



HAL
open science

Étude des mécanismes contrôlant l'expression des gènes HOX et implications pour la génération in vitro de tissus humains

Rémi Robert

► **To cite this version:**

Rémi Robert. Étude des mécanismes contrôlant l'expression des gènes HOX et implications pour la génération in vitro de tissus humains. Biologie cellulaire. Sorbonne Université, 2023. Français. NNT : 2023SORUS399 . tel-04705659

HAL Id: tel-04705659

<https://theses.hal.science/tel-04705659v1>

Submitted on 23 Sep 2024

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

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



Sorbonne Université

Ecole Doctorale Complexité du Vivant – ED515

Institut du Fer à Moulin, Inserm U1270

Equipe cellules souches et neurodéveloppement

Etude des mécanismes contrôlant la colinéarité temporelle des gènes *HOX* et application pour la génération de cellules humaines *in vitro*

Thèse de doctorat en Biologie

Présentée et soutenue publiquement le 22 septembre 2023 par

Rémi ROBERT

Composition du Jury :

Présidente - Mme Sylvie SCHNEIDER-MAUNOURY, directrice de recherche, Institut de Biologie Paris Seine

Rapporteur - M. Bertrand BENAZERAF, chargé de recherche, Centre de Biologie Intégrative Toulouse

Rapporteuse - Mme Marie KMITA, directrice de recherche, Institut de Recherches Cliniques de Montréal

Examinateur - M. Daan NOORDERMEER, chargé de recherche, Institut de Biologie Intégrative de la Cellule

Directeur de thèse - M. Stéphane NEDELEC, chargé de recherche, Sorbonne Université SIM (Sciences, Ingénierie, Médecine)

Co-directeur de thèse - M. Jacques VAN HELDEN, professeur, Aix Marseille Université

Remerciements:

Je remercie tout d'abord les membres de mon jury, Bertrand Bénazéraf, Marie Kmita, Daan Noordermeer et Sylvie Schneider-Maunoury pour avoir accepté de lire et d'évaluer mon travail.

Je souhaite également exprimer ma gratitude envers Jean-Antoine Girault et Fiona Francis, directeur et directrice de l'Institut du Fer à Moulin, pour m'avoir offert l'opportunité de réaliser ma thèse dans un environnement bien équipé et propice à l'entraide entre les équipes.

Evidemment, merci à Stéphane Nedelec, pour son implication continue tout au long de ma thèse. Je me souviendrais de notre première rencontre un jour de canicule où j'étais arrivé en courant (car en retard) et où je portais pour la première et dernière fois de ma vie une chemise blanche (qui devait être plutôt jaunâtre à ce moment-là). Malgré mes propos aléatoires à ce moment, tu as accepté que je rejoigne ton équipe et ces nombreuses années à tes côtés m'ont permis d'apprendre beaucoup. Je tiens également à remercier chaleureusement Jacques van Helden pour les discussions enrichissantes que nous avons eues. Bien que le projet n'ait pas été orienté autant que prévu vers un aspect bio-informatique, nos échanges sur ce sujet m'ont permis d'approfondir mes connaissances et d'élargir mon horizon scientifique.

Je remercie également les membres de mon comité de suivi de thèse, Fiona Francis, Claire Rougeulle et Morgane Thomas-Chollier, d'avoir accepté de suivre attentivement mes travaux au cours de ces quatre dernières années. Leurs contributions et nos discussions ont permis de faire mûrir ce projet et d'envisager de nouvelles expériences qui ont abouti à des résultats essentiels.

Je remercie tous les membres de mon équipe, passés et présents, qui ont contribué à créer un environnement chaleureux et bienveillant. Leur présence et leur soutien ont été essentiels tout au long de mon parcours. Merci à Hahn et Marion avec qui j'ai partagé mon premier bureau quand je suis arrivé ici avant qu'on se fasse virer par les patcheurs... Merci à Héléna, dont je suis convaincu qu'elle accomplira une thèse remarquable. Merci à Maëli pour la bonne humeur que tu apportes au quotidien malgré des playlists plus que douteuses. Merci à Vincent et Mathieu qui ont été des collègues mais surtout des partenaires de fêtes incroyables durant cette thèse. Entre les apéro-basket, les concerts et le Maroc (je crois je viens de faire un kamoulox sans faire exprès), je garde de très nombreux souvenirs que je n'oublierais jamais. Simona, j'ai adoré partager mes journées avec toi. Les choses n'avaient pourtant pas super bien commencé (le bain marie du L2 s'en rappelle encore) et puis on a appris à mieux se connaître jour après jour et je j'ai fini par t'adorer pour ton honnêteté en toutes situations et ta gentillesse.

Aussi, j'ai trouvé une alchimie de travail extraordinaire avec toi lorsque j'ai rejoint ton projet et j'aurais aimé avoir l'opportunité de travailler davantage à tes côtés durant ma thèse. Pour finir ces remerciements à mon équipe, je tiens tout particulièrement à exprimer ma profonde gratitude à Célia Vaslin qui m'a encadrée pendant tout mon stage de master 2. Tu es une personne extraordinaire, toujours prête à donner aux autres sans rien attendre en retour. Aujourd'hui, encore plus qu'avant, je réalise à quel point j'ai eu de la chance de débiter à travailler dans ce laboratoire à tes côtés.

Je remercie également l'équipe à l'Institut Jacques Monod à commencer par Vanessa Ribes toujours positive et débordante d'énergie. Merci à Pascale Gilardi sans qui mon projet de thèse ne serait pas ce qu'il est aujourd'hui. Merci à Robin pour m'avoir aidé à analyser les données de Cut&Tag ainsi qu'à Camil pour nos échanges trade deadline et draft. Je tiens également à remercier Benoit Sorre avec qui j'ai eu l'opportunité de collaborer et qui a apporté un regard différent sur mes résultats.

Merci également aux membres des autres équipes de l'Institut du Fer à Moulin qui ont également participé d'une façon ou d'une autre à la réussite de ce projet. En particulier merci à Catherine Béchade, Anne Roumier, Gress Kadaré, Tiago Mendes, Imane Moutkine, Marta Garcia et Richard Belvindrah. Merci aussi aux équipes administrative et technique qui ont été un support essentiel : Soumia, Laure, Sabrina, Baptiste, Alix, Patricia, Theano, Mythili.

Merci à tous les doctorants et post-doctorants avec qui j'ai passé de très bons moments ces dernières années : Billel, Anass, Amina, Eléonore, Facundo, Olivier, Benoit, Vincenzo, Donia, Taylor, Valeria, Florian, Zaha, Melody, Anne-Gaëlle, Emilie, Fanny, Giulia. Plus particulièrement, merci à Naya pour ton authenticité en toute circonstance et avec qui j'ai adoré partager chaque moment. Un immense merci à Erwan qui a apporté un vent nouveau à l'institut et a été d'une grande aide dans les moments les plus difficiles de ma thèse, que ce soit durant ces dernières semaines de rédaction ou bien au kilomètre 39 à la sortie du bois de Boulogne. Merci à Ruiyi, avec qui, malgré de nombreuses différences, je me suis sentie le plus proche pendant cette thèse. J'espère avoir été un ami et un soutien à la hauteur de ce que tu as été pour moi.

À Alexia, Célia et Claire : votre rencontre a été la meilleure chose qui me soit arrivé en venant à Paris et je suis heureux de pouvoir vous compter parmi mes amis.

Enfin à ma famille, merci pour votre présence et votre soutien sans faille depuis toujours. Merci en particulier à mes parents qui ont toujours été là pour moi et m'ont permis de réaliser tous mes objectifs.

Sommaire

Liste des figures.....	7
Liste des abréviations	8
Résumé:	9
Avant-propos:	11
Introduction	12
I. Morphogénèse de l'axe antéro-postérieur chez les Vertébrés.....	13
1. Premiers stades du développement embryonnaire	13
2. Formation de l'axe antéro-postérieur	14
a. L'axe antéro-postérieur émerge à partir des progéniteurs axiaux	14
b. Mécanismes et signaux contrôlant la genèse et le maintien des progéniteurs axiaux	17
i. Les voies Wnt, Fgf et Notch instruisent la prolifération des progéniteurs axiaux	18
ii. Autres facteurs participant à la spécification et au maintien des progéniteurs axiaux	20
c. Mécanismes et signaux contrôlant la différenciation des progéniteurs axiaux	22
i. Les facteurs Sox2 et Brachyury sont associés aux destins neuronaux et mésodermiques	22
ii. Les voies Wnt et Fgf instruisent la différenciation mésodermique.....	23
iii. La voie RA favorise la différenciation neuronale	24
d. Mécanismes et signaux contrôlant la terminaison de l'élongation axiale	26
II. Les gènes <i>Hox</i> contrôlent la spécification de l'axe antéro-postérieur.....	29
1. Découverte des gènes <i>Hox</i>	30
2. Organisation des domaines d'expression des <i>Hox</i> le long de l'axe antéro-postérieur	30
3. Rôle des facteurs <i>Hox</i> dans la spécification de l'axe antéro-postérieur.....	33
a. Le plan d'organisation des vertèbres dépend des facteurs de transcription <i>Hox</i>	33
b. La spécification des colonnes de motoneurones dépend des facteurs de transcription <i>Hox</i>	34
4. Induction temporelle des gènes <i>Hox</i> lors de l'élongation	37
5. Mécanismes participant à la régulation transcriptionnelle des gènes <i>Hox</i>	37

a.	Rôle des complexes polycomb et trithorax dans les complexes <i>Hox</i>	38
b.	Les facteurs de transcription et les enhancers au sein des complexes <i>Hox</i>	40
c.	Les enhancers en dehors des complexes <i>Hox</i>	42
d.	Topologie et dynamique génomique (TAD)	44
6.	Régulation des gènes <i>Hox</i> par les facteurs extrinsèques	47
a.	La voie Wnt régule l'induction des gènes <i>Hox</i> situés dans la partie 3' des complexes	47
b.	L'acide rétinoïque module l'expression des gènes <i>Hox1-5</i>	48
c.	Fgf coopèrent avec les facteurs Cdx pour activer l'expression des gènes <i>Hox</i> se trouvant dans la partie centrale des complexes.....	49
d.	La voie Gdf participe à l'induction des gènes <i>Hox</i> situés dans la partie 5' des complexes	50
7.	Contribution des signaux intrinsèques et extrinsèques dans la régulation des gènes <i>Hox</i>	52
III.	Cellules souches pluripotentes	55
1.	Généralités.....	55
2.	Modéliser le développement de l'axe antéro-postérieur <i>in vitro</i>	56
a.	Progéniteurs axiaux <i>in vitro</i>	56
b.	Spécification des sous-types cellulaires de l'axe antéro-postérieur <i>in vitro</i>	57
c.	Modèle <i>in vitro</i> de la morphogenèse antéro-postérieure basé sur l'auto-organisation des cellules souches pluripotentes.....	58
	Résultats	61
	Partie 1 - Rôle des facteurs extrinsèques dans le contrôle de l'induction des gènes <i>HOX</i>	62
	La régulation par les facteurs extrinsèques de l'horloge <i>HOX</i> dans les progéniteurs axiaux humains contrôle la diversification des motoneurones.....	63
a.	Contexte	63
b.	Résultats	64
c.	Contribution	65
	La dynamique temporelle des voies de signalisation rythme l'activation séquentielle des gènes <i>HOX</i>	88
a.	Contexte	88
b.	Résultats	89
	Partie 2 - Développement d'un modèle organoïde humain reproduisant la spécification de l'axe antéro-postérieur.....	116

Un modèle d'organoïde <i>in vitro</i> du tronc humain récapitule la co-morphogenèse de la moelle épinière et de la colonne vertébrale	117
a. Contexte	117
b. Résultats	118
c. Contributions	119
Discussion	147
Bibliographie	160

