (Figure 1.13). D'ailleurs, un taux élevé de souris *Hoxa5*^{-/-} meurent à la naissance de détresse respiratoire (Aubin *et al.*, 1997), alors que celles qui survivent montrent des anomalies affectant l'alvéogénèse pulmonaire (Mandeville *et al.*, 2006; Boucherat *et al.*, 2013). Afin de fournir de nouveaux outils pour étudier les cascades moléculaires impliquées dans la formation de ces différentes structures, j'ai généré deux lignées de souris dirigeant l'expression de la recombinaise Cre dans les domaines d'expression du gène *Hoxa5*. Le transgène 7.1-kb *Hoxa5/Cre* contient

les éléments temporel, BSC et MES, permettant une délétion au niveau du mésoderme paraxial et du LPM de la région cervico-thoracique. Quant au transgène 14.5-kb *Hoxa5/Cre*, il contient une séquence additionnelle de 7.4 kb englobant l'élément poumon/intestin. Ces deux lignées constituent de nouveaux outils pour l'analyse fonctionnelle de différents gènes à l'intérieur de tissus définis le long de l'axe antéro-postérieur. D'une part, elles diffèrent de la lignée de souris *Tg(Hoxa5-cre)1Fmr*, MGI:5474192 (Di Meglio *et al.*, 2013), où la séquence complète du complexe *HoxA* peut agir sur l'expression de la recombinase Cre insérée dans le locus *Hoxa5*. D'autre part, contrairement à la lignée *Dermo1^{+Cre}*, générée par ciblage de la recombinase Cre dans le locus *Twist2* (*Twist2^{tm1(cre)Dor}*; MGI:3044412) (Yu *et al.*, 2003) et qui permet également une délétion spécifique au niveau du mésenchyme, l'approche de transgénèse classique utilisée pour générer les souris transgéniques *Hoxa5/Cre* nous affranchi d'un contexte hétérozygote pour le gène *Twist2* pouvant influencer le phénotype observé suite à la délétion du gène d'intérêt.

L'efficacité et la spécificité de la recombinaison induite par les deux lignées de souris transgéniques *Hoxa5/Cre* ont été évaluées initialement en les croisant avec des souris rapportrices *Rosa26/lacZ* (*R26R*) (Soriano, 1999). En plus de confirmer que les deux transgènes *Hoxa5/Cre* sont exprimés spécifiquement dans les domaines d'expression du gène *Hoxa5* en fonction des séquences régulatrices en présence (Fig. 3.1), les croisements avec les souris *R26R* ont également fait ressortir d'autres informations très intéressantes. Tout d'abord, les données recueillies précédemment dans le laboratoire par transgénèse suggéraient que l'élément MES était en mesure de diriger l'expression du gène *Hoxa5* au niveau du système digestif en développement (Larochelle *et al.*, 1999). Nos prélèvements d'estomac au jour (J) 30 après la naissance de souris 7.1-kb *Hoxa5/cre; R26R* ont démontré une expression rostro-caudale croissante du transgène rappelant l'expression endogène du gène *Hoxa5*, suggérant que les séquences régulatrices appartenant à l'élément MES seraient responsables du gradient d'expression au niveau de l'estomac. Une approche par délétions successives dans un contexte de transgénèse *Hoxa5/lacZ* couplée à des EMSA pourrait être entreprise pour valider ce résultat et circonscrire davantage cette séquence. Nous avons également observé certaines cellules épithéliales positives au niveau de l'intestin, de la glande mammaire et de l'ovaire où les transgènes *Hoxa5/Cre* semblaient

être exprimés, alors qu'aucune expression du gène *Hoxa5* n'a jamais été rapportée dans ces types cellulaires pour ces tissus (Aubin *et al.*, 1999; Garin *et al.*, 2006; Gendronneau *et al.*, 2012). Trois hypothèses peuvent expliquer ce résultat. Premièrement, il peut s'agir d'une expression ectopique liée au site d'intégration du transgène. Deuxièmement, des séquences inhibitrices additionnelles qui ne font pas partie de nos deux transgènes sont nécessaires pour reproduire l'expression endogène du gène *Hoxa5*. Troisièmement, puisque la réaction entre le X-gal et la β -galactosidase génère un produit insoluble utilisé pour le lignage cellulaire, il est possible que les cellules épithéliales colorées soient en fait d'anciennes cellules mésenchymales ayant subi une transition mésenchyme-épithélium (MET). Par conséquent, un co-marquage avec un anticorps dirigé contre la recombinaison Cre pourrait valider que cette protéine n'est plus produite au sein de ces cellules.

Ces deux lignées de souris s'avèrent être des outils puissants pour l'étude fonctionnelle des différents gènes exprimés dans les domaines du gène *Hoxa5*. Par exemple, j'ai démontré que la protéine HOXA5 est en mesure de lier certaines séquences en 5' du gène *Fstll1* et ainsi potentiellement réguler son expression au niveau de la trachée (Fig. 6.4). Il pourrait être intéressant de déléter conditionnellement le gène *Fstll1* avec la lignée 14.5-kb *Hoxa5/Cre* et de comparer le phénotype au niveau du système respiratoire à celui des souris *Hoxa5*^{-/-}.

En plus des croisements avec les souris *R26R*, l'efficacité et la spécificité de la recombinaison induite par les deux lignées de souris transgéniques *Hoxa5/Cre* ont également été vérifiées en déléant le gène *Hoxa5* et en comparant les phénotypes obtenus avec ceux observés chez les souris *Hoxa5*^{-/-}. Dans un premier temps, le pourcentage de mortalité périnatale recensé chez les individus *Hoxa5*^{flx/flx};14.5-kb *Hoxa5/Cre* est comparable à celui décrit pour les individus *Hoxa5*^{-/-} (Fig. 2.2; Jeannotte *et al.*, 1993), alors qu'aucune mortalité n'est observée chez les *Hoxa5*^{flx/flx};7.1-kb *Hoxa5/Cre*, suggérant que l'expression du gène *Hoxa5* dirigée par l'élément poumon/intestin est nécessaire à la survie. Des données récentes dans le laboratoire montrent que la mortalité périnatale associée aux souris *Hoxa5*^{-/-} serait davantage associée à un défaut d'innervation du diaphragme par les neurones moteurs du nerf phrénique qu'à la dimorphogénèse des voies respiratoires à proprement dit (Chapitre VI ; Philippidou *et al.*, 2012; Landry-Truchon et Jeannotte,

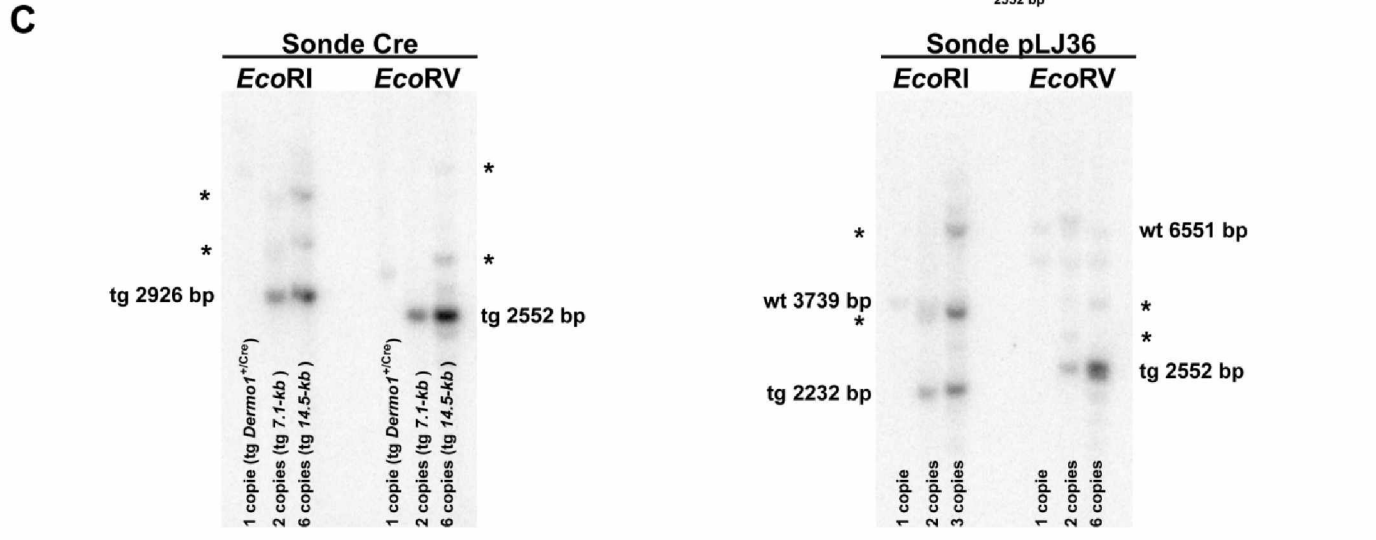
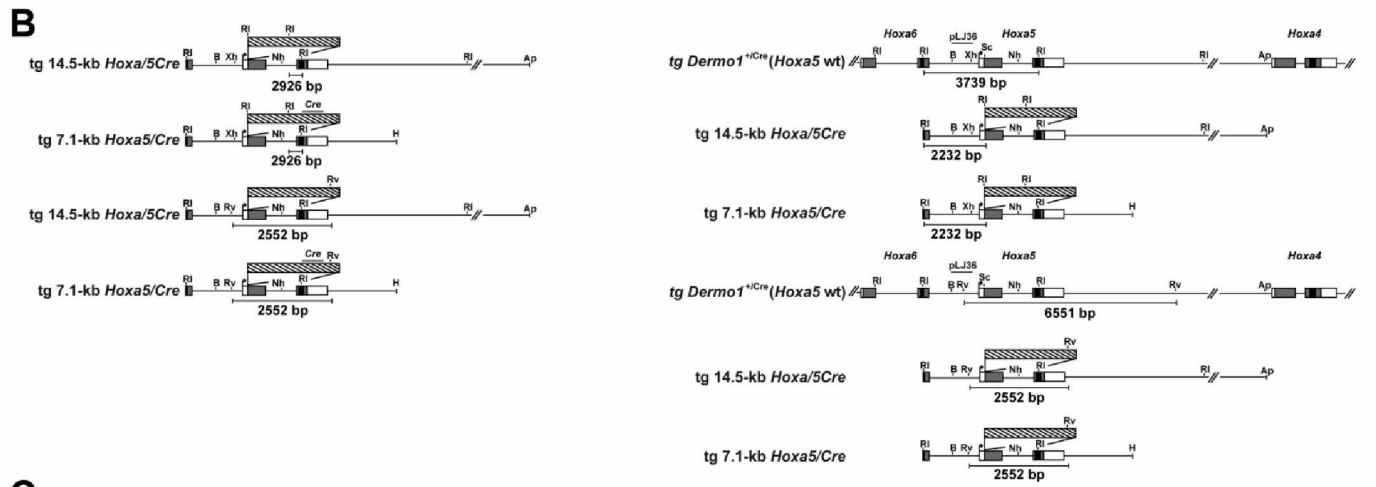
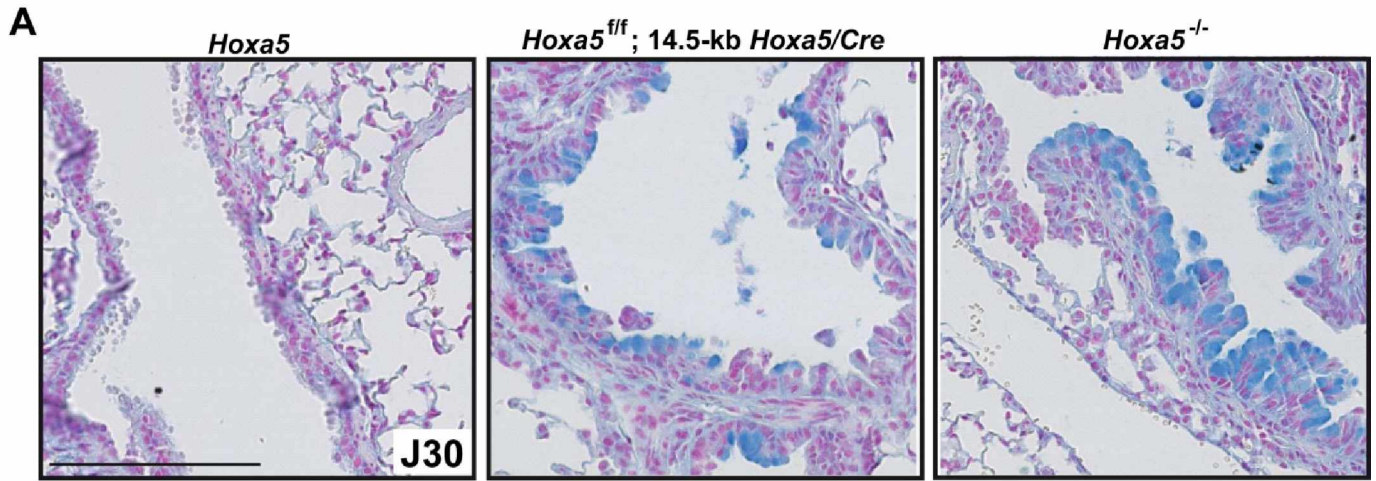
résultats non publiés). Comme le site RARE présent dans l'élément poumon/intestin semble important pour le positionnement de la limite antérieure d'expression du gène *Hoxa5* au niveau du tube neural, il est possible que ce soit l'absence d'expression au niveau de cette structure qui cause la mort à la naissance chez les individus *Hoxa5^{flox/flox}*, 14.5-kb *Hoxa5/Cre*, plutôt que la perte d'expression au niveau des poumons. La délétion conditionnelle du site RARE pourrait répondre à cette question. L'étude du gène *Hoxa5* au niveau du nerf phrénique innervant le diaphragme est présentement en cours dans le laboratoire à l'aide de la lignée de souris *Olig2^{tm1(cre)Tmj}* (Desseaud *et al.*, 2007). Ensuite, l'histologie des poumons prélevés chez les souris *Hoxa5^{flox/flox}*, 14.5-kb *Hoxa5/Cre* au jour embryonnaire E18.5, de même que ceux prélevés à J30, révèle un phénotype similaire aux poumons *Hoxa5^{-/-}* (Aubin *et al.*, 1997; Mandeville *et al.*, 2006). Toutefois, la délétion conditionnelle du gène *Hoxa5* à l'aide du transgène 14.5-kb *Hoxa5/Cre* ne récapitule pas complètement l'hyperplasie des cellules à mucus observée chez les souris *Hoxa5^{-/-}* adultes (Fig. 6.6; Mandeville *et al.*, 2006; Figure 8.5A). Ce résultat s'explique probablement par le fait que la délétion du gène *Hoxa5* est incomplète. En effet, le niveau résiduel d'ARNm *Hoxa5* mesuré par qPCR est de $11.4 \pm 3.1\%$, comparativement à un niveau qui tend vers 0 (1-3 %) chez les individus *Hoxa5^{-/-}*. Une des hypothèses pouvant expliquer ce problème d'efficacité rencontré avec les transgènes *Hoxa5/Cre* pour déléter le gène *Hoxa5* vient du potentiel mécanisme d'autorégulation contrôlant l'activité transcriptionnelle du gène *Hoxa5*. Effectivement, comme un site de liaison putatif pour la protéine HOXA5 a été identifié au niveau de son promoteur (Odenwald *et al.*, 1989; Zhao *et al.*, 1996) et que l'expression du transcrit de 1.8kb *Hoxa5* semble être affectée dans certaines circonstances par la mutation *Hoxa5* elle-même (Aubin *et al.*, 1998), il est possible que la protéine HOXA5 régule sa propre expression. Par conséquent, l'efficacité des transgènes *Hoxa5/Cre* peut diminuer parallèlement à la quantité de la protéine HOXA5 jusqu'à l'atteinte du seuil minimal nécessaire à son activité. Des expériences d'EMSA et de ChIP devraient être envisagées pour vérifier cette hypothèse, de même que l'utilisation d'une autre lignée porteuse d'une recombinaison Cre spécifique au mésenchyme pulmonaire comme les souris *Tbx4Cre* par exemple.

L'analyse des phénotypes squelettiques des souris ayant subi une délétion conditionnelle du gène *Hoxa5* induite par l'un ou l'autre des transgènes *Hoxa5/Cre* nous informe également sur l'efficacité et la spécificité de la recombinaison. Les transformations homéotiques induites par ces lignées de souris transgéniques ont toutes été rapportées avec une pénétrance comparable à celle observée chez les mutants *Hoxa5*^{-/-}, confirmant la spécificité de recombinaison des transgènes *Hoxa5/Cre* aux domaines d'expression endogène du gène *Hoxa5*. Cependant, certains phénotypes squelettiques observés initialement chez les souris *Hoxa5*^{-/-} n'ont pas été répertoriés dans nos souris *Hoxa5*^{flox/flox}, 7.1-kb *Hoxa5/Cre* et *Hoxa5*^{flox/flox}; 14.5-kb *Hoxa5/Cre* (absence/raccourcissement de l'acromion, nombre et apparence anormales de vertèbres sternales, présences de côtes en L1, *etc.*). Plusieurs facteurs non mutuellement exclusifs peuvent être envisagés. L'explication la plus simple serait que d'autres séquences importantes pour l'expression du gène *Hoxa5* dans ces structures ne sont pas présentes dans nos deux transgènes. Ensuite, même si divers environnements génétiques peuvent également expliquer en partie les différences de pénétrance des phénotypes liés à la mutation du gène *Hoxa5* (Jeannotte *et al.*, 1993; Aubin *et al.*, 1997, 1998), l'hypothèse d'une délétion incomplète demeure plus probable. D'ailleurs, le fait que la présence des côtes ectopiques sur C7 n'ait été observée que chez les individus *Hoxa5*^{flox/flox}; 14.5-kb *Hoxa5/Cre* pourrait s'expliquer par le fait que trois fois plus de copies du transgène sont intégrées dans le génome des souris 14.5-kb *Hoxa5/Cre* comparativement aux souris 7.1-kb *Hoxa5/Cre* (~6 copies vs ~2 copies, Figure 8.5B). En visualisant le niveau et les sites dans lesquels la protéine HOXA5 est présente, par immunofluorescence sur des coupes d'embryons à différents stades du développement, on pourrait tester cette hypothèse. L'insertion du transgène en elle-même pourrait également influencer l'efficacité de délétion. De plus, nous ne pouvons pas exclure que les fragments non attendus révélés par analyse de type Southern (Fig 2.1; Fig. 7.1) soient des insertions incomplètes du transgène limitant ainsi leur efficacité respective, bien que notre analyse exhaustive des différents transgènes nous suggère qu'il s'agirait plutôt de digestion incomplète.

En résumé, j'ai montré à l'intérieur du Chapitre III de ma thèse i) que l'expression des transgènes *Hoxa5/Cre* récapitule en partie le profil d'expression du gène *Hoxa5*, et ii) que la délétion conditionnelle du gène *Hoxa5* en utilisant ces transgènes reproduit les

Figure 8.5. Impact de la délétion conditionnelle du gène *Hoxa5* sur le poumon suite à l'action de la 14.5-kb *Hoxa5/Cre*

(A) Histologie comparative de poumons de souris *Hoxa5^{f/f}*, *Hoxa5^{f/f}*, 14.5-kb *Hoxa5/Cre* et *Hoxa5^{-/-}* au 30^{ième} jour après la naissance (j30) où les cellules de goblet sont colorées aux bleu d'alcian. Échelle: 200µm. (B) Cartes de restriction partielles des différents transgènes Cre. Les boîtes noires correspondent aux homéoboîtes, les grises, aux régions traduites et les blanches, aux régions transcrites. La flèche représente le site d'initiation de la transcription (TSS) du gène *Hoxa5*. Les fragments d'ADN, sur lesquels les sondes utilisées pour les analyses de type Southern blot s'hybrident (pLJ36 et Cre), sont représentés respectivement sur la carte de l'allèle *Hoxa5* sauvage (*Dermo1^{+Cre}*) et 7.1-kb *Hoxa5/Cre*. Les fragments d'ADN attendus suite aux digestions *EcoRI* et *EcoRV* sont indiqués sous les cartes des différents transgènes. Ap, *ApaI*; H, *HindIII*; Nh, *NheI*; RI, *EcoRI*; RV, *EcoRV*; Sc, *SacI*; Xh, *XhoI* (C) Trois échantillons d'ADN génomique différents extraits de la queue de souris transgénique *Dermo1^{+Cre}*, 7.1-kb *Hoxa5/Cre* et 14.5-kb *Hoxa5/Cre* ont été utilisés. L'analyse de type Southern blot combinée au calcul effectué par le programme informatique imageJ permettent de déterminer le nombre d'insertion des différents transgènes. Les astérisques représentent des fragments d'ADN non attendus (digestion incomplète, intégration incomplète et intégration en concatémère du transgène).



phénotype de défauts squelettiques, de mortalité périnatale et de malformations pulmonaires associés aux souris *Hoxa5*^{-/-}. Ces travaux ont été publiés dans la revue *genesis* (Bérubé-Simard *et al.*, 2014)

8.3 L'élément MES contient plusieurs séquences régulatrices importantes

Parallèlement aux informations acquises avec la lignée *7.1-kb Hoxa5/Cre*, l'étude de transgénèse menée dans le laboratoire (Fig. 4.1; Larochelle *et al.*, 1999; Tabariès *et al.*, 2005) a révélé que l'élément MES, localisé à l'intérieur d'un fragment *HindIII-HindIII* de 2084 pb entre les positions +2.9 kb et +4.9kb, dirige spécifiquement l'expression du gène *Hoxa5* dans plusieurs structures, notamment au niveau des dérivés mésodermiques de la région cervico-thoracique (entre pv3 et pv10), des bourgeons des membres, du système digestif et du système urogénital. Ces observations rappellent d'ailleurs que la régulation transcriptionnelle du gène *Hoxa5* est contrôlée par l'action combinée de plusieurs mécanismes régulateurs complexes. Une caractérisation approfondie de l'élément MES, réalisée conjointement par Sébastien Tabariès et moi a permis d'identifier deux séquences régulatrices responsables de l'établissement de la limite postérieure d'expression du gène *Hoxa5* au niveau de pv 10, dont un fragment *PstI-HindIII* de 331 pb (PsHi 331-bp), situé à l'extrémité 3'du MES (Fig. 5.1).

La dissection du fragment PsHi 331-bp à l'aide d'une approche de retard sur gel (EMSA) a démontré que des protéines extraites d'embryons au jour E9.5 sont en mesure de lier ce fragment via un motif de 16 pb (Oligo 16-bp) (Fig. 5.2). Cependant, le(s) facteur(s) de transcription liant cette séquence demeure(nt) inconnu(s). Le site de liaison potentiel pour les protéines CDX identifié au sein du motif de 16 pb par analyse bioinformatique ne s'est pas avéré important dans le cadre de nos EMSAs. De la même façon, aucune liaison n'a pu être observée entre ce motif et différentes protéines à domaines homéo produites *in vitro*, telles que HOXA5, HOXD9, HOXD11, PBX1A et PBX1B (Figure 8.6A-B). Deuxièmement, une approche combinant l'utilisation d'un appât moléculaire (ADN biotinylé) (Gabrielsen et Huet, 1993) à la spectrométrie de masse, ne

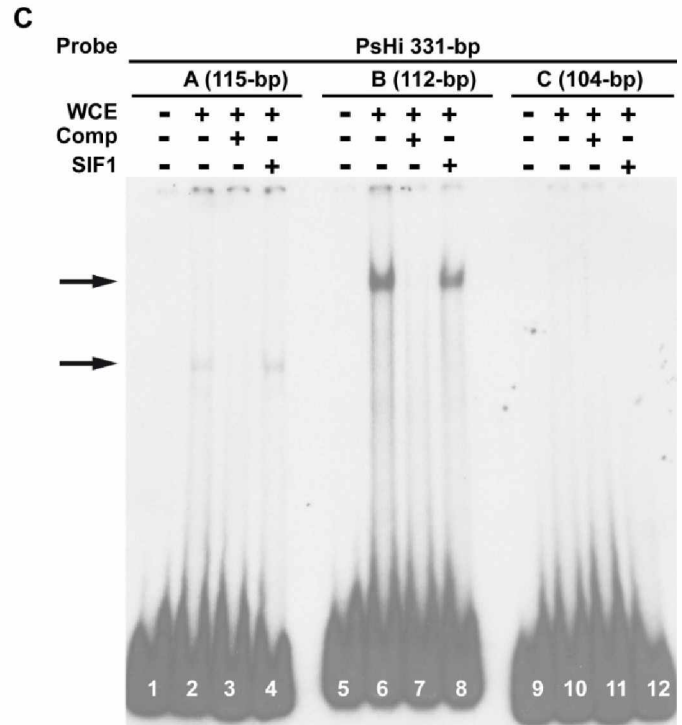
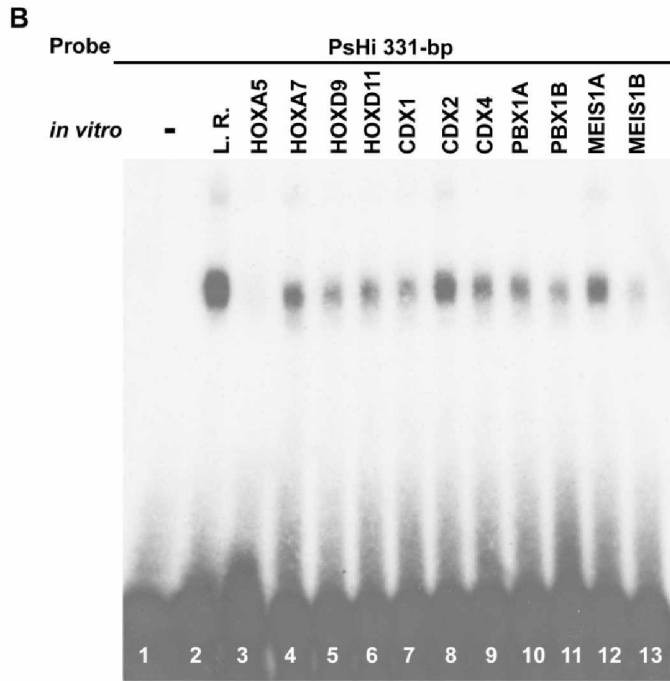
nous a pas permis d'identifier le facteur de transcription d'intérêt. Pour pallier aux problèmes techniques liés à la purification protéique biotine-streptavidine auxquels nous avons faits face, une alternative inspirée de la technique d'EMSA avec un fragment d'ADN non marqué pourrait être envisagée. Quant aux autres sites de liaison potentiels prédits par analyse TRANSFAC, ils pourraient être testés par EMSA pour la liaison de leur protéine cible produite *in vitro*.

Bien que la délétion du motif de 16 pb dans un contexte de transgène semble affecter la limite postérieure d'expression du gène *Hoxa5*, nous n'avons pas été en mesure de démontrer son action synergique avec le fragment AvEc 164-bp, contrairement à la délétion complète du fragment PsHi 331-bp, suggérant que d'autres séquences d'ADN sont importantes pour restreindre l'expression du gène *Hoxa5* entre pv3 et pv10. Comme une liaison entre le fragment PsHiA 115-bp et un extrait protéique d'embryon au jour E9.5 (Figure 8.6C) a été observée de façon sporadique, dépendamment de la préparation protéique, ce fragment pourrait être divisé en sous-fragments pour des analyses EMSA.

En ce qui concerne l'autre séquence régulatrice circonscrite dans l'élément MES et participant à l'établissement de la limite postérieure d'expression du gène *Hoxa5* au niveau de pv10, il s'agit du fragment *AvrII-Eco47III* de 164 pb (AvEc 164-bp) (Tabariès *et al.*, 2005). L'action répressive qu'exerce ce fragment sur l'expression postérieure du gène *Hoxa5* passe, du moins en partie, par la liaison des protéines CDX à deux sites putatifs de liaison situés dans un élément de 25 pb (Oligo1). De plus, CDX4 semble être la protéine membre de la famille CDX ayant le plus d'affinité pour les sites présents dans l'Oligo1 (Fig. 5.4B; Tabariès *et al.*, 2005). J'ai démontré que, bien que plusieurs protéines HOX de la famille AbdB produites *in vitro* peuvent lier le motif de 25 pb, HOXB9 est la seule protéine testée qui se fixe à l'Oligo1 *in vivo* par CHIP (Fig. 5.4). Ce résultat suggère donc que la protéine HOXB9 joue un rôle dans la répression de l'expression du gène *Hoxa5* dans la partie postérieure à pv10 le long de l'axe AP. D'ailleurs, cette observation concorde avec les concepts de "code *Hox*" (Kessel et Gruss, 1991; Duboule, 1998) et de "prévalence postérieure" (Duboule, 1991; Kmita et Duboule, 2003), voulant que la spécification des

Figure 8.6. Caractérisation du fragment d'ADN *PstI-HindIII* de 331-pb par EMSA

(A) Représentation schématique du fragment d'ADN *XboI-HindIII* de 892-pb situé entre les positions +4.04-kb et +4.93-kb à l'intérieur de la partie 3' de l'élément MES. La boîte représente la location de l'élément régulateur *PstI-HindIII* de 331-pb identifié par une approche de transgènes. La séquence du fragment PsHi 331-bp est indiquée. Les sous-fragments d'ADN PsHiA, PsHiB et PsHiC sont soulignés, de même que les sub-divisions Oligo 34-bp, 33-bp et 16-bp. Les nucléotides encadrés correspondent aux sites de liaison pour les protéines CDX (B) La capacité du fragment PsHi 331-pb à lier différentes protéines à boîte homéo produites *in vitro*. Il y a absence de liaison spécifique entre la sonde et toutes les protéines testées (C) La capacité des fragments d'ADN PsHiA, PsHiB et PsHiC à lier l'extrait entier d'embryons E9.5 a été testée par EMSA (puits 2, 6 et 10). La liaison spécifique est abolie par compétition avec un excès de sonde froide (100X) (puits 3, 9 et 11), mais pas par l'Oligo SIF1 contenant des sites de liaison aux protéines CDX (puits 4, 8 et 12). Les flèches indiquent les bandes correspondant aux facteurs de transcriptions non-identifiés présents dans l'extrait protéique. Comp, compétiteur; L.R., Lysat de Réticulocytes; SIF1, Sucrase-Isomaltase Factor 1; WCE, Extrait Cellulaire Entier d'Embryons E9.5.



différentes structures embryonnaires origine de coordonnées génétiques uniques, caractérisée par la combinaison spécifique de différentes protéines HOX, et que le produit protéique des gènes *Hox* postérieurs régule négativement l'expression des plus antérieurs. Enfin, mes expériences d'EMSA et de CoIP montrent que les protéines CDX4 et HOXB9 ne semblent pas interagir physiquement, qu'elles soient liées ou non au sites de liaison présents sur l'Oligo1 (Fig. 5.5).

L'expression du transgène *Hoxa5/lacZ* au niveau de l'axe AP entre les stades E8.5 et E12.5 n'est pas perturbée chez les souris *Cdx1^{-/-}*, *Cdx1^{-/-}; Cdx2^{+/-}* et *Cdx4^{-/-}*, suggérant que les protéines CDX n'auraient potentiellement pas un rôle important dans la répression du gène *Hoxa5* à ces stades (Tabariès *et al.*, 2005; Deschamps non-publié). Ces observations corrélerent avec l'effet sur la perte de la frontière postérieure au niveau de pv10 observé par transgénèse qui est plus important en retirant complètement le fragment AvEc 164-pb (ou l'Oligo1) qu'en mutant spécifiquement les deux sites de liaison aux CDX. Par conséquent, l'effet observé en transgénèse serait davantage attribuable à l'absence de liaison de la protéine HOXB9 qu'à celle des CDX. Toutefois, l'étude entre les différents mutants *Cdx* et l'expression du transgène *Hoxa5/lacZ* devrait être entreprise plus tôt, soit entre E7.5 et E8.5, pour vérifier si ces protéines empêchent l'expression du gène *Hoxa5* avant que les protéines AbdB soient exprimées. D'ailleurs, des expériences de ChIP sur des extraits de chromatine à ces différents stades devraient être envisagées afin de vérifier si l'on retrouve des protéines CDX liées au fragment AvEc 164-bp avant (ou suivant) l'activation du gène *Hoxa5*. Ensuite, afin de s'affranchir de la mortalité embryonnaire précoce associée à certains mutants *Cdx*, nous pourrions tirer avantage de la lignée de souris transgénique 7.1-kb *Hoxa5/Cre* et ainsi tester l'effet de la délétion conditionnelle de la famille *Cdx* au complet sur l'expression du transgène *Hoxa5/lacZ*. De cette manière, nous pourrions éviter la redondance fonctionnelle potentielle entre les différents membres CDX. Il serait également important de tester l'expression du transgène *Hoxa5/lacZ* dans un contexte *Hoxb9^{-/-}*, afin de valider son rôle dans l'établissement de la frontière d'expression du gène *Hoxa5* au niveau de pv 10. D'ailleurs, il serait intéressant de vérifier si certains membres de la famille PBX sont en mesure de lier le site de liaison qui chevauche l'un des deux sites CDX présent au niveau de l'Oligo1. Dans le cas échéant, une expérience de CoIP

permettrait de confirmer si la liaison de la protéine HOXB9 au fragment AvEc 164-bp est dépendante de l'interaction HOXB9-PBX.

Bien que certaines expériences doivent être faites pour confirmer notre hypothèse, nous proposons à la lumière de nos résultats un modèle où les protéines CDX (probablement CDX4) réguleraient négativement l'expression du gène *Hoxa5* entre les stades E8.0 et E8.5. Au jour E8.5, les protéines CDX activeraient l'expression du gène *Hoxb9* (van den Akker *et al.*, 2002). Les deux protéines auraient ensuite la possibilité de lier l'élément répresseur AvEc 164-bp au jour E9.5, alors que HOXB9 deviendrait la responsable de la régionalisation de l'expression du gène *Hoxa5* vers E10.75-E11.0, soit au moment où la protéine CDX4 cesse d'être exprimée.

Toutefois, nous ne pouvons pas exclure la possibilité que d'autres mécanismes interviennent dans la régionalisation de l'expression du gène *Hoxa5* au niveau cervico-thoracique. L'une des possibilités serait une régulation post-transcriptionnelle négative impliquant une déstabilisation du transcrit *Hoxa5* de 1.8 kb dans la partie caudale de l'embryon. Un mécanisme comparable a d'ailleurs été observé en ce qui concerne le gène *Hoxb4* (Brend *et al.*, 2003). Ensuite, l'action de miRNAs ciblant le transcrit *Hoxa5* de 1.8 kb pourrait aussi être impliquée (Pearson *et al.*, 2005). Tel que mentionné à la section 8.1, un exemple impliquant la régulation de l'expression du gène *Hoxa5* par *miR-130a* a été décrit dans un modèle de cellules endothéliales vasculaires (Chen *et al.*, 2008). Quant aux rôles que pourraient remplir les lncRNAs *Hoxa5*, ils seront discutés à la section 8.4.

En résumé, j'ai participé à la mise en évidence à l'intérieur du Chapitre IV i) que le MES fait intervenir un réseau complexe de séquences régulatrices différentes et ii) que la protéine HOXB9 serait également impliquée dans la restriction postérieure du domaine d'expression du *Hoxa5* au niveau de la colonne prévertébrale, supportant ainsi le concept de prévalence postérieure (section 1.3.3). Ces résultats font partie d'un manuscrit en préparation.

8.4 Caractérisation des lncRNAs du gène *Hoxa5*

Le promoteur distal (D2), situé dans la séquence intergénique *Hoxa6-Hoxa7* à la position -8.9 kb, génère trois longs transcrits de 5 kb, 9.5 kb et 11 kb (Fig. 5.1). De plus, les expériences d'amplification rapide aux extrémités 5' des ADNc ont identifié un transcrit *Hoxa6* de 5 kb, initié à la position -4.2kb à partir du promoteur du gène *Hoxa6* (D1) et couvrant à la fois les deux exons du gène *Hoxa6* et ceux du gène *Hoxa5*. D'ailleurs, les malformations squelettiques observées chez les souris *Hoxa6*^{-/-} présentent des similitudes avec celles des souris *Hoxa5*^{-/-} (Kostic et Capecchi, 1994). Nos résultats ont de plus démontré que la mutation du gène *Hoxa6* n'avait peu ou pas d'incidence sur la spécification de la vertèbre C7 (Table 5.1, Fig. 5.8). Par conséquent, le phénotype squelettique des souris *Hoxa6*^{-/-} serait attribuable à l'interférence transcriptionnelle sur le promoteur proximal du gène *Hoxa5* (P), créée par l'insertion de la cassette de sélection dans le locus *Hoxa6*, puisqu'on retrouve une diminution de l'expression du transcrit *Hoxa5* de 1.8 kb dans la région située entre les pv3 et pv10 chez ces souris.

Bien que j'ai démontré le potentiel de ces lncRNAs à produire une protéine HOXA5 dans un système de transcription/traduction *in vitro*, à partir de réticulocytes de lapin, l'expression *in vivo* de ces différents transcrits dans les cellules HEK293T n'a pas généré de protéine HOXA5. Cependant, nous ne pouvons pas exclure que dans certains systèmes, comme les cancers hématopoïétiques par exemple où les précurseurs des cellules sanguines peuvent s'apparenter aux réticulocytes, une traduction aberrante des long transcrits pourrait être responsable de l'augmentation de l'activité HOXA5. Ceci étant dit, puisque les longs transcrits *Hoxa5* ne codent pas pour la protéine en conditions normales et qu'il est peu probable qu'une cellule dépense de l'énergie inutilement à produire des transcrits poubelles, le laboratoire s'est intéressé à découvrir la fonction biologique des lncRNAs *Hoxa5*. J'ai ainsi participé à la création du vecteur de ciblage nécessaire à la génération de souris mutantes pour lesquelles la transcription des lncRNAs *Hoxa5* est altérée (Figure 8.7). Suite à la recombinaison homologue entre le vecteur de ciblage et l'allèle endogène, le croisement entre la souris porteuse de l'allèle ciblé et une souris exprimant la recombinaison Flipase permettra de vérifier si l'acte de transcrire en lui-même joue un rôle biologique, puisque la transcription sera initiée à partir du promoteur D2, mais sera interrompue

prématurément au niveau de la cassette IRES-*lacZ*-polyA (*Hoxa7-Hoxa5*^{D2-*IreslacZ*}). L'action de la Sox2Cre permettra, quant à elle, d'exciser le promoteur D2 afin de tester le rôle biologique des lncRNAs *Hoxa5* (*Hoxa7-Hoxa5*^{Δ-*IreslacZ*}). Les croisements générés jusqu'à maintenant suggèrent que les souris porteuses de ces mutations sont viables et fertiles.