Liste des figures

Figure 1. Formation de l'axe antéro-postérieur chez les Vertébrés au cours de l'embryogénèse	14
Figure 2. La formation de l'axe antéro-postérieur, les progéniteurs axiaux et leur descendance	17
Figure 3. Signalisation de la voie Wnt et Fgf durant l'élongation axiale.....	20
Figure 4. Les voies Wnt/Fgf et les facteurs Cdx/Brachyury coopèrent pour instruire la morphogénèse axiale	22
Figure 5. Destin cellulaire des progéniteurs axiaux et domaines d'expression des marqueurs Sox2 et Brachyury.....	23
Figure 6. Contribution du RA durant l'élongation axiale	26
Figure 7. Signalisation par la voie Gdf durant l'élongation axiale	28
Figure 8. Organisation des gènes <i>Hox</i> et leurs domaines d'expression chez la drosophile et les Vertébrés	31
Figure 9. Organisation de l'expression des protéines Hox dans la moelle épinière.....	32
Figure 10. Les gènes <i>Hox</i> contrôlent l'organisation du plan vertébral	34
Figure 11. Effet des pertes de fonctions des gènes <i>Hox</i> sur la formation des populations de motoneurones	36
Figure 12. Rôle des facteurs de transcription et enhancers dans la régulation des gènes <i>Hox</i>	42
Figure 13. Rôle des enhancers distaux dans la régulation des gènes <i>Hox</i>	44
Figure 14. Organisation 3D des complexes <i>Hox</i> au cours de l'élongation axiale.....	47
Figure 15. L'activation transcriptionnelle des complexes <i>Hox</i> est initiée par la voie Wnt.....	48
Figure 16. L'initiation des gènes <i>Hox</i> centraux dépend de la voie Fgf-Cdx.....	50
Figure 17. L'initiation des gènes <i>Hox</i> les plus caudaux dépend de la voie Gdf.....	52
Figure 18. Les facteurs environnementaux régulent l'expression des gènes <i>Hox</i>	54
Figure 19. Génération de cellules souches pluripotentes induites.....	56
Figure 20. Résumé de la différenciation ciblée <i>in vitro</i> en sous-types de neurones spinaux ..	58
Figure 21. Organisation cellulaire le long de l'axe antéro-postérieur des axioloïdes	60
Figure 22. Illustration d'un modèle théorique mettant en évidence l'importance de la dynamique temporelle des voies de signalisation dans le contrôle du rythme d'activation des gènes <i>HOX</i>	153

Liste des abréviations

3D tri-dimension

AKT protéine kinase B

BMP protéine osseuse morphogénétique

C-DOM domaine centromérique

CLE épiblaste latéral caudal antérieur

CNH limite du tube neural et de la notochorde néoformée

CSEm cellules souches embryonnaires murines

CSP cellule souche pluripotente

CSPi cellule souche pluripotente induite

CTCF facteur CCCTC-binding

EGF facteur de croissance épidermique

ERK extracellular signal-regulated kinases

FB cerveau antérieur

FGF fibroblast growth factors

GDF growth differentiation factor

HB cerveau postérieur

HH Hamburger–Hamilton

HMC colonne motrice hypaxiale

HOX boîte à homéodomaine

IM mésoderme intermédiaire

JNK jun N-terminal kinases

LMC colonne motrice latérale

LPMP progéniteur du mésoderme latéral et paraxial

LVM mésoderme ventral/latéral

MAPK mitogen activated protein kinase

MB mésencéphale

MMC colonne motrice médiane

MSP progéniteur du somite médian

NGF facteur de croissance nerveuse

NMP progéniteur neuromésodermal

NotoP progéniteur de la notochorde

NSB frontière nœud/ligne primitive

PGC colonne pré-ganglionnaire

Pi3K phosphoinositide 3-kinase

PMC colonne motrice phrénique

PNT tube pré-neural

PRC polycomb repressive complex

PRE polycomb response element

PS ligne primitive

PSM mésoderme pré-somitique

RA acide rétinoïque

RAR récepteur nucléaire à l'acide rétinoïque

RARE élément de réponse de l'acide rétinoïque

SBE smad binding element

SC moelle épinière

TAD domaine d'association topologique

TALE extension de boucle à trois acides aminés

T-DOM domaine télomérique

TGF- β transforming growth factor β

TrxG complexe trithorax

Wnt wingless type

Résumé:

La formation du schéma corporel des Vertébrés dépend, lors de l'embryogenèse, de l'organisation spatiale des différents types cellulaires le long de l'axe antéro-postérieur. Ce processus est orchestré par les facteurs de transcription HOX qui sont différentiellement exprimés le long de cet axe et confèrent leur identité positionnelle aux tissus en développement. Ces patrons d'expression sont initiés par l'activation séquentielle des gènes *HOX* dans les progéniteurs axiaux, une population de cellules souches nourrissant l'allongement progressif du corps des embryons en développement, notamment les somites et la moelle épinière. Parallèlement à l'induction progressive de cette famille de gènes, ces progéniteurs génèrent donc des structures de plus en plus caudales transformant la séquence temporelle d'activation en domaines d'expression spatiaux le long de l'axe antéro-postérieur. Néanmoins, les mécanismes qui rythment la temporalité de cette induction et sa transformation en domaines spatiaux restent mal connus, en particulier chez l'humain.

Pour aborder ces questions, au cours de ma thèse, j'ai produit à partir de cellules souches pluripotentes humaines des progéniteurs arborant les caractéristiques moléculaires et fonctionnelles des progéniteurs axiaux. En effet, nous avons montré que ces progéniteurs activent séquentiellement les gènes *HOX* et peuvent donner naissance à des organoïdes récapitulant de multiples aspects de la génération et de l'organisation de l'axe antéro-postérieur comme la formation de somites entourant un tube neural le long duquel les patrons d'expression spatiaux de plusieurs HOX sont récapitulés.

En utilisant ces progéniteurs axiaux produits *in vitro*, nous avons tout d'abord démontré que la vitesse d'induction des gènes *HOX* est dynamiquement modulée par l'activité graduelle de facteurs extrinsèques, FGFs et GDF11 qui sont séquentiellement exprimés dans la région caudale des embryons Vertébrés au cours du développement. Puis, nous avons montré 1) que les gènes *HOX* activés sont des cibles directes des voies de signalisation en aval de ces facteurs et 2) que la vitesse d'activation des gènes exprimés de plus en plus tardivement est déterminée par la durée d'activation des voies de signalisation dans les progéniteurs axiaux, une propriété des voies régulée par des mécanismes intrinsèques de rétrocontrôles négatifs.

Dans l'ensemble, mes résultats permettent de proposer un nouveau modèle dans lequel le rythme d'activation des gènes *HOX* est une propriété émergente des dynamiques de voies de signalisation en aval de facteurs extrinsèques. Parallèlement, mes études ont permis d'améliorer

l'ingénierie cellulaire et tissulaire des cellules du tronc à partir de cellules souches pluripotentes humains culminant en des protocoles de génération des différents sous-types de motoneurones présent le long de l'axe du corps et en un nouveau modèle d'organoïde mimant la morphogenèse et la formation de la diversité cellulaire le long de l'axe du corps humain.

Avant-propos:

Le temps est un élément central de l'embryogenèse. Les cellules doivent naître dans le bon ordre et au bon moment pour assurer l'organisation d'un organisme. Pourtant, les mécanismes moléculaires qui contrôlent la chronologie de la diversification cellulaire sont encore mal compris chez les animaux et encore moins chez l'homme, ce qui limite nos capacités d'ingénierie des cellules et des organoïdes. Ce problème est illustré par la régulation des gènes *HOX*, une famille de gènes homéotiques assurant, au cours de l'embryogenèse, l'organisation de la diversité cellulaire le long de l'axe antéro-postérieur du corps. Les gènes *HOX* sont regroupés dans 4 complexes génomiques (*HOXA/B/C/D*) et codent 39 facteurs de transcription différemment exprimés le long de l'axe principal du corps. Les patrons d'expression de ces facteurs de transcription sont initiés par l'activation séquentielle des gènes *HOX* le long des complexes, des gènes en 3', exprimés précocement, aux gènes en 5' exprimés de plus en plus tardivement. Cette activation progressive des gènes *HOX* a lieu dans les progéniteurs axiaux, une population de cellules situées dans la partie caudale de l'embryon et qui alimentent l'extension antéro-postérieure de celui-ci. Bien que le rythme de l'activation des *HOX* et la transformation de cette dynamique temporelle en patrons d'expression stéréotypés soit essentiel au bon développement, les mécanismes qui contrôlent ces processus restent encore mal compris, en particulier chez l'homme.

Lors de ma thèse, dans un but fondamental et d'ingénierie cellulaire, je me suis intéressé aux mécanismes contrôlant le rythme d'activation des gènes *HOX* chez l'humain et à la mise en place de modèles permettant d'étudier la formation de leurs patrons d'expression le long de l'axe antéro-postérieur. Pour aborder ces questions, j'ai utilisé la différenciation *in vitro* de cellules souches pluripotentes humaines en progéniteurs axiaux.

Grâce à cette approche, j'ai étudié le rôle de FGF2 et GDF11, deux facteurs extrinsèques qui contrôlent l'élongation axiale dans la régulation de la temporalité d'induction des gènes *HOX* exprimés différemment le long de l'axe du tronc. Parallèlement, j'ai contribué au développement d'un nouveau modèle d'organoïde capable de reproduire de nombreux aspects de la morphogenèse et de la formation de la diversité cellulaire du tronc humain

Introduction

I. Morphogénèse de l'axe antéro-postérieur chez les Vertébrés

La génération des éléments qui constituent l'axe antéro-postérieur du corps des embryons Vertébrés en développement, incluant l'embryon humain, s'effectue progressivement des régions antérieures aux régions postérieures lors de l'élongation axiale post-gastrulation. L'allongement axial est nourri par des progéniteurs axiaux, situés dans la région caudale de l'embryon, qui prolifèrent et se différencient pour former les tissus du tronc et de la queue (Steventon and Martinez Arias, 2017; Wilson et al., 2009). Dans ce chapitre, nous allons définir les mécanismes qui sous-tendent ces processus.

1. Premiers stades du développement embryonnaire

Pendant les deux premières semaines suivant la fécondation de l'ovule humain, l'œuf se segmente par une succession de divisions pour former un disque embryonnaire appelé blastocyste caractérisé par une couche externe correspondant au trophoctoderme et une masse cellulaire interne. Après l'implantation du blastocyste à la surface de cellules endométriales, la masse cellulaire interne se différencie en endoderme primitif et en épiblaste (**Fig. 1A**). Au niveau du disque embryonnaire, une réorganisation cellulaire se met ensuite en place caractérisée par l'épaississement de l'épiblaste qui s'invagine pour former la ligne primitive antérieure. Les cellules migrent le long de cette ligne pour donner naissance à l'endoderme (couche la plus profonde) et le mésoderme (couche médiane). La couche où se trouve originellement l'épiblaste (couche externe) forme l'ectoderme. Cette étape nommée gastrulation permet la formation des trois feuillets embryonnaires et des axes principaux de l'embryon.

A la troisième semaine du développement humain, le mésoderme commence à se réorganiser et se subdivise en mésoderme axial (précurseur de la notochorde), paraxial (précurseur des somites qui se différencient en muscles, cartilages, os et dermes), intermédiaire (précurseur des reins et des gonades) et latéral (précurseur du système cardiovasculaire). Lorsque la notochorde est formée à partir du mésoderme axial, cette nouvelle structure embryologique sécrète des facteurs qui induisent dans la partie dorsale la différenciation de l'ectoderme en neuroectoderme, également connue sous le nom de plaque neurale. Les bords latéraux de la plaque neurale sont alors exposés à des forces de contraction convergentes qui se rejoignent à la ligne médiane pour former le tube neural (**Fig. 1B**). Cette étape correspond à la neurulation et se termine à la 5ème semaine post-fertilisation.

La neurogenèse se poursuit avec l'apparition de trois vésicules primitives au niveau de l'extrémité crânienne du tube neural: le prosencéphale (cavité antérieure), le mésencéphale (cavité intermédiaire) et le rhombencéphale (cavité postérieure). L'ensemble de ces domaines donne naissance aux structures qui constituent le cerveau et le cervelet tandis que la partie caudale du tube neural va former l'axe du corps embryonnaire à partir du cerveau postérieur jusqu'à la queue.

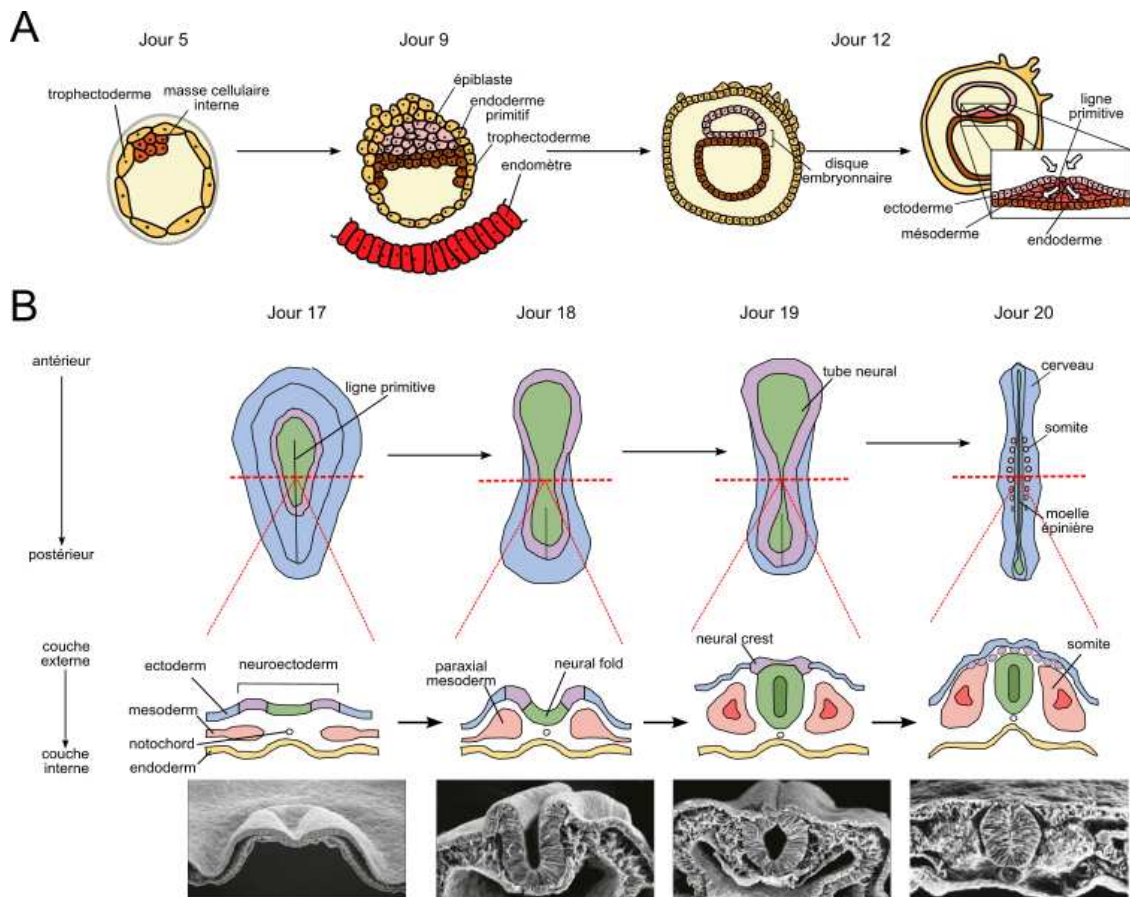


Figure 1. Formation de l'axe antéro-postérieur chez les Vertébrés au cours de l'embryogénèse. (a) Suite à la fécondation et la segmentation, il y a une réorganisation des cellules de la morula qui vont se différencier en trophoctoderme et masse cellulaire interne pour former le blastocyste (jour 4-5 post-fécondation humain). L'œuf va alors s'implanter dans l'endomètre et la gastrulation va être initiée par un épaissement au niveau de l'épiblaste pour former la ligne primitive. **(b)** Au cours de la neurulation, il y a une convergence des bords de la crête neurale pour former le tube neural dont la partie antérieure va former l'encéphale et la partie postérieure la moelle épinière.

2. Formation de l'axe antéro-postérieur

a. L'axe antéro-postérieur émerge à partir des progéniteurs axiaux

Chez les Vertébrés, l'élongation de l'axe antéro-postérieur dépend de progéniteurs se trouvant dans la partie caudale de l'embryon. Ces cellules génèrent séquentiellement des tissus neuronaux et mésodermiques, en commençant autour de la base postérieure du futur cerveau et en terminant à l'extrémité de la queue chez la souris ou du tronc chez l'homme. Connus sous le

nom de progéniteurs axiaux, ils sont les précurseurs donnant naissance à la moelle épinière, au système musculo-squelettique, à la paroi corporelle et aux organes mésodermiques (i.e. reins, gonades). Les progéniteurs axiaux et leurs contributions ont été identifiés par des études anatomiques, d'expériences de greffe et de lignages décrites ci-après.

Des expériences de lignage dans des modèles aviaires ont montré que les cellules dans la région de la ligne primitive de l'embryon donnent naissance à la fois au tube neural et à la notochorde (Selleck and Stern, 1991). Puis, des expériences de marquage sur des embryons de souris ont révélé que des cellules individuelles localisées dans le nœud primitif à différents stades de la gastrulation produisent des descendants localisés dans la notochorde et dans le nœud lui-même. (Forlani et al., 2003; Lawson et al., 1991). Ces résultats indiquent que certains des descendants des cellules marquées sont conservés dans la région des progéniteurs, ce qui suggère que les cellules d'origine ont des propriétés de cellules souches. Ces cellules participant à la formation de l'axe ont collectivement été nommées « progéniteurs axiaux ».

Des analyses rétrospectives du lignage de cellules uniques (par la réversion spontanée du gène *laacZ* en gène fonctionnel *lacZ*) ont été réalisées dans la région caudale proliférative de l'embryon de souris afin de déterminer précisément le destin de ces cellules progénitrices (Mathis and François Nicolas, 2000; Tzouanacou et al., 2009). Ces expériences ont mis en évidence la présence de clones donnant naissance à des cellules filles dans le mésoderme paraxial et dans la moelle épinière suggérant que certains des progéniteurs localisés dans cette région sont bipotents. Ces cellules ont été décrites comme des progéniteurs neuro-mésodermaux (NMPs). Cependant, ces analyses rétrospectives du lignage déduisent les propriétés d'un progéniteur à partir de ses descendants mais ne permettent pas d'identifier la localisation initiale de ces progéniteurs.

Des expériences de cartographie prospective ont été faites par injection de colorant ou greffe de tissus dans la ligne primitive d'embryons de souris à E8.5-E9.5 et ont permis d'identifier deux régions donnant naissance à des destins neuro-mésodermiques : la frontière nœud/ligne primitive (NSB) et l'épiblaste latéral caudal antérieur (CLE) (Cambray and Wilson, 2007, 2002; Kondoh and Takemoto, 2012; McGrew et al., 2008; Oginuma et al., 2017). Après E9.5, les progéniteurs se trouvant dans la NSB se regroupent à la limite du tube neural et de la notochorde néoformée (CNH) (Cambray and Wilson, 2002; Delfino-Machín et al., 2005; Kondoh and Takemoto, 2012; Wilson et al., 2009). Des études ont montré que les cellules du NSB contribuent de manière relativement égale aux dérivés neuraux et mésodermiques alors que le destin des cellules dans le CLE dépend de leur emplacement dans cette région

(Wymeersch et al., 2016). Ainsi, les cellules antérieures et latérales du CLE contribuent plus largement au tube neural, tandis que les cellules postérieures et médianes contribuent davantage aux tissus mésodermiques.

Néanmoins dans ces travaux, plusieurs cellules sont greffées ce qui ne permet pas de déterminer si des cellules uniques sont véritablement bipotentes ou bien s'il s'agit d'un ensemble de plusieurs types de progéniteurs ayant des identités légèrement différentes avec une propension plus ou moins importante à donner naissance à des cellules neurales ou du mésoderme. Des expériences d'imagerie live de traceurs de lignage (brainbow) chez le poulet ont démontré l'existence de progéniteurs uniques capables de donner naissance à des cellules dans les deux lignages à double destin (Guillot et al., 2021). Ces travaux ont également révélé que les cellules qui produisent des descendants neuraux et mésodermiques se trouvent à l'interface entre les cellules à destin purement neuronal ou mésodermique.

Malgré de nombreuses analyses transcriptomiques (Gouti et al., 2017; Guillot et al., 2021; Koch et al., 2017; Olivera-Martinez et al., 2014; Wymeersch et al., 2018), aucun marqueur exclusif des NMPs n'a été identifié à ce jour. Cependant, elles sont caractérisées par la co-expression de Sox2 qui est exprimé par les cellules pluripotentes ainsi que les progéniteurs neuraux, et Brachyury qui est exprimé par les progéniteurs mésodermiques (Henrique et al., 2015; Rivera-Pérez and Magnuson, 2005; Takemoto et al., 2006; Tsakiridis et al., 2014; Tzouanacou et al., 2009). Des cellules co-exprimant Sox2 et Brachyury ont été identifiées dans des régions analogues du bourgeon caudal des embryons de poulet, zebrafish et humains (Guillot et al., 2021; Martin and Kimelman, 2012; Olivera-Martinez et al., 2012).

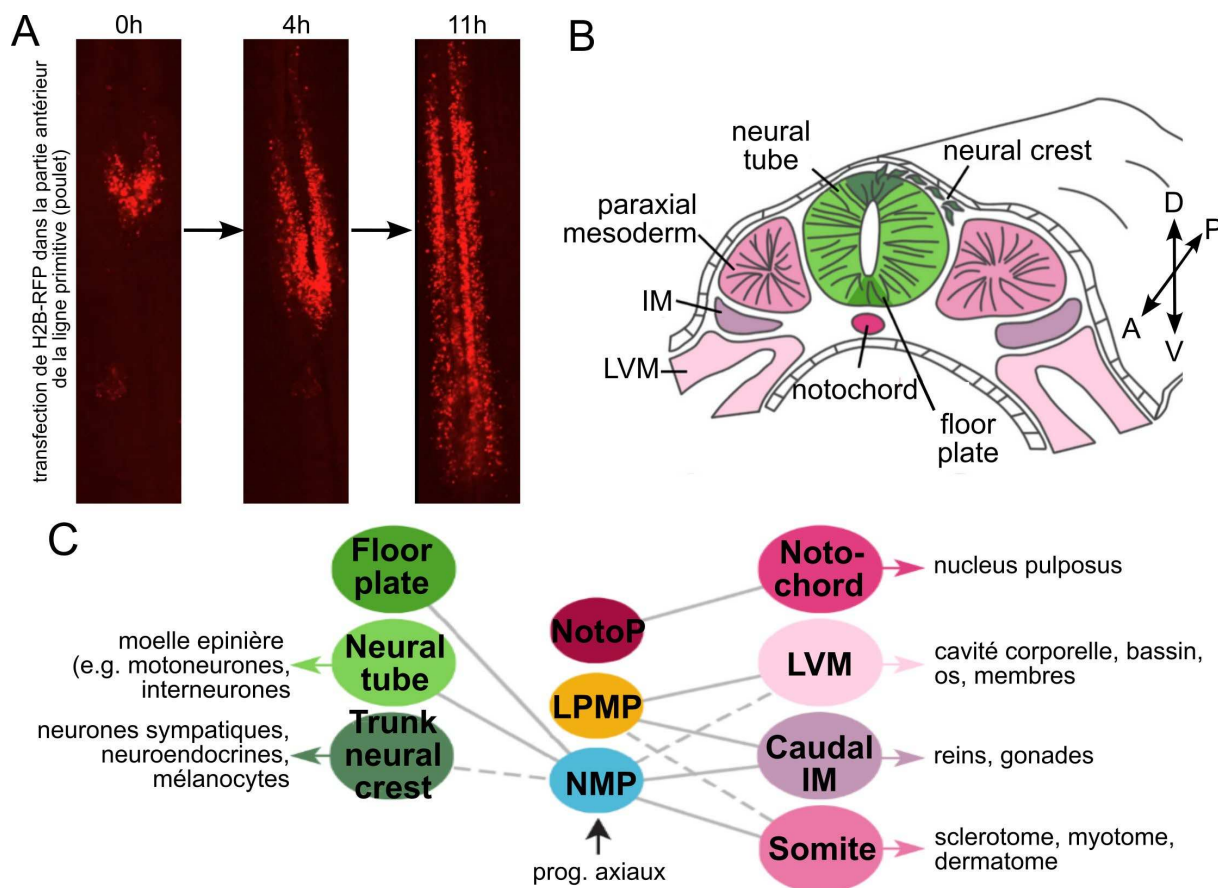


Figure 2. La formation de l'axe antéro-postérieur, les progéniteurs axiaux et leur descendance. (a) Imagerie live d'un embryon de poulet après électroporation de plasmides pCAGG-H2B-RFP dans la partie antérieure de la ligne primitive. (b) Schéma représentant l'organisation morphologique du tube neural avec la formation de la crête neurale (dorsal) et de la plaque du plancher (ventral). Les sous-types de mésoderme comprennent les somites, le mésoderme intermédiaire et le mésoderme latéral/ventral (c) Réseau représentant les NMPs/ LPMP/NotoP et les tissus axiaux qu'ils produisent (lignes pleines) ou qu'ils ont la capacité de former (lignes pointillées). LPMP = progéniteurs du mésoderme latéral et paraxial, NotoP = progéniteurs de la notochorde, LVM = mésoderme ventral/latéral, IM = mésoderme intermédiaire (Adapté d'après Oginuma et al., 2017; Wymeersch et al., 2021)

En résumé, ces résultats ont montré que des cellules progénitrices ayant des propriétés de cellules souches sont présentes dans la partie caudale de l'embryon et nourrissent progressivement l'élongation axiale en donnant naissance à des descendants peuplant l'axe antéro-postérieur (**Fig. 2A**). Parmi ces progéniteurs axiaux, certaines régions contiennent des NMPs qui donnent naissance aux cellules du mésoderme et du tube neural (**Fig. 2B/C**). La spécification de ces cellules et leur maintien est donc essentielle pour la formation de l'axe antéro-postérieur.

b. Mécanismes et signaux contrôlant la genèse et le maintien des progéniteurs axiaux

L'élongation de l'axe antéro-postérieur du corps est contrôlée par l'activité synergique de plusieurs voies de signalisation et de facteurs de transcription qui contrôlent la spécification,

le maintien et la différenciation des progéniteurs axiaux. Ces signaux agissent à des moments précis au cours du développement et nous allons les décrire dans cette partie.

i. Les voies Wnt, Fgf et Notch instruisent la prolifération des progéniteurs axiaux

Les ligands et récepteurs des voies de signalisation Wnt (e.g. *Wnt3a/5a/5b*, *Frizzled*) et Fgf (e.g. *Fgf8*, *Fgfr1*) sont exprimés dans la partie caudale de l'embryon et jouent un rôle crucial dans la spécification et le maintien des progéniteurs (Diez Del Corral and Morales, 2017; Echevarria et al., 2005; Takada et al., 1994; Wahl et al., 2007).