Une analyse plus poussée des deux différentes lignées sera entreprise. Par exemple, comme l'expression de ces longs transcrits est plus tardive et postérieure que celle du transcrit *Hoxa5* de 1.8 kb, soit une expression semblable à celle du gène *Hoxa7* (Larochelle *et al.*, 1999; Fig. 5.5), ceci suggère une action potentielle des séquences régulatrices du gène *Hoxa7* sur le promoteur D2. Donc, il serait intéressant de vérifier non seulement le profil d'expression des transcrits *Hoxa5* (longs transcrits interrompus, 5 kb *Hoxa6*, 1.8 kb *Hoxa5*, etc.) par hybridation *in situ*, mais également celui des gènes *Hoxa6* et *Hoxa7*. D'ailleurs, un transcrit non-codant *Mira* (*Mistral*) de 798 pb, initié à l'extrémité 5' de la séquence intergénique *Hoxa6-Hoxa7*, soit en amont du promoteur D2, est nécessaire au maintien de l'expression des gènes *Hoxa6* et *Hoxa7*, notamment en recrutant la protéine de la famille Trithorax MLL1 au niveau de leur promoteur, dans des ESC de souris traitées au RA (Bertani *et al.*, 2011). Par conséquent, il serait envisageable que les lncRNAs *Hoxa5* recrutent des protéines Trx, PcG, ou toutes autres protéines capables de modifier la structure de la chromatine, afin de réguler l'expression de certains gènes *Hox*.

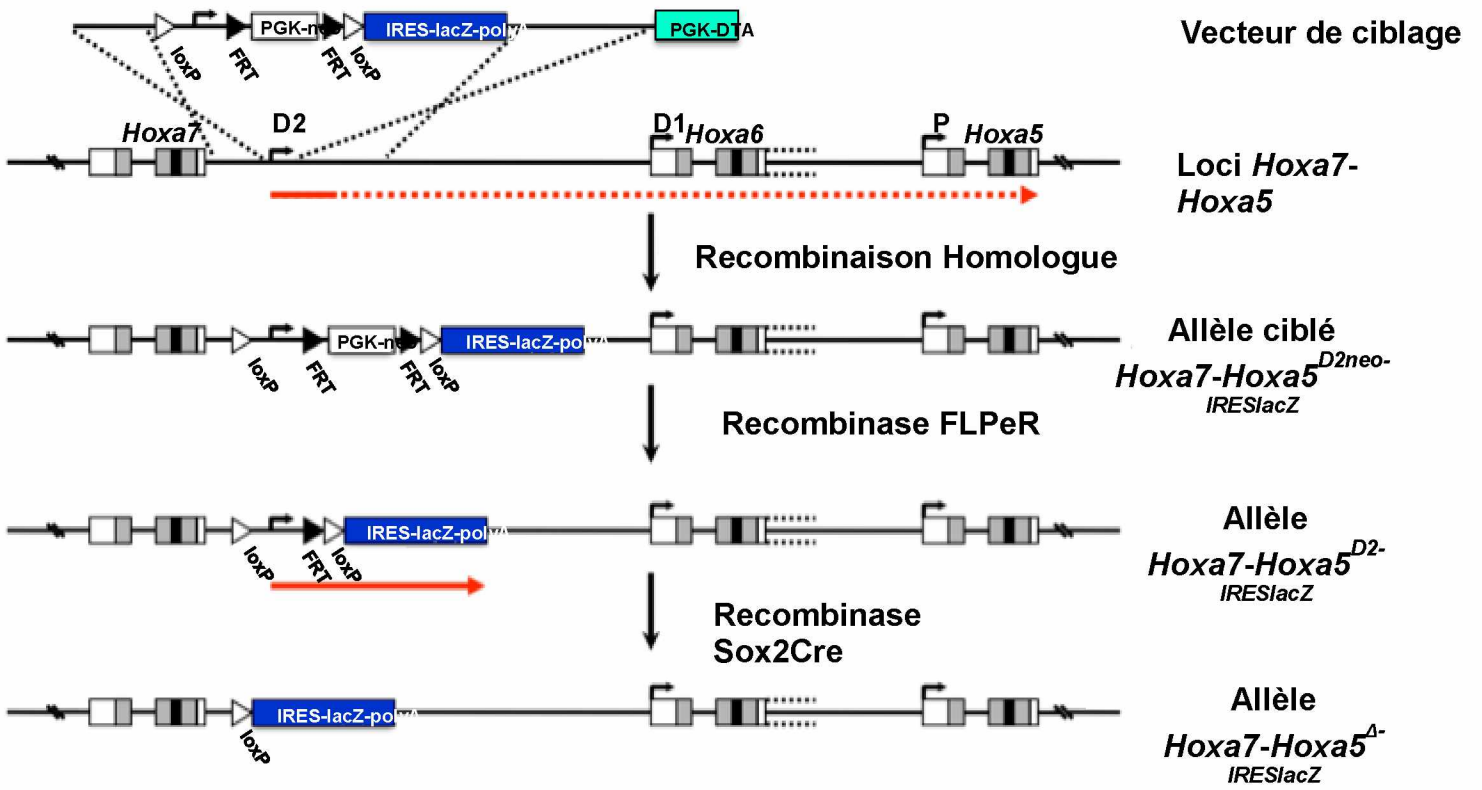
En résumé j'ai contribué aux résultats présentés dans le cadre du Chapitre V en démontrant que i) bien que tous ces transcrits ont le potentiel de produire une protéine HOXA5 *in vitro*, seul le transcrit de 1.8 kb est en mesure de le faire *in vivo*. Ces données ont été publiées dans la revue *PLOS ONE* (Coulombe *et al.*, 2010).

8.5 CONCLUSION

Initialement, ce fut sans prétention que nous avons suggéré que, de par sa position centrale au sein de son complexe ainsi de par son implication dans l'ontogénie de plusieurs structures et organes, l'étude de la régulation de l'expression du gène *Hoxa5* serait un bon modèle de l'étude des gènes *Hox* en général. À la lumière de mes travaux de doctorat, il

Figure 8.7 Stratégie utilisée pour générer les variants alléliques pour les lncRNAs *Hoxa5*

Représentation schématique du vecteur de ciblage des loci *Hoxa5-Hoxa7* et des différents allèles mutants obtenus suite aux différentes étapes de recombinaison. Les boîtes blanches correspondent aux exons des gènes, les grises aux régions codant pour les protéines et les noires représentent les domaines homéo. Les flèches noires indiquent les sites d'initiation de la transcription, alors que les rouges correspondent à la longueur des transcrits produits avant épissage. Suite à la recombinaison homologue, la cassette de sélection à la néomycine bordée de sites FRT est excisée en croisant des souris porteuses de l'allèle ciblé (*Hoxa7-Hoxa5^{D2neo-IRESlacZ}*) avec des souris exprimant la recombinaison Flipase (FLP_{er}). Chez les souris mutantes obtenues (*Hoxa7-Hoxa5^{D2-IRESlacZ}*), la transcription des lncRNAs *Hoxa5* est initiée à partir du promoteur D2, mais se termine prématurément au niveau de la cassette IRES-lacZ-polyA. De son côté, la recombinaison Sox2Cre, qui reconnaît les sites loxP flanquant le promoteur D2, excise ce dernier et génère des souris mutantes où la transcription des lncRNAs *Hoxa5* n'est pas initiée (*Hoxa7-Hoxa5^{Δ-IRESlacZ}*).



s'est avéré que le Dre Lucie Jeannotte et moi avons vu juste. Premièrement, grâce à mes résultats, nous avons I) identifié un des rares exemples où la protéine YY1 régule positivement l'expression d'un gène *Hox*, II) généré deux outils puissants pour étudier la perte de fonction de différents gènes au niveau des structures où le gène *Hoxa5* est exprimé, et III) mis en évidence un mécanisme moléculaire susceptible d'expliquer le concept de prévalence postérieure. Deuxièmement, j'ai démontré dans cette thèse toute la complexité de la régulation transcriptionnelle du gène *Hoxa5*, dont certaines séquences sont potentiellement partagées avec le gène *Hoxa4* et d'autres sont spécifiques à l'expression du gène *Hoxa5*. L'étude des différences séquences régulatrices entreprise de façon individuelle a permis de faire ressortir cette grande complexité, notamment avec l'implication de différentes sous-régions régulatrices associées à l'action de nombreux facteurs agissant en *trans*. D'ailleurs, l'implication du facteur de transcription YY1 dans la régulation transcriptionnelle du gène *Hoxa5* augmente d'autant plus le niveau de complexité, puisque qu'il est en mesure de réguler positivement et négativement une multitude de gènes, que ce soit en liant directement leur promoteur ou bien en s'associant à des protéines modifiant localement la structure de la chromatine. Si l'on ajoute à ces différents mécanismes de régulation un rôle potentiel des lncRNAs *Hoxa5* ainsi que de miRNA dans le contrôle spatio-temporel de l'expression du gène *Hoxa5*, il apparaît évident que l'étude de la régulation transcriptionnelle du gène *Hoxa5* doit être poursuivie pour tenter de mieux comprendre les différentes pathologies dans lesquelles ce gène est régulé de façon aberrante. De plus, la complexité des différents mécanismes régulateurs impliqués nous rappellent d'être prudent lors de l'élaboration d'avenues diagnostiques/thérapeutiques et qu'il est important de s'intéresser à la recherche dite plus fondamentale avant d'opter pour une approche dite plus translationnelle. Par exemple, plusieurs groupes s'intéressent en ce moment à produire des inhibiteurs pharmacologiques du gène *Yy1*, puisque celui-ci est surexprimé dans plusieurs types de cancer. Cependant, son expression s'avère également diminuée dans un bon nombre de tumeurs. Par conséquent, moduler l'expression de ce gène pourrait avoir des répercussions secondaires dramatiques. Enfin, ce dont je suis le plus fier est la capacité que j'ai acquise durant ma thèse à me remettre en question. En effet, c'est en voulant valider que le phénotype pulmonaire associé au souris *Yy1*^{flox/flox}; *Dermo1*^{+Cre} était bien spécifique à la délétion mésenchymale du gène *Yy1*, et non à la perte globale de son

expression dans le poumon, que j'ai décidé de déléter le gène *Yy1* au niveau de l'épithélium pulmonaire et ainsi découvert son importance capitale dans ce tissu. C'est donc grâce à mon incertitude face à mes résultats qu'un nouveau projet de recherche a vu le jour dans le laboratoire.

«Rien n'est plus dangereux que la certitude d'avoir raison.»

François Jacob (1920 - 2013). Biologiste, Médecin, Physiologiste, Scientifique. Il a également été récipiendaire du Prix Nobel 1965 de médecine/physiologie.

RÉFÉRENCES

- Abbott MA, Joksimovic M, Tuggle CK. (2005) Ectopic HOXA5 expression results in abnormal differentiation, migration and p53-independent cell death of superficial dorsal horn neurons. *Brain Res Dev Brain Res* 159, 87-97.
- Abu-Abed S, Dollé P, Metzger D, Wood C, MacLean G, Chambon P, Petkovich M. (2003) Developing with lethal RA levels: genetic ablation of Rarg can restore the viability of mice lacking Cyp26a1. *Development* 130, 1449-59.
- Affar el B, Gay F, Shi Y, Liu H, Huarte M, Wu S, Collins T, Li E, Shi Y. (2006) Essential dosage-dependent functions of the transcription factor yin yang 1 in late embryonic development and cell cycle progression. *Mol Cell Biol*. 26(9):3565-81.
- Affolter M, Percival-Smith A, Muller M, Leupin W, Gehring WJ. (1990) DNA binding properties of the purified Antennapedia homeodomain. *Proc Natl Acad Sci U S A*. 87, 4093-7.
- Afonja O, Smith JE, Jr., Cheng DM, Goldenberg AS, Amorosi E, Shimamoto T, Nakamura S, Ohyashiki K, Ohyashiki J, Toyama K, Takeshita K. (2000) MEIS1 and HOXA7 genes in human acute myeloid leukemia. *Leuk Res*. 24, 849-855.
- Akam M. (1987). The molecular basis for metameric pattern in the Drosophila embryo. *Development*. 101, 1-22.
- Akasaka T, Kanno M, Balling R, Mieza MA, Taniguchi M, Koseki H. (1996) A role for mel-18, a Polycomb group-related vertebrate gene, during theanterior-posterior specification of the axial skeleton. *Development* 122(5): 1513-22.
- Akasaka T, van Lohuizen M, van der LN, Mizutani-Koseki Y, Kanno M, Taniguchi M, Vidal M, Alkema M, Berns A, Koseki H. (2001) Mice doubly deficient for the Polycomb Group genes Mel18 and Bmi1 reveal synergy and requirement for maintenance but not initiation of Hox gene expression. *Development*. 128, 1587-1597.
- Allouche A, Nolens G, Tancredi A, Delacroix L, Mardaga J, Fridman V, Winkler R, Boniver J, Delvenne P, Begon DY. (2008) The combined immunodetection of AP-2 alpha and YY1 transcription factors is associated with ERBB2 gene overexpression in primary breast tumors. *Breast Cancer Res*. 10(1): R9.
- Amsellem S, Pflumio F, Bardinet D, Izac B, Charneau P, Romeo PH, Dubart Kupperschmitt A, Fichelson S. (2003) Ex vivo expansion of human hematopoietic stem cells by direct delivery of the HOXB4 homeoprotein. *Nat Med* 9, 1423-1427.
- Andrey G, Montavon T, Mascrez B, Gonzalez F, Noordermeer D, Leleu M, Trono D, Spitz F, Duboule D. (2013) A switch between topological domains underlies HoxD genes collinearity in mouse limbs. *Science* 340(6137):1234167.
- Arderiu G, Cuevas I, Chen A, Carrio M, East L, Boudreau NJ. (2007) HoxA5 stabilizes adherens junctions via increased Akt1. *Cell Adh Migr* 1, 185-95
- Argao EA, Kern MJ, Branford WW, Scott WJ Jr, Potter SS. (1995). Malformations of the heart, kidney, palate, and skeleton in alpha-MHC-Hoxb-7 transgenic mice. *Mech Dev* 52, 291-303.
- Artishevsky A, Wooden S, Sharma A, Resendez E Jr, Lee AS. (1987) Cell-cycle regulatory sequences in a hamster histone promoter and their interactions with cellular factors. *Nature* 328(6133): 823-7.

- Asli NS, Kessel M. (2010) Spatiotemporally restricted regulation of generic motor neuron programs by miR-196-mediated repression of Hoxb8. *Dev Biol.* 344(2): 857-68.
- Aubin J, Chailier P, Ménard D, Jeannotte L. (1999) Loss of Hoxa5 gene function in mice perturbs intestinal maturation. *Am J Physiol.* 277, C965-73.
- Aubin J, Déry U, Lemieux M, Chailier P, Jeannotte L. (2002a) Stomach regional specification requires Hoxa5-driven mesenchymal-epithelial signaling. *Development.* 129, 4075-87.
- Aubin J, Jeannotte L. (2001) Implication des gènes Hox dans les processus d'organogenèse chez les mammifères. *médecine/Sciences* 17, 54-62.
- Aubin J, Lemieux M, Moreau J, Lapointe J, Jeannotte L. (2002b) Cooperation of Hoxa5 and Pax1 genes during formation of the pectoral girdle. *Dev Biol.* 244, 96-113.
- Aubin J, Lemieux M, Tremblay M, Behringer RR, Jeannotte L. (1998) Transcriptional interferences at the Hoxa4/Hoxa5 locus: importance of correct Hoxa5 expression for the proper specification of the axial skeleton. *Dev Dyn.* 212, 141-56.
- Aubin J, Lemieux M, Tremblay M, Bérard J, Jeannotte L. (1997) Early postnatal lethality in Hoxa-5 mutant mice is attributable to respiratory tract defects. *Dev Biol.* 192, 432-45.
- Aulehla A, Wehrle C, Brand-Saberi B, Kemler R, Gossler A, Kanzler B, Herrmann BG. (2003) *Wnt3a* plays a major role in the segmentation clock controlling somitogenesis. *Dev. Cell* 4, 395-406.
- Ayoubi TA, Van De Ven WJ. (1996) Regulation of gene expression by alternative promoters. *FASEB J.* 10(4): 453-60.
- Azcoitia V, Aracil M, Martinez AC, Torres M. (2005) The homeodomain protein Meis1 is essential for definitive hematopoiesis and vascular patterning in the mouse embryo. *Devel Biol* 280: 307-320.
- Bachiller D, Macias A, Duboule D, Morata G. (1994) Conservation of a functional hierarchy between mammalian and insect Hox/HOM genes. *Embo J* 13, 1930-41.
- Bach C, Buhl S, Mueller D, Garcia-Cuellar MP, Maethner E, Slany RK. (2010) Leukemogenic transformation by HOXA cluster genes. *Blood* 115, 2910-8.
- Bagadi SA, Prasad CP, Kaur J, Srivastava A, Prashad R, Gupta SD, Ralhan R. (2008) Clinical significance of promoter hypermethylation of RASSF1A, RARbeta2, BRCA1 and HOXA5 in breast cancers of Indian patients. *Life Sci.* 282(25-26):1288-92.
- Bantignies F, Cavalli G. (2011) Polycomb group proteins: repression in 3D. *Trends Genet.* 27(11): 454-64.
- Bantignies F, Goodman RH, Smolik SM. (2000) Functional interaction between the coactivator Drosophila CREB-binding protein and ASH1, a member of the trithorax group of chromatin modifiers. *Mol Cell Biol.* 20(24): 9317-30.
- Bantignies F, Roure V, Comet I, Leblanc B, Schuettengruber B, Bonnet J, Tixier V, Mas A, Cavalli G. (2011) Polycomb-dependent regulatory contacts between distant Hox loci in Drosophila. *Cell* 144, 214-226.
- Baritaki S, Katsman A, Chatterjee D, Yeung KC, Spandidos DA, Bonavida B. (2007) Regulation of tumor cell sensitivity to TRAIL-induced apoptosis by the metastatic suppressor Raf kinase, inhibitor protein via Yin Yang 1 inhibition and death receptor 5 up-regulation. *J Immunol.* 179, 5441-5453.

- Barrow JR, Capecchi MR. (1996) Targeted disruption of the Hoxb-2 locus in mice interferes with expression of Hoxb-1 and Hoxb-4. *Development* 122, 3817-28.
- Barrow JR, Stadler HS, Capecchi MR. (2000) Roles of Hoxa1 and Hoxa2 in patterning the early hindbrain of the mouse. *Development* 127, 933-44.
- Bartel DP. (2009) MicroRNAs: target recognition and regulatory functions. *Cell* 136(2): 215-33.
- Bayha E, Jørgensen MC, Serup P, Grapin-Botton A. (2009) Retinoic acid signaling organizes endodermal organ specification along the entire antero-posterior axis. *PLoS One* 4(6):e5845.
- Beck F, Erlner T, Russell A, James R. (1995) Expression of Cdx-2 in the mouse embryo and placenta: possible role in patterning of the extra-embryonic membranes. *Dev Dyn* 204, 219-27.
- Becker KG, Jedlicka P, Templeton NS, Liotta L, Ozato K. (1994) Characterization of hUCRBP (YY1, NF-E1, delta): a transcription factor that binds the regulatory regions of many viral and cellular genes. *Gene* 150(2):259-66.
- Beisel C, Imhof A, Greene J, Kremmer E, Sauer F. (2002) Histone methylation by the Drosophila epigenetic transcriptional regulator Ash1. *Nature* 419(6909): 857-62
- Belak ZR, Ficzyz A, Ovsenek N. (2008) Biochemical characterization of Yin Yang 1-RNA complexes. *Biochem Cell Biol.* 86(1): 31-6.
- Belak ZR, Ovsenek N. (2007) Assembly of the Yin Yang 1 transcription factor into messenger ribonucleoprotein particles requires direct RNA binding activity. *J Biol Chem.* 282(52): 37913-20.
- Bell E, Wingate RJ, Lumsden A. (1999) Homeotic transformation of rhombomere identity after localized Hoxb1 misexpression. *Science* 284, 2168-71.
- Bel-Vialar S, Coré N, Terranova R, Goudot V, Boned A, Djabali M. (2000) Altered retinoic acid sensitivity and temporal expression of Hox genes in polycomb-M33-deficient mice. *Dev Biol.* 224(2): 238-49.
- Bel-Vialar S, Itasaki N, Krumlauf R. (2002) Initiating Hox gene expression: in the early chick neural tube differential sensitivity to FGF and RA signaling subdivides the HoxB genes in two distinct groups. *Development* 129, 5103-15.
- Bender W, Akam ME, Karch F, Beachy PA, Peifer M, Spierer P, Lewis EB. (1983) Molecular genetics of the *bithorax* complex in *Drosophila melanogaster*. *Science* 221, 23-29.
- Bender W, Fitzgerald DP. (2002) Transcription activates repressed domains in the *Drosophila* bithorax complex. *Development* 129, 4923-4930.
- Benson GV, Lim H, Paria BC, Satokata I, Dey SK, Maas RL. (1996). Mechanisms of reduced fertility in Hoxa-10 mutant mice: uterine homeosis and loss of maternal Hoxa-10 expression. *Development* 122, 2687-96.
- Benson GV, Nguyen TH, Maas RL. (1995) The expression pattern of the murine Hoxa-10 gene and the sequence recognition of its homeodomain reveal specific properties of Abdominal B-like genes. *Mol Cell Biol* 15, 1591-601.
- Berlet T, Burri M, Thoma G, Bopp D, Richstein S, Frigerio G, Noll M, Nusslein-Volhard C. (1988) The role of localization of bicoid RNA in organizing the anterior pattern of the Drosophila embryo. *EMBO J.* 7,1749-1756.
- Berns KI, Bohenzky RA. (1987) Adeno-associated viruses: an update. *Adv Virus Res.* 32, 243-306.

- Bertani S, Sauer S, Bolotin E, Sauer F. (2011) The noncoding RNA Mistral activates Hoxa6 and Hoxa7 expression and stem cell differentiation by recruiting MLL1 to chromatin. *Mol Cell*. 43(6): 1040-6.
- Berthelsen J, Zappavigna V, Mavilio F, Blasi F. (1998a) Prep1, a novel functional partner of Pbx proteins. *EMBO J*. 17: 1423-1433.
- Berthelsen J, Zappavigna V, Ferretti E, Mavilio F, Blasi F. (1998b) The novel homeoprotein Prep1 modulates Pbx-Hox protein cooperativity. *Embo J* 17, 1434-45.
- Beuchle D, Struhl G, Müller J. (2001) Polycomb group proteins and heritable silencing of Drosophila Hox genes. *Development* 128(6): 993-1004.
- Bhalla SS, Robitaille L, Nemer M. (2001) Cooperative activation by GATA-4 and YY1 of the cardiac B-type natriuretic peptide promoter. *J Biol Chem*. 276(14): 11439-45.
- Bhan A, Hussain I, Ansari KI, Kasiri S, Bashyal A, Mandal SS. (2013) Antisense Transcript Long Noncoding RNA (lncRNA) HOTAIR is Transcriptionally Induced by Estradiol. *J Mol Biol*. 425(19): 3707-22.
- Bhatia S, Pawar H, Dasari V, Mishra RK, Chandrashekar S, Brahmachari V. (2010) Chromatin remodeling protein INO80 has a role in regulation of homeotic gene expression in Drosophila. *Genes Cells* 15(7):725-35.
- Biggin MD, McGinnis W. (1997) Regulation of segmentation and segmental identity by Drosophila homeoproteins: the role of DNA binding in functional activity and specificity. *Development* 124, 4425-4433.
- Bogue CW, Gross I, Vasavada H, Dynia DW, Wilson CM, Jacobs HC. (1994) Identification of Hox genes in newborn lung and effects of gestational age and retinoic acid on their expression. *Am J Physiol*. 266(4 Pt 1): L448-54.
- Boucherat O, Chakir J, Jeannotte L. (2012) The loss of *Hoxa5* function in NOTCH-dependent goblet cell metaplasia in lung airways. *Biology Open* 1, 677-691.
- Boucherat O, Montaron S, Bérubé-Simard FA, Aubin J, Philippidou P, Wellik DM, Dasen JS, Jeannotte L. (2013) Partial functional redundancy between *Hoxa5* and *Hoxb5* paralog genes during lung morphogenesis. *Am J Physiol Lung Cell Mol Physiol*. 304(12),L817-30.
- Boudreau NJ, Varner JA. (2004) The Homeobox Transcription Factor Hox D3 Promotes Integrin $\alpha5\beta1$ Expression and Function during Angiogenesis. *J Biol Chem*. 279(6), 4862-
- Boulet AM, Capecchi MR. (1996) Targeted disruption of *hoxc-4* causes esophageal defects and vertebral transformations. *Dev Biol* 177, 232-49.
- Bradshaw MS, Shashikant CS, Belting HG, Bollekens JA, Ruddle FH. (1996) A long-range regulatory element of *Hoxc8* identified by using the pClasper vector. *PNAS* 93, 2426-30.
- Bracken AP, Helin K. (2009) Polycomb group proteins: navigators of lineage pathways led astray in cancer. *Nat Rev Cancer* 9(11): 773-84.
- Breitinger C, Maethner E, Garcia-Cuellar MP, Slany RK.(2012) The homeodomain region controls the phenotype of HOX-induced murine leukemia. *Blood* 120(19):4018-27
- Brennecke J, Stark A, Russell RB, Cohen SM. (2005) Principles of microRNA-target recognition. *PLoS Biol*. 3(3): e85.

- Brend, T., Gilthorpe, J., Summerbell, D. and Rigby, P. W. (2003). Multiple levels of transcriptional and post-transcriptional regulation are required to define the domain of Hoxb4 expression. *Development* 130, 2717-28.
- Brent AE, Tabin CJ. (2002) Developmental regulation of somite derivatives: muscle, cartilage and tendon. *Curr Opin Genet Dev* 12:548–557.
- Bridges CB, Morgan TH. (1923) The Third-Chromosome Group of Mutant Characters of *Drosophila melanogaster*. The Carnegie Institution, Washington, DC.
- Burri PH, Moschopoulos M. (1992) Structural analysis of fetal rat lung development. *Anat Rec* 234, 399–418.
- Bushey AM, Ramos E, Corces VG. (2009) Three subclasses of a *Drosophila* insulator show distinct and cell type-specific genomic distributions. *Genes Dev.* 23,1338–1350.
- Burke AC, Nelson CE, Morgan BA, Tabin C. (1995) Hox genes and the evolution of vertebrate axial morphology. *Development* 121, 333–46.
- Buske C, Feuring-Buske M, Antonchuk J, Rosten P, Hogge DE, Eaves CJ, Humphries RK. (2001) Overexpression of HOXA10 perturbs human lymphomyelopoiesis in vitro and in vivo. *Blood* 97, 2286-2292.
- Byrd KN, Shearn A. (2003) Byrd and Shearn, 2003ASH1, a *Drosophila* trithorax group protein, is required for methylation of lysine 4 residues on histone H3. *PNAS* 100(20): 11535-40.
- Cai Y, Jin J, Yao T, Gottschalk AJ, Swanson SK, Wu S, Shi Y, Washburn MP, Florens L, Conaway RC, Conaway JW. (2007) YY1 functions with INO80 to activate transcription. *Nat Struct Mol Biol.* 14(9):872-4.
- Cao R, Tsukada Y, Zhang Y. (2005) Role of Bmi-1 and Ring1A in H2A ubiquitylation and Hox gene silencing. *Mol Cell* 20(6): 845-54.
- Cao R, Wang L, Wang H, Xia L, Erdjument-Bromage H, Tempst P, Jones RS, Zhang Y. (2002) Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science* 298(5595): 1039-43.
- Capovilla M, Botas J. (1998). Functional dominance among Hox genes: repression dominates activation in the regulation of Dpp. *Development* 125, 4949-57.
- Capovilla M, Eldon ED, Pirrotta V. (1992) The giant gene of *Drosophila* encodes a b-ZIP DNA-binding protein that regulates the expression of other segmentation gap genes. *Development* 114, 99–112.
- Carapuco M, Novoa A, Bobola N, Mallo M. (2005) Hox genes specify vertebral types in the presomitic mesoderm. *Genes Dev* 19, 2116-21.
- Cardoso WV. (1995). Transcription factors and pattern formation in the developing lung. *Am J Physiol* 269, L429-42.
- Caronia G, Goodman FR, McKeown CM, Scambler PJ, Zappavigna V. (2003) An I47L substitution in the HOXD13 homeodomain causes a novel human limb malformation by producing a selective loss of function. *Development*, 130, 1701-1712.
- Carpenter EM. (2002) Hox genes and spinal cord development. *Dev. Neurosci.* 24, 24–34.

- Carpenter EM, Goddard JM, Davis AP, Nguyen TP, Capecchi MR. (1997) Targeted disruption of Hoxd-10 affects mouse hindlimb development. *Development* 124, 4505-14.
- Carr A, Biggin MD. (1999) A comparison of in vivo and in vitro DNA-binding specificities suggests a new model for homeoprotein DNA binding in *Drosophila* embryos. *EMBO J.* 18, 1598–1608.
- Carramolino L, Fuentes J, Garcia-Andres C, Azcoitia V, Riethmacher D, Torres M. (2010) Platelets play an essential role in separating the blood and lymphatic vasculatures during embryonic angiogenesis. *Circ Res* 106: 1197-201.
- Carrasco AE, McGinnis W, Gehring WJ, DeRobertis EM. (1984) Cloning of an *X. laevis* gene expressed during early embryogenesis coding for a peptide region homologous to *Drosophila* homeotic genes. *Cell* 37(2): 409-414.
- Casanova, J., Sanchez-Herrero, E., Busturia, A. and Morata, G. (1987). Double and triple mutant combinations of bithorax complex of *Drosophila*. *Embo J.* 6, 3103-9.
- Casanova, J., Sanchez-Herrero, E. and Morata, G. (1985). Prothoracic transformation and functional structure of the Ultrabithorax gene of *Drosophila*. *Cell.* 42, 663-9.
- Chambeyron S, Bickmore WA. (2004) Chromatin decondensation and nuclear reorganization of the HoxB locus upon induction of transcription. *Genes Dev* 18(10):1119-1130.
- Chambeyron S, Da Silva NR, Lawson KA, Bickmore WA. (2005) Nuclear re-organisation of the Hoxb complex during mouse embryonic development. *Development* 132(9):2215-23.
- Chambon P. (1996) A decade of molecular biology of retinoic acid receptors. *FASEB J.* 10, 940–54
- Chan CS, Rastelli L, Pirrotta V. (1994) A Polycomb response element in the Ubx gene that determines an epigenetically inherited state of repression. *EMBO J.* 13, 2553-2564.
- Chang LS, Shi Y, Shenk T. (1989) Adeno-associated virus P5 promoter contains an adenovirus E1A-inducible element and a binding site for the major late transcription factor. *J Virol.* 63(8): 3479-88.
- Chariot A, van Lint C, Chapelier M, Gielen J, Merville M, Bours V. (1999) CBP and histone deacetylase inhibition enhance the transactivation potential of the HOXB7 homeodomain-containing protein. *Oncogene* 18, 4007-14.
- Charité J, de Graaff W, Consten D, Reijnen MJ, Korving J, Deschamps J. (1998) Transducing positional information to the Hox genes: critical interaction of cdx gene products with position-sensitive regulatory elements. *Development* 125, 4349-58
- Charité J, de Graaff W, Shen S, Deschamps J. (1994) Ectopic expression of Hoxb-8 causes duplication of the ZPA in the forelimb and homeotic transformation of axial structures. *Cell* 78, 589-601.
- Charnay P. (1988) A gene encoding a protein with zinc fingers is activated during G0/G1 transition in cultured cells. *EMBO J.* 7, 29-35.
- Chawengsaksophak K, de Graaff W, Rossant J, Deschamps J, Beck F. (2004) Cdx2 is essential for axial elongation in mouse development. *PNAS* 101, 7641-5.
- Chawengsaksophak K, James R, Hammond VE, Kontgen F, Beck F. (1997) Homeosis and intestinal tumours in Cdx2 mutant mice. *Nature* 386, 84-7.

- Chen F, Capecchi MR. (1997) Targeted mutations in *hoxa-9* and *hoxb-9* reveal synergistic interactions. *Dev Biol* 181, 186-96.
- Chen F, Capecchi MR. (1999) Paralogous mouse Hox genes, *Hoxa9*, *Hoxb9*, and *Hoxd9*, function together to control development of the mammary gland in response to pregnancy. *PNAS* 96, 541–546.
- Chen F, Desai TJ, Qian J, Niederreither K, Lü J, Cardoso WV. (2007) Inhibition of Tgf beta signaling by endogenous retinoic acid is essential for primary lung bud induction. *Development* 134(16): 2969-76.
- Chen L, Cai Y, Jin J, Florens L, Swanson SK, Washburn MP, Conaway JW, Conaway RC. (2011) Subunit organization of the human INO80 chromatin remodeling complex: an evolutionarily conserved core complex catalyzes ATP-dependent nucleosome remodeling. *J Biol Chem.* 286(13): 11283-9.
- Chen L, Conaway RC, Conaway JW. Multiple modes of regulation of the human Ino80 SNF2 ATPase by subunits of the INO80 chromatin-remodeling complex. *PNAS* 110(51): 20497-502.
- Chen Y, Gorski DH. (2008) Regulation of angiogenesis through a microRNA (miR-130a) that down-regulates antiangiogenic homeobox genes *GAX* and *HOXA5*. *Blood* 111, 1217-26
- Chen F, Greer J, Capecchi MR. (1998) Analysis of *Hoxa7/Hoxb7* mutants suggests periodicity in the generation of the different sets of vertebrae. *Mech Dev* 77, 49-57.
- Chen Y, Wan B, Wang KC, Cao F, Yang Y, Protacio A, Dou Y, Chang HY, Lei M. (2011) Crystal structure of the N-terminal region of human Ash2L shows a winged-helix motif involved in DNA binding. *EMBO Rep.* 12(8):797-803.
- Chen JW, Zahid S, Shilts MH, Weaver SJ, Leskowitz RM, Habbsa S, Aronowitz D, Rokins KP, Chang Y, Pinnella Z, Holloway L, Mansfield JH. (2013) *Hoxa-5* acts in segmented somites to regulate cervical vertebral morphology. *Mech Dev.* 130(4- 5):226-40.
- Chiang A, O'Connor MB, Paro R, Simon J, Bender W. (1995) Discrete Polycomb-binding sites in each parasegmental domain of the bithorax complex. *Development* 121(6): 1681-9.
- Chinoy MR, Volpe MV, Cilley RE, Zgleszewski SE, Vosatka RJ, Martin A, Nielsen HC, Krummel TM. (1998) Growth factors and dexamethasone regulate *Hoxb5* protein in cultured murine fetal lungs. *Am J Pathol.* 274(4 pt 1); L610-20.
- Chisaka O, Capecchi MR. (1991) Regionally restricted developmental defects resulting from targeted disruption of the mouse homeobox gene *hox-1.5*. *Nature* 350, 473-9.
- Chisaka O, Musci TS, Capecchi MR. (1992) Developmental defects of the ear, cranial nerves and hindbrain resulting from targeted disruption of the mouse homeobox gene *Hox-1.6*. *Nature* 355, 516-20.
- Chisholm KM, Wan Y, Li R, Montgomery KD, Chang HY, West RB. (2012) Detection of long non-coding RNA in archival tissue: correlation with polycomb protein expression in primary and metastatic breast carcinoma. *PLoS ONE* 7: e47998.
- Choe A, Phun HQ, Tieu DD, Hu YH, Carpenter EM. (2006) Expression patterns of *Hox10* paralogous genes during lumbar spinal cord development. *Gene Expr. Patterns* 6, 730–737.

- Chouinard S, Kaufman TC. (1991). Control of expression of the homeotic labial (lab) locus of *Drosophila melanogaster*: evidence for both positive and negative autogenous regulation. *Development*. 113, 1267-80.
- Christ B, Huang R, Scaal M. (2007) Amniote somite derivatives. *Dev Dyn* 236:2382–96.
- Chu C, Qu K, Zhong FL, Artandi SE, Chang HY. (2011) Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. *Mol Cell* 44(4): 667-78.
- Ciruna B, Rossant J. (2001) FGF signaling regulates mesoderm cell fate specification and morphogenetic movement at the primitive streak. *Dev. Cell* 1, 37-49.
- Condie BG, Capecchi MR. (1993) Mice homozygous for a targeted disruption of *Hoxd-3* (*Hox-4.1*) exhibit anterior transformations of the first and second cervical vertebrae, the atlas and the axis. *Development* 119, 579-95.
- Conlon R, Rossant J. (1992) Exogenous retinoic acid rapidly induces anterior ectopic expression of murine *Hex-2* genes in vivo. *Development* 76, 357-368.
- Copeland JW, Nasiadka A, Dietrich BH, Krause HM. (1996) Patterning of the *Drosophila* embryo by a homeodomain-deleted *Ftz* polypeptide. *Nature* 379, 162–165.
- Copur Ö, Müller J. (2013) The histone H3-K27 demethylase *Utx* regulates *HOX* gene expression in *Drosophila* in a temporally restricted manner. *Development* 140(16):3478-85.
- Cordes R, Schuster-Gossler K, Serth K, Gossler A. (2004) Specification of vertebral identity is coupled to Notch signalling and the segmentation clock. *Development* 131, 1221-1233.
- Coré N, Bel S, Gaunt SJ, Aurrand-Lions M, Pearce J, Fisher A, Djabali M. (1997) Altered cellular proliferation and mesoderm patterning in Polycomb-M33-deficient mice. *Development* 124(3):721-9.
- Cribbs DL, Benassayag C, Randazzo FM, Kaufman TC. (1995) Levels of homeotic protein function can determine developmental identity: evidence from low-level expression of the *Drosophila* homeotic gene *proboscipedia* under *Hsp70* control. *Embo J* 14, 767-78.
- Crijns AP, de Graeff P, Geerts D, Ten Hoor KA, Hollema H, van der Sluis T, Hofstra RM, de Bock GH, de Jong S, van der Zee AG, de Vries EG. (2007) *MEIS* and *PBX* homeobox proteins in ovarian cancer. *Eur J Cancer* 43:2495-505.
- Crow KD, Stadler PF, Lynch VJ, Amemiya C, Wagner GP. (2006) The "fish-specific" *Hox* cluster duplication is coincident with the origin of teleosts. *Mol Biol Evol.* 23(1):121-36.
- Daga A, Podesta M, Capra MC, Piaggio G, Frassoni F, Corte G. (2000) The retroviral transduction of *HOXC4* into human CD34+ cells induces an in vitro expansion of clonogenic and early progenitors. *Experimental Hematology* 28, 569-574.
- Dahanukar A, Wharton RP. (1996) The *Nanos* gradient in *Drosophila* embryos is generated by translational regulation. *Genes Dev.* 10, 2610–2620.
- Dalton D, Chadwick R, McGinnis W. (1989) Expression and embryonic function of empty spiracles: a *Drosophila* homeo box gene with two patterning functions on the anterior–posterior axis of the embryo. *Genes Dev.* 3, 1940–1956.