La voie de signalisation canonique Wnt passe par l'activation du récepteur Frizzled puis la translocation nucléaire de la β -caténine. Lorsque la β -caténine se trouve dans le noyau, les facteurs Tcf/Lef sont co-activés et permettent l'expression de gènes cibles de la voie Wnt (**Fig. 3A**) (Brafman and Willert, 2017; Cadigan and Waterman, 2012). Chez la souris, l'expression de *Wnt3a* commence dans la ligne primitive à E7.5 et est ensuite maintenue dans les progéniteurs axiaux pendant toute l'élongation (**Fig. 3B**) (Takada et al., 1994).

Des expériences de perte de fonction ont montré que *Wnt3a* est nécessaire au maintien des NMPs dans un état indifférencié et à la survie de la population de progéniteurs (**Fig. 3C**) (Garriock et al., 2015; Takada et al., 1994; Wilson et al., 2009). Ainsi, une invalidation de *Wnt3a* chez la souris conduit à la perte des structures postérieures du corps (Takada et al., 1994). Un deuxième ligand de la voie Wnt, *Wnt8a*, est également exprimé dans le CLE jusqu'au stade E9.5. Les doubles mutants *Wnt8a/Wnt3a* présentent un phénotype de troncature plus sévère que les mutants *Wnt3a* seuls (Cunningham et al., 2015). Cela indique que *Wnt8a* et *Wnt3a* agissent de manière redondante pour maintenir les progéniteurs axiaux.

La voie Fgf est également impliquée dans la spécification et le maintien des progéniteurs axiaux. Ce signal est caractérisé par 23 ligands sécrétés chez les Vertébrés qui se lient à des récepteurs à tyrosine kinase afin d'induire la voie des Mitogen Activated Protein Kinases (MAPKs), mais aussi d'autres voies comme Pi3K, AKT ou JNK (**Fig. 3D**) (Diez Del Corral and Morales, 2017). Lorsque l'élongation axiale est initiée, les domaines d'expression pour *Fgf3/4/13/18* se limitent à la ligne primitive, à l'exception de *Fgf8* qui est exprimé plus largement incluant la région des progéniteurs axiaux et le mésoderme présomitique (**Fig. 3E**) (Delfino-Machín et al., 2005; Diez Del Corral and Morales, 2017; Karabagli et al., 2002).

L'expression de *Fgf8* est très dynamique dans cette région puisqu'il est d'abord fortement exprimé par les progéniteurs puis progressivement réprimé à mesure que ceux-ci se différencient. Les récepteurs *Fgfr1/2/3* sont originellement exprimés dans les progéniteurs axiaux. Progressivement, l'expression de *Fgfr1* est étendue à l'ensemble du neuroectoderme, alors que *Fgfr3* est localisé dans les cellules basales du tube neural et *Fgfr2* est restreint à la partie antérieure (Nishita et al., 2011).

La voie Fgf contrôle la spécification des progéniteurs axiaux et l'extension caudale. Chez la souris, les mutants conditionnels *Fgfr1* sont caractérisés par un axe tronqué dans les régions sacrées et une diminution de l'expression de gènes associés aux NMPs (**Fig. 3F**) (Wahl et al., 2007). L'axe des mutants *Fgf4* et *Fgf8* est également raccourci puisque l'élongation s'arrête après la génération de 15-20 somites au lieu de 55 (**Fig. 3G**) (Boulet and Capecchi, 2012; Naiche et al., 2011). En revanche, l'exposition continue des progéniteurs axiaux au Fgf les maintiens dans un état d'autorenouvellement au détriment de la différenciation neurale ou mésodermique (Wilcock et al., 2007).

La voie Fgf interagit avec la voie Notch pour assurer le maintien de cette population de cellules souches et contrôler leur prolifération (Akai et al., 2005). Les souris mutantes pour divers acteurs de la voie Notch tels que *Dll3* ou *Notch1* présentent une troncation postérieure du corps (Dunwoodie et al., 2002; Luis De La Pompa et al., 1997). De même, des pertes de fonction de Notch chez le poulet sont caractérisées par la diminution de cellules se trouvant dans la zone proliférative caudale de l'embryon (Akai et al., 2005). La voie Notch joue donc un rôle dans le maintien de la zone proliférative en bloquant la différenciation des cellules progénitrices axiales (Del Álamo et al., 2011). La voie Fgf interagit également avec la voie Wnt puisque chez les doubles mutants *Fgf4/Fgf8*, il y a une diminution de l'expression de *Wnt3a/5a* (Boulet and Capecchi, 2012; Naiche et al., 2011).

En conclusion, les voies Fgf, Wnt et Notch fonctionnent en synergie et participent à la formation de l'axe antéro-postérieur en assurant le renouvellement des progéniteurs axiaux.

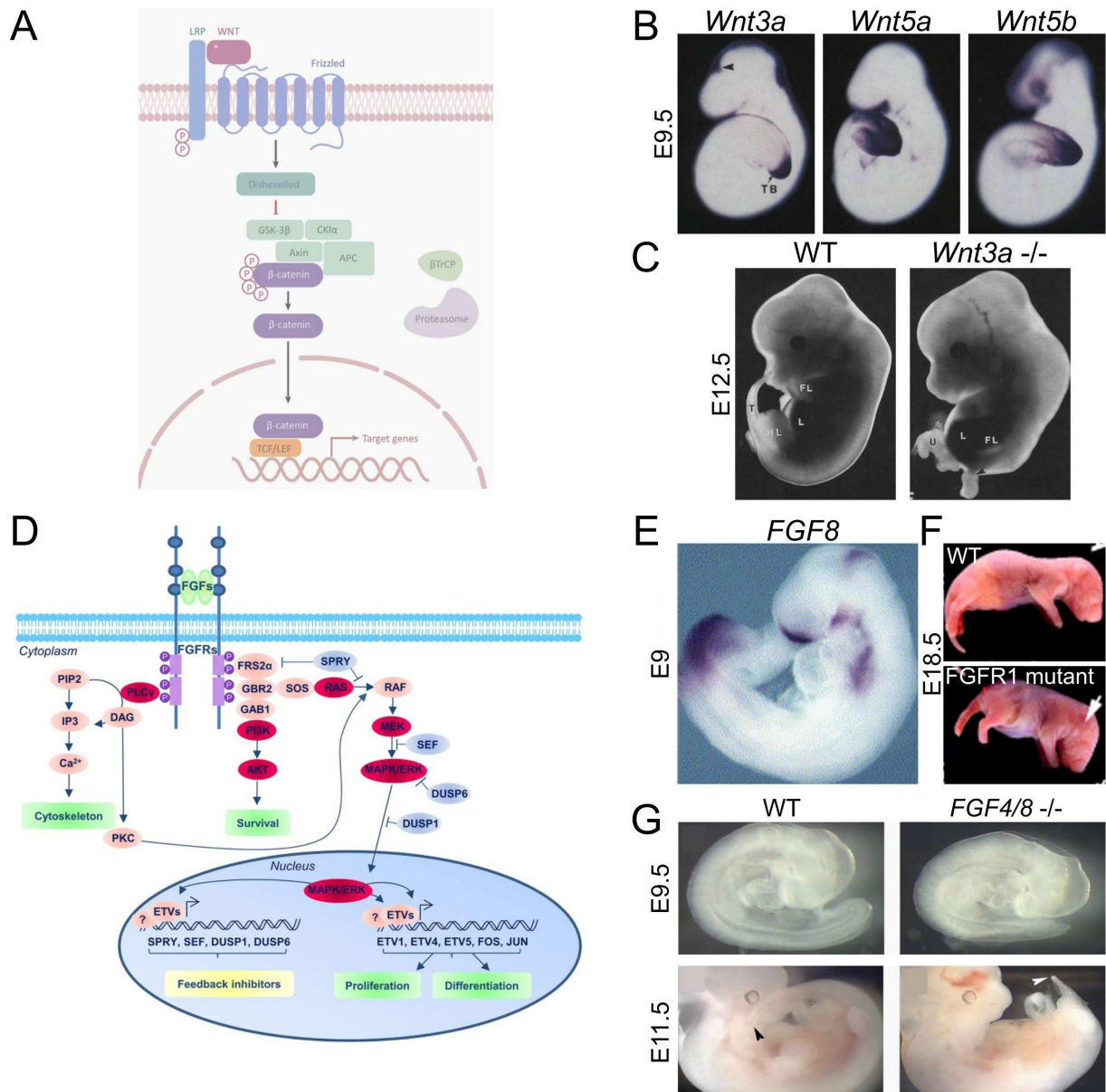


Figure 3. Signalisation de la voie Wnt et Fgf durant l'élongation axiale. (a) Schéma de la voie de signalisation Wnt canonique. La fixation de Wnt sur le récepteur Frizzled entraîne l'inhibition de la dégradation de la β -caténine par le complexe GSK3 et permet alors sa translocation nucléaire et l'activation des gènes Wnt-dépendants. (b) Hybridation *in situ* de *Wnt3a/5a/5b* dans embryon de souris. L'expression de *Wnt3a* est plus restreinte par rapport à *Wnt5a/5b*. (c) Morphologie d'embryons de type sauvage et mutant *Wnt3a*^{-/-} montrant la perte des structures postérieures chez les mutants. (d) Schéma d'un modèle de signalisation par le Fgf montrant l'activation de voies appartenant à la famille des MAPKs. (e) Hybridation *in situ* de *Fgf8* dans des embryons de souris. Le bourgeon caudal des embryons exprime *Fgf8*. (f) Morphologie de nouveaux nés sauvage et mutant *Fgfr1*^{-/-}. Il y a une troncation postérieure chez les mutants. (g) Embryons de souris de type sauvage et double mutant *Fgf4*^{-/-} et *Fgf3*^{-/-}. Les flèches pointent vers l'extrémité de la queue. Les doubles mutants présentent des défauts morphologiques avec un raccourcissement de la queue. (Adapté d'après Boulet and Capecchi, 2012; Diez Del Corral and Morales, 2017; Echevarria et al., 2005; Partanen et al., 1998; Takada et al., 1994)

ii. Autres facteurs participant à la spécification et au maintien des progéniteurs axiaux

1) Cdx : les protéines à homéo-boîte Cdx sont des facteurs de transcription exprimés tôt dans la ligne primitive puis plus tard dans le bourgeon de la queue, où ils sont nécessaires à la croissance des tissus embryonnaires postérieurs (Fig. 4A) (Ikeya and Takada, 2001; Metzis et

al., 2018; Pilon et al., 2006; Van Rooijen et al., 2012). Le rôle des gènes *Cdx* dans l'élongation axiale a été observé pour la première fois dans un modèle murin présentant une altération de l'élongation axiale suite à l'inactivation de *Cdx2* (**Fig. 4B**) (Chawengsaksophak et al., 1997). Les facteurs *Cdx1* et *Cdx4* coopèrent avec *Cdx2* car leur mutation individuelle n'est pas suffisante pour induire des défauts d'élongation. En revanche, le phénotype de troncation des mutants *Cdx2* est plus sévère en combinaison avec l'inactivation des deux autres gènes *Cdx* (Savory et al., 2009; Van Den Akker et al., 2002; Van Nes et al., 2006).

La perte de *Cdx2* chez la souris est caractérisée par des phénotypes similaires à ceux observés par la perte de *Wnt3a* (Ikeya and Takada, 2001; Shimizu et al., 2005; Takada et al., 1994) ou du récepteur *Fgfr1* (Partanen et al., 1998). En effet, les voies Wnt et Fgf induisent la transcription des gènes *Cdx*, ce qui indique qu'au moins une partie de l'action de ces voies sur la formation de l'axe antéropostérieur est médiée par ces facteurs de transcription (Bel-Vialar et al., 2002; Pilon et al., 2007, 2006; Shimizu et al., 2005). Puis lorsque les facteurs *Cdx* sont traduits, ils vont à leur tour réguler positivement l'activité des voies Wnt et Fgf (Amin et al., 2016; Neijts et al., 2017; Wymeersch et al., 2021). Il y a donc une boucle de rétroaction positive entre les facteurs *Cdx* et les voies Wnt/Fgf dans la partie caudale de l'embryon qui permet de maintenir les progéniteurs axiaux (**Fig. 4C**) (Savory et al., 2009; Van Den Akker et al., 2002; Van Rooijen et al., 2012; Young et al., 2009).

2) *Brachyury* : comme les facteurs *Cdx*, *Brachyury* est également nécessaire à l'élongation axiale. La mutation de *Brachyury* entraîne des défauts d'élongation axiale, et ce phénotype est encore plus sévère dans le cas de double mutant *Cdx2/Brachyury* (**Fig. 4B**) (Amin et al., 2016). *Brachyury* est un facteur de transcription qui régule directement l'expression de plusieurs gènes impliqués dans l'élongation axiale puisque l'expression des gènes *Wnt3a* et *Fgf8* est perdue chez les souris déficientes pour *Brachyury* (Amin et al., 2016; Evans et al., 2012). Ces données indiquent l'existence d'un rétrocontrôle positif de *Brachyury* sur l'expression de *Wnt3a* et *Fgf8* similaire à celui observé pour *Cdx* (**Fig. 4C**) (Amin et al., 2016; Martin and Kimelman, 2012; Yamaguchi et al., 1999).

3) *Oct4* : le facteur de pluripotence *Oct4* est un régulateur clé dans la formation du tronc des Vertébrés. Chez la souris, *Oct4* est largement exprimé dans les NMPs à E8.5 mais son expression diminue à partir de E9.5 (Osorno et al., 2012), ce qui correspond à la transition entre la formation du tronc et la queue. L'expression soutenue d'*Oct4* dans les progéniteurs axiaux pendant l'élongation embryonnaire en utilisant une approche transgénique provoque un prolongement du thorax au détriment de la queue (Aires et al., 2016).

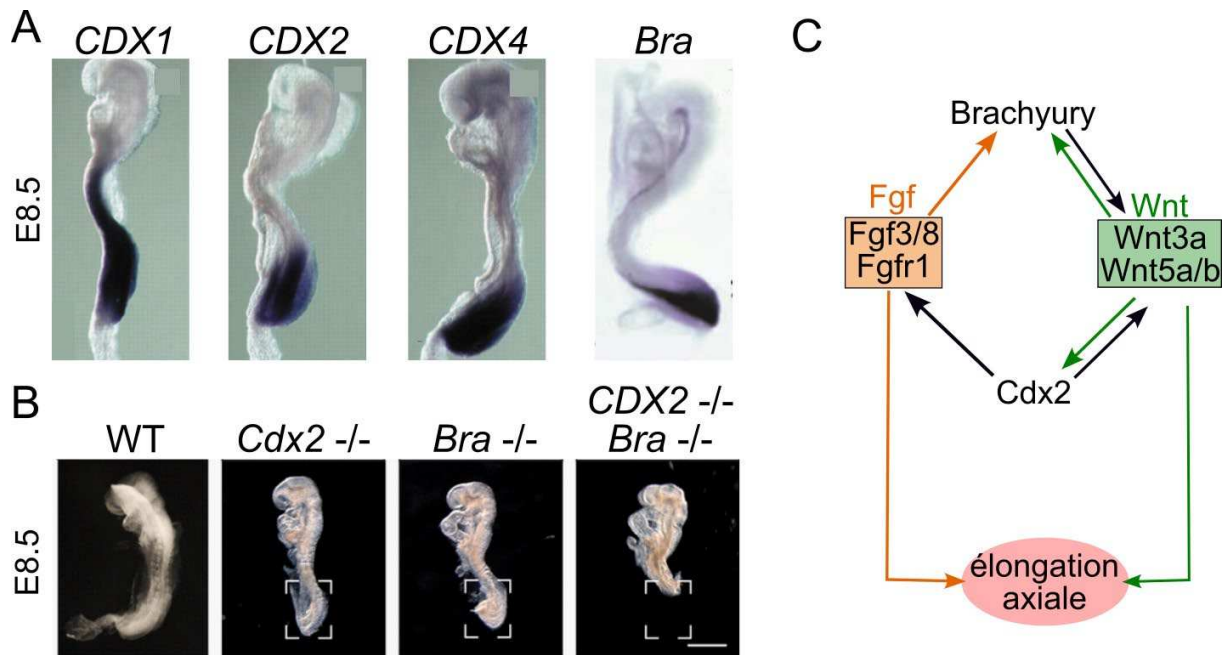


Figure 4. Les voies Wnt/Fgf et les facteurs Cdx/Brachyury coopèrent pour instruire la morphogénèse axiale. (a) hybridation *in situ* de *Cdx1/2/3* et *Brachyury* dans embryon de souris. L'expression de *Cdx1* s'étend à des positions plus antérieures que *Cdx2/4* et *Brachyury* qui sont restreints dans des régions postérieures. (b) Morphologie d'embryons de type sauvage et mutant *Cdx2*^{-/-}, *Brachyury*^{-/-} ou double mutant *Cdx2*^{-/-} et *Brachyury*^{-/-}. Les doubles mutants présentent un phénotype plus sévèrement que les mutants individuels. (c) Schéma illustrant le rôle des voies et facteurs dans l'élongation axiale. Les cibles *Brachyury* et *Cdx2* sont activées par Wnt et leur expression permet en retour le maintien de l'activité Wnt et Fgf. Cela permet le maintien des progéniteurs axiaux qui nourrissent l'élongation axiale. (adapté de Amin et al., 2016; Deschamps and Van nes, 2005)

En résumé, l'élongation axiale implique l'activité des voies Wnt et Fgf, ainsi que de facteurs de transcription tels que Brachyury, Cdx et Oct4 dans la région caudale de l'embryon. Outre la spécification et le maintien des progéniteurs axiaux, ces signaux participent également à la différenciation des progéniteurs axiaux vers des lignages neuraux ou mésodermiques.

c. Mécanismes et signaux contrôlant la différenciation des progéniteurs axiaux

i. Les facteurs Sox2 et Brachyury sont associés aux destins neuronaux et mésodermiques

Au cours de l'élongation axiale, les NMPs se différencient progressivement pour donner naissance à la moelle épinière et au mésoderme paraxial. Le niveau d'expression des marqueurs Sox2 et Brachyury, qui caractérisent les NMPs, est un paramètre important contrôlant la spécification d'un des deux destins cellulaires. En effet, dans la région caudale, les cellules les plus antérieures ou latérales de la ligne primitive expriment plus Sox2 que Brachyury et ont tendance à adopter un destin neural (**Fig. 5A/B**) (Wymeersch et al., 2016). A l'inverse, les

cellules se trouvant dans la partie plus caudale ou médiane vont exprimer plus Brachyury que Sox2 et ont tendance à se spécifier vers un lignage mésodermique. Sox2 et Brachyury sont donc exprimés selon un gradient antéro-postérieur et médio-latéral opposé. Des travaux complémentaires sur des embryons de caille ont ensuite révélé que des niveaux élevés de Brachyury confèrent une grande motilité aux progéniteurs, ce qui tend à les pousser à rejoindre le mésoderme paraxial, tandis que des niveaux d'expression élevés de Sox2 tendent à inhiber le mouvement des progéniteurs, favorisant leur intégration dans le lignage neural (Romanos et al., 2021).

En résumé, en fonction de leur position, les progéniteurs expriment plus ou moins Sox2 et Brachyury et montrent une préférence pour la différenciation vers le lignage neural ou le mésoderme présomitique. Par ailleurs, des expériences de greffes hétérotopiques ont montré que des NMPs provenant du NSB peuvent se différencier exclusivement en dérivés mésodermaux lorsqu'ils sont greffés dans la partie postérieure de la ligne primitive (Wymeersch et al., 2016). Ces résultats indiquent que les NMPs sont plastiques et leur différenciation est influencée par des facteurs environnementaux.

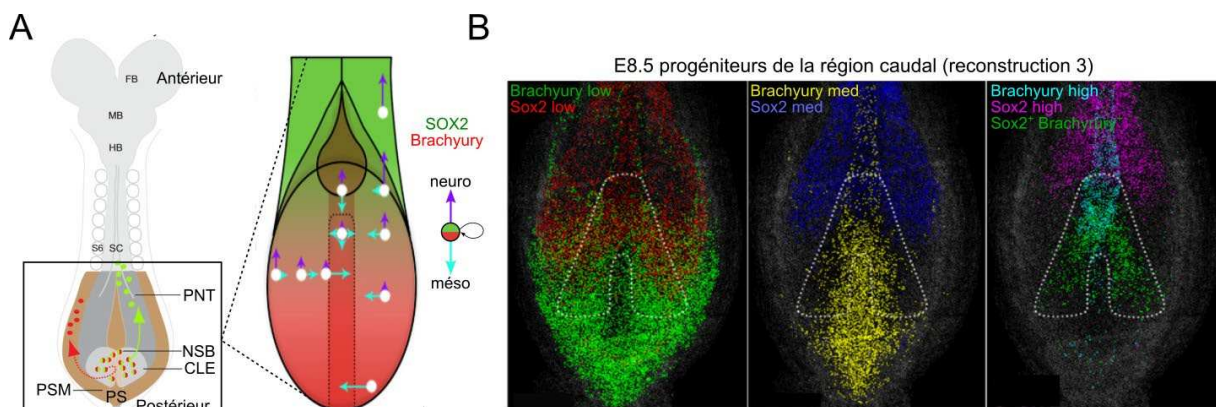


Figure 5. Destin cellulaire des progéniteurs axiaux et domaines d'expression des marqueurs Sox2 et Brachyury. (a) Schématisation de l'organisation du système nerveux en développement au stade E8.5 (embryon de souris). A ce stade les NMPs (Sox2 et Brachyury) donnent naissance aux progéniteurs mésodermaux et neuronaux. (b) Reconstruction 3D des marquages immunohistochimiques contre Sox2 et Brachyury dans la région caudale de l'embryon au stade E8.5. L'intensité de fluorescence est indiquée par une échelle de couleur graduée différente pour chaque marquage. FB = cerveau antérieur, MB = mésencéphale, HB = cerveau postérieur, SC = moelle épinière, PNT = tube pré-neural, NSB = frontière nœud/ligne primitive, CLE = épiblaste caudal latéral, PS = ligne primitive, PSM = mésoderme pré-somitique. (Adapté de Wymeersch et al., 2016)

ii. Les voies Wnt et Fgf instruisent la différenciation mésodermique

Les voies Wnt et Fgf, en plus de leur rôle dans le maintien des progéniteurs axiaux, jouent un rôle clé dans la différenciation des progéniteurs vers le destin mésodermique (Gouti et al., 2015; Takemoto et al., 2011; Wittler et al., 2007). En effet, la perte de *Wnt3a* chez la souris et le zebrafish conduit à une perte de progéniteurs du mésoderme paraxial (Garriock et al., 2015; Martin and Kimelman, 2012; Takada et al., 1994; Yoshikawa et al., 1997). Chez le

zebrafish, l'expression inductible d'un dominant négatif *Fgfr1* est caractérisée par un déficit dans la génération du mésoderme somitique postérieur (Goto et al., 2017). A l'inverse, l'activation prolongée de la voie Wnt et Fgf dans les NMPs dérivées de cellules souches pluripotentes humaines et de souris conduit à la différenciation mésodermique (Gouti et al., 2014).