- Darnell DK, Kaur S, Stanislaw S, Konieczka JH, Yatskievych TA, Antin PB. (2006) MicroRNA expression during chick embryo development. *Dev Dyn.* 235(11): 3156-65.
- Dasen JS, Jessell TM. (2009) Hox networks and the origins of motor neuron diversity. *Curr Top Dev Biol.* 88, 169-200.
- Dasen JS, Liu JP, Jessell TM. (2003) Motor neuron columnar fate imposed by sequential phases of Hox-c activity. *Nature* 425, 926–933.
- Dasen JS, Tice BC, Brenner-Morton S, Jessell TM. (2005) A Hox regulatory network establishes motor neuron pool identity and target-muscle connectivity. *Cell* 123, 477–491.
- Davenne M, Maconochie MK, Neun R, Pattyn A, Chambon P, Krumlauf R, Rijli FM. (1999). Hoxa2 and Hoxb2 control dorsoventral patterns of neuronal development in the rostral hindbrain. *Neuron* 22, 677-91.
- Davis AP, Witte DP, Hsieh-Li HM, Potter SS, Capecchi MR (1995) Absence of radius and ulna in mice lacking hoxa-11 and hoxd-11. *Nature* 375, 791-5.
- Decoville M, Giacomello E, Leng M, Locker D. (2001) DSP1, an HMG-like protein, is involved in the regulation of homeotic genes. *Genetics* 157(1):237-244.
- Delfino-Machín M, Lunn JS, Breikreuz DN, Akai J, Storey KG. (2005) Specification and maintenance of the spinal cord stem zone. *Development* 132, 4273-83.
- del Mar Lorente M, Marcos-Gutiérrez C, Pérez C, Schoorlemmer J, Ramírez A, Magin T, Vidal M. (2000) Loss- and gain-of-function mutations show a polycomb group function for Ring1A in mice. *Development* 127(23):5093-100.
- de Napoles M, Mermoud JE, Wakao R, Tang YA, Endoh M, Appanah R, Nesterova TB, Silva J, Otte AP, Vidal M, Koseki H, Brockdorff N. (2004) Polycomb group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation. *Dev Cell.* 7(5): 663-76.
- de Navas LF, Reed H, Akam M, Barrio R, Alonso CR, Sánchez-Herrero E. (2011) Integration of RNA processing and expression level control modulates the function of the Drosophila Hox gene Ultrabithorax during adult development. *Development* 138(1):107-16.
- Deng Z, Wan M, Sui G (2007) PIASy-mediated sumoylation of Yin Yang 1 depends on their interaction but not the RING finger. *Mol Cell Biol* 27, 3780–3792.
- DeRobertis EM. (2008). Evo-devo: variations on ancestral themes. *Cell* 132(2): 185-195.
- Desai TJ, Chen F, Lü J, Qian J, Niederreither K, Dollé P, Chambon P, Cardoso WV. (2006) Distinct roles for retinoic acid receptors alpha and beta in early lung morphogenesis. *Dev Biol.* 291(1):12-24
- Deschamps J, van Nes J. (2005) Developmental regulation of the Hox genes during axial morphogenesis in the mouse. *Development* 132(13):2931-42.
- Deschamps J, Wijgerde M. (1993) Two phases in the establishment of HOX expression domains. *Dev Biol* 156, 473-80.
- Desplan C, Theis J, O'Farrell PH. (1985) The Drosophila developmental gene, engrailed, encodes a sequence-specific DNA binding activity. *Nature.* 318, 630-5.
- Desplan C, Theis J, O'Farrell PH. (1988) The sequence specificity of homeodomain-DNA interaction. *Cell* 54,1081–1090.

- Dessaud E, Yang LL, Hill K, Cox B, Ulloa F, Ribeiro A, Mynett A, Novitch BG, Briscoe J (2007) Interpretation of the sonic hedgehog morphogen gradient by a temporal adaptation mechanism. *Nature* 450(7170):717-20.
- Diaz VM, Bachi A, Blasi F. (2007) Purification of Prep1 interactome identifies novel pathways regulated by Prep1. *Proteomics* 7:2617-23
- Diez del Corral R, Storey, K. G. (2004). Opposing FGF and retinoid pathways: a signalling switch that controls differentiation and patterning onset in the extending vertebrate body axis. *Bioessays* 26, 857-69.
- Di Meglio T, Kratochwil CF, Vilain N, Loche A, Vitobello A, Yonehara K, Hrycaj SM, Roska b, Peters AHFM, Eichmann A, Wellik D, Ducret S, Rijli FM. 2013. *Ezh2* orchestrates topographic migration and connectivity of mouse precerebellar neurons. *Science* 339: 204-207.
- Dintilhac A, Bihan R, Guerrier D, Deschamps S, Pellerin I. (2004) A conserved non-homeodomain Hoxa9 isoform interacting with CBP is co-expressed with the 'typical' Hoxa9 protein during embryogenesis. *Gene Expr Patterns* 4, 215-22.
- Di-Poi N, Montoya-Burgos JI, Miller H, Pourquoi O, Milinkovitch MC, Duboule D. (2010) Changes in Hox genes' structure and function during the evolution of the squamate body plan. *Nature* 464: 99–103.
- Di-Poi N, Zákány J, Duboule D. (2007) Distinct roles and regulations for HoxD genes in metanephric kidney development. *PLoS Genet.* 3(12):e232.
- Di Rocco G, Gavalas A, Pöpperl 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, 20506-15.
- Di Rosa P, Villaescusa JC, Longobardi E, Iotti G, Ferretti E, Diaz VM, Miccio A, Ferrari G, Blasi F. (2007) Homeodomain transcription factor *Prep1* (pKnox1) is required for hematopoietic stem cells and progenitors activity. *Devel Biol* 311:324-334.
- Doerksen LF, Bhattacharya A, Kannan P, Pratt D, Tainsky MA. (1996) Functional interaction between a RARE and an AP-2 binding site in the regulation of the human *Hoxa4* gene promoter. *Nucl. Acids Res.* 24 (14):2849-2856.
- Dolle P, Izpisua-Belmonte JC, Brown JM, Tickle C, Duboule D. (1991) HOX-4 genes and the morphogenesis of mammalian genitalia. *Genes Dev.* 5, 1767-1767.
- Dollé P, Lufkin T, Krumlauf R, Mark M, Duboule D, Chambon P. (1993b) Local alterations of Krox-20 and Hox gene expression in the hindbrain suggest lack of rhombomeres 4 and 5 in homozygote null Hoxa-1 (Hox-1.6) mutant embryos. *PNAS* 90, 7666-70.
- Dollé, P, Ruberte E, Kastner P, Petkovich M, Stoner CM, Gudas LJ, Chambon P. (1989). Differential expression of genes encoding alpha, beta and gamma retinoic acid receptors and CRABP in the developing limbs of the mouse. *Nature* 342, 702-705.
- Donohoe ME, Zhang X, McGinnis L, Biggers J, Li E, Shi Y. (1999) Targeted disruption of mouse Yin Yang 1 transcription factor results in peri-implantation lethality. *Mol Cell Biol.* 19(10):7237-44.
- Donohoe ME, Zhang LF, Xu N, Shi Y, Lee JT. (2007) Identification of a Ctfc cofactor, Yy1, for the X chromosome binary switch. *Mol Cell* 25(1): 43-56.
- Dorsett D. (1999) Distant liaisons: long-range enhancer-promoter interactions in *Drosophila*. *Curr Opin Genet Dev* 9, 505–514.

- Drabkin HA, Parsy C, Ferguson K, Guilhot F, Lacotte L, Roy L, Zeng C, Baron A, Hunger SP, Varella-Garcia M, Gemmill R, Brizard F, Brizard A, Roche J. (2002) Quantitative HOX expression in chromosomally defined subsets of acute myelogenous leukemia. *Leukemia* 16, 186-195.
- Drews D, Klar M, Dame C, Bräuer AU. (2009) Developmental expression profile of the YY2 gene in mice. *BMC Dev Biol.* 9, 45.
- Du H, Taylor HS. (2004). Molecular regulation of mullerian development by Hox genes. *Ann N Y Acad Sci* 1034, 152-65.
- Duboule D. (1998) Vertebrate hox gene regulation: clustering and/or colinearity? *Curr Opin Genet Dev* 8, 514-8.
- Duboule D. (2007) The rise and fall of Hox gene clusters. *Development* 134, 2549-60.
- Duboule, D. and Morata, G. (1994). Colinearity and functional hierarchy among genes of the homeotic complexes. *Trends Genet* 10, 358-64.
- Dubrulle J, McGrew MJ, Pourquié O. (2001) FGF signaling controls somite boundary position and regulates segmentation clock control of spatiotemporal Hox gene activation. *Cell* 106, 219-32.
- Dubrulle J, Pourquie O. (2004) Coupling segmentation to axis formation. *Development* 131:5783–5793.
- Dunham I, Kundaje A, Aldred SF, Collins PJ, Davis CA, Doyle F, Epstein CB, Frietze S, Harrow J, Kaul R, *et al.* (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature* 2012, 489:57–74.
- Dupé V, Davenne M, Brocard J, Dollé P, Mark M, Dierich A, Chambon P, Rijli FM. (1997). In vivo functional analysis of the Hoxa-1 3' retinoic acid response element (3'RARE). *Development* 124, 399-410.
- Durston AJ, Jansen HJ, In der Rieden P, Hooiveld MH. (2011) Hox collinearity - a new perspective. *Int J Dev Biol.* 55(10-12):899-908.
- Durston AJ, Timmermans JPM, Hage WJ, Hendricks HFJ, de Vries NJ, Heideveld M, Nieuwkoop PD. (1989) Retinoic acid causes an anteroposterior transformation in the developing central nervous system. *Nature* 340,140-144.
- Economides KD, Capecchi MR. (2003) Hoxb13 is required for normal differentiation and secretory function of the ventral prostate. *Development* 2003, 130, 2061–9
- Edgar BA, Datar SA. (1996) Zygotic degradation of two maternal Cdc25 mRNAs terminates Drosophila's early cell cycle program. *Genes Dev.* 10, 1966–1977.
- Eisenberg JC, Shilatifard A. (2010) Histone H3 lysine 4 (H3K4) methylation in development and differentiation. *Dev Biol.* 339(2):240-9.
- Ekker, SC, Jackson DG, von Kessler DP, Sun BI, Young KE, Beachy PA. (1994) The degree of variation in DNA sequence recognition among four Drosophila homeotic proteins. *EMBO J.* 13, 3551–3560.
- Eloy-Trinquet S, Nicolas JF. (2002) Cell coherence during production of the presomitic mesoderm and somitogenesis in the mouse embryo. *Development* 129, 3609-3619.
- Endoh M, Endo TA, Endoh T, Isono K, Sharif J, Ohara O, Toyoda T, Ito T, Eskeland R, Bickmore WA, Vidal M, Bernstein BE, Koseki H. (2012) Histone H2A monoubiquitination is a crucial step to mediate PRC1-dependent repression of developmental genes to maintain ES cell identity. *PLoS Genet.* 8(7): e1002774.
- Enright AJ, John B, Gaul U, Tuschl T, Sander C, Marks DS. (2003) MicroRNA targets in Drosophila. *Genome Biol.* 5(1): R1.

- Ephrussi A, Dickinson LK, Lehmann R. (1991) Oskar organizes the germ plasm and directs localization of the posterior determinant nanos. *Cell* 66, 37–50.
- Eskeland R, Freyer E, Leeb M, Wutz A, Bickmore WA. (2010) Histone acetylation and the maintenance of chromatin compaction by Polycomb repressive complexes. *Cold Spring Harb Symp Quant Biol.* 75, 71-8.
- Estrada B, Casares F, Busturia A, Sanchez-Herrero E. (2002) Genetic and molecular characterization of a novel iab-8 regulatory domain in the Abdominal-B gene of *Drosophila melanogaster*. *Development* 129, 5195–5204.
- Eun Kwon H, Taylor HS. (2004) The role of HOX genes in human implantation. *Ann NY Acad Sci* 1034, 1-18.
- Favier B, Rijli FM, Fromental-Ramain C, Fraulob V, Chambon P, Dollé P. (1996) Functional cooperation between the non-paralogous genes Hoxa-10 and Hoxd-11 in the developing forelimb and axial skeleton. *Development* 122, 449-60.
- Favot L, Hall SM, Haworth SG, Kemp PR. (2005) Cytoplasmic YY1 Is Associated with Increased Smooth Muscle-Specific Gene Expression : Implications for Neonatal Pulmonary Hypertension. *Am J Pathol.* 167(6):1497-509.
- Ferrai C, de Castro IJ, Lavitas L, Chotalia M, Pombo A. (2010) Gene positioning. *Cold Spring Harb Perspect Biol.* 6, a000588.
- Ferretti E, Cambronerio F, Tumpel S, Longobardi E, Wiedemann L, M Blasi F, Krumlauf R. (2005). Hoxb1 enhancer and control of rhombomere 4 expression: complex interplay between PREP1-PBX1-HOXB1 binding sites. *Mol Cell Biol* 25, 8541-52.
- Ferretti E, Villaescusa JC, Di Rosa P, Fernandez-Diaz LC, Longobardi E, Mazziere R, Miccio A, Micali N, Selleri L, Ferrari G, Blasi F. (2006) Hypomorphic mutation of the TALE gene *Prep1* (*pKnox1*) causes a major reduction of Pbx and Meis proteins and a pleiotropic embryonic phenotype. *Mol Cell Biol* 26:5650-5662.
- Fernandez-Diaz LC, Laurent A, Girasoli S, Turco M, Longobardi E, Iotti G, Jenkins NA, Fiorenza MT, Copeland NG, Blasi F. (2010) The absence of *Prep1* causes p53-dependent apoptosis of pluripotent epiblast cells. *Development* 137:3393-3403.
- Ferraiuolo MA, Rousseau M, Miyamoto C, Shenker S, Wang XQ, Nadler M, Blanchette M, Dostie J. (2010) The three-dimensional architecture of Hox cluster silencing. *Nucleic Acids Res.* 38(21):7472-84.
- Fischbach NA, Rozenfeld S, Shen W, Fong S, Chrobak D, Ginzinger D, Kogan SC, Radhakrishnan A, Le Beau MM, Largman C, Lawrence HJ. (2005) 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* 105, 1456-1466.
- Fitzpatrick VD, Percival-Smith A, Ingles CJ, Krause HM. (1992) Homeodomain-independent activity of the fushi tarazu polypeptide in *Drosophila* embryos. *Nature* 356, 610–612.
- Flanagan JR, Becker KG, Ennist DL, Gleason SL, Driggers PH, Levi BZ, Appella E, Ozato K. (1992) Cloning of a negative transcription factor that binds to the upstream conserved region of Moloney murine leukemia virus. *Mol Cell Biol.* 12(1): 38-44.

- Fognani C, Kilstrup-Nielsen C, Ferretti E, Zappavigna V, Blasi F. (2002) Human PREP-2, a novel interactor of PBX proto-oncogene, defines a novel sub-family of TALE homeodomain transcription factors. *Nucl Ac Res* 30:2043-2051.
- Folberg A, Kovacs EN, Featherstone MS. (1997) Characterization and retinoic acid responsiveness of the murine Hoxd4 transcription unit. *J Biol Chem* 272, 29151-7.
- Forlani S, Lawson KA, Deschamps J. (2003) Acquisition of *Hox* codes during gastrulation and axial elongation in the mouse embryo. *Development* 130, 3807-3819.
- Francis NJ, Kingston RE. (2001) Mechanisms of transcriptional memory. *Nat Rev Mol Cell Biol.* 2(6): 409-21.
- Franke A, DeCamillis M, Zink D, Cheng N, Brock HW, Paro R. (1992) Polycomb and polyhomeotic are constituents of a multimeric protein complex in chromatin of *Drosophila melanogaster*. *EMBO J.* 11(8): 2941-50.
- Freund JN, Domon-Dell C, Kedinger M, Duluc I. (1998) The Cdx-1 and Cdx-2 homeobox genes in the intestine. *Biochem Cell Biol* 76, 957-69.
- Friedmann Y, Daniel CA, Strickland P, Daniel CW. (1994) Hox genes in normal and neoplastic mouse mammary gland. *Cancer Res.* 54(22):5981-5.
- Fromental-Ramain C, Warot X, Lakkaraju S, Favier B, Haack H, Birling C, Dierich A, Dollé P, Chambon P. (1996a) Specific and redundant functions of the paralogous Hoxa-9 and Hoxd-9 genes in forelimb and axial skeleton patterning. *Development* 122, 461-72.
- Fujimoto S, Araki K, Chisaka O, Araki M, Takagi K, Yamamura K. (1998) Analysis of the murine Hoxa-9 cDNA: an alternatively spliced transcript encodes a truncated protein lacking the homeodomain. *Gene* 209, 77-85.
- Gabrielsen OS, Huet J. (1993) Magnetic DNA affinity purification of yeast transcription factor. *Methods Enzymol.* 218:508-25.
- Galli A, Robay D, Osterwalder M, Bao X, Bénazet JD, Tariq M, Paro R, Mackem S, Zeller R. (2010) Distinct roles of Hand2 in initiating polarity and posterior Shh expression during the onset of mouse limb bud development. *PLoS Genet.* 6(4):e1000901
- Galliot B, Dollé P, Vigneron M, Featherstone MS, Baron A, Duboule D. (1989) The mouse Hox-1.4 gene: primary structure, evidence for promoter activity and expression during development. *Development* 107, 343-59.
- Gamer LW, Wright CV. (1993) Murine Cdx-4 bears striking similarities to the *Drosophila* caudal gene in its homeodomain sequence and early expression pattern. *Mech Dev* 43, 71-81.
- Garcia-Fernandez, J. (2005). The genesis and evolution of homeobox gene clusters. *Nat Rev Genet.* 6, 881-92.
- Garcia-Gasca A, Spyropoulos DD. (2000). Differential mammary morphogenesis along the anteroposterior axis in *Hoxc6* gene targeted mice. *Development Dynamics,* 219, 261-276.
- Garin E, Lemieux, M, Coulombe Y, Robinson GW, Jeannotte L. (2006) Stromal *Hoxa5* function controls the growth and differentiation of mammary alveolar epithelium epithelium. *Dev Dyn* 235, 1858-71.
- Gaufo GO, Flodby P, Capecchi MR. (2000) Hoxb1 controls effectors of sonic hedgehog and Mash1 signaling pathways. *Development* 127, 5343-54.

- Gafo GO, Thomas KR, Capecchi MR (2003) Hox3 genes coordinate mechanisms of genetic suppression and activation in the generation of branchial and somatic motoneurons. *Development* 130, 5191-201.
- Gaunt SJ. (1994) Conservation in the Hox code during morphological evolution. *Int J Dev Biol* 38:549-52.
- Gaunt SJ. (2000) Evolutionary shifts of vertebrate structures and Hox expression up and down the axial series of segments: a consideration of possible mechanisms. *Int J Dev Biol*. 44(1):109-17.
- Gaunt SJ, Cockley A, Drage D (2004). Additional enhancer copies, with intact cdx binding sites, anteriorize Hoxa-7/lacZ expression in mouse embryos: evidence in keeping with an instructional cdx gradient. *Int J Dev Biol* 48, 613-22.
- Gaunt SJ, Krumlauf R, Duboule D. (1989) Mouse homeo-genes within a subfamily, Hox-1.4, -2.6 and -5.1, display similar anteroposterior domains of expression in the embryo, but show stage- and tissue dependent differences in their regulation. *Development* 107:131-41.
- Gavalas A, Krumlauf R. (2000) Retinoid signalling and hindbrain patterning. *Curr Opin Genet Dev* 10, 380-6.
- Gavalas A, Studer M, Lumsden A, Rijli FM, Krumlauf R, Chambon P. (1998). Hoxa1 and Hoxb1 synergize in patterning the hindbrain, cranial nerves and second pharyngeal arch. *Development* 125, 1123-36.
- Gehring WJ, Kloter U, Suga H. (2009) Evolution of the Hox Gene Complex from an Evolutionary Ground State. *Curr Top Dev Biol* 88, 35-61.
- Geisen MJ, Di Meglio T, Pasqualetti M, Ducret S, Brunet JF, Chedotal A, Rijli FM. (2008) Hox paralog group 2 genes control the migration of mouse pontine neurons through slit-robo signaling. *PLoS Biol*. 6, e142.
- Gendron-Maguire M, Mallo M, Zhang M, Gridley T. (1993) Hoxa-2 mutant mice exhibit homeotic transformation of skeletal elements derived from cranial neural crest. *Cell* 75, 1317-31.
- Gendronneau G, Boucherat O, Aubin J, Lemieux M, Jeannotte, L. (2012) The loss of *Hoxa5* function causes estrous acyclicity and ovarian epithelila inclusion cysts. *Endocrinology* 153:1484-1497.
- Gendronneau G, Lemieux M, Morneau M, Paradis J, Têtu B, Frenette N, Aubin J, Jeannotte L. (2010) Influence of Hoxa5 on p53 tumorigenic outcome in mice. *Am. J. Pathol.* 176: 995-1005.
- Giampaolo A, Sterpetti P, Bulgarini D, Samoggia P, Pelosi E, Valtieri M, Peschle C. (1994) Key functional role and lineage-specific expression of selected HOXB genes in purified hematopoietic progenitor differentiation. *Blood* 1994; 84, 3637-3647.
- Gillespie RF, Gudas LJ. (2007a) Retinoid regulated association of transcriptional co-regulators and the polycomb group protein SUZ12 with the retinoic acid response elements of Hoxa1, RARbeta(2), and Cyp26A1 in F9 embryonal carcinoma cells. *J Mol Biol.* 372(2):298-316.
- Gillespie RF, Gudas LJ. (2007b) Retinoic acid receptor isotype specificity in F9 teratocarcinoma stem cells results from the differential recruitment of coregulators to retinoic response elements. *J Biol Chem.* 282(46):33421-34.

- Gilthorpe J, Vandromme M, Brend T, Gutman A, Summerbell D, Totty N, Rigby PW. (2002) Spatially specific expression of Hoxb4 is dependent on the ubiquitous transcription factor NFY. *Development* 129(16): 3887-99.
- Glover JC, Renaud JS, Rijli, FM. (2006) Retinoic acid and hindbrain patterning. *J. Neurobiol.* 66, 705–725.
- Goddard JM, Rossel M, Manley NR, Capecchi MR. (1996) Mice with targeted disruption of Hoxb-1 fail to form the motor nucleus of the VIIIth nerve. *Development* 122, 3217-28.
- Godwin AR, Capecchi MR. (1998) Hoxc13 mutant mice lack external hair. *Genes Dev* 12, 11-20.
- Gokhman D, Livyatan I, Sailaja BS, Melcer S, Meshorer E. (2013) Multilayered chromatin analysis reveals E2f, Smad and Zfx as transcriptional regulators of histones. *Nat Struct Mol Biol.* 20(1): 119-26.
- Golebiowski FM, Górecki A, Bonarek P, Rapala-Kozik M, Kozik A, Dziedzicka-Wasylewska M. (2012) An investigation of the affinities, specificity and kinetics involved in the interaction between the Yin Yang 1 transcription factor and DNA. *FEBS J.* 279(17):3147-58.
- Golpon HA, Geraci MW, Moore MD, Miller HL, Miller GJ, Tuder RM, Voelkel NF. (2001) HOX genes in human lung: altered expression in primary pulmonary hypertension and emphysema. *Am J Pathol.* 158(3): 955-66.
- Gomez C, Ozbudak EM, Wunderlich J, Baumann D, Lewis J, Pourquie O. (2008) Control of segment number in vertebrate embryos. *Nature* 454:335–9.
- Gonzalez-Reyes A, Morata G. (1990) The developmental effect of overexpressing a Ubx product in Drosophila embryos is dependent on its interactions with other homeotic products. *Cell.* 61, 515-22.
- Goodrich ES. (1930) Studies on the structure and development of vertebrates. London: Macmillan.
- Gould A, Itasaki N, Krumlauf R. (1998). Initiation of rhombomeric Hoxb4 expression requires induction by somites and a retinoid pathway. *Neuron* 21, 39-51.
- Gould A, Morrison A, Sproat G, White RA, Krumlauf, R. (1997) Positive cross-regulation and enhancer sharing: two mechanisms for specifying overlapping Hox expression patterns. *Genes Dev* 11, 900-13.
- Greer JM, Capecchi MR. (2002). Hoxb8 is required for normal grooming behavior in mice. *Neuron* 33, 23-34.
- Greer J M, Puetz J, Thomas KR, Capecchi MR. (2000) Maintenance of functional equivalence during paralogous Hox gene evolution. *Nature.* 403, 661-5.
- Gregory RI, Chendrimada TP, Cooch N, Shiekhattar R. (2005) Human RISC couples microRNA biogenesis and posttranscriptional gene silencing. *Cell* 123, 631-40.
- Grewal SI. (2010) RNAi-dependent formation of heterochromatin and its diverse functions. *Curr Opin Genet Dev.*20(2): 134-41.
- Griffiths-Jones S, Hui JH, Marco A, Ronshaugen M. (2011) MicroRNA evolution by arm switching. *EMBO Rep.* 12(2): 172-7.
- Grimaud C, Bantignies F, Pal-Bhadra M, Ghana P, Bhadra U, Cavalli G. (2006) RNAi components are required for nuclear clustering of Polycomb group response elements. *Cell* 124(5):957-71.

- Grubach L, Juhl-Christensen C, Rethmeier A, Olesen LH, Aggerholm A, Hokland P, Ostergaard M. (2008) Gene expression profiling of Polycomb, Hox and Meis genes in patients with acute myeloid leukaemia. *Eur J Haematol* 81:112-22.
- Guenther MG, Young RA. (2010) Transcription. Repressive transcription. *Science* 329(5988): 150-1.
- Guerreiro I, Casaca1 A, Nunes1 A, Monteiro S, Nóvoal A, Ferreira RB, Bom J, Mallo M. (2012). Regulatory role for a conserved motif adjacent to the homeodomain of Hox10 proteins. *Development* 139, 0000-0000. doi:10.1242/dev.081448
- Guichet A, Copeland JW, Erdelyi M, Hlousek D, Zavorszky P, Ho J, Brown S, Percival-Smith A, Krause HM, Ephrussi A. (1997) The nuclear receptor homologue Ftz-F1 and the homeodomain protein Ftz are mutually dependent cofactors. *Nature* 385, 548–552.
- Guo Y, Mahony S, Gifford DK. (2012) High resolution genome wide binding event finding and motif discovery reveals transcription factor spatial binding constraints. *PLoS Comput Biol.* 8(8):e1002638.
- Gupta RA, Shah N, Wang KC, Kim J, Horlings HM, Wong DJ, Tsai MC, Hung T, Argani P, Rinn JL, Wang Y, Brzoska P, Kong B, Li R, West RB, van de Vijver MJ, Sukumar S, Chang HY. (2010) Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* 464, 1071–1076.
- Gutman A, Gilthorpe J, Rigby, PW. (1994) Multiple positive and negative regulatory elements in the promoter of the mouse homeobox gene Hoxb-4. *Mol Cell Biol* 14, 8143-54.
- Haerry TE, Gehring WJ. (1996) Intron of the mouse Hoxa-7 gene contains conserved homeodomain binding sites that can function as an enhancer element in *Drosophila*. *PNAS* 93, 13884-9.
- Hafen E, Levine M, Gehring WJ. (1984) Regulation of Antennapedia transcript distribution by the bithorax complex in *Drosophila*. *Nature.* 307, 287-9.
- Harding K, Levine M. (1988) Gap genes define the limits of antennapedia and bithorax gene expression during early development in *Drosophila*. *Embo J.* 7, 205-14.
- Harding K, Wedeen C, McGinnis W, Levine M. (1985) Spatially regulated expression of homeotic genes in *Drosophila*. *Science.* 229, 1236-42.
- Hariharan N, Kelley DE, Perry RP. (1991) Delta, a transcription factor that binds to downstream elements in several polymerase II promoters, is a functionally versatile zinc finger protein. *PNAS* 88(21):9799-803
- Harfe BD, Scherz PJ, Nissim S, Tian H, McMahon AP, Tabin CJ (2004) Evidence for an expansion-based temporal Shh gradient in specifying vertebrate digit identities. *Cell* 118(4): 517-28.
- Modified after Hartenstein V. (1993) Atlas of *Drosophila* Development, p 52, *Cold Spring Harbor Laboratory Press.*
- Hayashi S, Scott MP. (1990) What determines the specificity of action of *Drosophila* homeodomain proteins? *Cell* 63, 883–894.
- Hayashida T, Takahashi F, Chiba N, Brachtel E, Takahashi M, Godin-Heymann N, Gross KW, Vivanco Md, Wijendran V, Shioda T, Sgroi D, Donahoe PK, Maheswaran S. (2010) HOXB9, a gene overexpressed in breast cancer, promotes tumorigenicity and lung metastasis. *Proc Natl Acad Sci U S A.* 107(3):1100-5

- He X, Yan YL, Eberhart JK, Herpin A, Wagner TU, Scharl M, Postlethwait JH. (2011) miR-196 regulates axial patterning and pectoral appendage initiation. *Dev Biol.* 357(2): 463-77.
- Hérault Y, Rassoulzadegan M, Cuzin F, Duboule D. (1998) Engineering chromosomes in mice through targeted meiotic recombination (TAMERE). *Nat Genet* 20, 381-384.
- Herriges JC, Yi L, Hines EA, Harvey JF, Xu G, Gray PA, Ma Q, Sun X. (2012) Genome-scale study of transcription factor expression in the branching mouse lung. *Dev Dyn* 241, 1432–1453
- Hiromura M, Choi CH, Sabourin NA, Jones H, Bachvarov D, Usheva A. (2003) YY1 is regulated by O-linked N-acetylglucosaminylation (O-glcNAcylation). *J Biol Chem.* 278(16): 14046-52.
- Ho MC, Schiller BJ, Akbari OS, Bae E, Drewell RA. (2011) Disruption of the abdominal-B promoter tethering element results in a loss of long-range enhancer-directed Hox gene expression in *Drosophila*. *Plos One* 6(1): e16283
- Hodgson JW, Brock HW. (2011) Are polycomb group bodies gene silencing factories? *Cell* 144(2):170-1.
- Hoegg S, Meyer A. (2005) Hox clusters as models for vertebrate genome evolution. *Trends Genet.* 21, 421-4.
- Hoey T, Levine M. (1988) Divergent homeo box proteins recognize similar DNA sequences in *Drosophila*. *Nature.* 332, 858-61.
- Hogan BL, Holland PW, Lumsden A. (1988) Expression of the homeobox gene, Hox 2.1, during mouse embryogenesis. *Cell Differ Dev* 25 Suppl, 39-44.
- Hogga I, Karch F. (2002) Transcription through the *iab-7 cis*-regulatory domain of the bithorax complex interferes with maintenance of Polycomb-mediated silencing. *Development* 129, 4915–4922
- Holland PW, Takahashi T. (2005) The evolution of homeobox genes: Implications for the study of brain development. *Brain Res Bull.* 66, 484-90.
- Hongo F, Garban H, Huerta-Yepez S, Vega M, Jazirehi AR, Mizutani Y, Miki T, Bonavida B. (2005) Inhibition of the transcription factor Yin Yang 1 activity by S-nitrosation. *Biochem Biophys Res Commun.* 336(2):692-701.
- Hosler BA, LaRosa GJ, Grippo JF, Gudas LJ. (1989) Expression of REX-1, a gene containing zinc finger motifs, is rapidly reduced by retinoic acid in F9 teratocarcinoma cells. *Mol Cell Biol.* 9(12):5623-9.
- Horan GS, Ramirez-Solis R, Featherstone MS, Wolgemuth DJ, Bradley A, Behringer RR. (1995) Compound mutants for the paralogous *hoxa-4*, *hoxb-4*, and *hoxd-4* genes show more complete homeotic transformations and a dose-dependent increase in the number of vertebrae transformed. *Genes Dev* 9:1667–1677.
- Horan GS, Wu K, Wolgemuth DJ, Behringer RR. (1994) Homeotic transformation of cervical vertebrae in *Hoxa-4* mutant mice. *PNAS* 91, 12644-8.
- Hornstein E, Mansfield JH, Yekta S, Hu JK, Harfe BD, McManus MT, Baskerville S, Bartel DP, Tabin CJ. (2005) The microRNA miR-196 acts upstream of *Hoxb8* and *Shh* in limb development. *Nature* 438(7068): 671-4.
- Hostikka SL, Capecchi MR. (1998) The mouse *Hoxc11* gene: genomic structure and expression pattern. *Mech Dev* 70, 133–4

- Houle M, Sylvestre JR, Lohnes D. (2003) Retinoic acid regulates a subset of Cdx1 function in vivo. *Development* 130, 6555-67.
- Hsieh-Li HM, Witte DP, Weinstein M, Branford W, Li H, Small K, Potter SS. (1995) Hoxa 11 structure, extensive antisense transcription, and function in male and female fertility. *Development* 121, 1373-85.
- Huang DH, Chang YL. (2004) Isolation and characterization of CHRASCH, a polycomb-containing silencing complex. *Methods Enzymol.* 377, 267-82.
- Huang D, Chen SW, Gudas LJ. (2002) Analysis of two distinct retinoic acid response elements in the homeobox gene Hoxb1 in transgenic mice. *Dev Dyn* 223(3):353-370.
- Huang L, Pu Y, Hepps D, Danielpour D, Prins GS. (2007) Posterior Hox gene expression and differential androgen regulation in the developing and adult rat prostate lobes. *Endocrinology* 148, 1235-45
- Hutvagner G. (2005) Small RNA asymmetry in RNAi: function in RISC assembly and gene regulation. *FEBS Lett* 579, 5850-7
- Hyde-DeRuyscher RP, Jennings E, Shenk T. (1995) DNA binding sites for the transcriptional activator/repressor YY1. *Nucleic Acids Res.* 23(21): 4457-65.
- Hyduk D, Percival-Smith A. (1996) Genetic characterization of the homeodomain-independent activity of the *Drosophila* fushi tarazu gene product. *Genetics* 142, 481-492.
- Ikeya M, Takada S. (2001). Wnt-3a is required for somite specification along the anteroposterior axis of the mouse embryo and for regulation of cdx-1 expression. *Mech Dev* 103, 27-33.
- Ingham PW. (1983) Differential expression of bithorax complex genes in the absence of the extra sex combs and trithorax genes. *Nature* 306, 591-593.
- Ingham PW. (1998) trithorax and the regulation of homeotic gene expression in *Drosophila*: a historical perspective. *Int J Dev Biol.* 42(3): 423-9.
- Inouye CJ, Seto E. (1994) Relief of YY1-induced transcriptional repression by protein-protein interaction with the nucleolar phosphoprotein B23. *J Biol Chem.* 269(9): 6506-10.
- Irving C, Mason I. (2000) Signalling by FGF8 from the isthmus patterns anterior hindbrain and establishes the anterior limit of Hox gene expression. *Development* 127, 177-86.
- Isaacs HV, Pownall ME, Slack JM. (1998) Regulation of Hox gene expression and posterior development by the *Xenopus* caudal homologue Xcad3. *Embo J* 17, 3413-27.
- Iwasaki M, Kuwata T, Yamazaki Y, Jenkins NA, Copeland NG, Osato M, Ito Y, Kroon E, Sauvageau G, Nakamura T. (2005) Identification of cooperative genes for NUP98-HOXA9 in myeloid leukemogenesis using a mouse model. *Blood* 105, 784-793
- Izpisua-Belmonte JC, Falkenstein H, Dollé P, Renucci A, Duboule, D. (1991) Murine genes related to the *Drosophila* AbdB homeotic genes are sequentially expressed during development of the posterior part of the body. *Embo J* 10, 2279-89.
- Jäckle HD, Tautz D, Schuh E, Seifert E, Lehmann, R. (1986). Cross-regulatory interactions among gap genes of *Drosophila*. *Nature* 324, 668-670.