Le changement d'activité de la voie Wnt, passant d'un signal favorisant le maintien des progéniteurs axiaux à un signal conduisant à leur différenciation, est étroitement lié à l'activité du complexe TALE/HOX (Pbx, Meis2, Hox1) (Mariani et al., 2021). En effet, le complexe TALE/HOX rend l'expression des gènes impliqués dans le destin mésodermique permissifs à la voie Wnt. De même, le niveau d'expression du facteur Sox2 joue un rôle crucial, car lorsque son activité diminue, la voie Wnt cesse de contribuer au maintien des gènes de la pluripotence et induit l'expression des gènes participant à la différenciation mésodermique (Blassberg et al., 2022).

iii. La voie RA favorise la différenciation neuronale

Lors de leur différenciation vers le destin neuronal, les NMPs migrent vers la zone préneurale du tube (PNT). La PNT, située dans la partie caudale de l'embryon, immédiatement rostrale par rapport aux NMPs, représente une population moléculairement distincte qui contribue largement à la moelle épinière (Bertrand et al., 2000; Brown and Storey, 2000; Delfino-Machín et al., 2005; Olivera-Martinez et al., 2014; Selleck and Stern, 1991). Les cellules PNT expriment *Nkx1.2* en réponse à la signalisation Fgf et répriment les facteurs de transcription des progéniteurs neuraux tels que *Pax6* (Bertrand et al., 2000; Delfino-Machín et al., 2005; Olivera-Martinez and Storey, 2007). L'exposition des cellules PNT à la signalisation de l'acide rétinoïde (RA) provenant des somites favorise leur transition vers une identité de progéniteur neural (Del Corral et al., 2003; Kumar and Duester, 2014; Lunn et al., 2007).

Le RA est synthétisé à partir de la vitamine A qui est transformée par la rétinaldéhyde déshydrogénase (**Fig. 6A**) (*Raldh1/2/3*). Moléculairement, le RA agit en se fixant aux récepteurs nucléaires à l'acide rétinoïque (RARs) qui se lient de manière constitutive aux régions génomiques appelées éléments de réponse de l'acide rétinoïque (RAREs) (Niederreither and Dollé, 2008). Les RARs peuvent se lier au génome en absence de RA mais recruteront dans ce cas des cofacteurs répresseurs comme PRC2 qui agit comme compacteur de la chromatine (Kumar and Duester, 2014; Studer et al., 1994). Au contraire lorsque les RARs ont lié le RA,

le complexe agit comme activateur en recrutant des cofacteurs induisant la transcription (Gillespie and Gudas, 2007).

Le rôle de RA a d'abord été identifié au cours de la formation de l'axe antéro-postérieur du zebrafish. L'augmentation du temps d'exposition au RA ou de sa concentration conduit à un raccourcissement de l'axe (**Fig. 6B**) (Sive et al., 1990). Dans les mutants murin *Raldh2* ou dans des modèles de privation de vitamine A, qui permettent de mimer une diminution de l'activité de la voie, on observe une extension du domaine d'expression de *Fgf8* dans des régions plus antérieures (**Fig. 6D**). A l'inverse, la co-inhibition des voies Wnt et Fgf a pour conséquence l'augmentation du niveau d'expression de *Raldh2* (Del Corral et al., 2003; Diez Del Corral and Morales, 2017; Olivera-Martinez and Storey, 2007; Vermot and Pourquié, 2005). Les voies Fgf et RA semblent donc être antagonistes, mais les mécanismes moléculaires qui sous-tendent cette régulation restent mal compris.

D'autres études montrent que RA inhibe la voie Fgf en se fixant sur les éléments régulateurs de *Fgf8* (Del Corral et al., 2003; Kumar and Duester, 2014). Cette inhibition de la voie Fgf par RA induit alors la perte d'expression de *Nkx1.2* ainsi que l'induction de l'expression des gènes *Pax6* et *Irx3* qui sont importants pour la différenciation neuronale (Del Corral et al., 2003). De plus, dans des explants de la région caudale d'embryon de poulet, le nombre de cellules exprimant le gène proneural *NeuroM* augmente suite à des expositions au RA (Del Corral et al., 2003). A l'inverse, les mutants *Raldh2* sont caractérisés par une expansion des marqueurs du mésoderme au dépens des marqueurs neuraux (Ribes et al., 2009). En résumé, le RA contrôle la formation de l'axe antéro-postérieur en réprimant les voies Wnt/Fgf et en induisant la différenciation des progéniteurs vers le lignage neural (**Fig. 6E**).

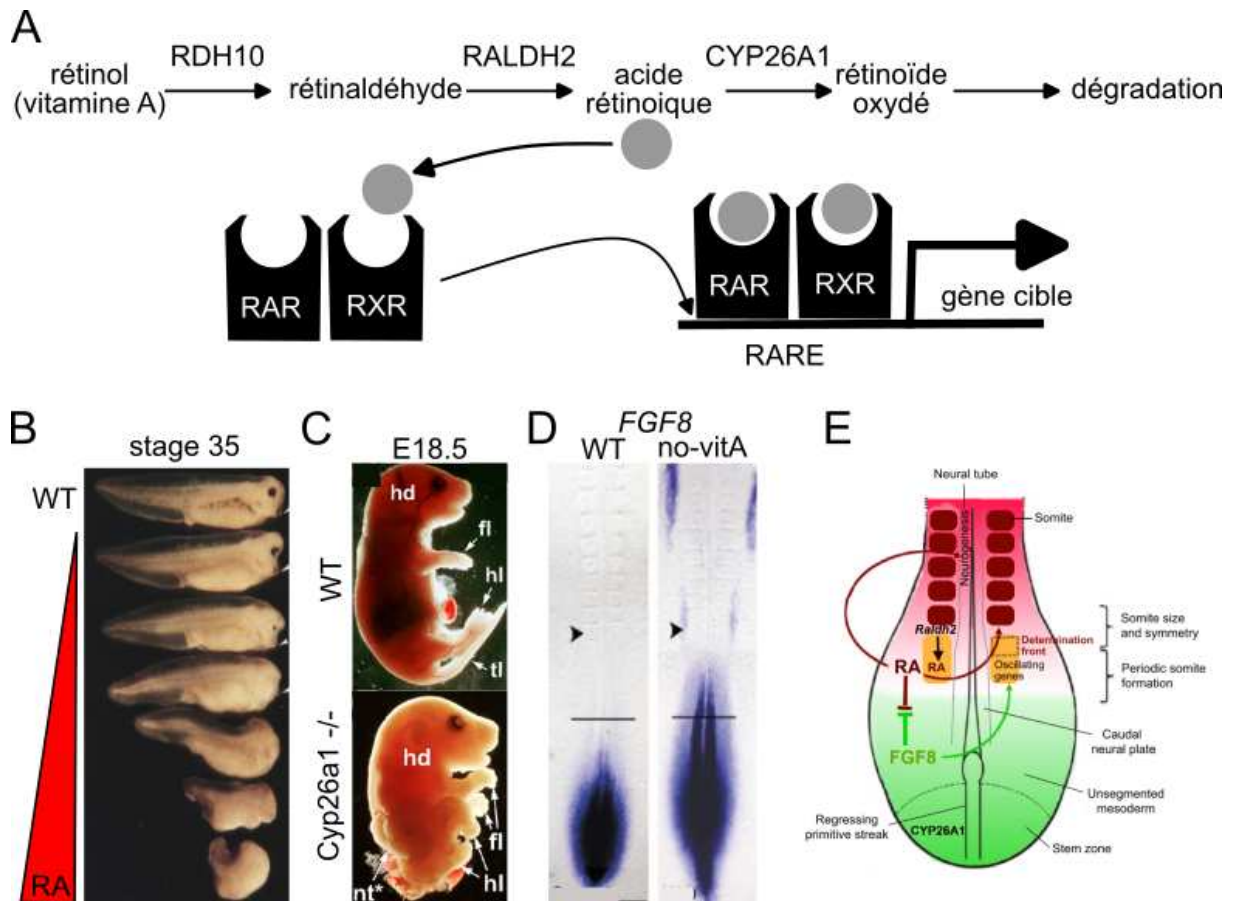


Figure 6. Contribution du RA durant l'élongation axiale. (a) Schéma de la biosynthèse et de la signalisation de l'acide rétinolique. L'acide rétinolique venant du milieu extracellulaire vient se fixer aux récepteurs RARs localisés dans le noyau au niveau des gènes cibles, permettant le recrutement de co-activateurs rendant possible la transcription. (b) Morphologie d'embryons de zebrafish exposés pendant des périodes de temps variables à 1 μ M de RA. RA est appliqué au stade 10 pendant 5min à 2h. Les embryons résultants sont présentés au stade 35. Un embryon control est représenté en haut. (c) Phénotypes de souris de type sauvage et double mutant *Cyp26a1*^{-/-} au stade prénatal. Les mutants *Cyp26a1* sont caractérisés par un tube ouvert et une troncation des segments postérieurs. (d) Hybridation *in situ* de *Fgf8* dans embryon de poulet de type sauvage et déplété de vitamine A (absence de synthèse RA). L'expression de *Fgf8* est prolongée dans les régions antérieures chez les embryons sans vitamine A. (e) Le RA produit au niveau des somites va s'opposer au gradient de Fgf venant de la région la plus caudale de l'embryon et participer ainsi à l'équilibre entre induction neurale et maintien d'un état prolifératif des progéniteurs. (Adapté d'après Del Corral et al., 2003; Rhinn and Dollé, 2012; Sive et al., 1990)

d. Mécanismes et signaux contrôlant la terminaison de l'élongation axiale

L'élongation axiale est stoppée par plusieurs événements moléculaires coordonnées. Le facteur sécrété *Gdf11*, un membre de la famille TGF- β , joue un rôle clé dans ce processus. Dans les embryons de souris, l'expression de *Gdf11* est initiée dès E7.5 puis son expression augmente progressivement à E8.0 dans la région de la ligne primitive. *Gdf11* est ensuite exprimé dans le bourgeon caudal à E9.5, une période qui initie la transition de la formation du tronc à la queue et qui précède la terminaison de l'élongation axiale (**Fig. 7B**) (McPherron et al., 1999; Nakashima et al., 1999; Tam and Tan, 1992).

Comme pour les autres membres de la famille TGF- β , la forme active de *Gdf11* est générée par clivage protéolytique grâce à la proprotéine convertase PCSK5 (PC5/6) (Essalmani

et al., 2008; Szumska et al., 2008). Le Gdf11 signale en se liant d'abord au récepteur II de l'activine (ActRIIA/B) qui recrute ensuite le récepteur I de l'activine (Alk4/5/7). Cette activation provoque la phosphorylation de Smad2/3 qui recrute alors Smad4 (**Fig. 7A**) (Massagué et al., 2005). Lorsque les complexes Smad2/4 et Smad3/4 sont formés, ils transloquent vers le noyau et régulent l'expression de gènes en se liant directement sur des promoteurs caractérisés par des séquences "SMAD-binding elements" (Shi and Massagué, 2003; Zawel et al., 1998) ou en s'associant à des cofacteurs (e.g. Foxh1) (Attisano et al., 2001). En plus de la voie canonique, Gdf11 active d'autres signaux non-Smad comme les MAPKs (p38, ERK et JNK) ou AKT (Zhang, 2009).

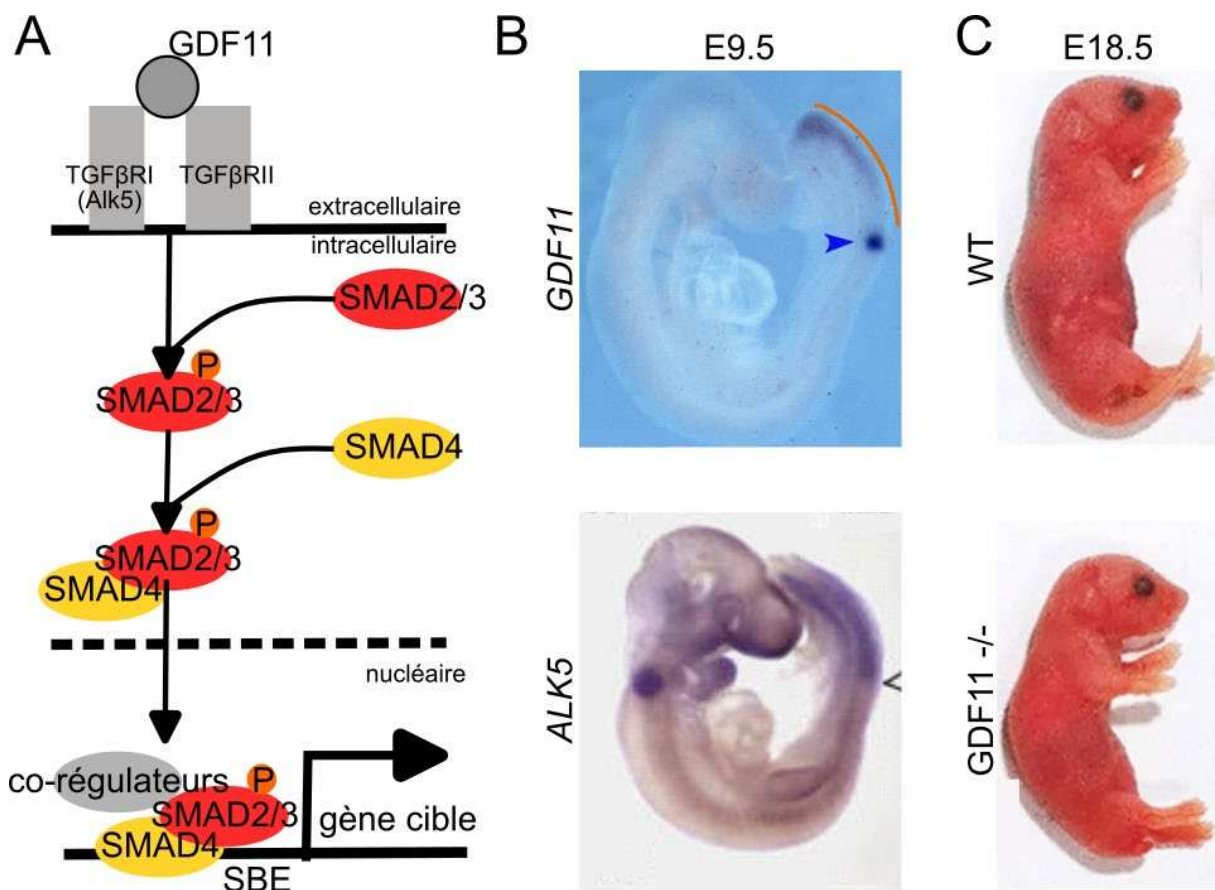
Les mutants murins nuls pour *Gdf11* sont caractérisés par des membres inférieurs décalés caudalement de plusieurs vertèbres ainsi qu'une queue tronquée (**Fig. 7C**) (McPherron et al., 1999). Les mutants *Alk5* et *ActRIIB* présentent des malformations similaires aux défauts trouvés chez les mutants *Gdf11* avec un raccourcissement de la queue (Andersson et al., 2006; Oh et al., 2002). Ces résultats indiquent que la voie Gdf11 joue un rôle important dans la détermination de la position des membres postérieurs et la transition du tronc à la queue.

Moléculairement, les mutants *Gdf11* sont caractérisés par une diminution d'expression de *Cyp26a1* (enzyme dégradant RA), et par l'augmentation concomitante de l'activité de RA dans la région caudale des embryons (Lee et al., 2010). De manière cohérente, le phénotype de queue tronquée est partiellement restauré par l'inhibition de la signalisation RA et à l'inverse, une augmentation de cette signalisation aggrave ce phénotype (Jurberg et al., 2013; Lee et al., 2010). Ces résultats indiquent que le Gdf11 est important pour prévenir la signalisation RA dans la partie caudale afin de permettre la génération de la queue.

Par ailleurs, d'autres résultats indiquent que la voie Gdf est importante pour stopper l'élongation axiale. En effet, le maintien des NMPs dépend de *Lin28/let-7* et *Hox13* qui sont régulés par la voie Gdf (Aires et al., 2019). Les gènes *Lin28* codent pour des protéines de liaison à l'ARN qui ont été impliquées dans le contrôle de l'activité des cellules souches (Yang et al., 2015). En particulier, l'inactivation de *Lin28a* entraîne un raccourcissement significatif de la queue, tandis que la stimulation excessive de l'expression du gène *Lin28* augmente le nombre de vertèbres de la queue (Aires et al., 2019; Robinton et al., 2019). D'autres études génétiques ont montré que l'activité de *Lin28* passe par le blocage de la maturation de *let-7* (Viswanathan and Daley, 2010). Les gènes *HoxB13/C13* sont quant à eux exprimés tardivement dans le bourgeon de la queue et leur activation prématurée conduit à une réduction de la taille de la queue (Aires et al., 2019; Young et al., 2009). Lorsque les facteurs *Hox13* sont exprimés, ils

interfèrent également avec l'activité de *Cdx2* ce qui entraîne la répression des gènes impliqués dans la signalisation Wnt et Fgf (Amin et al., 2016; Young et al., 2009). A l'inverse, lorsque *Gdf11* est muté, alors *Hox13* n'est pas induit et la population de NMPs dans le bourgeon caudal est plus importante (Aires et al., 2019). Par conséquent, le facteur Gdf11 contrôle l'équilibre entre les activités de Lin28 et Hox13, qui répriment et activent respectivement l'élongation axiale en régulant le maintien des NMPs.

L'arrêt de l'élongation axiale est également associé à l'augmentation de l'activité de la voie RA dans le bourgeon caudal qui entraîne une baisse de la signalisation FGF, ce qui provoque la perte de l'expression de Brachyury et induit localement la mort cellulaire. De manière intéressante, la diminution de l'activité de la voie RA, indépendamment de la voie FGF, atténue cet effet. À ce stade, la voie FGF ne semble plus s'opposer à la synthèse du RA et à son activité, cependant, les mécanismes par lesquels le RA échappe au contrôle de la voie FGF ne sont pas encore connus (Olivera-Martinez et al., 2012)



Comme nous avons pu le voir dans ce chapitre, la formation de l'axe antéro-postérieur du corps dépend de progéniteurs axiaux qui se renouvellent ou se différencient vers des lignages neuraux ou mésodermiques. Le comportement de renouvellement ou de différenciation des progéniteurs axiaux, incluant les NMPs, est régulé par les voies de signalisation Wnt, Fgf et RA ainsi que des facteurs de transcription Sox2 et Brachyury. Ce réseau génétique fonctionne tout au long de l'élongation axiale. L'état bipotent neuro-mésodermique des progéniteurs est maintenu lorsque les signaux Wnt, Fgf et RA sont équilibrés. Puis, le choix du destin cellulaire est déterminé par un changement de l'équilibre entre ces signaux avec les voies Wnt et FGF qui promeuvent une différenciation mésodermique, tandis que RA favorise la différenciation neurale. Ces signaux pourraient agir sur la balance entre les deux facteurs Sox2 et Brachyury qui promeuvent respectivement la différenciation neurale ou mésodermique (Henrique et al., 2015). La terminaison de l'élongation axiale est quant à elle instruite par 1) l'augmentation de l'activité de RA dans le bourgeon caudal qui provoque une baisse de la signalisation Fgf et 2) l'activation de Hox13 et la répression de Lin28 par la voie Gdf.

Parallèlement à cette élongation antéro-postérieure, les progéniteurs générés acquièrent une identité positionnelle qui détermine la formation appropriée du schéma corporel en permettant la spécification de différents types de cellules au bon endroit et au bon moment. L'identité positionnelle des cellules composant les différents segments du corps est orchestrée par la grande famille des gènes *Hox*.

II. Les gènes *Hox* contrôlent la spécification de l'axe antéro-postérieur

Les gènes *Hox* constituent une large famille codant pour des facteurs de transcription caractérisés par un homéo-domaine, une séquence très conservée de liaison à l'ADN formée de 3 hélices (Keynes and Krumlauf, 1994; McGinnis et al., 1984). Par leur action transcriptionnelle, les facteurs Hox instruisent la morphogenèse de l'axe en fournissant des informations de position aux tissus axiaux embryonnaires émergents. Dans ce chapitre, nous allons discuter comment les gènes *Hox* participent à la spécification des identités cellulaires de l'axe antéro-postérieur principalement chez les Vertébrés, et décrire les mécanismes régulant leur expression.

1. Découverte des gènes *Hox*

Des expériences de génétique menées sur la drosophile ont démontré l'existence de gènes impliqués dans l'attribution d'identités morphologiques distinctes (Lewis, 1978; Postlethwait and Schneiderman, 1971). Les mutations de ces gènes conduisent en effet à des transformations homéotiques où un segment corporel adopte la morphologie d'un autre segment. Il a ensuite été démontré que la drosophile possède 8 gènes homéotiques organisés dans deux complexes : *Antennapedia* (comprenant 5 gènes : *Labial*, *Peoboscipedi*, *Deformed*, *Sex comb reduced* et *Antennapedia*) et *Bithorax* (comprenant 3 gènes : *Ultrabithorax*, *Abdominal A* et *Abdominal B*) (Lewis, 1998, 1978). Par la suite, le clonage d'homologues de ces gènes a permis d'identifier les gènes *Hox* chez les Vertébrés (Krumlauf, 1994). Les gènes *Hox* sont conservés chez tous les animaux bilatériens mais leur nombre varient selon les espèces.

2. Organisation des domaines d'expression des *Hox* le long de l'axe antéro-postérieur

Chez les Vertébrés, 39 gènes *Hox* différents ont été découverts. Ils sont regroupés dans le génome dans 4 complexes paralogues (*HoxA/B/C/D*). Le nombre de gènes *Hox* dans chaque complexe varie de 9 à 11, et les paralogues sont présents dans au moins deux complexes différents (Pang and Thompson, 2011). Ces complexes résultent probablement d'une duplication successive d'un complexe ancestral commun (**Fig. 8A**).

Grâce à des expériences d'hybridation *in situ*, les domaines d'expression des gènes *Hox* ont été caractérisés au cours du développement embryonnaire chez le poulet et la souris (Dasen et al., 2005; Duboule and Dollé, 1989; Gaunt, 1988; Gaunt et al., 1989; Liu et al., 2001). Ils sont exprimés dans les dérivés des trois feuilletts embryonnaires avec des patrons d'expression qui reflètent leur position dans le génome (Beck, 2002; Carapuço et al., 2005; Deschamps and Van Nes, 2005; Grapin-Botton and Melton, 2000; Yoshida et al., 2017). Les gènes sont exprimés à partir de la région caudale de l'embryon, cependant, plus un gène se trouve en position 3' des complexes *Hox*, plus sa frontière d'expression est antérieure (**Fig. 8B**).

Ainsi dans le système nerveux, les gènes se trouvant à l'extrémité 3' des complexes jusqu'aux paralogues 4 ont des limites antérieures d'expression séquentielle au niveau du tronc cérébral. Les gènes *HoxA* et *HoxC* à partir des paralogues 5 jusqu'aux 11 sont exprimés séquentiellement dans la moelle épinière (Dasen et al., 2005; Liu et al., 2001; Philippidou et al., 2012). Pour le complexe *HoxD*, *HoxD3/D4* sont exprimés toute le long de la moelle tandis que

HoxD8/9 sont exprimés dans la région thoracique et *HoxD10/D11* commencent à s'exprimer dans la région lombaire (Dasen et al., 2005; Tschopp et al., 2012). En revanche, les gènes du complexe *HoxB* ne présentent pas une expression régionalisée dans la moelle épinière (Carpenter, 2002; Dasen et al., 2003; Graham et al., 1989).