- Javed S, Langley SEM. (2013) Importance of *HOX* genes in normal prostate gland formation, prostate cancer development and its early detection. *BJU Int*. DOI: 10.1111/bju.12269
- Jeannotte L, Lemieux M, Charron J, Poirier F, Robertson EJ. (1993) Specification of axial identity in the mouse: role of the *Hoxa-5* (*Hox1.3*) gene. *Genes Dev*. 7, 2085-96.
- Jeannotte L, Ruiz JC, Robertson EJ. (1991) Low level of *Hox1.3* gene expression does not preclude the use of promoterless vectors to generate a targeted gene disruption. *Mol Cell Biol*. 11, 5578-85.
- Jegalian BG, Miller RW, Wright CV, Blum M, De Robertis EM. (1992) A *Hox 3.3-lacZ* transgene expressed in developing limbs. *Mech Dev* 39, 171-80.
- Jeon Y, Lee JT. (2011) YY1 tethers Xist RNA to the inactive X nucleation center. *Cell* 146(1): 119-33.
- Jones CA, Ng J, Peterson AJ, Morgan K, Simon J, Jones RS. (1999) The *Drosophila* *esc* and *E(z)* proteins are direct partners in polycomb group-mediated repression. *Mol Cell Biol*. 18(5): 2825-34
- Joshi R, Sun L, Mann R. (2010) Dissecting the functional specificities of two *Hox* proteins. *Genes Dev* 24, 1533-45.
- Joksimovic M, Jeannotte L, Tuggle CK. (2005) Dynamic expression of murine *HOXA5* protein in the central nervous system. *Gene Expr Patterns*. 5, 792-800.
- Jungbluth S, Bell E, Lumsden A. (1999) Specification of distinct motor neuron identities by the singular activities of individual *Hox* genes. *Development* 126, 2751-8.
- Jürgens G. (1985) A group of genes controlling the spatial expression of the bithorax complex in *Drosophila*. *Nature* 316, 153-155.
- Kalionis B, O'Farrell PH. (1993) A universal target sequence is bound in vitro by diverse homeodomains. *Mech. Dev*. 43, 57-70.
- Kamps MP, Murre C, Sun XH, Baltimore D. (1990) A new homeobox gene contributes the DNA binding domain of the t(1;19) translocation protein in pre-B ALL. *Cell* 60, 547-55.
- Karch F, Weiffenbach B, Peifer M, Bender W, Duncan I, Celniker S, Crosby M, Lewis EB. (1985) The abdominal region of the bithorax complex. *Cell* 43, 81-96.
- Kashyap V, Gudas LJ, Brenet F, Funk P, Viale A, Scandura JM. (2011) Epigenomic reorganization of the clustered *Hox* genes in embryonic stem cells induced by retinoic acid. *J Biol Chem*. 286(5):3250-60.
- Kassardjian A, Rizkallah R, Riman S, Renfro SH, Alexander KE, Hurt MM. (2012) The transcription factor YY1 is a novel substrate for Aurora B kinase at G2/M transition of the cell cycle. *PLoS One* 7(11):e50645.
- Kastner P, Mark M, Ghyselinck N, Krezel W, Dupe V, Grondona JM, Chambon P. (1997) Genetic evidence that the retinoid signal is transduced by heterodimeric RXR/RAR functional units during mouse development. *Development* 124, 313-26
- Kaufman TC, Lewis R, Wakimoto BT. (1980) Cytogenetic analysis of chromosome 3 in *Drosophila melanogaster*: The homeotic gene complex in polytene chromosome 84A-B. *Genetics* 94, 115-133.
- Kawagoe H, Humphries RK, Blair A, Sutherland HJ, Hogge DE. (1999) Expression of *HOX* genes, *HOX* cofactors, and *MLL* in phenotypically and functionally

- defined subpopulations of leukemic and normal human hematopoietic cells. *Leukemia* 13, 687-698.
- Kawazoe Y, Sekimoto T, Araki M, Takagi K, Araki K, Yamamura K. (2002) Region-specific gastrointestinal Hox code during murine embryonal gut development. *Dev. Growth Differ.* 44, 77-84.
- Kennison JA. (1993) Transcriptional activation of Drosophila homeotic genes from distant regulatory elements. *Trends Genet.* 9(3): 75-9.
- Keren H, Lev-Maor G, Ast G. (2010) Alternative splicing and evolution: diversification, exon definition and function. *Nat Rev Genet.* 11(5):345-55.
- Kessel M, Balling R, Gruss P. (1990) Variations of cervical vertebrae after expression of a Hox-1.1 transgene in mice. *Cell* 61, 301-8.
- Kessel M, Gruss P. (1991) Homeotic transformations of murine vertebrae and concomitant alteration of Hox codes induced by retinoic acid. *Cell* 67, 89-104.
- Kiecker C, Lumsden A. (2005) Compartments and their boundaries in vertebrate brain development. *Nat. Rev.* 6, 553–564.
- Kim JD, Faulk C, Kim J (2007) Retroposition and evolution of the DNA-binding motifs of YY1, YY2 and REX1. *Nucleic Acids Res* 35, 3442-52.
- Kim JD, Hinz AK, Bergmann A, Huang JM, Ovcharenko I, Stubbs L, Kim J. (2006) Identification of clustered YY1 binding sites in imprinting control regions. *Genome Res.* 16(7): 901-11.
- Kim J, Kim JD. (2008) In vivo YY1 knockdown effects on genomic imprinting. *Hum Mol Genet* 17, 391-401.
- Kim DS, Kim MJ, Lee JY, Lee SM, Choi JY, Yoon GS, Na YK, Hong HS, Kim SG, Choi JE, Lee SY, Park JY. (2009) Epigenetic inactivation of Homeobox A5 gene in nonsmall cell lung cancer and its relationship with clinicopathological features. *Mol Carcinog.* 48(12): 1109-15.
- Kim-Ha J, Smith JL, Macdonald PM. (1991) oskar mRNA is localized to the posterior pole of the Drosophila oocyte. *Cell* 66, 23–35.
- Kinthead R, LeBlanc M, Gulemetova R, Lalancette-Hébert M, Lemieux M, Mandeville I, Jeannotte L. (2004) Respiratory adaptations to lung morphological defects in adult mice lacking Hoxa5 gene function. *Pediatr Res.* 56, 553-562.
- Klar M, Bode J. (2005) Enhanceosome formation over the beta interferon promoter underlies a remote-control mechanism mediated by YY1 and YY2. *Mol Cell Biol.* 25(22): 10159-70.
- Kloosterman WP, Steiner FA, Berezikov E, de Bruijn E, van de Belt J, Verheul M, Cuppen E, Plasterk RH. (2006) Cloning and expression of new microRNAs from zebrafish. *Nucleic Acids Res.* 34(9): 2558-69.
- Klug J, Beato M. (1996) Binding of YY1 to a site overlapping a weak TATA box is essential for transcription from the uteroglobin promoter in endometrial cells. *Mol Cell Biol.* 16(11): 6398-407
- Kmita M, Fraudeau N, Herault Y, Duboule D. (2002a) Serial deletions and duplications suggest a mechanism for the collinearity of Hoxd genes in limbs. *Nature* 420, 145-150.
- Kmita M, Kondo T, Duboule D. (2000) Targeted inversion of a polar silencer within the HoxD complex re-allocates domains of enhancer sharing. *Nat. Genet.* 26, 451-454.

- Kmita M, Tarchini B, Zákány J, Logan M, Tabin CJ, Duboule D. (2005) Early developmental arrest of mammalian limbs lacking HoxA/HoxD gene function. *Nature* 435, 1113-6.
- Kmita M, van Der Hoeven F, Zákány J, Krumlauf R, Duboule D. (2000b) Mechanisms of Hox gene colinearity: transposition of the anterior Hoxb1 gene into the posterior HoxD complex. *Genes Dev* 14, 198-211.
- Knezevic V, De Santo R, Schughart K, Huffstadt U, Chiang C, Mahon KA, Mackem S. (1997) Hoxd-12 differentially affects preaxial and postaxial chondrogenic branches in the limb and regulates Sonic hedgehog in a positive feedback loop. *Development* 124, 4523-4536.
- Knipple DC, Seifert E, Rosenberg UB, Preiss A, Jackle H. (1985) Spatial and temporal patterns of Kruppel gene expression in early Drosophila embryos. *Nature* 317, 40-44.
- Knittel T, Kessel M, Kim MH, Gruss T. (1995) A conserved enhancer of the human and murine *Hoxa-7* genes specifies the anterior boundary of expression during embryonal development. *Development* 121: 1077-1088.
- Kobrossy L, Featherstone M (2007) Chromatin and the control of Hox gene expression. *Hox gene expression* 91-101. DOI: 10.1007/978-0-387-68990-6_6
- Kobrossy L, Rastegar M, Featherstone M. (2006) Interplay between chromatin and transacting factors regulating the Hoxd4 promoter during neural differentiation. *Biol Chem.* 281(36):25926-39.
- Kogo R, Shimamura T, Mimori K, Kawahara K, Imoto S, Sudo T, Tanaka F, Shibata K, Suzuki A, Komune S, Miyano S, Mori M. (2011) Long noncoding RNA HOTAIR regulates polycomb-dependent chromatin modification and is associated with poor prognosis in colorectal cancers. *Cancer Res.* 71, 6320-6326.
- Kolb A (2003) The first intron of the murine b-casein gene contains a functional promoter. *Biochem Biophys Res Commun.* 306(4):1099-105.
- Kondo T, Dolle P, Zákány J, Duboule D. (1996) Function of posterior HoxD genes in the morphogenesis of the anal sphincter. *Development* 122, 2651-2659.
- Kondo T, Hérault Y, Zákány J, Duboule D. (1998a) Genetic control of murine limb morphogenesis: relationships with human syndromes and evolutionary relevance. *Mol Cell Endocrinol* 140, 3-8.
- Kondrashov N, Pusic A, Stumpf CR, Shimizu K, Hsieh AC, Xue S, Ishijima J, Shiroishi T, Barna M. (2011) Ribosome-Mediated Specificity in Hox mRNA Translation and Vertebrate Tissue Patterning. *Cell* 145(3): 383-97.
- Kosak ST, Groudine M. (2004) Gene order and dynamic domains. *Science* 306(5696):644-7.
- Kostic D, Capecchi MR. (1994) Targeted disruptions of the murine Hoxa-4 and Hoxa-6 genes result in homeotic transformations of components of the vertebral column. *Mech Dev* 46, 231-47.
- Kuziora MA, McGinnis W. (1988). Autoregulation of a Drosophila homeotic selector gene. *Cell.* 55, 477-85.
- Krieger KE, Abbott MA, Joksimovic M, Lueth PA, Sonea IM, Jeannotte L, Tuggle CK. (2004). Transgenic mice ectopically expressing HOXA5 in the dorsal spinal cord show structural defects of the cervical spinal cord along with sensory and motor defects of the forelimb. *Brain Res Dev Brain Res.* 150, 125-39.

- Krause HM, Klemenz R, Gehring WJ. (1988) Expression, modification, and localization of the fushi tarazu protein in Drosophila embryos. *Genes Dev.* 2, 1021–1036.
- Krumlauf R. (1994) Hox genes in vertebrate development. *Cell* 78, 191-201.
- Krumlauf R, Holland PW, McVey JH, Hogan BL. (1987) Developmental and spatial patterns of expression of the mouse homeobox gene, Hox 2.1. *Development* 99(4):603-17.
- Kurihara Y, Watanabe Y. (2004) Arabidopsis micro-RNA biogenesis through Dicer-like 1 protein functions. *PNAS* 101, 12753-8.
- 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(1): e1003184.
- Lafontaine CA, Grainger S, Hess BL, Béland M, Lohnes D. (2012) Cdx1 interacts physically with a subset of Hox proteins. *Biochemistry* 51(48): 9698-705.
- Lampe X, Picard JJ, Rezsöházy R. (2004). The Hoxa2 enhancer 2 contains a critical Hoxa2 responsive regulatory element. *Biochem Biophys Res Commun* 316, 898-902.
- Lance-Jones C, Omelchenko N, Bailis A, Lynch S, Sharma K. (2001) Hoxd10 induction and regionalization in the developing lumbosacral spinal cord. *Development* 128, 2255–2268.
- Landeira D, Sauer S, Poot R, Dvorkina M, Mazzarella L, Jørgensen HF, Pereira CF, Leleu M, Piccolo FM, Spivakov M, Brookes E, Pombo A, Fisher C, Skarnes WC, Snoek T, Bezstarosti K, Demmers J, Klose RJ, Casanova M, Tavares L, Brockdorff N, Merckenschlager M, Fisher AG. (2010) Jarid2 is a PRC2 component in embryonic stem cells required for multi-lineage differentiation and recruitment of PRC1 and RNA Polymerase II to developmental regulators. *Nat Cell Biol.* 12(6): 618-24.
- Langlais KK, Brown JL, Kassis JA. (2012) Polycomb group proteins bind an engrailed PRE in both the "ON" and "OFF" transcriptional states of engrailed. *PLoS One* 7(11): e48765.
- Langston AW, Gudas LJ. (1992) Identification of a retinoic acid responsive enhancer 3' of the murine homeobox gene Hox-1.6. *Mech Dev* 38(3):217-227.
- Larochelle C, Tremblay M, Bernier D, Aubin J, Jeannotte L. (1999). Multiple cis-acting regulatory regions are required for restricted spatio-temporal Hoxa5 gene expression. *Dev Dyn.* 214, 127-40.
- LaRonde-LeBlanc NA, Wolberger C. (2003) Structure of HoxA9 and Pbx1 bound to DNA: Hox hexapeptide and DNA recognition anterior to posterior. *Genes Dev* 17, 2060-72.
- Laughon A, Boulet AM, Bermingham JR Jr, Laymon RA, Scott MP. (1986) Structure of transcripts from the homeotic Antennapedia gene of Drosophila melanogaster: two promoters control the major protein-coding region. *Mol Cell Biol.* 6(12): 4676-89.
- Laursen KB, Mongan NP, Zhuang Y, Ng MM, Benoit YD, Gudas LJ. (2013) Polycomb recruitment attenuates retinoic acid-induced transcription of the bivalent NR2F1 gene. *Nucleic Acids Res.* 41(13): 6430-43.
- Lawrence HJ, Christensen J, Fong S, Hu YL, Weissman I, Sauvageau G, Humphries RK, Largman C. (2005) Loss of expression of the Hoxa-9 homeobox gene impairs the

- proliferation and repopulating ability of hematopoietic stem cells. *Blood* 106, 3988-3994.
- Lawrence PA (1992) *The Making of a Fly: The Genetics of Animal Design*, Blackwell Scientific Publications, Oxford.
- Lee JS, Galvin KM, See RH, Eckner R, Livingston D, Moran E, Shi Y. (1995) Relief of YY1 transcriptional repression by adenovirus E1A is mediated by E1A-associated protein p300. *Genes Dev* 9, 1188–1198.
- Lee Y, Kim M, Han J, Yeom KH, Lee S, Baek SH, Kim VN. (2004) MicroRNA genes are transcribed by RNA polymerase II. *EMBO J* 23, 4051-60.
- Lee SW, Youn H, Kim EJ, Um SJ. (2013) Histone H2B Ubiquitination Regulates Retinoic Acid Signaling through the Cooperation of ASXL1 and BAP1. *Mol Cell*. 51(2):200-10.
- Lee TC, Zhang Y, Schwartz RJ. (1994) Bifunctional transcriptional properties of YY1 in regulating muscle actin and c-myc gene expression during myogenesis. *Oncogene* 9(4): 1047-52.
- Lemons D, Paré A, McGinnis W. (2012) Three Drosophila Hox Complex microRNAs Do Not Have Major Effects on Expression of Evolutionarily Conserved Hox Gene Targets during Embryogenesis. *PLoS One* 7(2): e31365.
- Le Mouellic H, Lallemand Y, Brulet P. (1992) Homeosis in the mouse induced by a null mutation in the Hox-3.1 gene. *Cell* 69, 251-64.
- Lengerke C, Schmitt S, Bowman TV, Jang IH, Maouche-Chretien L, McKinney-Freeman S, Davidson AJ, Hammerschmidt M, Rentzsch F, Green JB, Zon LI, Daley GQ. (2008) BMP and Wnt specify hematopoietic fate by activation of the Cdx-Hox pathway. *Cell Stem Cell* 2, 72–82
- Levin M, Pagan S, Roberts DJ, Cooke J, Kuehn MR, Tabin CJ. (1997) Left/right patterning signals and the independent regulation of different aspects of situs in the chick embryo. *Dev. Biol.* 189, 57-67.
- Lewis EB. (1978) A gene complex controlling segmentation in Drosophila. *Nature*. 276, 565-70.
- Lewis MT. (2000). Homeobox genes in mammary gland development and neoplasia. *Breast Cancer Research*, 2, 158–169.
- Lewis RA, Kaufman T, Denell RE, Tillerico P. (1980a) Genetic analysis of the Antennapedia Gene Complex (ANT-C) and adjacent chromosomal regions of *Drosophila melanogaster*. I. Polytene chromosome segments 84B-D. *Genetics* 95, 367–381.
- Lewis BA, Tullis G, Seto E, Norikoshi N, Weinmann R, Shenk T. (1995) Adenovirus E1A proteins interact with the cellular YY1 transcription factor. *J Virol*. 69(3): 1628-36.
- Lewis RA, Wakimoto BT, Denell RE, Kaufman T. (1980b) Genetic analysis of the Antennapedia Gene Complex (ANT-C) and adjacent chromosomal regions of *Drosophila melanogaster*. II. Polytene chromosome segments 84A-84B1,2. *Genetics* 95, 383–397.
- Li D, Feng J, Wu T, Wang Y, Sun Y, Ren J, Liu M. (2013) Long intergenic noncoding RNA HOTAIR is overexpressed and regulates PTEN methylation in laryngeal squamous cell carcinoma. *Am. J. Pathol.* 182, 64–70.

- Li G, Margueron R, Ku M, Chambon P, Bernstein BE, Reinberg D. (2010) Jarid2 and PRC2, partners in regulating gene expression. *Genes Dev.* 24, 368–380.
- Li X, McGinnis, W. (1999) Activity regulation of Hox proteins, a mechanism for altering functional specificity in development and evolution. *PNAS* 96, 6802-7.
- Li X, Murre C, McGinnis W. (1999) Activity regulation of a Hox protein and a role for the homeodomain in inhibiting transcriptional activation. *EMBO J.* 18, 198–211.
- Liang Z, Biggin MD. (1998) Eve and ftz regulate a wide array of genes in blastoderm embryos: the selector homeoproteins directly or indirectly regulate most genes in Drosophila [published erratum appears in Development 1999 Feb;126(5):following table of contents]. *Development* 125, 4471–4482.
- Lin YH, Kakadia PM, Chen Y, Li YQ, Deshpande AJ, Buske C, Zhang KL, Zhang Y, Xu GL, Bohlander SK. (2009) Global reduction of the epigenetic H3K79 methylation mark and increased chromosomal instability in CALM-AF10–positive leukemias. *Blood* 114(3):651-8
- Lin YW, Slape C, Zhang Z, Aplan PD. (2005) NUP98-HOXD13 transgenic mice develop a highly penetrant, severe myelodysplastic syndrome that progresses to acute leukemia. *Blood* 106, 287-295.
- Liu JP, Laufer E, Jessell TM. (2001) Assigning the positional identity of spinal motor neurons: rostrocaudal patterning of Hox-c expression by FGFs, Gdf11, and retinoids. *Neuron* 32, 997-1012.
- Liu XH, Lu KH, Wang KM, Sun M, Zhang EB, Yang JS, Yin DD, Liu ZL, Zhou J, Liu ZJ, De W, Wang ZX. (2012) MicroRNA-196a promotes non-small cell lung cancer cell proliferation and invasion through targeting HOXA5. *BMC Cancer* 12, 348.
- Lohnes D, Kastner P, Dierich A, Mark M, LeMeur M, Chambon P. (1993) Function of the retinoic acid receptor γ in the mouse. *Cell* 73, 643-658.
- Longobardi E, Iotti G, Di Rosa P, Meietta S, Nuciforo P, Ponzoni M, Doglioni C, Caniatti P, Bianchi F, Di Fiore PP, Blasi F. (2010) The homeodomain transcription factor *Prep1* gene (*pKnox1*) is a haploinsufficient tumor suppressor in man and mice. *Mol Oncol* 4:226-234.
- Lopez A, Higuete D, Rosset R, Deutsch J, Peronnet F. (2001) *corto* genetically interacts with *Pc-G* and *trx-G* genes and maintains the anterior boundary of *Ultrabithorax* expression in Drosophila larvae. *Mol Genet Genomics* 266(4):572-583.
- Lorente M, Pérez C, Sánchez C, Donohoe M, Shi Y, Vidal M. (2006) Homeotic transformations of the axial skeleton of YY1 mutant mice and genetic interaction with the Polycomb group gene *Ring1/Ring1A*. *Mech Dev.* 123(4): 312-20.
- Lufkin T. (1996) Transcriptional control of Hox genes in the vertebrate nervous system. *Curr Opin Genet Dev* 6, 575-80.
- Lufkin T, Dierich A, LeMeur M, Mark M, Chambon P. (1991) Disruption of the Hox-1.6 homeobox gene results in defects in a region corresponding to its rostral domain of expression. *Cell* 66, 1105-19.
- Luke MP, Sui G, Liu H, Shi Y. (2006) Yin Yang 1 physically interacts with Hoxa11 and represses Hoxa11-dependent transcription. *J Biol Chem.* 281(44): 33226-32.
- Lumsden A, Krumlauf R. (1996) Patterning the vertebrate neuraxis. *Science* 274, 1109-15.

- Luria V, Krawchuk D, Jessell TM, Laufer E, Kania A. (2008) Specification of motor axon trajectory by ephrin-B:EphB signaling: symmetrical control of axonal patterning in the developing limb. *Neuron* 60 (6): 1039-53
- Macdonald PM, Struhl G. (1986) A molecular gradient in early Drosophila embryos and its role in specifying the body pattern. *Nature* 324, 537–545.
- Mace KA, Hansen SL, Myers C, Young DM, Boudreau N. (2005) HOXA3 induces cell migration in endothelial and epithelial cells promoting angiogenesis and wound repair. *J Cell Sci.* 118(Pt 12), 2567-77
- Mack JA, Abramson SR, Ben Y, Coffin JC, Rothrock JK, Maytin EV, Hascall VC, Largman C, Stelnicki EJ. (2003). Hoxb13 knockout adult skin exhibits high levels of hyaluronan and enhanced wound healing. *Faseb J* 17, 1352-4.
- Mack, J. A., Li, L., Sato, N., Hascall, V. C. and Maytin, E. V. (2005). Hoxb13 up-regulates transglutaminase activity and drives terminal differentiation in an epidermal organotypic model. *J Biol Chem* 280, 29904-11.
- Maconochie M, Nonchev S, Morrison A., Krumlauf, R. (1996) Paralogous Hox genes: Function and regulation. *Annu. Rev. Genet.* 30, 529–556.
- Maconochie MK, Nonchev S, Studer M, Chan SK, Pöpperl H, Sham MH, Mann RS, Krumlauf R. (1997) Cross-regulation in the mouse HoxB complex: the expression of Hoxb2 in rhombomere 4 is regulated by Hoxb1. *Genes Dev* 11, 1885-95.
- Maden M, Holder M. (1992). Retinoic acid and development of the central nervous system. *Bioessays* 74, 431-438.
- Maezawa Y, Binnie M, Li C, Thorner P, Hui CC, Alman B, Taketo MM, Quaggin SE. (2012) A new Cre driver mouse line, Tcf21/Pod1-Cre, targets metanephric mesenchyme. *PLoS One* 7(7): e40547.
- Mahdipour E. (2013) Regulatory crosstalk between Hox genes and miRNAs during angiogenesis. *Microvasc Res.* 87,1-6.
- Mahmoudi T, Zuijderduijn LM, Mohd-Sarip A, Verrijzer CP. (2003) GAGA facilitates binding of Pleiohomeotic to a chromatinized Polycomb response element. *Nucleic Acids Res.* 31(14):4147-56.
- Mahony S, Mazzone EO, McCuine S, Young RA, Wichterle H, Gifford DK. (2011) Ligand dependent dynamics of retinoic acid receptor binding during early neurogenesis. *Genome Biol.* 12(1):R2.
- Mainguy G, Koster J, Woltering J, Jansen H, Durston A. (2007) Extensive polycistronism and antisense transcription in the Mammalian Hox clusters. *PLoS ONE* 2(4): e356.
- Mallo M, Vinagre T, Carapuco M. (2009) The road to the vertebral formula. *Int J Dev Biol* 53:1469–81.
- Mallo M, Wellik DM, Deschamps J. (2010) Hox genes and regional patterning of the vertebrate body plan. *Dev Biol* 344:7–15.
- Malpel S, Mendelsohn C, Cardoso WV. (2000) Regulation of retinoic acid signaling during lung morphogenesis. *Development* 127, 3057– 3067.
- Mandeville I, Aubin J, LeBlanc M, Lalancette-Hébert M, Janelle MF, Tremblay GM, Jeannotte L. (2006) Impact of the loss of *Hoxa5* function on lung alveogenesis. *Am J Physiol.* 169, No 4 1312-1327.
- Manley NR, Capecchi MR. (1995) The role of Hoxa-3 in mouse thymus and thyroid development. *Development* 121, 1989-2003.

- Manley NR, Capecchi MR. (1997) Hox group 3 paralogous genes act synergistically in the formation of somitic and neural crest-derived structures. *Dev Biol* 192, 274-88.
- Manley NR, Capecchi MR. (1998) Hox group 3 paralogs regulate the development and migration of the thymus, thyroid, and parathyroid glands. *Dev Biol* 195, 1-15.
- Mann RS, Hogness DS. (1990). Functional dissection of Ultrabithorax proteins in *D. melanogaster*. *Cell* 60, 597-610.
- Mansfield JH, Harfe BD, Nissen R, Obenaus J, Srineel J, Chaudhuri A, Farzan-Kashani R, Zuker M, Pasquinelli AE, Ruvkun G, Sharp PA, Tabin CJ, McManus MT. (2004) MicroRNA-responsive 'sensor' transgenes uncover Hox-like and other developmentally regulated patterns of vertebrate microRNA expression. *Nat Genet.* 36(10): 1079-83.
- Manzanares M, Bel-Vialar S, Ariza-McNaughton L, Ferretti E, Marshall H, Maconochie MM, Blasi F, Krumlauf R. (2001) Independent regulation of initiation and maintenance phases of *Hoxa3* expression in the vertebrate hindbrain involve auto- and cross-regulatory mechanisms. *Development* 128, 3595-607.
- Margalit Y, Yarus S, Shapira E, Gruenbaum Y, Fainsod A. (1993) Isolation and characterization of target sequences of the chicken *CdxA* homeobox gene. *Nucleic Acids Res* 21, 4915-22.
- Marshall H, Nonchev S, Sham MH, Muchamore I, Lumsden A, Krumlauf R. (1992) Retinoic acid alters the hindbrain Hox code and induces the transformation of rhombomeres 2/3 into a rhombomere 4/5 identity. *Nature* 360, 737-741.
- Marshall H, Studer M, Pöpperl H, Aparicio S, Kuroiwa A, Brenner S, Krumlauf R. (1994). A conserved retinoic acid response element required for early expression of the homeobox gene *Hoxb-1*. *Nature* 370, 567-71.
- Martinez-Arias A, Ingham PW, Scott MP, Akam ME. (1987) The spatial and temporal deployment of *Dfd* and *Scr* transcripts throughout development of *Drosophila*. *Development* 100, 673-683.
- Martinez-Arias A, Lawrence PA. (1985) Parasegments and compartments in the *Drosophila* embryo. *Nature*. 313, 639-42.
- Masliyah-Planchon J, Pasmant E, Luscan A, Laurendeau I, Ortonne N, Hivelin M, Varin J, Valeyrie-Allanore L, Dumaine V, Lantieri L, Leroy K, Parfait B, Wolkenstein P, Vidaud M, Vidaud D, Bièche I. (2013) MicroRNAome profiling in benign and malignant neurofibromatosis type 1-associated nerve sheath tumors: evidences of PTEN pathway alterations in early NF1 tumorigenesis. *BMC Genomics* 14, 473.
- Matis C, Oury F, Remacle S, Lampe X, Gofflot F, Picard JJ, Rijli FM, Rezsöházy R. (2007) Identification of *Lmo1* as part of a Hox-dependent regulatory network for hindbrain patterning. *Dev. Dyn.* 236, 2675-2684.
- Mazzoni EO, Mahony S, Peljto M, Patel T, Thornton SR, McCuine S, Reeder C, Boyer LA, Young RA, Gifford DK, Wichterle H. (2013) Saltatory remodeling of Hox chromatin in response to rostrocaudal patterning signals. *Nat Neurosci.* 16(9):1191-8.
- McGinnis W, Krumlauf R. (1992) Homeobox genes and axial patterning. *Cell* 68, 283-302.

- McGinnis W, Levine MS, Hafen E, Kuroiwa A, Gehring WJ. (1984) A conserved DNA sequence in homoeotic genes of the *Drosophila* Antennapedia and bithorax complexes. *Nature*. 308, 428-33.
- McGlenn E, Yekta S, Mansfield JH, Soutschek J, Bartel DP, Tabin CJ. (2009) In ovo application of antagomiRs indicates a role for miR-196 in patterning the chick axial skeleton through Hox gene regulation. *PNAS* 106(44): 18610-5.
- McIntyre DC, Rakshit S, Loken L, Jeannotte L, Capecchi MR, Wellik DM. (2007) Hox patterning of the rib cage: Loss of Hox paralogous group 5, 6 and 9 results in anterior homeotic vertebral transformations and non-collinear sternal defects. *Development* 134, 2981-2989
- McNulty CL, Peres JN, Bardine N, van den Akker WM, Durston AJ. (2005) Knockdown of the complete Hox paralogous group 1 leads to dramatic hindbrain and neural crest defects. *Development* 132, 2861-71.
- Meccia E, Bottero L, Felicetti F, Peschle C, Colombo MP, Carè A. (2003) HOXB7 expression is regulated by the transcription factors NF-Y, YY1, Sp1 and USF-1. *Biochim Biophys Acta*. 1626(1-3): 1-9.
- Medina-Martinez O, Bradley A, Ramírez-Solis R. (2000). A large targeted deletion of Hoxb1-Hoxb9 produces a series of single-segment anterior homeotic transformations. *Dev Biol* 222, 71-83.
- Mendenhall EM, Koche RP, Truong T, Zhou VW, Issac B, Chi AS, Ku M, Bernstein BE. (2010) GC-rich sequence elements recruit PRC2 in mammalian ES cells. *PLoS Genet*. 6(12):e1001244.
- Merrill VK, Turner FR, Kaufman TC. (1987) A genetic and developmental analysis of mutations in the Deformed locus in *Drosophila melanogaster*. *Dev. Biol.* 122, 379-395.
- Meunier D, Aubin J, Jeannotte L. (2003) Perturbed thyroid morphology and transient hypothyroidism symptoms in Hoxa5 mutant mice. *Dev Dyn*. 227, 367-78.
- Meyer BI, Gruss P. (1993) Mouse Cdx-1 expression during gastrulation. *Development* 117, 191-203.
- Mic FA, Haselbeck RJ, Cuenca AE, Duester G. (2002) Novel retinoic acid generating activities in the neural tube and heart identified by conditional rescue of Raldh2 null mutant mice. *Development* 129, 2271-2282.
- Milne TA, Briggs SD, Brock HW, Martin ME, Gibbs D, Allis CD, Hess JL. (2002) MLL targets SET domain methyltransferase activity to Hox gene promoters. *Mol Cell* 10 (5):1107-17.
- Minoo P, King RJ. (1994) Epithelial-mesenchymal interaction in lung development. *Annu Rev Physiol* 56, 13-45.
- Misteli T. (2004) Spatial positioning; a new dimension in genome function. *Cell* 119(2):153-6.
- Miura S, Nozawa M, Nei M. (2011) Evolutionary Changes of the Target Sites of Two MicroRNAs Encoded in the Hox Gene Cluster of *Drosophila* and Other Insect Species. *Genome Biol Evol*. 3, 129-39.
- Mlodzik M, Gehring WJ. (1987) Expression of the caudal gene in the germ line of *Drosophila*: formation of an RNA and protein gradient during early embryogenesis. *Cell* 48, 465-47
- Moens, CB, Prince VE. (2002) Constructing the hindbrain: Insights from the zebrafish.

- Dev. Dyn.* 224, 1–17.
- Moens CB, Selleri L. (2006). Hox cofactors in vertebrate development. *Dev Biol* 291, 193-206.
- Mohan M, Lin C, Guest E, Shilatifard A. (2010) Licensed to elongate: a molecular mechanism for MLL-based leukaemogenesis. *Nat Rev Cancer.* 10(10): 721-8.
- Mohd-Sarip A, Venturini F, Chalkley GE, Verrijzer CP. (2002) Pleiohomeotic can link polycomb to DNA and mediate transcriptional repression. *Mol Cell Biol.* 22(21):7473-83.
- Mohler J, Eldon ED, Pirrotta V. (1989) A novel spatial transcription pattern associated with the segmentation gene, *giant*, of *Drosophila*. *EMBO J.* 8, 1539–1558.
- Mollard R, Ghyselinck NB, Wendling O, Chambon P, Mark M. (2000) Stage-dependent responses of the developing lung to retinoic acid signaling. *Int. J. Dev. Biol.* 44, 457– 462.
- Montavon T, Soshnikova N, Mascrez B, Joye E, Thevenet L, Splinter E, de Laat W, Spitz F, Duboule D. (2011) A regulatory archipelago controls Hox genes transcription in digits. *Cell* 147(5):1132-45.
- Moreau J, Jeannotte L. (2002) Sequence analysis of a Hoxa4-Hoxa5 intergenic region including shared regulatory elements. *DNA Seq.* 13, 203-9.
- Morris-Kay G. (1993) Retinoic acid and craniofacial development: molecules and morphogenesis. *Bioessays* 75, 9-15.
- Morrison A, Moroni MC, Ariza-McNaughton L, Krumlauf R, Mavilio F. (1996) In vitro and transgenic analysis of a human HOXD4 retinoid-responsive enhancer. *Development* 122, 1895-907.
- Moskow JJ, Bullrich F, Huebner K, Daar IO, Buchberg AM. (1995) Meis1, a PBX1-related homeobox gene involved in myeloid leukemia in BXH-2 mice. *Mol Cell Biol* 5:5434-43.
- Mukherjee K, Burglin TR. (2007) Comprehensive analysis of animal TALE homeobox genes: new conserved motifs and cases of accelerated evolution. *J Mol Evol* 65:137-153.
- Mulero MC, Ferres-Marco D, Islam A, Margalef P, Pecoraro M, Toll A, Drechsel N, Charneco C, Davis S, Bellora N, Gallardo F, López-Arribillaga E, Asensio-Juan E, Rodilla V, González J, Iglesias M, Shih V, Mar Albà M, Di Croce L, Hoffmann A, Miyamoto S, Villà-Freixa J, López-Bigas N, Keyes WM, Domínguez M, Bigas A, Espinosa L. (2013) Chromatin-Bound IκBα Regulates a Subset of Polycomb Target Genes in Differentiation and Cancer. *Cancer Cell* 24(2):151-66.
- Muller J, Bienz M. (1991) Long range repression conferring boundaries of Ultrabithorax expression in the *Drosophila* embryo. *EMBO J.* 10, 3147--3155.
- Muller J, Bienz M. (1992) Sharp anterior boundary of homeotic gene expression conferred by the fushi tarazu protein. *EMBO J.* 11(10): 3653-61.
- Myers C, Charboneau A, Cheung I, Hanks D, Boudreau N. (2002) Sustained expression of homeobox D10 inhibits angiogenesis. *Am J Pathol.* 161(6):2099-109.
- Nakamura T. (2005) NUP98 fusion in human leukemia: dysregulation of the nuclear pore and homeodomain proteins. *Int J Hematol.* 82 (1):21-27.
- Nasiadka A, Grill A, Krause HM. (2000) Mechanisms regulating target gene selection by the homeodomain-containing protein Fushi tarazu. *Development* 127, 2965–2976.

- Nam J, Nei M. (2005) Evolutionary change of the numbers of homeobox genes in bilateral animals. *Mol Biol Evol.* 22, 2386-94.
- Nègre N, Brown CD, Shah PK, Kheradpour P, Morrison CA, Henikoff JG, Feng X, Ahmad K, Russell S, White RA. (2010) A comprehensive map of insulator elements for the *Drosophila* genome. *PLoS Genet.* 6:e1000814.
- Ng J, Hart CM, Morgan K, Simon JA. (2000) A *Drosophila* ESC-E(Z) protein complex is distinct from other polycomb group complexes and contains covalently modified ESC. *Mol Cell Biol.* 20(9): 3069-78.
- Nguyen N, Zhang X, Olashaw N, Seto E. (2004) Molecular Cloning and Functional Characterization of the Transcription Factor YY2. *J Biol Chem.* 279(24): 25927-34.
- Niederreither K, McCaffery P, Dräger UC, Chambon P, Dollé P. (1997) Restricted expression and retinoic acid-induced downregulation of the retinaldehyde dehydrogenase type 2 (RALDH-2) gene during mouse development. *Mech. Dev.* 62(1): 67-78.
- Niederreither K, Vermot J, Schuhbaur B, Chambon P, Dollé P. (2000) Retinoic acid synthesis and hindbrain patterning in the mouse embryo. *Development* 127, 75-85.
- Niinuma T, Suzuki H, Nojima M, Nosho K, Yamamoto H, Takamaru H, Yamamoto E, Maruyama R, Nobuoka T, Miyazaki Y, Nishida T, Bamba T, Kanda T, Ajioka Y, Taguchi T, Okahara S, Takahashi H, Nishida Y, Hosokawa M, Hasegawa T, Tokino T, Hirata K, Imai K, Toyota M, Shinomura Y. (2012) Upregulation of miR-196a and HOTAIR drive malignant character in gastrointestinal stromal tumors. *Cancer Res.* 72(5): 1126-36.
- Nolte C, Krumlauf R. (2006) Expression of Hox genes in the nervous system of vertebrates. In "HOX Gene Expression" (S. Papageorgiou, Ed.), pp. 14-41. Landes Bioscience and Springer, Austin, TX.
- Noordermeer D, Leleu M, Splinter E, Rougemont J, De Laat W, Duboule D. (2011) The Dynamic Architecture of Hox Gene Clusters. *Science* 334(6053):222-5.
- Nordström U, Maier E, Jessell TM, Edlund T. (2006) An early role for WNT signaling in specifying neural patterns of Cdx and Hox gene expression and motor neuron subtype identity. *PLoS Biol.* 4, e252.
- Nourse J, Mellentin JD, Galili N, Wilkinson J, Stanbridge E, Smith SD, Cleary ML. (1990) Chromosomal translocation t(1;19) results in synthesis of a homeobox fusion mRNA that codes for a potential chimeric transcription factor. *Cell* 60, 535-45.
- Novak RL, Harper DP, Caudell D, Slape C, Beachy SH, Aplan PD. (2012) Gene expression profiling and candidate gene resequencing identifies pathways and mutations important for malignant transformation caused by leukemogenic fusion genes. *Exp Hematol.* 40(12), 1016-27
- Nowicki JL, Burke AC. (2000). Hox genes and morphological identity: axial versus lateral patterning in the vertebrate mesoderm. *Development* 127, 4265-75.
- Nowling T, Zhou W, Krieger KE, Larochelle C, Nguyen-Huu MC, Jeannotte L, Tuggle CK. (1999). Hoxa5 gene regulation: A gradient of binding activity to a brachial spinal cord element. *Dev Biol.* 208, 134-46.
- Nusslein-Volhard C, Wieschaus E. (1980). Mutations affecting segment number and polarity in *Drosophila*. *Nature.* 287, 795-801.

- O'Carroll D, Erhardt S, Pagani M, Barton SC, Surani MA, Jenuwein T. (2001) The polycomb-group gene *Ezh2* is required for early mouse development. *Mol Cell Biol.* 21(13): 4330-6.
- Odenwald WF, Garbern J, Arnheiter H, Tournier-Lasserre E, Lazzarini RA. (1989) The Hox-1.3 homeo box protein is a sequence-specific DNA-binding phosphoprotein. *Genes Dev* 3, 158-72.
- Odenwald WF, Taylor CF, Palmer-Hill FJ, Friedrich V Jr, Tani M, Lazzarini RA. (1987) Expression of a homeo domain protein in noncontact-inhibited cultured cells and postmitotic neurons. *Genes Dev* 1, 482-96.
- Oei SL, Griesenbeck J, Schweiger M. (1997) The role of poly(ADP-ribosylation). *Rev Physiol Biochem Pharmacol* 131, 127–173.
- Oei SL, Griesenbeck J, Schweiger M, Ziegler M. (1998) Regulation of RNA polymerase II-dependent transcription by poly(ADP-ribosylation) of transcription factors. *J Biol Chem* 273, 31644–31647.
- Okada Y, Jiang Q, Lemieux M, Jeannotte L, Su L, Zhang Y. (2006) Leukaemic transformation by CALM–AF10 involves upregulation of *Hoxa5* by hDOT1L. *Nat Cell Biol.* 8(9):1017-24.
- Olson EN, Arnold HH, Rigby PW, Wold BJ. (1996) Know your neighbors: three phenotypes in null mutants of the myogenic bHLH gene *MRF4*. *Cell* 85, 1-4.
- Oosterveen T, Niederreither K, Dollé P, Chambon P, Meijlink F, Deschamps J. (2003a) Retinoids regulate the anterior expression boundaries of 5' *Hoxb* genes in posterior hindbrain. *Embo J.* 22, 262-9.
- Oriente F, Fernandez Diaz LC, Miele C, Iovino S, Mori S, Diaz VM, Troncone G, Cassese A, Formisano P, Blasi F and Beguinot F. (2008) *Prep1* deficiency induces protection from diabetes and increased insulin sensitivity through a p160a-mediated mechanism. *Mol Cell Biol* 28: 5634- 5645.
- Ota T, Choi KB, Gilks CB, Leung PCK, Auersperg N. (2006) Cell type- and stage-specific changes in *HOXA7* protein expression in human ovarian folliculogenesis: possible role of *GDF-9*. *Differentiation* 74, 1–10 (2006).
- Oury F, Murakami Y, Renaud JS, Pasqualetti M, Charnay P, Ren SY, Rijli FM. (2006) *Hoxa2*- and rhombomere-dependent development of the mouse facial somatosensory map. *Science* 313, 1408–1413.
- Packer, A. I., Crotty, D. A., Elwell, V. A. and Wolgemuth, D. J. (1998). Expression of the murine *Hoxa4* gene requires both autoregulation and a conserved retinoic acid response element. *Development* 125, 1991-8.
- Palmqvist L, Pineault N, Wasslavik C, Humphries RK. (2007) Candidate genes for expansion and transformation of hematopoietic stem cells by NUP98-HOX fusion genes. *PLoS One* 2:e768.
- Panganiban GE, Reuter R, Scott MP, Hoffmann FM. (1990) A *Drosophila* growth factor homolog, decapentaplegic, regulates homeotic gene expression within and across germ layers during midgut morphogenesis. *Development* 110, 1041-50.
- Papadopoulos DK, Skouloudaki K, Adachi Y, Samakovlis C, Gehring WJ. (2012) Dimer formation via the homeodomain is required for function and specificity of sex combs reduced in *Drosophila*. *Dev Biol.* 367(1):78-89.

- Papoulas O, Beek SJ, Moseley SL, McCallum CM, Sarte M, Shearn A, Tamkun JW. (1998) The *Drosophila trithorax* group proteins BRM, ASH1 and ASH2 are subunits of distinct protein complexes. *Development* 125(20):3955-66.
- Park K, Atchison ML. (1991) Isolation of a candidate repressor/activator, NF-E1 (YY-1, delta), that binds to the immunoglobulin kappa 3' enhancer and the immunoglobulin heavy-chain mu E1 site. *PNAS* 88(21): 9804-8.
- Paro R. (1990) Imprinting a determined state into the chromatin of *Drosophila*. *Trends Genet.* 6 (12):416-21.
- Parrish M., Nolte C, Krumlauf R. (2009) Hox genes expression. In “Encyclopedia of Neuroscience” (L.R. Squire, Ed.), Vol. 4, pp. 1221–1231. Academic Press, Oxford, UK.
- Partanen J, Schwartz L, Rossant J. (1998) Opposite phenotypes of hypomorphic and Y766 phosphorylation site mutations reveal a function for Fgfr1 in anteroposterior patterning of mouse embryos. *Genes Dev* 12, 2332-2344.
- Pasini D, Bracken AP, Jensen MR, Lazzarini Denchi E, Helin K. (2004) Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity. *EMBO J.* 23(20): 4061-71.
- Pattatucci AM, Otteson DC, Kaufman TC. (1991) A functional and structural analysis of the Sex combs reduced locus of *Drosophila melanogaster*. *Genetics.* 129, 423-41.
- Patterson LT, Pembaur M, Potter SS. (2001). Hoxa11 and Hoxd11 regulate branching morphogenesis of the ureteric bud in the developing kidney. *Development* 128, 2153-61.
- Patterson LT, Potter SS. (2004) Atlas of Hox gene expression in the developing kidney. *Dev Dyn* 229, 771–779
- Pazin MJ, Kadonaga JT. (1997) What's up and down with histone deacetylation and transcription? *Cell* 89(3): 325-8.
- Pearson JC, Lemons D, McGinnis W. (2005) Modulating Hox gene functions during animal body patterning. *Nat. Rev. Genet.* 6, 893–904
- Pederson JD, Kiehart DP, Mahaffey JW. (1996) The role of HOM-C genes in segmental transformations: reexamination of the *Drosophila* Sex combs reduced embryonic phenotype. *Dev Biol.* 180, 131-42.
- Peichel CL, Prabhakaran B, Vogt TF. (1997) The mouse *Ulnaless* mutation deregulates posterior HoxD gene expression and alters appendicular patterning. *Development* 124, 3481-92.
- Peifer M, Karch F, Bender W. (1987). The Bithorax complex: control of segment identity. *Genes Dev* 1, 891-898.
- Peinado H, Olmeda D and Cano A. (2007) Snail, Zeb and bHLH factors in tumour progression: an alliance against the epithelial phenotype? *Nat Rev Cancer* 7, 415-428.
- Peljto M, Dasen JS, Mazzoni EO, Jessell TM, Wichterle H. (2010) Functional diversity of ESC-derived motor neuron subtypes revealed through intraspinal transplantation. *Cell Stem Cell* 7, 355–366.
- Peng JC, Valouev A, Swigut T, Zhang J, Zhao Y, Sidow A, Wysocka J. (2009) Jarid2/Jumonji coordinates control of PRC2 enzymatic activity and target gene occupancy in pluripotent cells. *Cell* 139(7): 1290-302.

- Penkov D, Mateos San Martin D, Fernandez-Diaz LC, Rosello CA, Torroja C, Sanchez-Cabo F, Warnatz HJ, Sultan M, Yaspo ML, Gabrieli A, Tkachuk V, Brendolan A, Blasi F, Torres M. (2013) Analysis of the DNA-Binding Profile and Function of TALE Homeoproteins Reveals Their Specialization and Specific Interactions with Hox Genes/Proteins, *Cell Reports* 25:1321-33
- Perrimon N, Mahowald AP. (1987) Multiple functions of segment polarity genes in *Drosophila*. *Dev. Biol.* 119, 587–600.
- Perrimon N, Smouse D. (1989) Multiple functions of a *Drosophila* homeotic gene, *zeste-white 3*, during segmentation and neurogenesis. *Dev. Biol.* 135, 287–305.
- Peterson CL, Tamkun JW. (1995) The SWI-SNF complex: a chromatin remodeling machine. *Trends Biochem Sci.* 20(4):143-6.
- Petruk S, Sedkov Y, Smith S, Tillib S, Kraevski V, Nakamura T, Canaani E, Croce CM, Mazo A. (2001) Trithorax and dCBP acting in a complex to maintain expression of a homeotic gene. *Science* 294(5545): 1331-4.
- 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(12): 1636-44.
- Phua SL, Sivakamasundari V, Shao Y, Cai X, Zhang LF, Lufkin T, Featherstone M. (2011) Nuclear accumulation of an uncapped RNA produced by Drosha cleavage of a transcript encoding miR-10b and HOXD4. *PLoS One* 6(10):e25689.
- Pignoni F, Baldarelli RM, Steingrimsson E, Diaz RJ, Patapoutian A, Merriam JR, Lengyel JA. (1990) The *Drosophila* gene *tailless* is expressed at the embryonic termini and is a member of the steroid receptor superfamily. *Cell* 62, 151–163.
- Pilon N, Oh K, Sylvestre JR, Bouchard N, Savory J, Lohnes D. (2006) *Cdx4* is a direct target of the canonical Wnt pathway. *Dev Biol* 289, 55-63.
- Pineault N, Helgason CD, Lawrence HJ, Humphries RK. (2002) Differential expression of *Hox*, *Meis1*, and *Pbx1* genes in primitive cells throughout murine hematopoietic ontogeny. *Experimental Hematology* 30, 49-57.
- Pitera JE, Smith VV, Thorogood P, Milla PJ. (1999). Coordinated expression of 3' *hox* genes during murine embryonal gut development: an enteric *Hox* code. *Gastroenterology* 117, 1339-1351.
- Plaza S, Prince F, Adachi Y, Punzo C, Cribbs DL, Gehring WJ. (2008) Cross-regulatory protein-protein interactions between *Hox* and *Pax* transcription factors. *PNAS* 105(36): 13439-13444.
- Plummer PN, Freeman R, Taft RJ, Vider J, Sax M, Umer BA, Gao D, Johns C, Mattick JS, Wilton SD, Ferro V, McMillan NA, Swarbrick A, Mittal V, Mellick AS. (2012) MicroRNAs regulate tumor angiogenesis modulated by endothelial progenitor cells. *Cancer Res.* 73(1): 341-52.
- Pollock RA, Jay G, Bieberich CJ. (1992) Altering the boundaries of *Hox3.1* expression: evidence for antipodal gene regulation. *Cell* 71, 911-923.
- Pollock RA, Sreenath T, Ngo L, Bieberich CJ. (1995) Gain of function mutations for paralogous *Hox* genes: implications for the evolution of *Hox* gene function. *PNAS* 92, 4492-6.
- Pöpperl H, Featherstone MS. (1993) Identification of a retinoic acid response element upstream of the murine *Hex-4.2* gene. *Mol. Cell. Biol.* 73, 257-265

- Pöpperl H, Bienz M, Studer M, Chan SK, Aparicio S, Brenner S, Mann RS, Krumlauf R. (1995). Segmental expression of Hoxb-1 is controlled by a highly conserved autoregulatory loop dependent upon *exd/pbx*. *Cell* 81, 1031-42.
- Pownall ME, Isaacs HV, Slack JM. (1998) Two phases of Hox gene regulation during early *Xenopus* development. *Curr Biol* 8, 673-6.
- Pownall ME, Tucker AS, Slack JM, Isaacs HV (1996) eFGF, *Xcad3* and *Hox* genes form a molecular pathway that establishes the anteroposterior axis in *Xenopus*. *Development* 122, 3881-3892.
- Radtke F, Wilson A, Stark G, Bauer M, van Meerwijk J, MacDonald HR, Aguet M. (1999) Deficient T cell fate specification in mice with an induced inactivation of Notch1. *Immunity* 10, 547-58.
- Raines AM, Adam M, Magella B, Meyer SE, Grimes HL, Dey SK, Potter SS. (2013) Recombineering based dissection of flanking and paralogous Hox gene functions in mouse reproductive tracts. *Development* 140(14):2942-52.
- Raman V, Martensen SA, Reisman D, Evron E, Odenwald WF, Jaffee E, Marks J, Sukumar S. (2000a) Compromised HOXA5 function can limit p53 expression in human breast tumours. *Nature* 405, 974-8.
- Raman V, Tamori A, Vali M, Zeller K, Korz D, Sukumar S. (2000b) HOXA5 regulates expression of the progesterone receptor. *J Biol Chem* 275, 26551-5.
- Ramírez-Solis R, Zheng H, Whiting J, Krumlauf R, Bradley A. (1993) Hoxb-4 (Hox-2.6) mutant mice show homeotic transformation of a cervical vertebra and defects in the closure of the sternal rudiments. *Cell* 73, 279-94
- Rancourt DE, Tsuzuki T, Capecchi MR. (1995) Genetic interaction between *hoxb-5* and *hoxb-6* is revealed by nonallelic noncomplementation. *Genes Dev* 9, 108-22.
- Rank G, Prestel M, Paro R. (2002) Transcription through intergenic chromosomal memory elements of the *Drosophila* bithorax complex correlates with an epigenetic switch. *Mol. Cell. Biol.* 22, 8026–8034.
- Rappailles A, Decoville M, Locker D. (2005) DSP1, a *Drosophila* HMG protein, is involved in spatiotemporal expression of the homeotic gene *Sex combs reduced*. *Biol Cell.* 97(10): 779-85.
- Rastegar M, Kobrossy L, Kovacs EN, Rambaldi I, Featherstone M. (2004) Sequential histone modifications at *Hoxd4* regulatory regions distinguish anterior from posterior embryonic compartments. *Mol Cell Biol.* 24(18): 8090-103.
- Ratovitski EA (2010) LKB1/PEA3/ Δ Np63 pathway regulates PTGS-2 (COX-2) transcription in lung cancer cells upon cigarette smoke exposure. *Oxid Med Cell Longev.* 3(5):317-24.
- Ren SY, Angrand PO, Rijli FM. (2002) Targeted insertion results in a rhombomere specific *Hoxa2* knockdown and ectopic activation of *Hoxa1* expression. *Dev Dyn* 225, 305-315.
- Reuter R, Panganiban GE, Hoffmann FM, Scott MP. (1990) Homeotic genes regulate the spatial expression of putative growth factors in the visceral mesoderm of *Drosophila* embryos. *Development* 110, 1031-40.
- Reuter R, Scott MP. (1990). Expression and function of the homeotic genes *Antennapedia* and *Sex combs reduced* in the embryonic midgut of *Drosophila*. *Development.* 109, 289-303.

- Rhoads K, Arderiu G, Charboneau A, Hansen SL, Hoffman W, Boudreau N. (2005) A role for Hox A5 in regulating angiogenesis and vascular patterning. *Lymphat Res Biol* 3, 240-52
- Riddle RD, Johnson RL, Laufer E, Tabin C. (1993) Sonic hedgehog mediates the polarizing activity of the ZPA. *Cell* 75, 1401-1416.
- Riggs KJ, Saleque S, Wong KK, Merrell KT, Lee JS, Shi Y, Calame K. (1993) Yin-yang 1 activates the c-myc promoter. *Mol Cell Biol*. 13(12):7487-95.
- Rieger E, Bijl JJ, van Oostveen JW, Soyer HP, Oudejans CB, Jiwa NM, Walboomers JM, Meijer CJ. (1994) Expression of the homeobox gene HOXC4 in keratinocytes of normal skin and epithelial skin tumors is correlated with differentiation. *J Invest Dermatol*. 103(3):341-6.
- Rijli FM, Chambon P. (1997) Genetic interactions of Hox genes in limb development: learning from compound mutants. *Curr Opin Genet Dev* 7, 481-7.
- Rijli FM, Dollé P, Fraulob V, LeMeur M, Chambon P. (1994) Insertion of a targeting construct in a Hoxd-10 allele can influence the control of Hoxd-9 expression. *Dev Dyn* 201, 366-377.
- Rijli FM, Mark M, Lakkaraju S, Dierich A, Dollé P, Chambon P. (1993) A homeotic transformation is generated in the rostral branchial region of the head by disruption of Hoxa-2, which acts as a selector gene. *Cell* 75, 1333-49.
- Rijli FM, Matyas R, Pellegrini M, Dierich A, Gruss P, Dollé P, Chambon P. (1995) Cryptorchidism and homeotic transformations of spinal nerves and vertebrae in Hoxa-10 mutant mice. *PNAS* 92, 8185-9.
- Riman S, Rizkallah R, Kassardjian A, Alexander KE, Luscher B, Hurt MM. (2012) Phosphorylation of the transcription factor YY1 by CK2alpha prevents cleavage by caspase 7 during apoptosis. *Mol Cell Biol* 32, 797-807.
- Ringrose L, Ehret H, Paro R. (2004) Distinct contributions of histone H3 lysine 9 and 27 methylation to locus-specific stability of polycomb complexes. *Mol Cell* 16(4):641-53.
- Ringrose L, Paro R. (2004) Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. *Annu Rev Genet*. 38,413-43.
- Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Bruggmann SA, Goodnough LH, Helms JA, Farnham PJ, Segal E, Chang HY. (2007) Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* 129, 1311-1323.
- Rizkallah R, Alexander KE, Hurt MM. (2011a) Global mitotic phosphorylation of C2H2 zinc finger protein linker peptides. *Cell Cycle* 10, 3327-3336.
- Rizkallah R, Alexander KE, Kassardjian A, Luscher B, Hurt MM. (2011b) The transcription factor YY1 is a substrate for Polo-like kinase 1 at the G2/M transition of the cell cycle. *PLoS One* 6: e15928.
- Rizkallah R, Hurt MM. (2009) Regulation of the transcription factor YY1 in mitosis through phosphorylation of its DNA-binding domain. *Mol Biol Cell* 20, 4766-4776.
- Rodini CO, Xavier FC, Paiva KB, De Souza Setúbal Destro MF, Moyses RA, Michaluart P, Carvalho MB, Fukuyama EE; Head And Neck Genome Project Gencapo, Tajara EH, Okamoto OK, Nunes FD. (2012) Homeobox gene

- expression profile indicates HOXA5 as a candidate prognostic marker in oral squamous cell carcinoma. *Int J Oncol.* 40(4): 1180-8.
- Roelen BA, de Graaff W, Forlani S, Deschamps J. (2002) Hox cluster polarity in early transcriptional availability: a high order regulatory level of clustered Hox genes in the mouse. *Mech Dev* 119, 81-90.
- Rossant J, Zirngibl R, Cado D, Shago M, Giguere V. (1991) Expression of a retinoic acid response element-hsplacZ transgene defines specific domains of transcriptional activity during mouse embryogenesis. *Genes Dev.* 5, 1333–1344.
- Rossel M, Capecchi MR. (1999) Mice mutant for both Hoxa1 and Hoxb1 show extensive remodeling of the hindbrain and defects in craniofacial development. *Development* 126, 5027-40.
- Rousseau M, Crutchley JL, Miura H, Suderman M, Blanchette M, Dostie J. (2014) Hox in motion: tracking HoxA cluster conformation during differentiation. *Nucleic Acids Res.* 42(3):1524-40.
- Ruddle FH, Bentley KL, Murtha MT, Risch N. (1994) Gene loss and gain in the evolution of the vertebrates. *Dev Suppl*, 155-61.
- Sade H, Holloway K, Romero IA, Male D. (2009) Transcriptional control of occludin expression in vascular endothelia: Regulation by Sp3 and YY1. *Biochim Biophys Acta.* 1789(3):175-84.
- Sakai Y, Meno C, Fujii H, Nishino J, Shiratori H, Saijoh Y, Rossant J, Hamada H. (2001) The retinoic acid-inactivating enzyme CYP26 is essential for establishing an uneven distribution of retinoic acid along the antero-posterior axis within the mouse embryo. *Genes Dev* 15, 213-25.
- Salvaing J, Decoville M, Mouchel-Vielh E, Bussièrè M, Daulny A, Boldyreva L, Zhimulev I, Locker D, Peronnet F. (2006) Corto and DSP1 interact and bind to a maintenance element of the *Scr Hox* gene: understanding the role of *Enhancers of trithorax and Polycomb*. *BMC Biol* 4, 9
- Sanchez-Elsner T, Gou D, Kremmer E, Sauer F. (2006) Noncoding RNAs of trithorax response elements recruit Drosophila Ash1 to Ultrabithorax. *Science* 311(5764): 1118-23.
- Sanchez-Herrero E, Vernas I, Marco R, Morata G. (1985) Genetic organization of *Drosophila* bithorax complex. *Nature* 313, 108–113.
- Santagati F, Minoux M, Ren SY, Rijli FM. (2005) Temporal requirement of Hoxa2 in cranial neural crest skeletal morphogenesis. *Development* 132, 4927-36.
- Santagati F, Rijli FM. (2003) Cranial neural crest and the building of the vertebrate head. *Nat Rev Neurosci* 4, 806-18.
- Sasaki YT, Sano M, Kin T, Asai K, Hirose T. (2007) Coordinated expression of ncRNAs and HOX mRNAs in the human HOXA locus. *Biochem Biophys Res Commun.* 357(3): 724-30.
- Satokata I, Uchiyama M. (1995) Mice carrying null mutations of the homeotic genes Hoxa10 and Msx1. *Tanpakushitsu Kakusan Koso* 40, 2134-43.
- Saurin AJ, Shao Z, Erdjument-Bromage H, Tempst P, Kingston RE. (2001) A Drosophila polycomb group complex includes Zeste and dTAFII proteins. *Nature* 412(6847): 655-60.
- Sauvageau G, Thorsteinsdottir U, Eaves CJ, Lawrence HJ, Largman C, Lansdorp PM, Humphries RK. (1995) Overexpression of HOXB4 in hematopoietic cells causes

- the selective expansion of more primitive populations in vitro and in vivo. *Genes Dev.* 9(14):1753-65.
- Sauvageau G, Thorsteinsdottir U, Hough MR, Hugo P, Lawrence HJ, Largman C, Humphries RK. (1997). Overexpression of HOXB3 in hematopoietic cells causes defective lymphoid development and progressive myeloproliferation. *Immunity* 6, 13-22.
- Savory JG, Bouchard N, Pierre V, Rijli FM, De Repentigny Y, Kothary R, Lohnes D. (2009) Cdx2 regulation of posterior development through non-Hox targets. *Development* 136(24):4099-110.
- Sawaya PL, Stripp BR, Whitsett JA, Luse DS (1993) The Lung-Specific CC10 Gene Is Regulated by Transcription Factors from the AP-1, Octamer, and Hepatocyte Nuclear Factor 3 Families. *Mol Cell Biol.* 13(7):3860-71.
- Schwalie PC, Ward MC, Cain CE, Faure AJ, Gilad Y, Odom DT, Flicek P. (2013) Co-binding by YY1 identifies the transcriptionally active, highly conserved set of CTCF-bound regions in primate genomes. *Genome Biol.* 14(12):R148.
- Schmitt S, Prestel M, Paro R. (2005) Intergenic transcription through a polycomb group response element counteracts silencing. *Genes Dev.* 19, 697-708
- Schorderet P, Duboule D. (2011) Structural and functional differences in the long non-coding RNA hotair in mouse and human. *PLoS Genet.* 7(5): e1002071.
- Schuettengruber B, Chourrout D, Vervoort M, Leblanc B, Cavalli G. (2007) Genome regulation by polycomb and trithorax proteins. *Cell* 128(4):735-45.
- Schuettengruber B, Martinez AM, Iovino N, Cavalli G. (2011) Trithorax group proteins: switching genes on and keeping them active. *Nat Rev Mol Cell Biol.* 12(12): 799-814.
- Schughart K, Bieberich CJ, Eid R, Ruddle FH. (1991) A regulatory region from the mouse Hox-2.2 promoter directs gene expression into developing limbs. *Development* 112, 807-11.
- Schwartz YB, Pirrotta V. (2007) Polycomb silencing mechanisms and the management of genomic programmes. *Nat Rev Genet.* 8(1): 9-22.
- Schwartz YB, Pirrotta V. (2008) Polycomb complexes and epigenetic states. *Curr Opin Cell Biol.* 20(3):266-73.
- Scott MP, Weiner AJ. (1984) Structural relationships among genes that control development: sequence homology between the Antennapedia, Ultrabithorax, and fushi tarazu loci of Drosophila. *Proc Natl Acad Sci U S A.* 81, 4115-9.
- Scott V, Morgan EA, Stadler HS. (2005) Genitourinary functions of Hoxa13 and Hoxd13. *J Biochem (Tokyo)* 137, 671-6.
- Scotti M, Kmita M. (2012) Recruitment of 5' Hoxa genes in the allantois is essential for proper extra-embryonic function in placental mammals. *Development* 139(4):731-9.
- Sekimoto T, Yoshinobu K, Yoshida M, Kuratani S, Fujimoto S, Araki M, Tajima N, Araki K, Yamamura K. (1998) Region-specific expression of murine Hox genes implies the Hox code-mediated patterning of the digestive tract. *Genes Cells* 3, 51-64.
- Selleri L, DiMartino J, Van Deursen J, Brendolan A, Sanyal M, Boon E, Capellini T, Smith KS, Rhee J, Popperl H, Grosveld G, Cleary ML. (2004) The TALE

- homeodomain protein Pbx2 is not essential for development and long-term survival. *Mol Cell Biol* 24:5324-5331.
- Shah N, Jin K, Cruz LA, Park S, Sadik H, Cho S, Goswami CP, Nakshatri H, Gupta R, Chang HY, Zhang Z, Cimino-Mathews A, Cope L, Umbricht C, Sukumar S. (2013) HOXB13 Mediates Tamoxifen Resistance and Invasiveness in Human Breast Cancer by Suppressing ER α and Inducing IL-6 Expression. *Cancer Res* DOI: 10.1158/0008-5472.CAN-13-1178
- Shao Z, Raible F, Mollaaghababa R, Guyon JR, Wu CT, Bender W, Kingston RE. (1999) Stabilization of chromatin structure by PRC1, a polycomb complex. *Cell* 98(1):37-46.
- Sharpe J, Nonchev S, Gould A, Whiting J, Krumlauf R. (1998) Selectivity, sharing and competitive interactions in the regulation of Hoxb genes. *Embo J* 17, 1788-98.
- Shashikant CS, Bieberich CJ, Belting HG, Wang JC, Borbely MA, Ruddle FH. (1995) Regulation of *Hoxc-8* during mouse embryonic development: identification and characterization of critical elements involved in early neural tube expression. *Development* 121, 4339-4347.
- Shaut, C. A., Keene, D. R., Sorensen, L. K., Li, D. Y. and Stadler, H. S. (2008). HOXA13 Is essential for placental vascular patterning and labyrinth endothelial specification. *PLoS Genet.* 4, e1000073.
- 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 Res* 24, 898-906.
- Shen X, Kim W, Fujiwara Y, Simon MD, Liu Y, Mysliwiec MR, Yuan GC, Lee Y, Orkin SH. (2009) Jumonji modulates Polycomb activity and self-renewal versus differentiation of stem cells. *Cell* 139(7):1303-14.
- Shen WF, Montgomery JC, Rozenfeld S, Moskow JJ, Lawrence HJ, Buchberg AM, Largman C. (1997a) AbdB-like Hox proteins stabilize DNA binding by the Meis1 homeodomain proteins. *Mol Cell Biol* 17, 6448-58.
- Shen WF, Rozenfeld S, Kwong A, Kom ves LG, Lawrence HJ, Largman C. (1999) HOXA9 forms triple complexes with PBX2 and MEIS1 in myeloid cells. *Mol Cell Biol* 19, 3051-61.
- Sheth R, Bastida MF, Kmita M, Ros M. (2013) 'Self-Regulation', a new facet of *Hox* genes' function. *Dev Dyn.* doi: 10.1002/dvdy.24019.
- Sheth R, Marcon L, Bastida MF, Junco M, Quintana L, Dahn R, Kmita M, Sharpe J, Ros MA. (2012) Hox genes regulate digit patterning by controlling the wavelength of a Turing-type mechanism. *Science* 338(6113):1476-80.
- Shi Y, Seto E, Chang LS, Shenk T. (1991) Transcriptional repression by YY1, a human GLI-Krüppel-related protein, and relief of repression by adenovirus E1A protein. *Cell* 67(2): 377-88.
- Shimeld SM, Gaunt SJ, Coletta PL, Geada AM, Sharpe PT. (1993) Spatial localisation of transcripts of the Hox-C6 gene. *J Anat* 183 (Pt 3), 515-23.
- Shimizu T, Bae YK, Hibi M. (2006) Cdx-Hox code controls competence for responding to Fgfs and retinoic acid in zebrafish neural tissue. *Development* 133, 4709-4719.
- Shrivastava A, Saleque S, Kalpana GV, Artandi S, Goff SP, Calame K. (1993) Inhibition of transcriptional regulator Yin-Yang-1 by association with c-Myc. *Science* 262(5141): 1889-92.

- Shumacher A, Faust C, Magnuson T. (1996) Positional cloning of a global regulator of anterior-posterior patterning in mice. *Nature* 383(6597): 250-3.
- Sing A, Pannell D, Karaiskakis A, Sturgeon K, Djabali M, Ellis J, Lipshitz HD, Cordes SP. (2009) A vertebrate Polycomb response element governs segmentation of the posterior hindbrain. *Cell* 138 (5):885-97.
- Sio A, Chehal MK, Tsai K, Fan X, Roberts ME, Nelson BH, Grembecka J, Cierpicki T, Krebs DL, Harder KW. (2013) Dysregulated hematopoiesis caused by mammary cancer is associated with epigenetic changes and Hox gene expression in hematopoietic cells. *Cancer Res.* doi: 10.1158/0008-5472.CAN-13-0842
- Sipos L, Gyurkovics H. (2005) Long-distance interactions between enhancers and promoters. The case of the Abd-B domain of the *Drosophila* bithorax complex. *FEBS J.* 272, 3253–3259.
- Skromne I, Thorsen D, Hale M, Prince VE, Ho RK. (2007) Repression of the hindbrain developmental program by Cdx factors is required for the specification of the vertebrate spinal cord. *Development* 134, 2147–2158.
- Smale ST, Baltimore D. (1989) The "initiator" as a transcription control element. *Cell* 57(1): 103-13.
- Smith ST, Petruk S, Sedkov Y, Cho E, Tillib S, Canaani E, Mazo A. (2004) Modulation of heat shock gene expression by the TAC1 chromatin-modifying complex. *Nat Cell Biol.* 6(2):162-7.
- Soshnikova N, Duboule D. (2008) Epigenetic regulation of Hox gene activation: the waltz of methyls. *Bioessays* 30(3):199-202.
- Soulier J, Clappier E, Cayuela JM, Regnault A, García-Peydró M, Dombret H, Baruchel A, Toribio ML, Sigaux F. (2005) HOXA genes are included in genetic and biologic networks defining human acute T-cell leukemia (T-ALL). *Blood* 106 (1):274-286.
- Soutiere SE, Tankersley CG, Mitzner W. (2004) Differences in alveolar size in inbred mouse strains. *Respir Physiol Neurobiol* 140, 283–291, 2004.
- Spitz F, Gonzalez F, Duboule D. (2003) A global control region defines a chromosomal regulatory landscape containing the *HoxD* cluster. *Cell* 113, 405-417.
- Spitz F, Gonzalez F, Peichel C, Vogt TF, Duboule D, Zakany J. (2001) Large scale transgenic and cluster deletion analysis of the *HoxD* complex separate an ancestral regulatory module from evolutionary innovations. *Genes Dev* 15, 2209-14.
- Spitz F, Herkenne C, Morris MA, Duboule D (2005). Inversion-induced disruption of the *Hoxd* cluster leads to the partition of regulatory landscapes. *Nat. Genet.* 37, 889-893.
- Speleman F, Cauwelier B, Dastugue N, Cools J, Verhasselt B, Poppe B, Van Roy N, Vandesompele J, Graux C, Uyttebroeck A, Boogaerts M, De Moerloose B, Benoit Y, Selleslag D, Billiet J, Robert A, Huguet F, Vandenberghe P, De Paepe A, Marynen P, Hagemeijer A. (2005) A new recurrent inversion, *inv(7)(p15q34)*, leads to transcriptional activation of HOXA10 and HOXA11 in a subset of T-cell acute lymphoblastic leukemias. *Leukemia* 19 (3):358-366.

- Srivastava S, Puri D, Garapati HS, Dhawan J, Mishra RK. (2013) Vertebrate GAGA factor associated insulator elements demarcate homeotic genes in the HOX clusters. *Epigenetics Chromatin*. 6(1):8.
- Stark A, Bushati N, Jan CH, Kheradpour P, Hodges E, Brennecke J, Bartel DP, Cohen SM, Kellis M. (2008) A single Hox locus in *Drosophila* produces functional microRNAs from opposite DNA strands. *Genes Dev*. 22(1): 8-13.
- Stark A, Kheradpour P, Parts L, Brennecke J, Hodges E, Hannon GJ, Kellis M. (2007) Systematic discovery and characterization of fly microRNAs using 12 *Drosophila* genomes. *Genome Res*. 17(12): 1865-79.
- Steelman S, Moskow JJ, Muzynski K, North C, Druck T, Montgomery JC, Huebner K, Daar IO, Buchberg AM. (1997) Identification of a conserved family of Meis1-related homeobox genes. *Genome Res* 7:142-56.
- Stein S, Fritsch R, Lemaire L, Kessel M. (1996) Checklist : vertebrate homeobox genes. *Mech Dev* 55 : 91-108.
- Struhl G. (1982). Genes controlling segmental specification in the *Drosophila* thorax. *Proc Natl Acad Sci U S A*. 79, 7380-4.
- Struhl G. (1983) Role of the *esc+* gene product in ensuring the selective expression of segment-specific homeotic genes in *Drosophila*. *J Embryol Exp Morphol*. 76, 297-331.
- Struhl G, White RA. (1985) Regulation of the Ultrabithorax gene of *Drosophila* by other bithorax complex genes. *Cell*. 43, 507-19.
- Studer M, Gavalas A, Marshall H, Ariza-McNaughton L, Rijli FM, Chambon P, Krumlauf R. (1998) Genetic interactions between *Hoxa1* and *Hoxb1* reveal new roles in regulation of early hindbrain patterning. *Development* 125, 1025-36.
- Studer M, Lumsden A, Ariza-McNaughton L, Bradley A, Krumlauf R. (1996). Altered segmental identity and abnormal migration of motor neurons in mice lacking *Hoxb-1*. *Nature* 384, 630-4.
- Studer M, Pöpperl H, Marshall H, Kuroiwa A, Krumlauf R. (1994). Role of a conserved retinoic acid response element in rhombomere restriction of *Hoxb-1*. *Science* 265, 1728-32.
- Su DM, Ellis S, Napier A, Lee K, Manley NR. (2001) *Hoxa3* et *Pax1* Regulate Epithelial Cell Death and Proliferation during Thymus and Parathyroid Organogenesis. *Dev Biol*. 236, 316-29.
- Subramanian V, Meyer BI, Gruss P. (1995) Disruption of the murine homeobox gene *Cdx1* affects axial skeletal identities by altering the mesodermal expression domains of Hox genes. *Cell* 83, 641-53.
- Suemori H. (1995) Function of Hox genes in formation of skeletons. *Tanpakushitsu Kakusan Koso* 40, 2144-51.
- Suemori H, Noguchi S. (2000) Hox C cluster genes are dispensable for overall body plan of mouse embryonic development. *Dev Biol* 220, 333-42.
- Tabariès S, Lemieux M, Aubin J, Jeannotte L. (2007) Comparative analysis of *Hoxa5* allelic series. *Genesis* 245(4):218-28.
- Tabariès S, Lapointe J, Besch T, Carter M, Woollard J, Tuggle CK, Jeannotte L. (2005) *Cdx* protein interaction with *Hoxa5* regulatory sequences contributes to *Hoxa5* regional expression along the axial skeleton. *Mol Cell Biol*. 25, 1389-401.

- Taghli-Lamallem O, Hsia C, Ronshaugen M, McGinnis W. (2008) Context-dependent regulation of Hox protein functions by CK2 phosphorylation sites. *Dev Genes Evol* 218, 321–332.
- Taylor HS, Arici A, Olive D, Igarashi P. (1998) HOXA10 is expressed in response to sex steroids at the time of implantation in the human endometrium. *J Clin Invest* 101, 1379-84.
- Taylor JK, Levy T, Suh ER, Traber PG. (1997) Activation of enhancer elements by the homeobox gene Cdx2 is cell line specific. *Nucl. Acids Res.* 25, 2293-2300
- Taylor HS, Vanden Heuvel GB, Igarashi P. (1997) A conserved *Hox* axis in the mouse and human female reproductive system: late establishment and persistent adult expression of the *Hoxa* cluster genes. *Biol Reprod* 57, 1338–1345
- Tarchini B, Duboule D. (2006). Control of Hoxd genes' collinearity during early limb development. *Dev Cell* 10, 93-103.
- Ten Have-Opbroek AAW. (1991) Lung development in the mouse embryo. *Exp Lung Res* 17, 111–30.
- Tie F, Furuyama T, Harte PJ. (1999) The Drosophila polycomb group proteins ESC and E(Z) bind directly to each other and co-localize at multiple chromosomal sites. *Development* 125(17): 3483-96.
- Tiong S, Bone LM, Whittle JR. (1985) Recessive lethal mutations within the bithorax complex in Drosophila. *Mol. Gen. Genet.* 200, 335–342.
- Tiret L, Le Mouellic H, Lallemand Y, Maury M, Brulet P. (1993) Altering the spatial determinations in the mouse embryos by manipulating the Hox genes. *C R Acad Sci III* 316, 1009-24.
- Tiret L, Le Mouellic H, Maury M, Brulet P. (1998) Increased apoptosis of motoneurons and altered somatotopic maps in the brachial spinal cord of Hoxc-8-deficient mice. *Development* 125, 279-91.
- Tiwari VK, McGarvey KM, Licchesi JD, Ohm JE, Herman JG, Schübeler D, Baylin SB. (2008) PcG proteins, DNA methylation, and gene repression by chromatin looping. *PLoS Biol.* 6(12): 2911-27.
- Thorsteinsdottir U, Mamo A, Kroon E, Jerome L, Bijl J, Lawrence HJ, Humphries K, Sauvageau G. (2002) Overexpression of the myeloid leukemia-associated Hoxa9 gene in bone marrow cells induces stem cell expansion. *Blood* 99, 121-129.
- Trainor PA, Krumlauf R. (2000) Patterning the cranial neural crest: hinbrain segmentation and Hox gene plasticity. *Nat Rev Neurosci.* 1(2):116-24.
- Trainor PA, Ariza-McNaughton L, Krumlauf R. (2002) Role of the isthmus and FGFs in resolving the paradox of neural crest plasticity and prepatterning. *Science* 295, 1288-91.
- Trask MC, Tremblay KD, Mager J. (2012) Yin-Yang1 is required for epithelial-to-mesenchymal transition and regulation of Nodal signaling during mammalian gastrulation. *Dev Biol.* 368(2):273-82
- Tsai MC, Manor O, Wan Y, Mosammaparast N, Wang JK, Lan F, Shi Y, Segal E, Chang HY. (2010) Long noncoding RNA as modular scaffold of histone modification complexes. *Science* 329, 689–693.
- Tschopp P, Christen AJ, Duboule D. (2012) Bimodal control of Hoxd gene transcription in the spinal cord defines two regulatory subclusters. *Development* 139(5):929-39.

- Tschopp P, Duboule D. (2011) A regulatory 'landscape effect' over the HoxD cluster. *Dev Biol.* 351(2):288-96.
- Tschopp P, Fraudeau N, Béna F, Duboule D. (2011) Reshuffling genomic landscapes to study the regulatory evolution of Hox gene clusters. *PNAS* 108(26):10632-7.
- Tuggle CK, Zákány J, Cianetti L, Peschle C, Nguyen-Huu MC. (1990) Region-specific enhancers near twomammalian homeo box genes define adjacent rostrocaudal domains in the central nervous system. *Genes Dev* 4, 180-9.
- Usheva A, Shenk T. (196) YY1 transcriptional initiator: protein interactions and association with a DNA site containing unpaired strands. *PNAS* 93(24): 13571-6.
- Utsch B, Becker K, Brock D, Lentze MJ, Bidlingmaier F, Ludwig M. (2002) A novel stable polyalanine [poly(A)] expansion in the HOXA13 gene associated with hand-foot-genital syndrome: proper unctioin of poly(A)-harbouring transcription factors depends on a critical repeat length? *Hum Genet* 110, 488-494.
- Uyeno LA, Newman-Keagle JA, Cheung I, Hunt TK, Young DM, Boudreau N. (2001) Hox D3 expression in normal and impaied wound healing. *J Surg Res* 100, 46-56.
- Valentine JW, Erwin DH, Jablonski D. (1996) Developmental evolution of metazoan bodyplans: the fossil evidence. *Dev Biol* 173, 373-81.
- van den Akker E, Forlani S, Chawengsaksophak K, de Graaff W, Beck F, Meyer BI, Deschamps J. (2002) Cdx1 and Cdx2 have overlapping functions in anteroposterior patterning and posterior axis elongation. *Development.* 129, 2181-93.
- van den Akker E, Fromental-Ramain C, de Graaff W, Le Mouellic H, Brulet P, Chambon P, Deschamps J. (2001) Axial skeletal patterning in mice lacking all paralogous group 8 Hox genes. *Development* 128, 1911-21.
- van den Akker, E., Reijnen, M., Korving, J., Brouwer, A., Meijlink, F. and Deschamps, J. (1999). Targeted inactivation of Hoxb8 affects survival of a spinal ganglion and causes aberrant limb reflexes. *Mech Dev* 89, 103-14.
- van der Hoeven F, Zákány J, Duboule D. (1996) Gene transpositions in the HoxD complex reveal a hierarchy of regulatory controls. *Cell* 85, 1025-35.
- van der Lugt NM, Domen J, Linders K, van Roon M, Robanus-Maandag E, te Riele H, van der Valk M, Deschamps J, Sofroniew M, van Lohuizen M, Berns A. (1994) Posterior transformation, neurological abnormalities, and severe hematopoietic defects in mice with a targeted deletion of the bmi-1 proto-oncogene. *Genes Dev.* 8(7):757-69.
- van Nes J, de Graaff W, Lebrin F, Gerhard M, Beck F, Deschamps J. (2006) The Cdx4 mutation affects axial development and reveals an essential role of Cdx genes in the ontogenesis of the placental labyrinth in mice. *Development* 133, 419-28.
- Vieux-Rochas M, Mascrez B, Krumlauf R, Duboule D. (2013) Combined function of HoxA and HoxB clusters in neural crest cells. *Dev Biol.* pii: S0012-1606(13)00360-6.
- Vinagre T, Moncaut N, Carapuco M, Novoa A, Bom J, Mallo M. (2010) Evidence for a Myotomal Hox/Myf Cascade Governing Nonautonomous Control of Rib Specification within Global Vertebral Domains. *Dev Cell.*

- Vogels R, Charité J, de Graaff W, Deschamps J. (1993) Proximal cis-acting elements cooperate to set Hoxb-7 (Hox-2.3) expression boundaries in transgenic mice. *Development* 118, 71-82.
- Volpe MV, Archavachotikul K, Bhan I, Lessin MS, Nielsen HC. (2000a) Association of bronchopulmonary sequestration with expression of the homeobox protein Hoxb-5. *J Pediatr Surg* 35, 1817-9.
- Volpe MV, Martin A, Vosatka RJ, Mazzoni CL, Nielsen HC. (1997) Hoxb-5 expression in the developing mouse lung suggests a role in branching morphogenesis and epithelial cell fate. *Histochem Cell Biol* 108, 495-504.
- Volpe MV, Pham L, Lessin M, Ralston SJ, Bhan I, Cutz E, Nielsen HC. (2003) Expression of Hoxb-5 during human lung development and in congenital lung malformations. *Birth Defects Res A Clin Mol Teratol* 67, 550-6.
- Volpe MV, Vosatka RJ, Nielsen HC. (2000b) Hoxb-5 control of early airway formation during branching morphogenesis in the developing mouse lung. *Biochim Biophys Acta* 1475, 337-45.
- Wacker SA, Jansen HJ, McNulty CL, Houtzager E, Durston AJ. (2004) Timed interactions between the *Hox* expressing non-organiser mesoderm and the Spemann organiser generate positional information during vertebrate gastrulation. *Dev. Biol.* 268, 207-219.
- Wagner K, Mincheva A, Korn B, Lichter P, Pöpperl H. (2001) Pbx4, a new Pbx family member on mouse chromosome 8, is expressed during spermatogenesis. *Mech Dev* 103, 127-31.
- Wahba GM, Hostikka SL, Carpenter EM. (2001) The paralogous Hox genes Hoxa10 and Hoxd10 interact to pattern the mouse hindlimb peripheral nervous system and skeleton. *Dev Biol* 231, 87-102.
- Wakimoto BT, Kaufman TC. (1981) Analysis of larval segmentation in lethal genotypes associated with the antennapedia gene complex in *Drosophila melanogaster*. *Dev. Biol.* 81, 51-64.
- Wall NA, Jones CM, Hogan BL, Wright CV. (1992) Expression and modification of Hox 2.1 protein in mouse embryos. *Mech Dev.* May;37(3):111-20.
- Walshe J, Maroon H, McGonnell IM, Dickson C, Mason I. (2002) Establishment of hindbrain segmental identity requires signaling by FGF3 and FGF8. *Curr Biol* 12, 1117-23.
- Wang Z, Dollé P, Cardoso WV, Niederreither K. (2006) Retinoic acid regulates morphogenesis and patterning of posterior foregut derivatives. *Dev Biol.* 297(2):433-45.
- Wang C, Lehmann R. (1991) Nanos is the localized posterior determinant in *Drosophila* [published erratum appears in *Cell* 1992 Mar. 20;68(6):1177]. *Cell* 66, 637-647.
- Wang P, Lin C, Smith ER, Guo H, Sanderson BW, Wu M, Gogol M, Alexander T, Seidel C, Wiedemann LM, Ge K, Krumlau R, Shilatifard A. (2009) Global analysis of H3K4 methylation defines MLL family member targets and points to a role for MLL1-mediated H3K4 methylation in the regulation of transcriptional initiation by RNA polymerase II. *Mol Cell Biol.* 29(22):6074-85.
- Wang CC, Tsai MF, Dai TH, Hong TM, Chan WK, Chen JJ, Yang PC. (2007) Synergistic activation of the tumor suppressor, HLJ1, by the transcription factors YY1 and activator protein 1. *Cancer Res.* 67(10): 4816-26.

- Wang CC, Tsai MF, Hong TM, Chang GC, Chen CY, Yang WM, Chen JJ, Yang PC. (2005) The transcriptional factor YY1 upregulates the novel invasion suppressor HLJ1 expression and inhibits cancer cell invasion. *Oncogene* 24(25): 4081-93.
- Wang H, Wang L, Erdjument-Bromage H, Vidal M, Tempst P, Jones RS, Zhang Y. (2004) Role of histone H2A ubiquitination in Polycomb silencing. *Nature* 431(7010): 873-8.
- Wang KC, Yang YW, Liu B, Sanyal A, Corces-Zimmerman R, Chen Y, Lajoie BR, Protacio A, Flynn RA, Gupta RA, Wysocka J, Lei M, Dekker J, Helms JA, Chang HY. (2011) A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature* 472(7341):120-4.
- Warburton D, El-Hashash A, Carraro G, Tiozzo C, Sala F, Rogers O, De Langhe S, Kemp PJ, Riccardi D, Torday J, Bellusci S, Shi W, Lubkin SR, Jesudason E. (2010) Lung organogenesis. *Curr Top Dev Biol* 90, 73-158.
- Warburton D, Schwarz M, Tefft D, Flores-Delgado G, Anderson KD, Cardoso WV. (2000) The molecular basis of lung morphogenesis. *Mech Dev* 92, 55-81.
- Warot X, Fromental-Ramain C, Fraulob V, Chambon P, Dollé P. (1997) Gene dosage-dependent effects of the Hoxa-13 and Hoxd-13 mutations on morphogenesis of the terminal parts of the digestive and urogenital tracts. *Development* 124, 4781-91.
- Weerkamp F, Luis TC, Naber BA, Koster EE, Jeannotte L, van Dongen JJ, Staal FJ. (2006) Identification of Notch target genes in uncommitted T-cell progenitors: No direct induction of a T-cell specific gene program. *Leukemia* 20(11):1967-77.
- Wellik DM. (2007) Hox patterning of the vertebrate axial skeleton. *Dev Dyn* 236: 2454-2463.
- Wellik DM. (2009) Hox genes and vertebrate axial pattern. *Curr Top Dev Biol* 88:257-78.
- Wellik DM. (2011) *Hox* genes and kidney development. *Pediatr Nephrol* 26, 1559-1565
- Wellik DM, Capecchi MR. (2003) Hox10 and Hox11 genes are required to globally pattern the mammalian skeleton. *Science* 301:363-367.
- Wellik DM, Hawkes PJ, Capecchi MR. (2002) Hox11 paralogous genes are essential for metanephric kidney induction. *Genes Dev* 16, 1423-1432
- Weng AP, Ferrando AA, Lee W, Morris JP t, Silverman LB, Sanchez-Irizarry C, Blacklow SC, Look AT, Aster JC. (2004) Activating mutations of NOTCH1 in human T cell acute lymphoblastic leukemia. *Science* 306, 269-71.
- Wert SE, Glasser SW, Korfhagen TR, Whitsett JA. (1993) Transcriptional elements from the human SP-C gene direct expression in the primordial respiratory epithelium of transgenic mice. *Dev. Biol.* 156, 426-443
- Weston AD, Blumberg B, Underhill TM. (2003a) Active repression by unliganded retinoid receptors in development: Less is sometimes more. *J Cell Biol* 161(2):223-228.
- Weston AD, Hoffman LM, Underhill TM. (2003b) Revisiting the role of retinoid signaling in skeletal development. *Birth Defects Res C Embryo Today* 69(2):156-173.
- White RA, Wilcox M. (1985) Regulation of the distribution of Ultrabithorax proteins in *Drosophila*. *Nature* 318, 563-567

- Whiting J, Marshall H, Cook M, Krumlauf R, Rigby PW, Stott D, Allemann RK. (1991) Multiple spatially specific enhancers are required to reconstruct the pattern of Hox-2.6 gene expression. *Genes Dev* 5, 2048-59.
- Wichterle H, Lieberam I, Porter JA, Jessell TM. (2002). Directed differentiation of embryonic stem cells into motor neurons. *Cell* 110, 385–397 (2002).
- Wienholds E, Kloosterman WP, Miska E, Alvarez-Saavedra E, Berezikov E, de Bruijn E, Horvitz HR, Kauppinen S, Plasterk RH. (2005) MicroRNA expression in zebrafish embryonic development. *Science* 309(5732): 310-1.
- Wilkinson D, Bhatt S, Chavrier P, Bravo R, Charnay P. (1969a) Segment-specific expression of a zinc finger gene in the developing nervous system of the mouse. *Nature* 337, 461-464.
- Williams ME, Lehoczky JA, Innis JW. (2006) A group 13 homeodomain is neither necessary nor sufficient for posterior prevalence in the mouse limb. *Dev Biol*, 297, 493-507.
- Williamson I, Eskeland R, Lettice LA, Hill AE, Boyle S, Grimes GR, Hill RE, Bickmore WA. (2012) Anterior-posterior differences in HoxD chromatin topology in limb development. *Development* 139(17):3157-67.
- Wilson JG, Roth CB, Warkany J. (1953) An analysis of the syndrome of malformations induced by maternal vitamin A deficiency. Effects of restoration of vitamin A at various times during gestation. *Am. J. Anat.* 92, 189– 217.
- Winnik S, Klinkert M, Kurz H, Zoeller C, Heinke J, Wu Y, Bode C, Patterson C, Moser M. (2009) HoxB5 induces endothelial sprouting in vitro and modifies intussusceptive angiogenesis in vivo involving angiopoietin-2. *Cardiovasc Res* 2009; 83:558-65
- Wolgemuth DJ, Behringer RR, Mostoller MP, Brinster RL, Palmiter RD. (1989) Transgenic mice overexpressing the mouse homoeobox-containing gene Hox-1.4 exhibit abnormal gut development. *Nature* 337, 464-7.
- Woltering JM, Durston AJ. (2006) The zebrafish hoxDb cluster has been reduced to a single microRNA. *Nat Genet.* 38(6): 601-2.
- Woltering JM, Durston AJ. (2008). MiR10 represses HoxB1a and HoxB3a in Zebrafish. *PLoS ONE* 3(1): e1396.
- Woltering JM, Vonk FJ, Muller H, Bardine N, Tuduze IL, de Bakker MA, Knochel W, Sirbu IO, Durston AJ, Richardson MK. (2009) Axial patterning in snakes and caecilians: evidence for an alternative interpretation of the Hox code. *Dev Biol* 332:82–89.
- Wong EY, Wang XA, Mak SS, Sae-Pang JJ, Ling KW, Fritsch B, Sham MH. (2011) Hoxb3 negatively regulates Hoxb1 expression in mouse hindbrain patterning. *Dev. Biol.* 352(2): 382-92.
- Wong P, Iwasaki M, Somerville TC, So CW, Cleary ML. (2007) Meis1 is an essential and rate-limiting regulator of MLL leukemia stem cell potential. *Genes Dev* 21:2762-74
- Wongtrakool C, Malpel S, Gorenstein J, Sedita J, Ramirez MI, Underhill TM, Cardoso WV. (2003) Down-regulation of retinoic acid receptor alpha signaling is required for sacculation and type I cell formation in the developing lung. *J. Biol. Chem.* 278, 46911 –46918.

- Woo CJ, Kharchenko PV, Daheron L, Park PJ, Kingston RE. (2010) A region of the human HOXD cluster that confers polycomb-group responsiveness. *Cell* 140(1):99-110.
- Wu Y, Moser M, Bautch VL, Patterson C. (2003) HoxB5 is an upstream transcriptional switch for differentiation of the vascular endothelium from precursor cells. *Mol Cell Biol* 23, 5680-91
- Wu S, Shi Y, Mulligan P, Gay F, Landry J, Liu H, Lu J, Qi HH, Wang W, Nickoloff JA, Wu C, Shi Y. (2007) A YY1-INO80 complex regulates genomic stability through homologous recombination-based repair. *Nat Struct Mol Biol.* 14(12):1165-72.
- Würtele H, Chartrand P. (2006) Genome-wide scanning of HoxB1-associated loci in mouse ES cells using an open-ended Chromosome Conformation Capture methodology. *Chromosome Res.* 14, 477-495.
- Würtele H, Chartrand P (2006) Genome-wide scanning of HoxB1-associated loci in mouse ES cells using an open-ended Chromosome Conformation Capture methodology. *Chromosome Res.* 14(5): 477-95.
- Xu B, Chen H, Xu W, Zhang W, Buckley S, Zheng SG, Warburton D, Kolb M, Gaudie J, Shi W. (2012) Molecular mechanisms of MMP9 overexpression and its role in emphysema pathogenesis of Smad3-deficient mice. *Am J Physiol Lung Cell Mol Physiol.* 303(2): L89-96.
- Xu B, Wellik DM. (2011) Axial Hox9 activity establishes the posterior field in the developing forelimb. *PNAS* 108(12):4888-91.
- Yang WM, Inouye C, Zeng Y, Bearss D, Seto E. (1996) Transcriptional repression by YY1 is mediated by interaction with a mammalian homolog of the yeast global regulator RPD3. *PNAS* 93(23): 12845-50.
- Yang F, Miao L, Mei Y, Wu M. (2013) Retinoic acid-induced HOXA5 expression is co-regulated by HuR and miR-130a. *Cell Signal.* 25(6):1476-85.
- Yant SR, Zhu W, Millinoff D, Slightom JL, Goodman M, Gumucio DL. (1995) High affinity YY1 binding motifs: identification of two core types (ACAT and CCAT) and distribution of potential binding sites within the human beta globin cluster. *Nucleic Acids Res* 23(21): 4353-62.
- Yao YL, Yang WM, Seto E. (2001) Regulation of transcription factor YY1 by acetylation and deacetylation. *Mol Cell Biol* 21, 5979-5991.
- Yasunaga S, Ohtsubo M, Ohno Y, Saeki K, Kurogi T, Tanaka-Okamoto M, Ishizaki H, Shirai M, Mihara K, Brock HW, Miyoshi J, Takihara Y. (2013) Scmh1 has E3 ubiquitin ligase activity for geminin and histone H2A and regulates geminin stability directly or indirectly via transcriptional repression of Hoxa9 and Hoxb4. *Mol Cell Biol.* 33(4):644-60.
- Yau TO, Kwan CT, Jakt LM, Stallwood N, Cordes S, Sham MH. (2002) Auto/cross-regulation of Hoxb3 expression in posterior hindbrain and spinal cord. *Dev Biol* 252, 287-300.
- Yekta S, Shih H, Bartel DP. (2004) MicroRNA-directed cleavage of HOXB8. mRNA. *Science* 304 (5670): 594-596.
- Yekta S, Tabin CJ, Bartel DP. (2008) MicroRNAs in the Hox network: an apparent link to posterior prevalence. *Nat Rev Genet* 9(10): 789-796.
- Young T, Rowland JE, van de Ven C, Bialecka M, Novoa A, Carapuco M, van Nes J, de Graaff W, Duluc I, Freund JN, Beck F, Mallo M, Deschamps J. (2009) Cdx and

- Hox genes differentially regulate posterior axial growth in mammalian embryos. *Dev Cell*. 17(4):516-26.
- Yu BD, Hanson RD, Hess JL, Horning SE, Korsmeyer SJ. (1998) MLL, a mammalian trithorax-group gene, functions as a transcriptional maintenance factor in morphogenesis. *PNAS* 95(18):10632-6.
- Yu K, Xu J, Liu Z, Susic D, Shao J, Olson EN, Towler DA, Ornitz DM (2003) Conditional inactivation of FGF receptor 2 reveals an essential role for FGF signaling in the regulation of osteoblast function and bone growth. *Development* 130: 3063-3074.
- Yue R, Kang J, Zhao C, Hu W, Tang Y, Liu X, Pei G. (2009) Beta-arrestin1 regulates zebrafish hematopoiesis through binding to YY1 and relieving polycomb group repression. *Cell* 139(3): 535-46.
- Zákány J, Duboule D. (1999a). Hox genes and the making of sphincters. *Nature* 401, 761-2.
- Zákány J, Duboule D. (1999b) Hox genes in digit development and evolution. *Cell Tissue Res* 296, 19-25.
- Zákány J, Duboule D. (2007) The role of *Hox* genes during vertebrate limb development. *Curr Opin Genet Dev*. 17, 359-66
- Zákány J, Fromental-Ramain C, Warot X, Duboule D. (1997a) Regulation of number and size of digits by posterior Hox genes: a dose-dependent mechanism with potential evolutionary implications. *PNAS* 94, 13695-700.
- Zákány J, Gérard M, Favier B, Duboule D. (1997b) Deletion of a HoxD enhancer induces transcriptional heterochrony leading to transposition of the sacrum. *Embo J* 16, 4393-402.
- Zákány J, Kmita M, Alarcon P, de la Pompa JL, Duboule D. (2001) Localized and transient transcription of Hox genes suggests a link between patterning and the segmentation clock. *Cell* 106, 207-17.
- Zakany J, Kmita M, Duboule D. (2004) A dual role for Hox genes in limb anterior-posterior asymmetry. *Science* 304, 1669-1672.
- Zákány J, Tuggle CK, Patel MD, Nguyen-Huu MC. (1988) Spatial regulation of homeobox gene fusions in the embryonic central nervous system of transgenic mice. *Neuron* 1, 679-91.
- Zaravinos A, Kanellou P, Lambrou GI, Spandidos DA (2014) Gene set enrichment analysis of the NF- κ B/Snail/YY1/RKIP circuitry in multiple myeloma. *Tumor Biol*. DOI: 10.1007/s13277-014-1659-9
- Zavortink M, Sakonju S. (1989) The morphogenetic and regulatory functions of the *Drosophila* Abdominal-B gene are encoded in overlapping RNAs transcribed from separate promoters. *Genes Dev*. 3(12A): 1969-81.
- Zeng Y, Cai X, Cullen BR. (2005) Use of RNA polymerase II to transcribe artificial microRNAs. *Methods Enzymol* 392,371-80.
- Zhang F, Nagy Kovacs E, Featherstone MS. (2000) Murine *hoxd4* expression in the CNS requires multiple elements including a retinoic acid response element. *Mech Dev* 96(1):79-89.
- Zhao H, Zhang X, Frazão JB, Condino-Neto A, Newburger PE. (2013) HOX antisense lincRNA HOXA-AS2 is an apoptosis repressor in all Trans retinoic acid treated NB4 promyelocytic leukemia cells. *J Cell Biochem*. 114(10):2375-83.

- Zhao JJ, Lazzarini RA, Pick L. (1993) The mouse Hox-1.3 gene is functionally equivalent to the Drosophila Sex combs reduced gene. *Genes Dev* 7, 343-54.
- Zhao JJ, Lazzarini RA, Pick L. (1996) Functional dissection of the mouse *Hoxa-5* gene. *EMBO J.* 15,1313–1322.
- Zhao Y, Potter SS (2002) Functional comparison of the Hoxa 4, Hoxa 10, and Hoxa 11 homeoboxes. *Dev Biol* 244, 21-36.
- Zhao Z, Boyle TJ, Liu Z, Murray JI, Wood WB, Waterston RH. (2010) A Negative Regulatory Loop between MicroRNA and Hox Gene Controls Posterior Identities in *Caenorhabditis elegans*. *PLoS Genet.* 6(9): e1001089.
- Zhu Y, Cuevas IC, Gabriel RA, Su H, Nishimura S, Gao P, Fields A, Hao Q, Young WL, Yang GY, Boudreau NJ. (2009) Restoring HOXA5 Expression Inhibits the Growth of Experimental Hemangiomas in the Brain. *J Neuropathol Exp Neurol.* 68(6): 626-32.

ANNEXE I

Functional resolution of duplicated *hoxb5* genes in teleosts

Olga Jarinova^{1,2}, Gary Hatch¹, Luc Poitras¹, Christelle Prudhomme³, Magdalena Grzyb¹,
Josée Aubin³, Félix-Antoine Bérubé-Simard³, Lucie Jeannotte³ and Marc Ekker^{1,2,*}

1 Center for Advanced Research in Environmental Genomics, Department of Biology,
University of Ottawa, Ottawa, Ontario K1N 6N5, Canada.

2 Department of Cellular and Molecular Medicine, University of Ottawa, Ottawa, Ontario
K1H 8M5, Canada.

3 Centre de recherche en cancérologie de l'Université Laval, Centre Hospitalier
Universitaire de Québec, L'Hôtel-Dieu de Québec, Québec G1R 2J6, Canada.

*Author for correspondence (e-mail: mekker@uottawa.ca)

Article publié en 2008 dans : *Development* **135**, 3543-3553

AVANT-PROPOS

Plusieurs événements de duplication génique ont émergé au fur et à mesure que les organismes vivants se sont complexifiés au cours de l'évolution, afin de traduire la pression qu'exerçait leur environnement en une nouvelle fonction facilitant leur survie. Dans cette optique, le modèle de duplication-dégénération-complémentation (DCC) stipule que le mécanisme par lequel les gènes dupliqués se subdivisent les différentes fonctions du gène ancestral (subfonctionnalisation), afin que la nouvelle copie du gène soit conservée dans le génome, est assez fréquent. Puisque les poissons téléostéens possèdent davantage de complexes *Hox* que les mammifères, ils constituent un bon modèle d'étude pour tester la validité de l'hypothèse avancé par le DCC. En effet, les différences entre les profils d'expression des gènes *Hoxb5a* et *Hoxb5b* chez le poisson zèbre, situés sur deux complexes différents, suggèrent une subfonctionnalisation du gène *Hoxb5* ancestral. D'ailleurs, la comparaison des séquences intergéniques *Hoxb4-Hoxb5* appartenant à différentes espèces a mis en évidence la conservation de trois éléments régulateurs (J1-J3), où l'élément J3, qui est propre au gène *Hoxb5a*, pourrait être à l'origine des différences fonctionnelles entre *Hoxb5a* et *Hoxb5b*.

Tout d'abord, le haut niveau d'homologie de séquence entre les éléments J1, J2 et J3 pourrait expliquer le chevauchement de leur activité activatrice/amplificatrice lorsque testés individuellement en transgénèse chez la souris et le poisson zèbre. Bien que ces résultats ne suggéraient pas qu'il y ait eu subfonctionnalisation entre *Hoxb5a* et *Hoxb5b*, le transgène portant la délétion de l'élément J3 dans un contexte *Hoxb5a* récapitule quant à lui un profil d'expression similaire à celui du transgène *Hoxb5b* chez la souris et le poisson zèbre.

Pour ma part, j'ai participé à l'expérience inverse, au cours de laquelle j'ai généré des souris transgénique où l'élément J3 a été inséré au niveau du locus du gène *Hoxb5b*. Cette expérience m'a permis d'observer un profil d'expression comparable à celui du transgène *Hoxb5a*, confirmant l'importance de l'élément J3 et des interactions entre les différents éléments régulateurs dans l'exécution des subfonctions de gènes dupliqués (Fig. 4E-F, G, K, M-N)

ABSTRACT

The duplication-degeneration-complementation (DDC) model predicts that subfunctionalization of duplicated genes is a common mechanism for their preservation. The additional Hox complexes of teleost fish constitute a good system in which to test this hypothesis. Zebrafish have two *hoxb* complexes, with two *hoxb5* genes, *hoxb5a* and *hoxb5b*, the expression patterns of which suggest subfunctionalization of an ancestral *hoxb5* gene. We characterized conserved non-coding elements (CNEs) near the zebrafish *hoxb5* genes. One CNE, J3, is only retained in the *hoxb5a* locus, whereas the others, J1 and J2, are present in both *hoxb5* loci. When tested individually, the enhancer activity of individual CNEs, including J3, extensively overlapped and did not support a role in subfunctionalization. By contrast, reporter transgene constructs encompassing multiple CNEs were able to target reporter gene expression to unique domains of *hoxb5a* and *hoxb5b* expression. The deletion of J3 from the *hoxb5a* locus resulted in expression that approached that of *hoxb5b*, whereas its insertion in the *hoxb5b* locus increased reporter expression and rendered it more similar to that of *hoxb5a*. Our results highlight the importance of interactions between CNEs in the execution of complementary subfunctions of duplicated genes.

INTRODUCTION

Duplication events, which can affect anything from single genes to the whole genome, are thought to contribute to adaptive evolution and promote increased animal complexity (Ohno, 1970). Duplicated genes can be lost by accumulating deleterious mutations (nonfunctionalization), or they may acquire a new adaptive function (neofunctionalization) or divide an ancestral function (subfunctionalization). The duplication-degeneration-complementation (DDC) model predicts that subfunctionalization is the most common mechanism for the preservation of duplicated genes (Force et al., 1999). According to the DDC model, functional modules of duplicated genes are likely to undergo complementary degeneration. The division of a single functional unit into separate subunits makes the complete loss of the original function through nucleotide substitutions, deletions or other types of DNA alterations, an unlikely event. Moreover, the subfunctionalized duplicates are under relaxed evolutionary constraints compared with the ancestral gene. Therefore, subfunctionalization of the duplicates increases an organism's chances of survival and, in rare cases, may be a transitional state leading to neofunctionalization (Lynch and Force, 2000).

The Hox gene family underwent significant expansion during evolution and is ideally suited to address the question of how genes acquire new functions. Hox genes are involved in the control of regional identity of the embryonic axes of metazoans. They are organized in clusters with the most anteriorly expressed genes located at the 3' end of the cluster and the posteriorly expressed genes at the 5' end. All vertebrates contain several Hox clusters that are thought to have resulted from the sequential duplications of a proHox cluster early in metazoan evolution. Thus, mammals have four Hox gene clusters (*HoxA*, *HoxB*, *HoxC* and *HoxD*), whereas zebrafish possess seven (*hoxaa*, *hoxab*, *hoxba*, *hoxbb*, *hoxca*, *hoxcb* and *hoxd*). The additional clusters in zebrafish are likely to have resulted from a duplication event that happened after the divergence of the fish and tetrapod lineages, somewhere between 300 and 450 million years ago (Amores et al., 1998; Taylor et al., 2001). The presence of additional Hox clusters is common, but not ubiquitous, in teleosts, including those species with a small genome such as *Takifugu rubripes* and *Spheroides nephelus* (Amores et al., 2004; Chiu et al., 2004). Comparison of the structure

of the teleost Hox clusters with those of mammals reveals the loss of individual genes within some of the duplicated clusters. For instance, the genome of the pufferfish, *Takifugu rubripes*, contains two *hoxb3* genes, *hoxb3a* and *hoxb3b*, whereas only *hoxb3a* is found in the zebrafish genome, suggesting the loss of one of the *hoxb3* duplicates following the divergence of zebrafish and *Takifugu* lineages (Amores et al., 2004).

There are also examples in which both paralogous genes were retained after the duplication event, as is the case for the *hoxb5* genes in both zebrafish and *Takifugu* (Amores et al., 1998; Amores et al., 2004). Bruce and colleagues (Bruce et al., 2001) determined that two zebrafish *hoxb5* genes, *hoxb5a* and *hoxb5b*, are expressed in overlapping yet distinct domains in the notochord, neural tube and somites. Thus, their combined expression domains are strikingly similar to that of the single *Hoxb5* gene in the mouse. Furthermore, the same study reported the biochemical equivalence of the Hoxb5a and Hoxb5b proteins (Bruce et al., 2001). Combined, these results suggest that the *hoxb5a* and *hoxb5b* genes underwent subfunctionalization through loss of cis-acting regulatory elements. To determine whether the subfunctionalization of *Hoxb5* genes is reflected in changes in non-coding regions, we have compared the *Hoxb5* loci of human, mouse, zebrafish and *Takifugu*. This allowed us to identify blocks of highly conserved non-coding elements (CNEs). We compared the regulatory activity of the CNEs in transgenic assays in which they were tested collectively and individually. This led to the finding that one CNE specific to the *hoxb5a* locus, named J3, accounts for the differences in expression between the paralogs, but that interactions between J3 and other CNEs are essential to achieve correct expression. This has important implications for the experimental approaches chosen to determine the patterns of regulatory evolution of duplicated genes.

MATERIALS AND METHODS

Sequence alignments and phylogenetic analysis

Sequences surrounding the *Hoxb5* genes of *Danio rerio*, *Takifugu rubripes*, *Homo sapiens* and *Mus musculus* were obtained from Ensembl Genome Data Resources (<http://www.sanger.ac.uk>). *Danio rerio hoxba*, contig 11157.4 (zv.3), position 40334-

67614; *Danio rerio hoxbb*, contig BX001014.6 (zv.4), position 25612-57312; *Takifugu rubripes hoxba*, contig Scaffold_706, position 25001-50001; *Takifugu rubripes hoxbb*, contig Scaffold_1245, position 50965-70027; *Mus musculus Hoxb*, contig SuperContig NT_165773, position 96162969-96184091; *Homo sapiens HOXB* GenBank accession AC103702.3.1.187386, position 84742-109241.

Multiple sequence alignments were computed with MultiPipMaker (<http://bio.cse.psu.edu/pipmaker>) or with the Pretty Results program of the GCG Wisconsin package. The boundary of each CNE was determined as the longest region of high sequence similarity between mammalian and teleost sequences. Phylogenetic and molecular evolutionary analyses were conducted with the Kimura two-parameter model using the neighbor-joining method and bootstrapping (Kumar et al., 2001). Searches for potential transcription factor binding sites in the CNEs were performed with Patch_Search (<http://www.gene-regulation.com/pub/programs.html>).

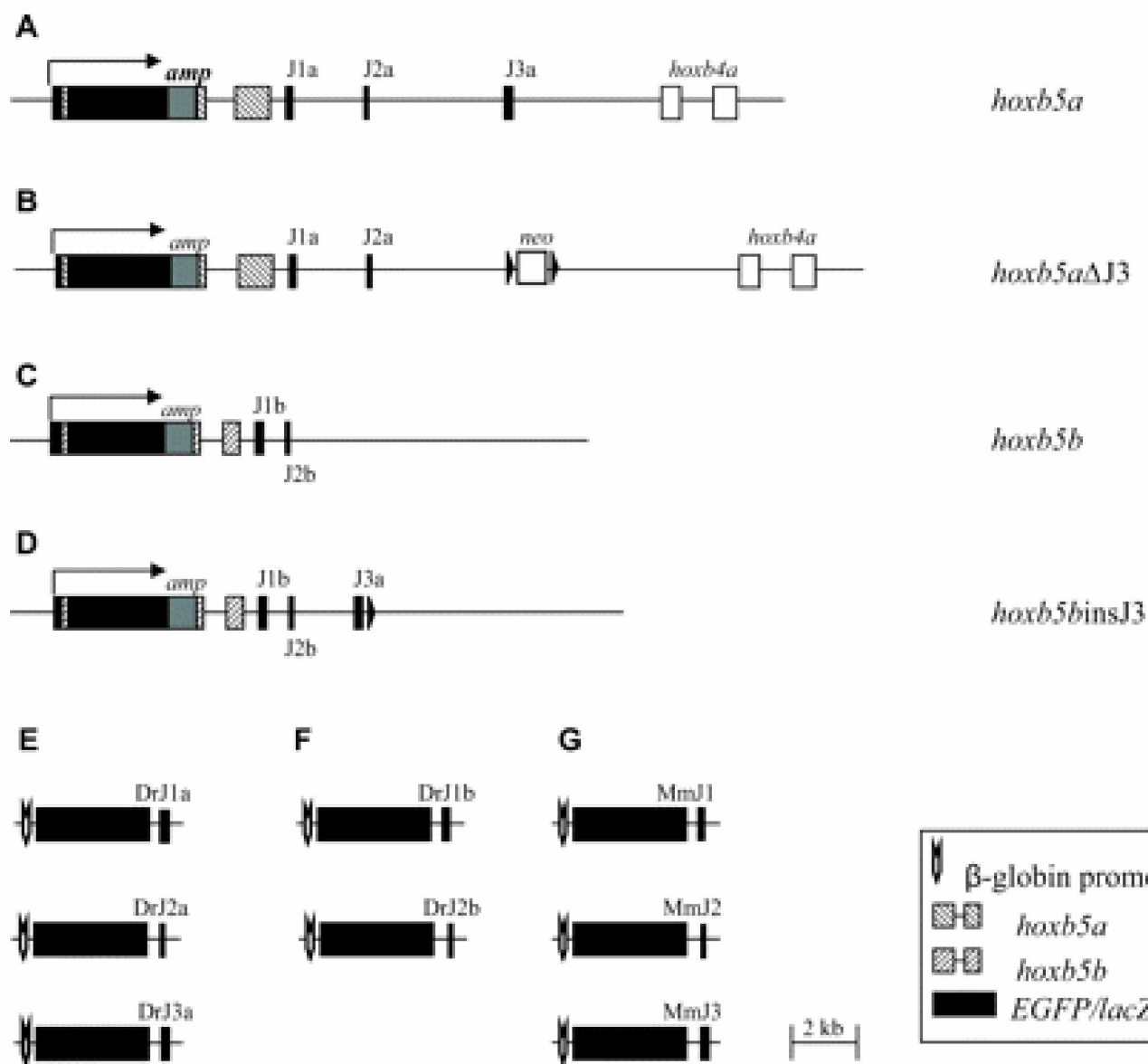
Constructs

Reporter constructs encompassing large regions of the zebrafish *hoxb5a* and *hoxb5b* loci were generated by homologous recombination in bacteria (Copeland et al., 2001). Recombination was carried out in *Escherichia coli* DY380 and EL350 strains provided by Dr Neal Copeland (US National Cancer Institute, Frederick, MD).

PACs containing the *hoxb5a* and *hoxb5b* zebrafish loci were obtained from RZPD (plate positions 254O17 and 227H9, respectively). For *hoxb5a*, a PAC fragment spanning from 1353 bp upstream of the *hoxb5a* initiation codon to 1316 bp downstream of the *hoxb4a* termination codon (total size 18,495 bp), was first inserted into a plasmid vector using homology arms of ~200 bp. Similarly, we created a plasmid containing a 12,169 bp fragment from the *hoxb5b* locus that spans from 1154 bp upstream of the *hoxb5b* initiation codon to 9763 bp downstream of the *hoxb5b* termination codon. We inserted the *lacZ* and *EGFP* reporter genes, encoding β -galactosidase and enhanced green fluorescent proteins, respectively, 33 bp downstream of either the *hoxb5a* or the *hoxb5b* initiation codon, in frame, to generate the *hoxb5alacZ*, *hoxb5aEGFP*, *hoxb5blacZ* and *hoxb5bEGFP* constructs (Fig. 1A, C). Constructs containing *lacZ* and *EGFP* were used for the production

Fig. 1. Constructs used to generate transgenic animals.

(A-D) The reporter transgene was inserted into the first exon of the zebrafish *hoxb5a* (A) or *hoxb5b* (C) gene. (B) The J3 element was deleted from the *hoxb5a* construct. (D) J3 was inserted into the *hoxb5b* construct. (E-G) Individual conserved non-coding elements (CNEs) from mouse and zebrafish *Hoxb5* loci were coupled to a β -globin minimal promoter and the *EGFP* or *lacZ* reporter genes (black rectangles). a and b represent the *hoxba* and *hoxbb* zebrafish complexes, respectively. In A-D, hatched rectangles indicate exon sequences; black triangles, loxP sites; *amp*, ampicillin resistance gene; *neo*, neomycin resistance gene. Mm, *Mus musculus*; Dr, *Danio rerio*.



of transgenic mice and zebrafish, respectively.

Additional reporter constructs derived from *hoxb5alacZ* and *hoxb5aEGFP*, from which the J3 enhancer sequence was deleted, were produced through an additional round of homologous recombination. The homology arms in the targeting cassette corresponded to sequences located directly upstream and downstream of J3. The deletion constructs are referred to as *hoxb5aΔJ3lacZ* and *hoxb5aΔJ3EGFP* (Fig. 1B).

Finally, the J3 sequence from *hoxb5a* was introduced into the *hoxb5blacZ* and *hoxb5bEGFP* plasmids to generate *hoxb5binsJ3lacZ* and *hoxb5binsJ3EGFP*. The J3 sequence was inserted at a position that would correspond to that of J3 if this enhancer existed in the *hoxb5b* locus (Fig. 1D). Although the *hoxb5b* locus appears to have diminished in size after duplication, elements existing in both zebrafish loci (J1, J2 and *mir10a*) are present in the same order and at the same relative positions. By extrapolation, if present, J3 would be located ~3.1 kb downstream of the *hoxb5b* termination codon. To excise the *neo* gene adjacent to the introduced J3 element, EL350 cells hosting *hoxb5binsJ3lacZ* or *hoxb5binsJ3EGFP* were grown in arabinose-containing medium to induce recombination between loxP sites flanking the *neo* gene.

To produce transgene constructs containing individual CNEs, these were PCR-amplified and inserted downstream of the reporter gene in a way that mimics their position and orientation in their original genomic context (Fig. 1E-G). The *pI229* and *βpEGFP* vectors, which contain a human β-globin minimal promoter coupled to either *lacZ* or *EGFP*, respectively, were used for the production of transgenic mice and zebrafish (Cormack et al., 1996; Yee and Rigby, 1993; Zerucha et al., 2000).

Production and genotyping of transgenic animals

All experiments were performed according to the guidelines of the Canadian Council on Animal Care and were approved by the institutional animal care committees. Transgene preparation and microinjection into fertilized eggs were according to standard procedures (Hogan, 1986). Genotyping of animals and analysis of *lacZ* expression were as previously described (Fraidenraich et al., 1998; Larochelle et al., 1999). Production of

transgenic zebrafish and monitoring of EGFP expression were as described (Amsterdam et al., 1995). *EGFP* reporter constructs were microinjected at 100 ng/μl into more than 100 zebrafish embryos and those showing at least two *EGFP*-expressing cells were referred to as primary transgenic embryos.

In situ hybridization

In situ hybridization was performed as described on whole-mount preparations (Thisse et al., 1993) or on paraffin sections (Jaffe et al., 1990).

Quantitative RT-PCR

One hundred transient-transgenic zebrafish embryos were generated for each of the constructs and were randomly divided in three pools of 33 embryos. Total RNA was extracted from each pool using the RNeasy Kit (Qiagen) and cDNA synthesis was carried out using oligo(dT) primers. Relative amounts of *EGFP*, *hoxb5a* and *hoxb5b* transcripts were estimated by PCR amplification of ~150 bp regions of each gene using the QuantiTect SYBR Green PCR Kit (Qiagen) with standard curves inferred from five different concentrations of a *hoxb5aEGFP* cDNA standard. Relative *EGFP* transcript numbers were quantified for each group and normalized to endogenous *hoxb5a* expression.

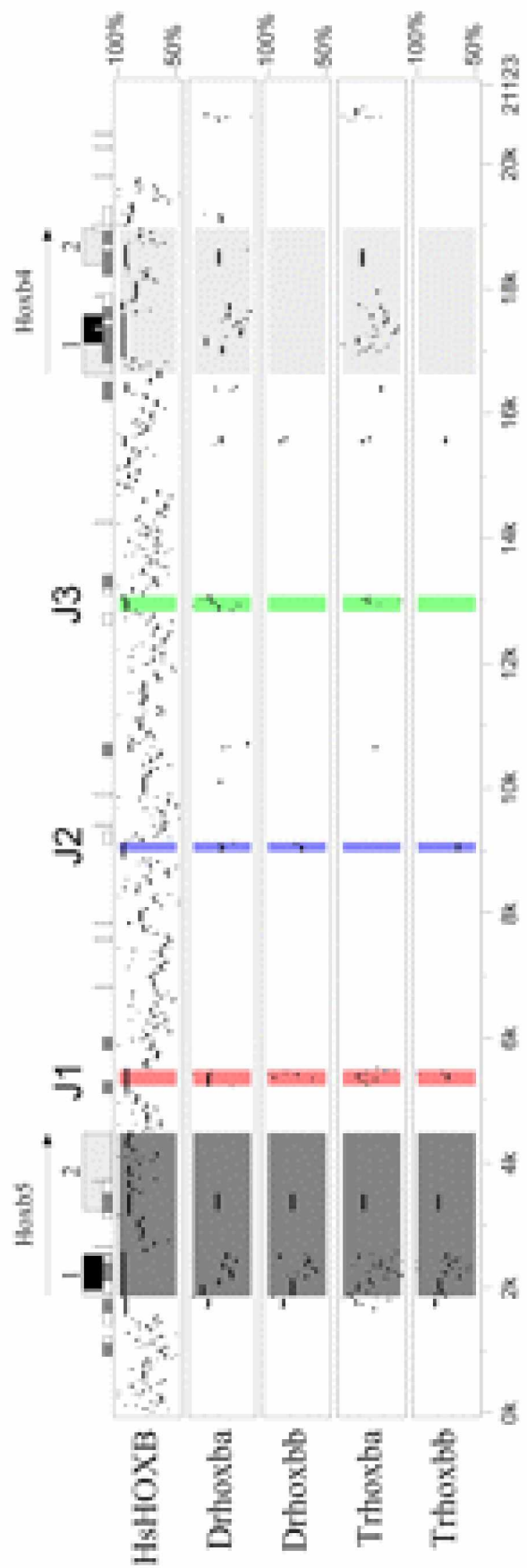
RESULTS

Identification of putative regulatory elements in the non-coding regions of paralogous group 5 *Hox* genes

In order to determine how the regulatory mechanisms that control the expression of the two teleost *hoxb5* genes can be compared with that of the single *Hoxb5* gene of mammals, we carried out a phylogenetic footprinting analysis of the sequences flanking the *hoxb5* transcription units from two mammalian species, mouse and human, and from the two teleosts, zebrafish and *Takifugu*. Using this approach, we identified three conserved sequences, named J1, J2 and J3, located in the 3' flanking region of *Hoxb5* loci (Fig. 2 and see Fig. S1 in the supplementary material; GenBank accession numbers BK006730-BK006737). The length of these sequences varied from ~120 to ~260 bp. J1 and J2

Fig. 2. Conserved non-coding elements in the *Hoxb5* genes of vertebrates.

Sequences from mouse, human, zebrafish and *Takifugu* were compared using Pipmaker. The mouse *Hoxb5* locus is used as a reference and is shown on the horizontal axis. Sequences from human (Hs), zebrafish (Dr) and *Takifugu rubripes* (Tr) are compared and the similarity is shown on the vertical axis. CNE positions are indicated by shading: red, J1; blue, J2; green, J3. A sequence similarity lower than 50% was observed between the J2 elements from *Takifugu hoxb5a* and mouse *Hoxb5* using a different algorithm. Identical results to those shown were obtained when teleost sequences were used as a reference. The annotations at the top refer to the mouse reference sequence: exons (black boxes), simple repeats (vertical rectangles) and CpG islands (white and gray rectangles).



sequences are both present in all the *Hoxb5* loci we examined. In mouse and human, the J1 and J2 sequences are 1.7 kb and 5.5 kb downstream of the *Hoxb5* stop codon, respectively (Fig. 2).

The sequence identity varied from 56 to 99% for the J1 element, depending on the species compared (see Table S1 in the supplementary material). Identity for J2 sequences varied between 46 and 90% in pairwise comparisons. The J2 sequences from *Takifugu* were more divergent and not initially detected by the Pipmaker algorithm (Fig. 2 and see Table S1 in the supplementary material). The third conserved sequence, J3, was found further downstream of the *Hoxb5* genes of mammals and teleosts, but was absent from the teleost *hoxb5b* loci (Fig. 2). This element showed identity that varied from 47 to 91% in pairwise comparisons.

We identified two additional conserved non-coding sequences in the *Hoxb5* loci. The first, located between J2 and J3, is found in all the mammalian and teleost species examined. It is found in *hoxb5a*, but not in *hoxb5b* loci (Fig. 2). This conserved sequence is smaller than the others (<50 bp). Another conserved region is located downstream of J3 (Fig. 2) and represents a previously identified microRNA gene (*mir10a*) that codes for a 22 nt non-coding RNA molecule that is suggested to be involved in the regulation of Hox gene expression (Mansfield et al., 2004). Finally, the immediate 5' flanking region of *Hoxb5* genes was well conserved in all species examined (69 to 94%, Fig. 2; see Table S1 in the supplementary material).

The conserved non-coding elements found in the 3' flanking region of *Hoxb5* fall within three large fragments (3-4.5 kb) of the mouse *Hoxb5/Hoxb4* intergenic region previously shown to have regulatory activity that recapitulates aspects of endogenous *Hoxb5* and *Hoxb4* expression (Sharpe et al., 1998). The zebrafish J1, J2 and J3 CNEs also correspond to sequences reported by Hadrys and colleagues in their comparative analysis of the *Hoxb* clusters in mammals and teleosts (Hadrys et al., 2004).

Transgenic analysis of cis-regulatory sequences found near the zebrafish *hoxb5* loci

We analyzed the activity of non-coding elements found near the zebrafish *hoxb5* loci by inserting the *lacZ* or *EGFP* reporter genes in frame with the first exon of either

hoxb5a or *hoxb5b* (see Materials and methods). The constructs encompassed 18.5 kb and 12.2 kb of the respective loci (Fig. 1A,C) and included ~1.2-1.3 kb of 5' flanking sequences. In the 3' direction, the *hoxb5a* construct spanned ~12 kb and encompassed the J1-J3 conserved non-coding sequences. It also contained the *hoxb4a* exons and intron, as well as 1.3 kb of *hoxb4a* 3' flanking sequence. Similarly, the *hoxb5b* construct extended over a distance of ~8 kb downstream of the last exon and encompassed the J1 and J2 conserved non-coding sequences.

Expression of the *hoxb5alacZ* and *hoxb5blacZ* transgenes was examined in mouse embryos at E10.5, E12.5 and E13.5. Transgenic mouse embryos, either primary or from established lines, showed *lacZ* expression in regions that coincide with domains of mouse *Hoxb5* expression (Fig. 3A-D, Fig. 4I-L) (Hogan et al., 1988). At E10.5, *hoxb5alacZ* embryos exhibited β -galactosidase activity along the embryonic neural tube, with the anterior border of reporter gene expression reaching the caudal hindbrain (Fig. 4A,B, arrows). By E12.5, the anterior boundary of *lacZ* expression shifted rostrally, extending to the caudal limit of the otic vesicle (Fig. 4I,J, arrows). This rostral extension of the expression domain might be mediated by a retinoic acid response element (RARE), previously characterized by Sharpe and colleagues (Sharpe et al., 1998), that resides within the J2 CNE. Furthermore, the *hoxb5alacZ* construct consistently targeted reporter gene expression to the paraxial mesoderm/anterior somites at all developmental stages examined (Fig. 4A, B, I, J, asterisks). Analysis of the *hoxb5alacZ* expression patterns on sagittal sections of E13.5 embryos revealed β -galactosidase activity in the cartilage primordia of embryonic ribs (Fig. 5A, B, cp).

At E10.5, the anterior boundary of *hoxb5blacZ* reporter expression in the central nervous system (CNS) was located at the caudal limit of the embryonic forelimbs, that is, more posteriorly than the CNS expression boundary observed for *hoxb5alacZ* (Fig. 4A-D, arrows). By E12.5, the rostral boundary of *hoxb5blacZ* reporter gene expression in the CNS extended up to the caudal limit of the otic vesicle, similar to that detected in the *hoxb5alacZ* transgenic embryos (Fig. 4J,L, arrows). In addition to the CNS, *hoxb5blacZ* was expressed in neural crest derivatives such as cranial and dorsal root ganglia and associated nerve

Fig. 3. Embryonic expression of mouse and zebrafish *Hoxb5* genes.

(A-D) Sagittal sections of E12.5 mouse embryos. The antisense *Hoxb5* probe consisted of a 430 bp fragment from the *Hoxb5* cDNA that includes the Nterminal part of the Hoxb5 protein (Krumlauf et al., 1987). (E-J) Whole-mount in situ hybridization on (E,F) 24 hpf, (G,I) 48 hpf and (H,J) 72 hpf zebrafish embryos using antisense riboprobes for *hoxb5a* (b5a, 550 bp probe) and *hoxb5b* (b5b, 600 bp probe). Zebrafish embryos are shown as lateral views. Arrows point to the embryonic neural tube. ba, branchial arches; g, gut; gn, trigeminal ganglion; l, lung; pd, pronephric duct; s, somites. Scale bars: 500 μm in A-D; 250 μm in E-J.

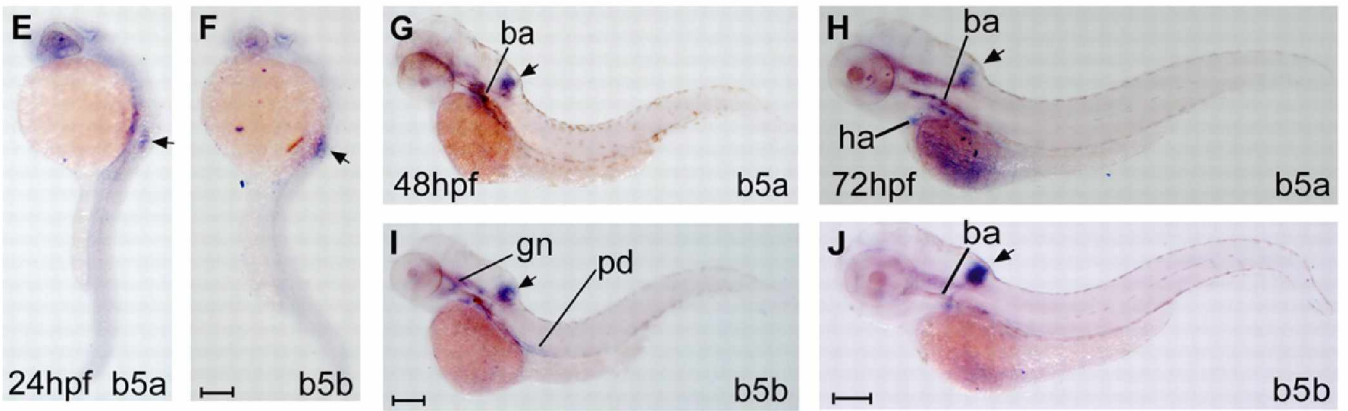
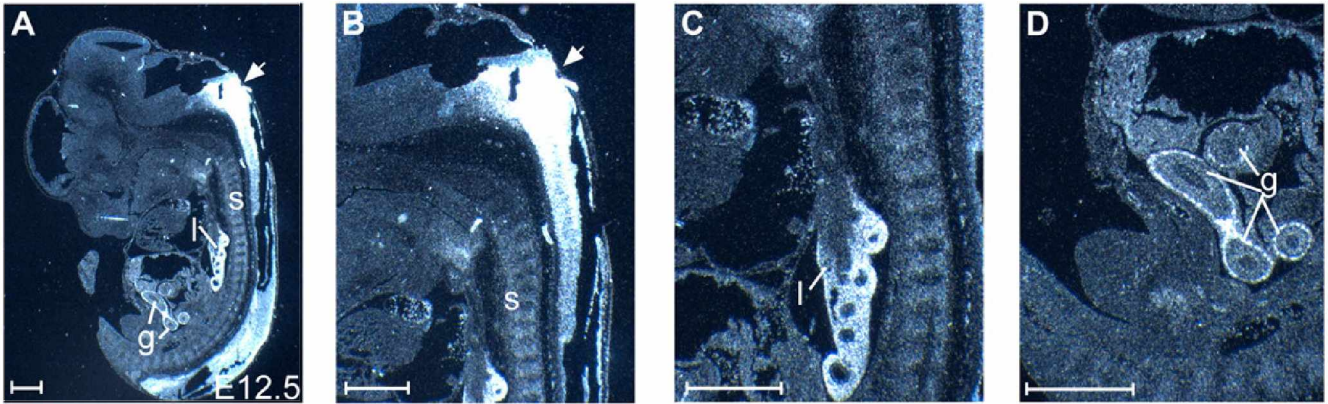


Fig. 4. Expression of *lacZ* reporter transgenes in mouse embryos.

E10.5 (A-H) and E12.5 (I-P) transgenic mouse embryos, as indicated. A,C,E,G,I,K,M,O are lateral views; B,D,F,H,J,L,N,P are dorsal views. Arrows indicate the neural tube and asterisks indicate somites. fg, facioacoustic ganglion; ng, nodose ganglion. Scale bars: 500 μm .

fibers. At E10.5 *hoxb5lacZ* expression was detected in the nodose, tenth cranial nerve, facio-acoustic and dorsal root ganglia (Fig. 4C, ng, fg). Expression persisted in these domains until at least E13.5 (Fig. 5D,E).

Dissection of internal organs from *hoxb5alacZ* mouse embryos at E13.5 revealed β -galactosidase activity in the midgut in a punctuate distribution (Fig. 5C,I). This might reflect a dot-like pattern of *Hoxb5* expression in the enteric nervous system similar to that reported for the *Hoxa5* gene (Larochelle et al., 1999). In addition, *hoxb5alacZ* was expressed in the stomach, meta- and mesonephros and adrenal gland (Fig. 5I, st, k, ag). A β -galactosidase signal of lower intensity was detected in the stomach, midgut and adrenal gland in *hoxb5blacZ* embryos (Fig. 5F,J).

Expression of *hoxb5alacZ* was stable and consistent between embryos from established lines at all embryonic stages examined (Fig. 4 and data not shown). By contrast, embryos obtained from *hoxb5blacZ* males showed patterns of transgene expression that differed between embryos collected from the same mating (see Fig. S2 in the supplementary material). Only embryos with the strongest X-Gal staining are shown in Figs 4, 5. This phenomenon was observed until at least the F2 generation obtained from an outcross of transgenic F1 males to wild-type females. Similar constructs containing *EGFP* as the reporter were tested in primary transgenic zebrafish embryos. The *hoxb5aEGFP* and *hoxb5bEGFP* constructs recapitulated aspects of the endogenous *hoxb5a* and *hoxb5b* expression. In most cases, the anterior border of reporter gene expression corresponded to the posterior hindbrain, reflecting the rostral restriction of endogenous *hoxb5* gene expression detected by in situ hybridization (Fig. 3E-J, Fig. 6A-I).

At 24-36 hours post-fertilization (hpf), *hoxb5aEGFP* expression was observed in cells of the developing CNS (Fig. 6A, arrows) and somites (Fig. 6A, arrowheads), mirroring *hoxb5a* expression (Fig. 3 and data not shown). At later stages, strong reporter expression persisted in the developing CNS and was also observed in muscle cells along the anterior-posterior axis (Fig. 6B,C). At 48 hpf, 10% of primary *hoxb5aEGFP* embryos expressed the transgene in the heart and the cardinal veins (data not shown).

In zebrafish, the *hoxb5b* sequences directed reporter gene expression preferentially to cells of neuronal origin at all developmental stages examined (Fig. 6D,E), similar to what was found in transgenic mouse embryos. This contrasted with the mesodermal activity recorded for the paralogous *hoxb5a* locus (Fig. 6A-C, arrowheads). Occasionally, *hoxb5bEGFP* expression was observed in cells of hematopoietic origin (data not shown).

Consistent with data obtained in transgenic mice, the levels of *hoxb5aEGFP* expression generally appeared to be higher than those seen in *hoxb5bEGFP* embryos, based on visual inspection and on the total number of cells expressing the transgenes. Furthermore, the *hoxb5aEGFP* construct yielded a much larger proportion of *EGFP* positive embryos (defined as embryos containing two or more *EGFP*-expressing cells) than the *hoxb5bEGFP* construct (60% as compared with 30%). To further address these differences, we performed real-time quantitative RT-PCR on RNA extracted from 100 primary transgenic embryos for each construct. The relative levels of *EGFP* transcripts in each group were estimated and normalized to those of endogenous *hoxb5a*. As shown in Fig. 6J, *hoxb5aEGFP* embryos had 104-fold higher *EGFP* transcript levels than *hoxb5bEGFP* embryos.

Combined, the results obtained from the mouse and zebrafish transgenesis experiments demonstrate that the DNA sequences included in the constructs are able to recapitulate most of the *hoxb5a* and *hoxb5b* expression domains except for the notochord and branchial arches. Furthermore, the differences in expression of the *hoxb5a* and *hoxb5b* constructs appear to be associated with those aspects of expression that differ between paralogs. For example, one of the most striking differences observed between *hoxb5a* and *hoxb5b* expression is the exclusive expression of *hoxb5b* in structures of the peripheral nervous system, such as the nerve ganglia, and this was mimicked by *hoxb5blacZ* but not by *hoxb5alacZ* (Fig. 4C, Fig. 5D,E). Similarly, the high expression levels of *hoxb5a* in somites and their derivatives (Bruce et al., 2001) were recapitulated by *hoxb5aEGFP* (Fig. 4B,J, asterisks; Fig. 5A,B, cp; Fig. 6A,C, arrowheads).

Fig. 5. Analysis of reporter transgene expression on sagittal sections and in the internal organs of mouse embryos.

(A-H) Sagittal sections from E13.5 transgenic mouse embryos, as indicated. Arrows indicate the neural tube. (I-L) Internal organs from E13.5 transgenic embryos, as indicated. cp, cartilage primordium; DRG, dorsal root ganglia; g, gut; k, kidney; st, stomach; ag, adrenal glands; X, trigeminal ganglia. Scale bars: 330 μm in A-D,F,G; 1000 μm in E,H; 125 μm in I,J,L; 250 μm in K.

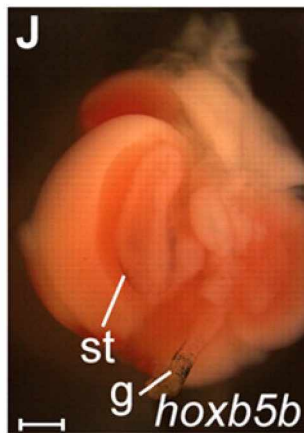
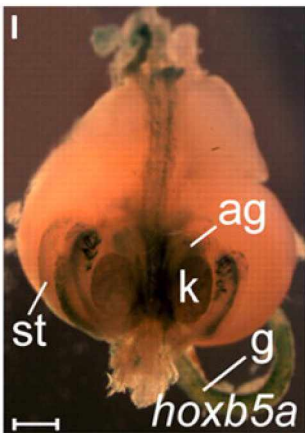
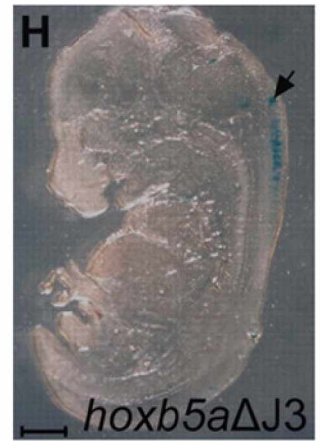
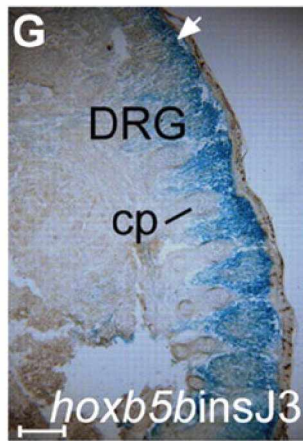
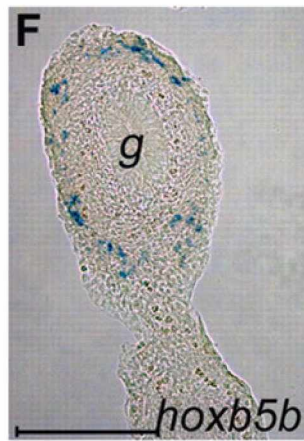
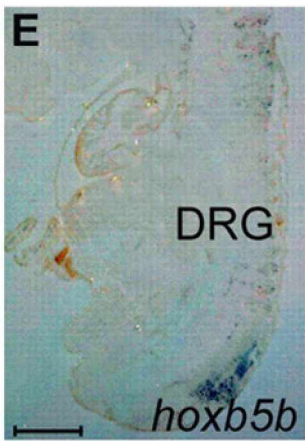
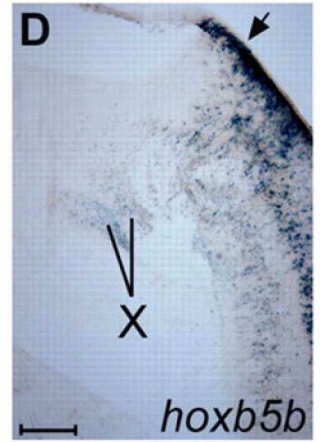
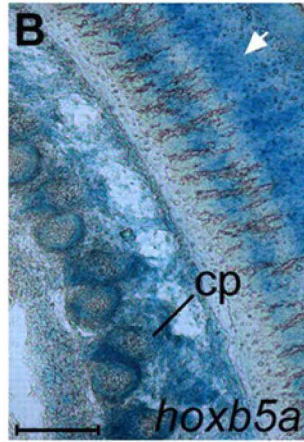
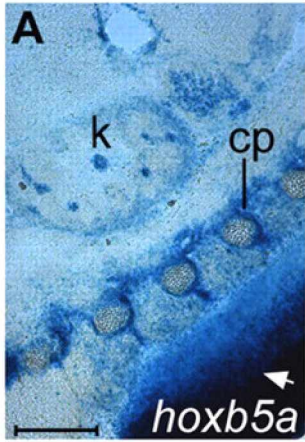
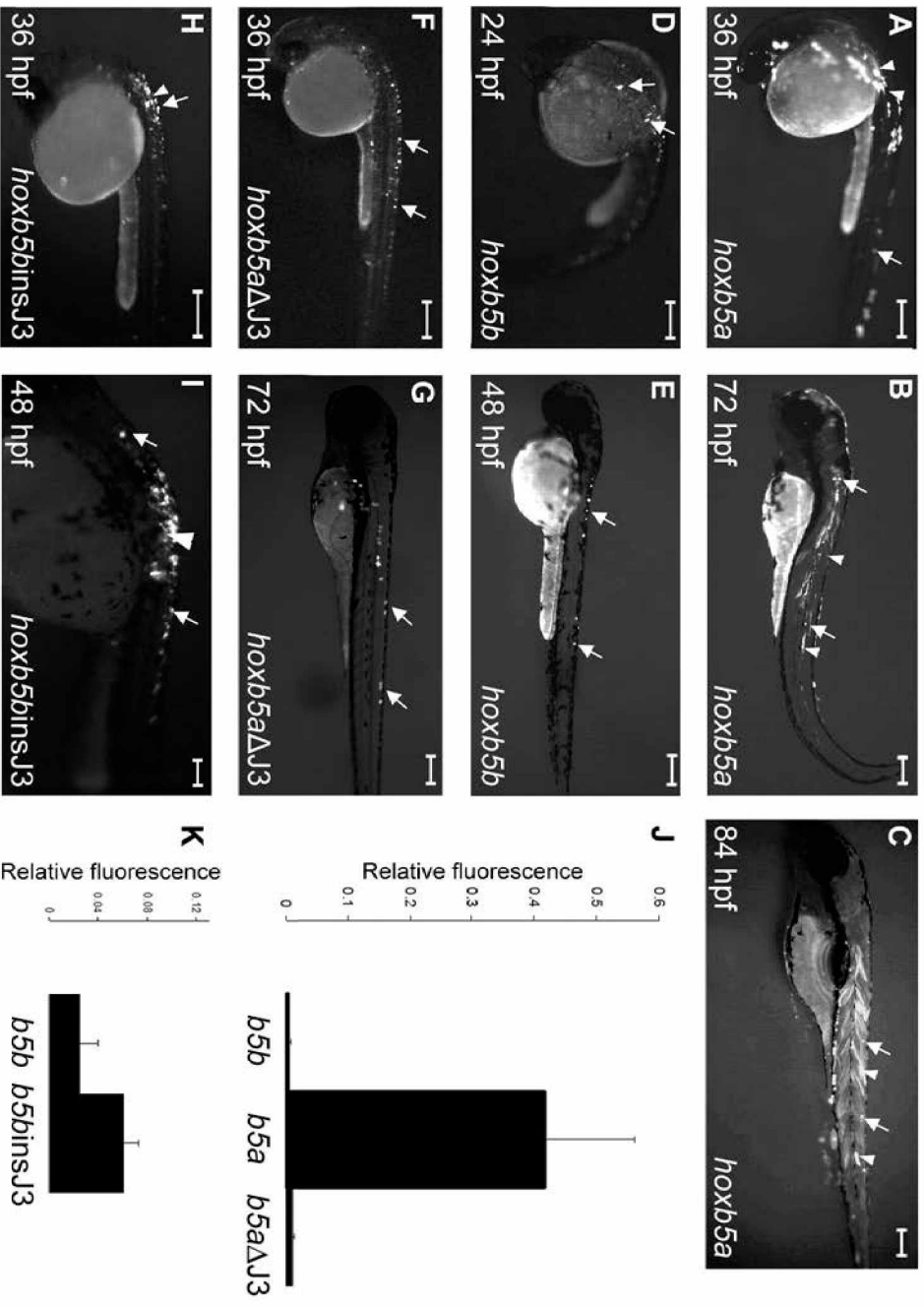


Fig. 6. Expression of *EGFP* reporter transgenes in zebrafish embryos.

(A-I) Lateral views (anterior to the left) of primary transgenic zebrafish embryos injected with (A-C) *hoxb5aEGFP*, (D,E) *hoxb5bEGFP*, (F,G) *hoxb5aΔJ3EGFP* or (H,I) *hoxb5binsJ3EGFP*. Developmental stages are indicated. Arrows indicate cells of the developing nervous system; arrowheads mark somites and muscle cells. Approximately 60% of the embryos injected with *hoxb5bEGFP* and 50% of embryos injected with *hoxb5binsJ3EGFP* showed *EGFP* expression, as compared with only 30% of *hoxb5bEGFP* and 37.5% of *hoxb5aΔJ3EGFP* embryos. Scale bars: 250 μm. (J,K) Relative copy numbers of *EGFP* transcripts detected by RT-PCR in primary transgenic embryos injected with *hoxb5aEGFP* (*b5a*), *hoxb5bEGFP* (*b5b*), *hoxb5aΔJ3EGFP* (*b5aΔJ3*) or *hoxb5binsJ3EGFP* (*b5binsJ3*) constructs. Column heights represent relative copy numbers of *EGFP* transcripts after normalization to endogenous *hoxb5a* expression. Each bar represents the average of three groups of transgenic embryos and error bars indicate s.e.m.



As the J3 element is only present in *hoxb5a*, and presumably degenerated in *hoxb5b*, we hypothesized that it might be involved in directing unique aspects of *hoxb5a* expression. To test this, we deleted J3 from the *hoxb5a**lacZ* and *hoxb5a**EGFP* reporter constructs, yielding *hoxb5a* Δ J3*lacZ* and *hoxb5a* Δ J3*EGFP*. Conversely, we also inserted J3 into the *hoxb5b**lacZ* and *hoxb5b**EGFP* constructs to examine whether this would render *hoxb5b* expression more similar to that of *hoxb5a*; the resulting constructs were named *hoxb5bins*J3*lacZ* and *hoxb5bins*J3*EGFP*.

In mouse, the *hoxb5a*_J3*lacZ* construct targeted *lacZ* expression to several domains of *Hoxb5* expression where its counterpart, *hoxb5a**lacZ*, was also expressed. However, expression levels were markedly diminished. Thus, at E10.5, *hoxb5a* Δ J3*lacZ* was expressed in the spinal cord, with anterior limits of expression reaching the base of the hindbrain (Fig. 4G,H, arrows), but the expression was reduced compared with that observed in *hoxb5a**lacZ* embryos (Fig. 4A,B; data not shown). Similar to *hoxb5a**lacZ*, the *hoxb5a* Δ J3*lacZ* construct was expressed in somites and their derivatives at E10.5 and E12.5 (Fig. 4G,H,O,P, asterisks).

Similar observations were made in transgenic zebrafish embryos. Thus, the *hoxb5a* Δ J3*EGFP* construct (Fig. 6F,G) exhibited both neuronal and mesodermal expression similar to that recorded for its intact counterpart, *hoxb5a**EGFP* (Fig. 6B; data not shown). However, considerably fewer *EGFP*-expressing cells were observed. Furthermore, transgene expression appeared significantly diminished in cells of mesodermal origin as compared with that observed in *hoxb5a**EGFP* zebrafish embryos (data not shown). When quantified by RT-PCR, transcript levels for *hoxb5a* Δ J3*EGFP* resembled those seen in *hoxb5b**EGFP* transgenic embryos, rather than those characteristic of *hoxb5a**EGFP* embryos (Fig. 6J).

The rostral boundary of *hoxb5bins*J3*lacZ* expression in the E10.5 mouse neural tube was shifted anteriorly compared with that of *hoxb5b**lacZ* (Fig. 4E,F). Furthermore, high levels of *hoxb5bins*J3*lacZ* expression were detected in the gut and stomach (Fig. 5K, st, g), similar to *hoxb5a**lacZ*. Finally, the *hoxb5bins*J3*lacZ* construct drove reporter gene

expression in somites and their derivatives (Fig. 4E,F,M,N, asterisks), whereas the *hoxb5lacZ* did not (Fig. 4C,D,K,L).

Levels of the *hoxb5binsJ3lacZ* reporter were apparently higher than those driven by the *hoxb5lacZ* construct (Fig. 4K-N). Moreover, the *hoxb5binsJ3lacZ* construct drove more consistent patterns of reporter gene expression than its parental counterpart, and the variegated patterns of expression seen with *hoxb5lacZ* were only observed in 10% of *hoxb5binsJ3lacZ* embryos (data not shown). When injected into zebrafish embryos, the *hoxb5binsJ3EGFP* construct drove reporter expression in both neural and mesodermal cells (Fig. 6I, arrows and arrowhead), whereas *hoxb5bEGFP* only exhibited neural expression (Fig. 6D,E, arrows). The *hoxb5binsJ3EGFP* construct also produced greater numbers of *EGFP*-positive embryos (Fig. 6H,I) and levels of *EGFP* expression were increased 2.4-fold compared with *hoxb5bEGFP* (Fig. 6K).

Taken together, the above results suggest that the J3 element is necessary to maintain higher levels of *hoxb5* gene expression and is at least partially responsible for the exclusive expression of *hoxb5a* in cells of mesodermal origin (Bruce et al., 2001). Deleting the J3 sequence from the *hoxb5a* locus reduced overall reporter transgene expression, whereas introduction of J3 into the *hoxb5b* locus resulted in higher expression levels and induced reporter expression in cells uniquely associated with *hoxb5a* expression.

Transgenic analysis of individual non-coding elements

To assess the individual contributions of the J1, J2 and J3 CNEs to overall *Hoxb5* expression, we generated reporter constructs carrying individual CNEs directing expression of *lacZ* or *EGFP* from a β -globin minimal promoter (Fig. 1E-G). Primary transgenic mouse embryos were examined at E12.5-14.5 and primary transgenic zebrafish embryos were examined at 24-96 hpf.

When tested in transgenic mouse embryos, each of the conserved sequences from the mouse *Hoxb5* locus or from the zebrafish *hoxb5a* and *hoxb5b* loci drove reporter transgene expression in domains representative of endogenous *Hoxb5* expression (Oosterveen et al., 2003; Sakach and Safaei, 1996). A transgene containing the mouse J1

element (MmJ1) directed *lacZ* expression in structures of the paraxial mesoderm (Fig. 7A) and neural tube (Table 1, data not shown). The mouse J2 element, MmJ2, induced reporter gene expression in the neural tube and in prevertebrae along the entire anterior-posterior axis (Fig. 7D, Table 1). Finally, the construct containing MmJ3 drove *lacZ* expression in the neural tube, developing vertebrae and posterior somites (Fig. 7G, Table 1). Individual zebrafish elements from either *hoxb5a* or from *hoxb5b* induced transgene expression in prevertebrae, in the neural tube and, occasionally, in the dorsal root ganglia and/or associated nerves (Fig. 7, Table 1). No major differences were observed between the activity of the zebrafish CNEs and their mouse counterparts, except for an inability of zebrafish J3 (DrJ3) to target expression to the somites/prevertebrae (Table 1).

When tested in transgenic zebrafish, the murine CNEs showed a spectrum of activities comparable to those observed in transgenic mice (Fig. 8, Table 1). MmJ1 targeted reporter expression to neural and mesodermal cells (Fig. 8C, Table 1; data not shown). MmJ2 drove expression almost exclusively in cells of mesodermal origin, in contrast to the strong neural activity detected in transgenic mice (Fig. 7D, Table 1). The mouse J3 sequence also targeted transgene expression almost exclusively to cells of mesodermal origin. The individual zebrafish CNEs targeted *EGFP* expression in transgenic zebrafish with patterns resembling the endogenous *hoxb5a* and *hoxb5b* genes. Activity was generally similar to that of the mouse CNEs. Exceptions included the observation that zebrafish J1A was less active in neural cells than mouse J1 or zebrafish J1B, and that zebrafish J2 sequences targeted expression to neural cells much better than did the orthologous mouse sequence. No other major differences were observed between the activity of an individual element from the *hoxb5a* locus and its paralogous sequence from the *hoxb5b* locus.

The above comparative analysis revealed that regulatory activity conferred by the cognate CNEs from duplicated zebrafish loci extensively overlapped and, in general, corresponded to expression patterns common to both *hoxb5* orthologs.

Table 1. Expression of reporter constructs in primary transgenic mouse and zebrafish embryos

Element tested	Mouse injections				Zebrafish injections	
	# Expressing/ # Tg	Expression sites			Neural	Mesodermal
		Neural tube	DRG	Somites/ prevertebrae		
MmJ1	5/5	4	2	4	20% (11/55)	100% (55/55)
DrJ1a)	4/7	3	2	2	2% (1/48)	100% (48/48)
DrJ1b	3/10	2	1	2	30% (22/74)	80% (59/74)
MmJ2	3/6	3	2	3	0% (0/42)	100% (42/42)
DrJ2a	2/2	2	1	1	13% (7/54)	93% (50/54)
DrJ2b	4/8	3	1	1	71% (52/73)	67% (49/73)
MmJ3	4/8	2	1	2	2% (1/49)	100% (49/49)
DrJ3a	6/8	5	2	0	4% (2/54)	98% (53/54)

DRG, dorsal root ganglion

Fig. 7. Enhancer activity of individual *Hoxb5* non-coding elements in transgenic mouse embryos.

(A-H) Reporter transgene expression was analyzed in primary transgenic embryos whose age varied from E12.5 to E14.5. The name and origin of the tested CNE are shown at the bottom right-hand corner of each panel. Arrowheads mark somite derivatives; arrows point to the neural tube. Mm, *Mus musculus*; Dr, *Danio rerio*; a and b refer to the *hoxba* and *hoxbb* zebrafish complexes, respectively. Scale bar: 1000 μm .

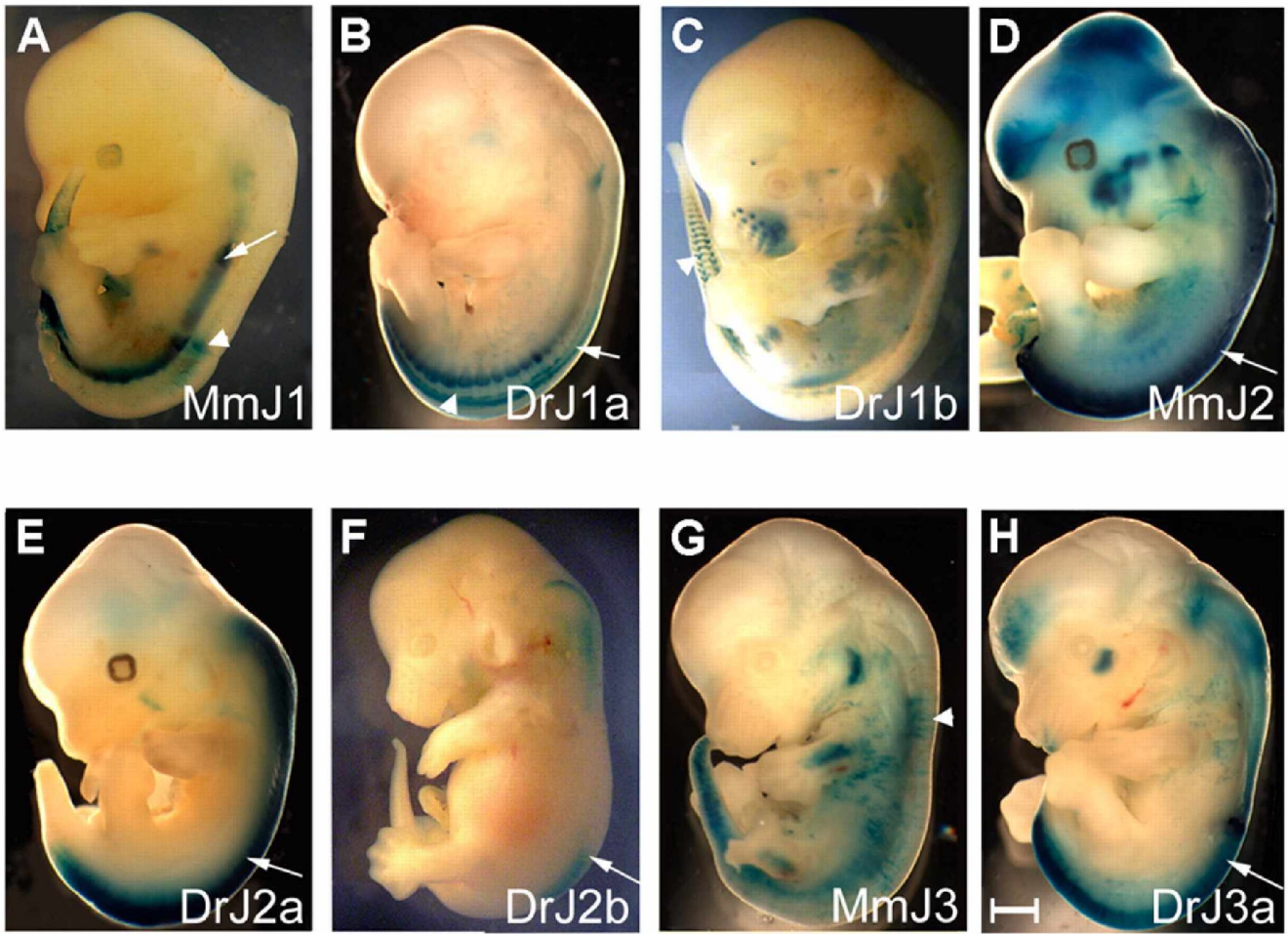
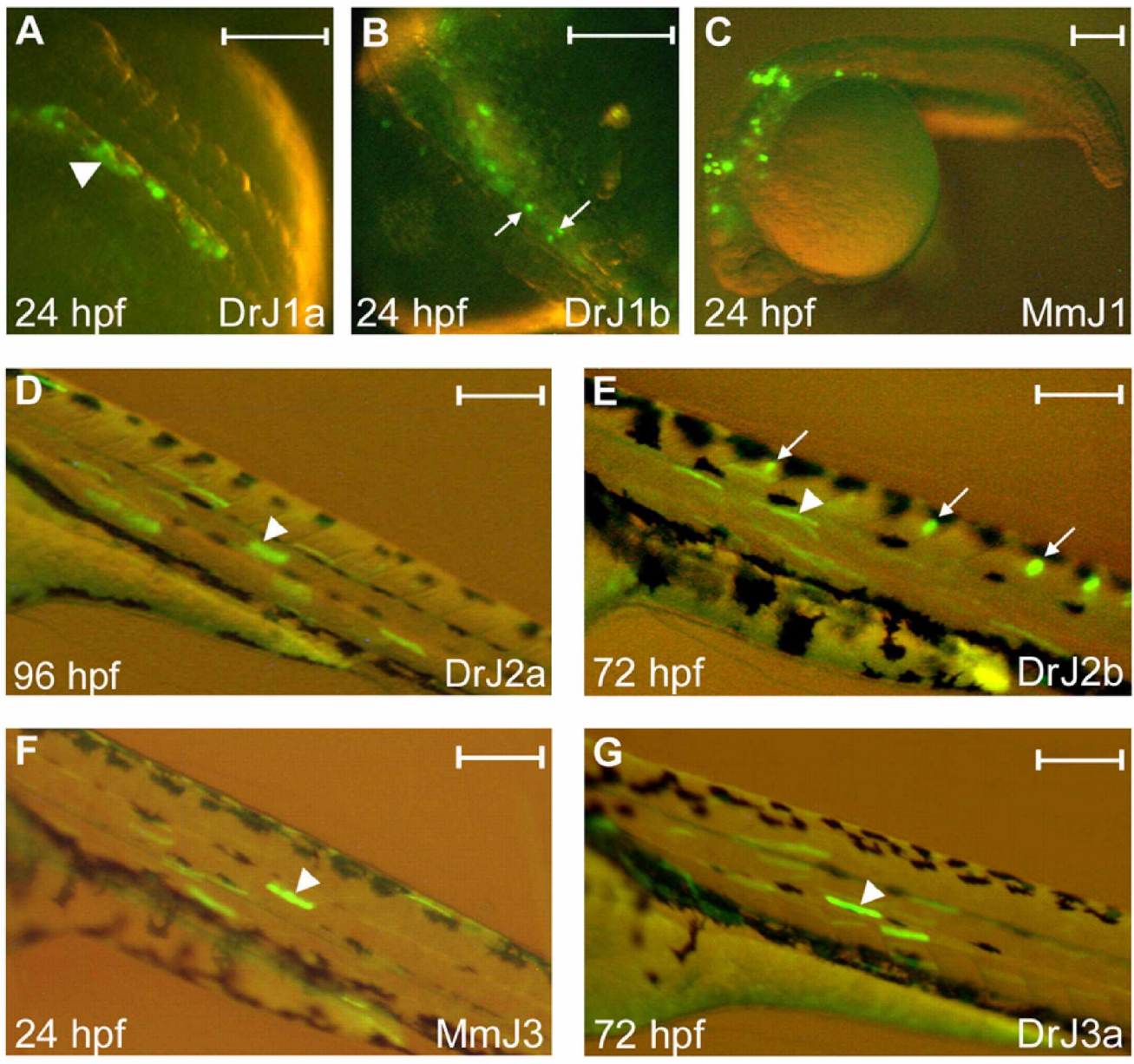


Fig. 8. Enhancer activity of individual *Hoxb5* non-coding elements in transgenic zebrafish.

(A-G) Lateral views, anterior to the left. Arrows in B and E indicate cells of the developing CNS. Arrowheads indicate somites in A and muscle cells in D-G. The name and origin of each regulatory element are indicated as in Fig. 7. Age, in hpf, is indicated. Scale bars: 500 μm .



DISCUSSION

Gene duplication has long been considered a major source of evolutionary novelty. Paralogous genes resulting from duplication events have the opportunity to diverge and acquire new functions. Divergence in the expression patterns of duplicated genes has been proposed to be the first step in this process. The DDC model originally proposed by Force and colleagues states that subfunctionalization is a common mechanism of duplicated gene preservation (Force et al., 1999). The model makes the following predictions with regards to molecular mechanisms of subfunctionalization. (1) Subfunctionalization happens by means of changes in the non-coding regions of duplicated genes. (2) Resolution of functional redundancy occurs through complementary degeneration of individual regulatory elements. This requires the regulation of an ancestral gene to be modular in nature; the gene subfunctions have to be independent from each other and executed by discrete regulatory elements. (3) The process of subfunctionalization should be completed within 12.5 million years after a duplication event.

This model has received a lot of attention in the field of evolutionary developmental biology. The fates of several gene duplicates were investigated in an attempt to bridge the mathematical predictions of the DDC model with experimental observations. Studies of the paralogous *Nudt10* and *Nudt11* genes in mammals, the *sox9a* and *sox9b* and the *mitf-m* and *mitf-b* genes in teleost (Altschmied et al., 2002; Hua et al., 2003; Kluver et al., 2005) confirmed that complementary loss of independent subfunctions (subfunctionalization) may indeed constitute a common mechanism of resolution of functional redundancy between duplicated genes. However, these studies did not address the molecular changes that occurred at the level of the regulatory elements that instigated subfunctionalization.

To clarify how the dynamics of cis-regulatory sequence evolution support the DDC model, Santini et al. (Santini et al., 2003) and Wolfe and Elgar (Wolfe and Elgar, 2007), performed comparative sequence analyses of multiple gene loci that are duplicated in teleosts but are present at single copy in mammals. They found that paralogous genes in teleost often retain differential subsets of putative regulatory elements, consistent with the notion of regulatory subfunctionalization (Santini et al., 2003; Woolfe and Elgar, 2007).

We characterized the structure and function of regulatory elements from the subfunctionalized zebrafish *hoxb5a* and *hoxb5b* genes to determine whether the evolution of these paralogs occurred in accordance with the predictions of the DDC model. We identified conserved non-coding elements and tested their activity collectively in a context that is close to their natural environment. Changes occurring in CNEs from duplicated genes were previously shown to be associated with the differential domains of expression of coparalogs (Kleinjan et al., 2008; Tumpel et al., 2006). In these two studies, CNEs were tested individually in reporter constructs that were used to produce transgenic animals in the endogenous species or in more-amenable model species. Although this approach can reveal changes in cis-regulatory elements that might account for some aspects of the differential expression patterns of the co-paralogs, it does have distinct limitations, such as a failure to reveal regulatory interactions as demonstrated by the results of the present study.

Execution of complementary subfunctions of the *hoxb5a* and *hoxb5b* genes may rely on interactions between multiple cis-regulatory elements

When tested in transgenic animals, large fragments from the zebrafish *hoxb5a* and *hoxb5b* loci targeted reporter gene expression to domains of endogenous *Hoxb5* expression that were consistent with the differences detected in the expression patterns of the two paralogs (Figs 3, 4, 5) (Bruce et al., 2001). For example, the rostral boundary of *hoxb5blacZ* expression in the neural tube was shifted posteriorly at E10.5 as compared with *hoxb5alacZ* embryos (Fig. 4). We then questioned if, consistent with the DDC model, complementary degenerative changes within individual elements are responsible for changes in expression. The J3 element is only retained in the *hoxb5a* locus and is lost from *hoxb5b*. Transgenic data obtained in both zebrafish and mice suggest that the loss of J3 might have contributed to divergence in expression between the zebrafish *hoxb5a* and *hoxb5b* genes (Figs 4, 6).

By contrast, functional tests of individual regulatory elements from zebrafish *hoxb5a* and *hoxb5b* did not reveal clear differences in activity (Figs 7, 8). MmJ2, DrJ2a and DrJ2b targeted transgeneexpression to similar domains in the mouse neural tube with a correct anterior border of expression. In addition, when tested individually, *hoxb5a* elements occasionally showed regulatory activity associated with domains exclusive to

hoxb5b. For example, the DrJ3a and DrJ1a elements occasionally drove *lacZ* expression in the dorsal root ganglia and associated nerve fibers (Table 1, data not shown), an expression domain that correlates with zebrafish *hoxb5b* rather than *hoxb5a* expression.

Thus, contrary to the results of experiments involving large fragments from the zebrafish *hoxb5* loci, the functional analysis of individual enhancer elements did not reveal clear complementation in activities between CNEs from the duplicated zebrafish *hoxb5* genes. Therefore, we propose that some complementary *hoxb5* subfunctions rely on interactions between regulatory elements, rather than being executed by individual enhancers. Alternatively, the activity of J1, J2 and J3 might be fine-tuned by other regulatory elements that are included in the *hoxb5a* and *hoxb5b* reporter constructs but which have not been identified in this study.

The dynamics of change in the cis-acting regulatory elements of *hoxb5* duplicates are more complex than predicted by the DDC model

The DDC model predicts that regulatory elements of duplicated genes will undergo rapid complementary degeneration. Partition of an original CNE regulatory function, although not suggested by our transgenic analysis, could also be supported by phylogenetic analysis.

We aligned the J1, J2 and the immediate 5' flanking sequences from all *Hoxb5* loci and examined the alignments for the presence of short (6-12 bp) highly conserved (>91%) sequences. Such sequences, which are evolving at an exceptionally slow rate, are likely to be functionally important. Although we identified short DNA sequences that were conserved in cognate elements of mammals and teleosts (Table 2), we were unable to find sequences that were specific for one of the two teleost paralogs (e.g. *hoxb5a* from zebrafish or *Takifugu*), but divergent or absent in the other teleost paralogs (e.g. *hoxb5b*) (data not shown). Thus, the data do not provide evidence for a simple complementary degeneration of regulatory subfunctions within the elements present in both teleost *hoxb5* loci.

We also scanned *Hoxb5* loci for putative transcription factor binding sites (Table 2). Among those that were found, some corresponded to transcription factors known to be

Table 2. Potential transcription factor binding sites identified in the conserved *Hoxb5* non-coding sequences

Sequence (5' to 3')	Transcription factor	CNS	Position on Mm <i>Hoxb</i> contig (NT_165773)
TTACCT	Rar β	5' region of <i>hoxb5a</i>	96164708-96164713
AGAATTT	c/EBP	5' region of <i>hoxb5a</i>	96164730-96164736
TTTACGA	Hoxa9	5' region of <i>hoxb5a</i>	96164734-96164740
CACGTG	Smad3/c-Myc	5' region of <i>hoxb5a</i>	96164755-96164760
CCATATTTGG	Mcm1/Srf	5' region of <i>hoxb5a</i>	96164798-96164807
TGACATT	Hoxa9/Meis1a	J1	96168219-96168225
TCGTAAA	Hoxa9/Meis1a	J1	96168265-96168271
CACGTG	c-Myc	J1	96168318-96168323
TTTATGG	Hoxa9/Meis1a	J1	96168328-96168334
CATAAAGTG	Pax2.1	J1	96168329-96168337
TTTTATGGTTTA	Ubx	J1	96168339-96168351
TTTATGG	Hoxa9/Meis1a	J1	96168340-96168347
TAAGTG	c-Myb	J1	96168408-96168413
ATGAGA	XPF (Ercc4)	J2	96172038-96172044
TGATCC	GR (Nr3c1)	J2	96172057-96172062
AGGTCA	Rar α 1	J2	96172060-96172065
GGGTGA	Rxr α	J3	96175913-96175918
TGAACC	Rxr α	J3	96175916-96175921
AGGTCA	Rar α 1	J3	96175924-96175929
TAAAAT	Pou1F1a	J3	96175947-96175952
AACAAAG	Lef1, Sox3, Sox5	J3	96175984-96175990
AGATTA	Gata1/4	J3	96176005-96176010

The 6-12 bp sequences listed are highly conserved in all the species examined (mouse, human, zebrafish and *Takifugu*) and coincide with known transcription factor binding sites.

involved in Hox gene regulation. For example, a RARE known to be essential for correct anterior expression of both *Hoxb8* and *Hoxb5* genes in the mouse neural tube (Oosterveen et al., 2003) was found within J2. Another putative RARE was found in one of the highly conserved regions of J3 (Table 2).

The DDC model also predicts that degeneration of redundant subfunctions should occur within 4.0-12.5 million years after the duplication event (Force et al., 1999). Thus, functional specialization of the *hoxb5a* and *hoxb5b* genes in teleosts would have preceded the divergence of the zebrafish and *Takifugu* lineages, and one might expect higher sequence conservation between orthologous functional modules (i.e. *hoxb5a* from zebrafish and *hoxb5a* from *Takifugu*) than between paralogous modules (i.e. zebrafish *hoxb5a* and *hoxb5b*) (Fig. 9). We examined phylogenies based on the analysis of coding sequences of *Hoxb5* genes from human, mouse, zebrafish and *Takifugu* (Fig. 9B), as well as sequences from the 5' flanking regions of *Hoxb5* genes (Fig. 9C) and of the J1 and J2 elements (Fig. 9D,E). The tree built for the J1 CNE matched the branching pattern of the hypothetical model (Fig. 9A,D), whereas phylogeny for other functional modules did not (Fig. 9A-C,E). We also joined the sequences of the CNEs in a continuous sequence for each locus. The topology of the resulting tree (Fig. 9F) was similar to that calculated for the coding sequences and the 5' flanking regions of *Hoxb5* genes, suggesting that the coding region and most regulatory regions of *Hoxb5* loci are under common selective constraints or are drifting at the same rate. Overall, the dynamics of the molecular changes involved in the evolution of *hoxb5a* and *hoxb5b* teleost genes differ from those anticipated for duplicated genes that undergo subfunctionalization (Force et al., 1999). Alternatively, the DDC model might be imperfect in its timing aspects. In fact, the regulatory elements of *hoxb5a* and *hoxb5b* duplicates appear to diverge slower than anticipated.

Deviations from the predicted rate of divergence have also been reported for the duplicated teleost *sox9a* and *sox9b* genes (Cresko et al., 2003). Differences in the expression patterns of *sox9a* and *sox9b* in zebrafish and stickleback led to the conclusion

that, even though partitioning of most *sox9* subfunctions occurred before the divergence of the teleost lineages, some gene subfunctions might have assorted differently in the two teleost species.

In addition to its enhancer activity, the J3 element may be required for proper maintenance of *hoxb5* expression

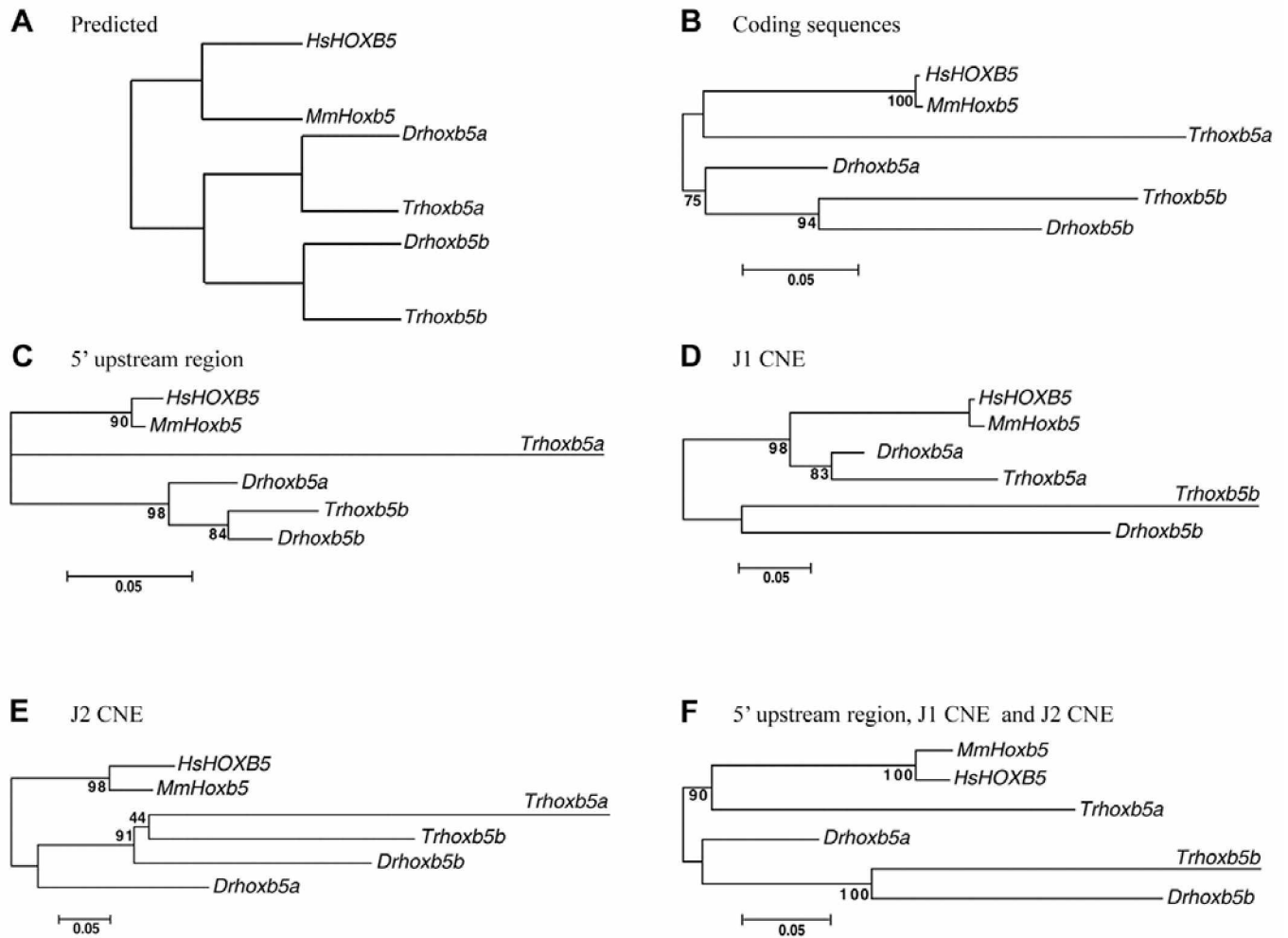
The absence of the J3 regulatory element in one of the two *hoxb5* loci is consistent with the DDC model. Sharpe and colleagues have previously referred to a large fragment of the mouse *Hoxb5/Hoxb4* intergenic region, encompassing J3, as the ‘mesodermal enhancer region’ (Sharpe et al., 1998). Our experiments revealed that the J3 element is not only responsible for the mesodermal *Hoxb5* expression but may also be required for maintaining high levels of *Hoxb5* expression. Indeed, reporter constructs containing J3 (*hoxb5a* and *hoxb5binsJ3*) appear to show higher expression levels than their J3-deleted counterparts (*hoxb5b* and *hoxb5aΔJ3*). This might also be associated with the intriguing observation that mouse embryos carrying *hoxb5blacZ* and *hoxb5aΔJ3lacZ*, the two transgenes lacking J3, showed variegated patterns of transgene expression among embryos from the same mating (see Fig. S2 in the supplementary material). Primary transgenic embryos also showed a similar variability in the transgene expression pattern (data not shown). As the number of independent integration events and the use of established lines argue against positional effects or mosaicism, we propose that one of the functions of the J3 element is to maintain gene expression. This additional function of J3 might involve interactions with proteins such as members of the Polycomb and Trithorax groups, which are known to regulate and maintain the transcriptional status of Hox genes through modification of chromatin structure.

CONCLUSIONS

Our data suggest that the patterns of regulatory evolution of teleost *hoxb5* duplicates involve mechanisms additional to those suggested by the DDC model (Force et al., 1999). Although phylogenetic filtering and functional tests of individual elements from the *hoxb5a* and *hoxb5b* loci did not reveal clear signs of complementation between the

Fig. 9. Phylogenetic analysis of *Hoxb5* loci from human, mouse, zebrafish and *Takifugu*.

(A) Phylogenetic tree predicted by the DDC model. Its topology is based on the assumption that *hoxba* and *hoxbb* teleost complexes resulted from a duplication event that happened after the separation of the fish and tetrapod lineages and before the divergence of *Takifugu* and zebrafish lineages. (B-F) Phylogenetic trees built with MEGA3 based on the following sequences: (B) *Hoxb5* coding sequences (850 bp); (C) 5' upstream region of *Hoxb5* genes (130 bp); (D) J1 CNE (255 bp); (E) J2 CNE (150 bp); (F) CNEs present in all *Hoxb5* loci (535 bp), for which the 5' upstream region, J1 and J2 were joined in a continuous single sequence for each *Hoxb5* locus. As the J3 CNE was only found in the *hoxb5a* complexes of zebrafish and *Takifugu*, it was not included in this analysis. Hs, human; Mm, mouse; Dr, zebrafish; Tr, *Takifugu*. The scale bar indicates an estimated evolutionary distance of 0.05 nucleotide substitutions per site. Numbers at nodes indicate the bootstrap values.



regulatory elements retained in both zebrafish loci, our results highlight the importance of interactions between CNEs in the execution of complementary subfunctions of duplicated genes.

We thank Adrianna Gambarotta, Marion Alfero, Marcelle Carter, Lucille Joly, Margot Lemieux and Vishal Saxena for technical help; Fabien Avaron for useful discussions; and Marie-Andrée Akimenko and David Lohnes for comments on the manuscript. This work was supported by a grant from NSERC, by OGS and NSERC awards to O.J., by a FRSQ Chercheur National Award to L.J. and by an Investigator award from CIHR to M.E.

REFERENCES

- Altschmied, J., Delfgaauw, J., Wilde, B., Duschl, J., Bouneau, L., Volff, J. N. and Scharl, M. (2002). Subfunctionalization of duplicate *mitf* genes associated with differential degeneration of alternative exons in fish. *Genetics* 161, 259-267.
- Amores, A., Force, A., Yan, Y. L., Joly, L., Amemiya, C., Fritz, A., Ho, R. K., Langeland, J., Prince, V., Wang, Y. L. et al. (1998). Zebrafish *hox* clusters and vertebrate genome evolution. *Science* 282, 1711-1714.
- Amores, A., Suzuki, T., Yan, Y. L., Pomeroy, J., Singer, A., Amemiya, C. and Postlethwait, J. H. (2004). Developmental roles of pufferfish *Hox* clusters and genome evolution in ray-fin fish. *Genome Res.* 14, 1-10.
- Amsterdam, A., Lin, S. and Hopkins, N. (1995). The *Aequorea victoria* green fluorescent protein can be used as a reporter in live zebrafish embryos. *Dev. Biol.* 171, 123-129.
- Bruce, A. E., Oates, A. C., Prince, V. E. and Ho, R. K. (2001). Additional *hox* clusters in the zebrafish: divergent expression patterns belie equivalent activities of duplicate *hoxB5* genes. *Evol. Dev.* 3, 127-144.
- Chiu, C. H., Dewar, K., Wagner, G. P., Takahashi, K., Ruddle, F., Ledje, C., Bartsch, P., Scemama, J. L., Stellwag, E., Fried, C. et al. (2004). Bichir *HoxA* cluster sequence reveals surprising trends in ray-finned fish genomic evolution. *Genome Res.* 14, 11-17.
- Copeland, N. G., Jenkins, N. A. and Court, D. L. (2001). Recombineering: a powerful new tool for mouse functional genomics. *Nat. Rev. Genet.* 2, 769-779.
- Cormack, B. P., Valdivia, R. H. and Falkow, S. (1996). FACS-optimized mutants of the green fluorescent protein (GFP). *Gene* 173, 33-38.
- Cresko, W. A., Yan, Y. L., Baltrus, D. A., Amores, A., Singer, A., Rodriguez-Mari, A. and Postlethwait, J. H. (2003). Genome duplication, subfunction partitioning, and lineage divergence: *Sox9* in stickleback and zebrafish. *Dev. Dyn.* 228, 480-489.
- Force, A., Lynch, M., Pickett, F. B., Amores, A., Yan, Y. L. and Postlethwait, J. (1999). Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* 151, 1531-1545.

- Fraidenraich, D., Lang, R. and Basilico, C. (1998). Distinct regulatory elements govern Fgf4 gene expression in the mouse blastocyst, myotomes, and developing limb. *Dev. Biol.* 204, 197-209.
- Hadrys, T., Prince, V., Hunter, M., Baker, R. and Rinkwitz, S. (2004). Comparative genomic analysis of vertebrate Hox3 and Hox4 genes. *J. Exp. Zool. B Mol. Dev. Evol.* 302, 147-164.
- Hogan, B., Constantini, F. and Lacy, E. (1986). *Manipulating the Mouse Embryo: a Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
- Hogan, B. L., Holland, P. W. and Lumsden, A. (1988). Expression of the homeobox gene, Hox 2.1, during mouse embryogenesis. *Cell Differ. Dev.* 25 Suppl, 39-44.
- Hua, L. V., Hidaka, K., Pesesse, X., Barnes, L. D. and Shears, S. B. (2003). Paralogous murine Nudt10 and Nudt11 genes have differential expression patterns but encode identical proteins that are physiologically competent diphosphoinositol polyphosphate phosphohydrolases. *Biochem. J.* 373, 81-89.
- Jaffe, L., Jeannotte, L., Bikoff, E. K. and Robertson, E. J. (1990). Analysis of beta 2-microglobulin gene expression in the developing mouse embryo and placenta. *J. Immunol.* 145, 3474-3482.
- Kleinjan, D. A., Bancewicz, R. M., Gautier, P., Dahm, R., Schonhaler, H. B., Damante, G., Seawright, A., Hever, A. M., Yeyati, P. L., van Heyningen, V. et al. (2008). Subfunctionalization of duplicated zebrafish pax6 genes by cisregulatory divergence. *PLoS Genet.* 4, e29.
- Klüber, N., Kondo, M., Herpin, A., Mitani, H. and Schartl, M. (2005). Divergent expression patterns of Sox9 duplicates in teleosts indicate a lineage specific subfunctionalization. *Dev. Genes Evol.* 215, 297-305.
- Krumlauf, R., Holland, P. W., McVey, J. H. and Hogan, B. L. (1987). Developmental and spatial patterns of expression of the mouse homeobox gene, Hox 2.1. *Development* 99, 603-617.
- Kumar, S., Tamura, K., Jakobsen, I. B. and Nei, M. (2001). MEGA2: molecular evolutionary genetics analysis software. *Bioinformatics* 17, 1244-1245.
- Larochelle, C., Tremblay, M., Bernier, D., Aubin, J. and Jeannotte, L. (1999). Multiple cis-acting regulatory regions are required for restricted spatio-temporal Hoxa5 gene expression. *Dev. Dyn.* 214, 127-140.
- Lynch, M. and Force, A. (2000). The probability of duplicate gene preservation by subfunctionalization. *Genetics* 154, 459-473.
- Mansfield, J. H., Harfe, B. D., Nissen, R., Obenaus, J., Srineel, J., Chaudhuri, A., Farzan-Kashani, R., Zuker, M., Pasquinelli, A. E., Ruvkun, G. et al. (2004). MicroRNA-responsive 'sensor' transgenes uncover Hox-like and other developmentally regulated patterns of vertebrate microRNA expression. *Nat. Genet.* 36, 1079-1083.
- Ohno, S. (1970). *Evolution by Gene Duplication*. Heidelberg, Germany: Springer-Verlag.
- Oosterveen, T., Niederreither, K., Dolle, P., Chambon, P., Meijlink, F. and Deschamps, J. (2003). Retinoids regulate the anterior expression boundaries of 5' Hoxb genes in posterior hindbrain. *EMBO J.* 22, 262-269.
- Sakach, M. and Safaei, R. (1996). Localization of the HoxB5 protein in the developing CNS of late gestational mouse embryos. *Int. J. Dev. Neurosci.* 14, 567-573.
- Santini, S., Boore, J. L. and Meyer, A. (2003). Evolutionary conservation of regulatory elements in vertebrate Hox gene clusters. *Genome Res.* 13, 1111-1122.

- Sharpe, J., Nonchev, S., Gould, A., Whiting, J. and Krumlauf, R. (1998). Selectivity, sharing and competitive interactions in the regulation of Hoxb genes. *EMBO J.* 17, 1788-1798.
- Taylor, J. S., Van de Peer, Y., Braasch, I. and Meyer, A. (2001). Comparative genomics provides evidence for an ancient genome duplication event in fish. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 356, 1661-1679.
- Thisse, C., Thisse, B., Schilling, T. F. and Postlethwait, J. H. (1993). Structure of the zebrafish snail1 gene and its expression in wild-type, spadetail and no tail mutant embryos. *Development* 119, 1203-1215.
- Tumpel, S., Cambroneiro, F., Wiedemann, L. M. and Krumlauf, R. (2006). Evolution of cis elements in the differential expression of two Hoxa2 coparalogous genes in pufferfish (Takifugu rubripes). *Proc. Natl. Acad. Sci. USA* 103, 5419-5424.
- Woolfe, A. and Elgar, G. (2007). Comparative genomics using Fugu reveals insights into regulatory subfunctionalization. *Genome Biol.* 8, R53.
- Yee, S. P. and Rigby, P. W. (1993). The regulation of myogenin gene expression during the embryonic development of the mouse. *Genes Dev.* 7, 1277-1289.
- Zerucha, T., Stuhmer, T., Hatch, G., Park, B. K., Long, Q., Yu, G., Gambarotta, A., Schultz, J. R., Rubenstein, J. L. and Ekker, M. (2000). A highly conserved enhancer in the Dlx5/Dlx6 intergenic region is the site of cross-regulatory interactions between Dlx genes in the embryonic forebrain. *J. Neurosci.* 20, 709-721.