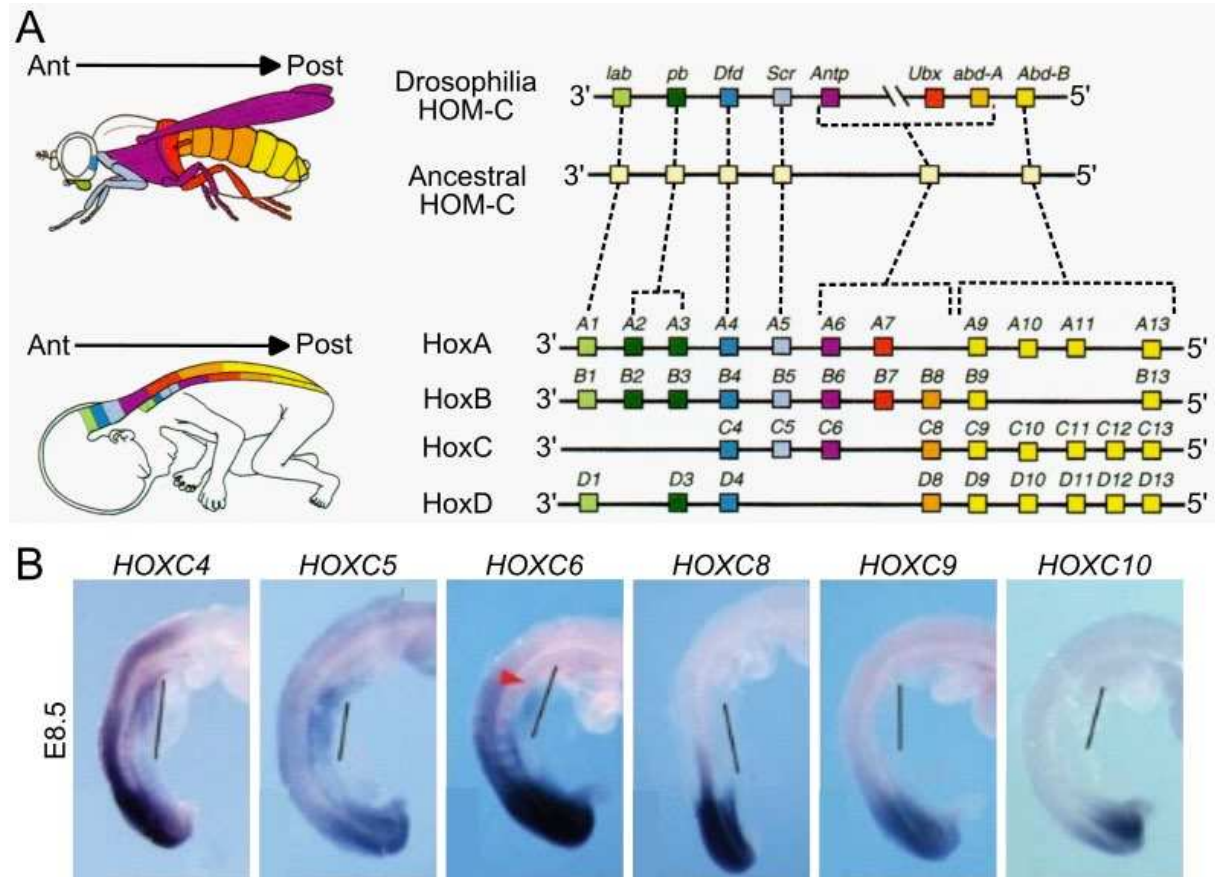


Figure 8. Organisation des gènes *Hox* et leurs domaines d'expression chez la drosophile et les Vertébrés. (a) Chez la drosophile, représentation du cluster C, de 3' vers 5' et domaines d'expression selon le plan d'organisation antéro-postérieur. Chez les Vertébrés, représentation de l'organisation des quatre complexes A, B, C et D et leurs domaines d'expression dans la moelle épinière. **(b)** Hybridation *in situ* sur des embryons de souris au stade E8.5. Les gènes *Hox* 3' induits précocement dans la partie caudale de l'embryon sont exprimés dans la partie antérieure de la moelle, née dans les premiers stades de développement. Les gènes *Hox* plus localisés en 5' sont exprimés à des stades plus tardifs et leur expression se limite aux zones caudales de l'embryon. (Adapté d'après Mark et al., 1997; Nishimoto et al., 2014)

Initialement, les gènes *Hox* ont donc des patrons d'expression imbriqués depuis les régions caudales de l'embryon jusqu'aux extrémités antérieures ce qui conduit à un chevauchement considérable entre les domaines d'expression des gènes *Hox*. Par la suite l'expression des gènes *Hox* est modulée par répression des gènes antérieurs par les facteurs *Hox* plus postérieurs, un phénomène nommé prévalence du postérieur qui donne naissance à des domaines d'expression plus restreints le long de l'axe antéro-postérieur et la formation d'un code *Hox* régionalisé (**Fig. 9A**) (Duboule and Morata, 1994; Kmita and Duboule, 2003; Schöck et al., 2000). Ainsi dans la moelle épinière de souris, les motoneurons *HoxA5/C4/C5* sont

exprimés dans le domaine cervical, HoxC4/A5/C5/A6/C6/A7 dans le domaine brachial antérieur et HoxC4/A6/C6/A7/C8 dans le domaine brachial postérieur. Les motoneurones thoraciques se distinguent par la co-expression de HoxC8/C9 dans la partie la plus antérieure et HoxC9/D9 dans la partie la plus caudale. HoxA10/C10/D10 sont exprimés par les motoneurones lombaires antérieurs alors que HoxA11/C11/D11 par les lombaires postérieurs (Fig. 9B/C) (Philippidou and Dasen, 2013).

L'expression des gènes *Hox* est similairement régionalisée dans les dérivés mésodermiques (par exemple les somites), mais les limites d'expression ne sont pas nécessairement alignées avec celles du tissu nerveux (Imura et al., 2009). Ainsi, la limite d'un gène *Hox* dans les cellules neuronales va généralement être plus antérieure que la limite de ce même gène dans le mésoderme. Il y a donc un décalage des domaines d'expression des *Hox* entre les cellules neuronales et mésodermiques.

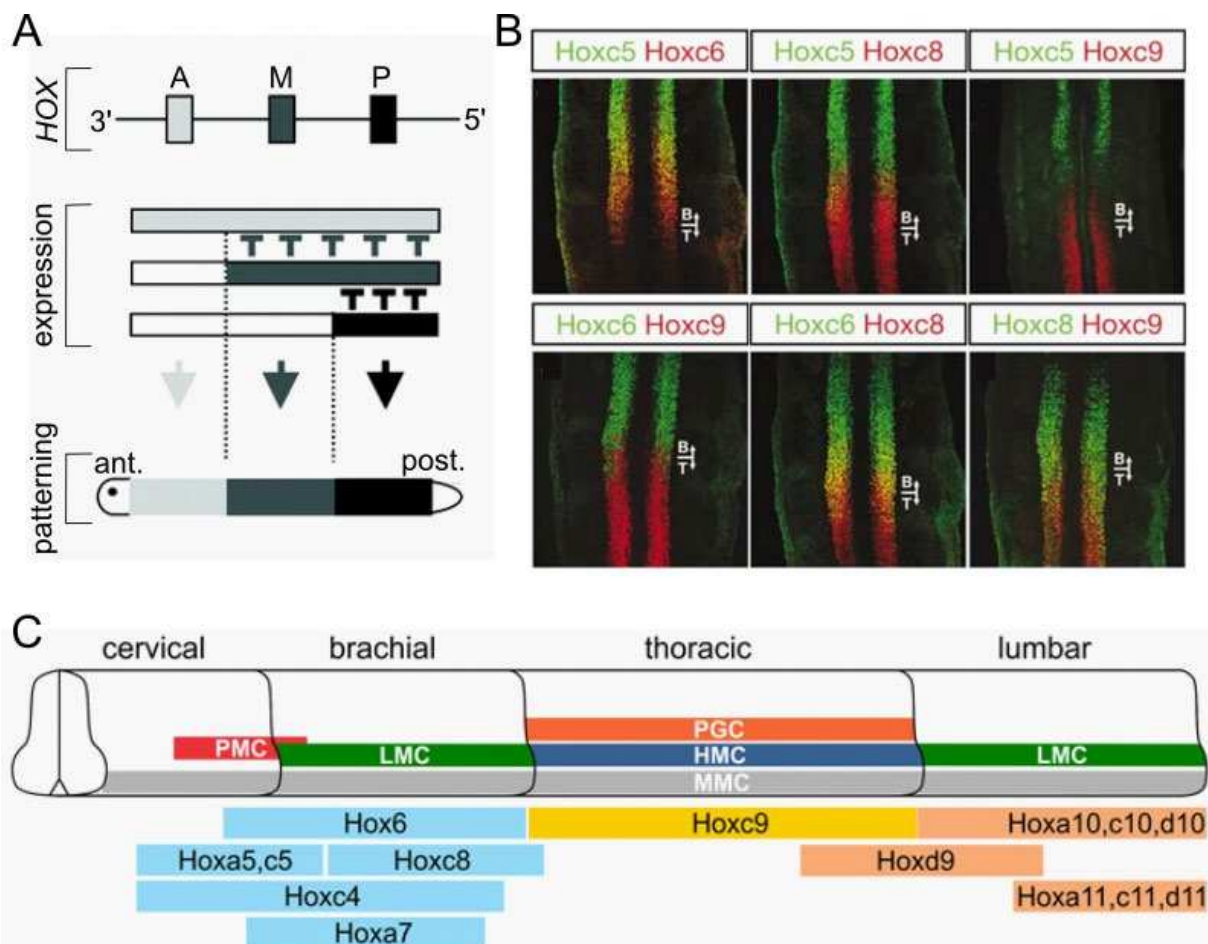


Figure 9. Organisation de l'expression des protéines Hox dans la moelle épinière. (a) Illustration de l'établissement du code Hox le long de l'axe antéro-postérieur en fonction d'un gène antérieur (A), moyen (M) et postérieur (P). L'établissement du code Hox le long de l'axe antéro-postérieur dépend de la prévalence de la protéine la plus postérieure exprimée qui va réprimer les protéines qui lui sont antérieures. (b) Immunohistochimie mettant en évidence les domaines d'expression des facteurs Hox dans la moelle épinière. HoxC5 est exprimé dans le domaine cervical et la partie la plus antérieure du domaine brachial. La frontière entre les domaines brachial et thoracique est définie par la limite postérieure d'expression de HoxC6 et

la limite antérieure du domaine HoxC9. HoxC8 est co-exprimé avec HoxC6 dans la partie postérieure du domaine brachial et co-exprimé avec HoxC9 dans la partie antérieure du domaine thoracique. (c) Schéma représentant l'organisation antéro-postérieure des motoneurones en colonnes dans la moelle épinière avec les domaines d'expression des Hox associés. PMC = colonne motrice phrénique, LMC = Colonne motrice latérale, HMC = Colonne motrice Hypaxiale, PGC = Colonne préanglionnaire, MMC = colonne motrice médiane. (Adapté d'après Dasen et al., 2005; Kmita and Duboule, 2003; Philippidou and Dasen, 2013)

3. Rôle des facteurs Hox dans la spécification de l'axe antéro-postérieur

L'importance des facteurs Hox pour établir l'identité positionnelle a été illustrée dans différentes espèces, types cellulaires et segments de l'axe antéro-postérieur. Chez les vertébrés, le rôle des facteurs Hox comme encodeur de l'information positionnelle a d'abord été démontré par l'impact des pertes et gains de fonction sur la formation du squelette. Puis ces mutations ont été caractérisées par rapport à l'acquisition de l'identité positionnelle des motoneurones le long de l'axe antéro-postérieur.

a. Le plan d'organisation des vertèbres dépend des facteurs de transcription Hox

Au cours de l'embryogenèse, chaque vertèbre dérive d'une paire de somites. De plus, le nombre de vertèbres et leur morphologie sont conservés au sein de chaque espèce. Par exemple chez la souris on retrouve : 7 vertèbres cervicales, 13 thoraciques caractérisées par une paire de côtes (les 7 antérieures sont rattachées au sternum et les 6 postérieures sont flottantes), 6 lombaires, 4 sacrales et la queue qui est formée d'environ 30 vertèbres caudales.

Plusieurs études ont montré que le changement d'expression des gènes *Hox* a des conséquences sur le plan d'organisation des vertèbres. Chez la souris, l'expression ectopique de *HoxC6* dans des régions plus postérieures que sa zone d'expression normale induit la formation d'une vertèbre thoracique supplémentaire là où se forme normalement des vertèbres lombaire (Jegalian and De Robertis, 1992). La surexpression de *HoxC8* dans le domaine postérieur induit également une antériorisation des vertèbres (Pollock et al., 1992). La surexpression de *HoxA10* dans l'ensemble de l'axe antéro-postérieur induit une perte des vertèbres thoraciques et l'absence de formation du sacrum, issu de la fusion des vertèbres sacrées (**Fig. 10**) (Carapuço et al., 2005). A l'inverse des expériences par gain de fonction, l'absence de *HoxC8* (Akker et al., 2001) aboutit à une conversion de la première vertèbre thoracique en une huitième vertèbre cervicale alors que les souris n'exprimant pas *HoxA11* (Small and Potter, 1993) ont une vertèbre lombaire à la place de la dernière vertèbre thoracique.

La modification de l'activité d'un seul gène *Hox* peut donc provoquer une antériorisation d'une partie des vertèbres, et dans certains cas même modifier leur nombre.

Néanmoins, les phénotypes dans les souris mutées pour des gènes *Hox* paralogues ont parfois des perturbations de l'identité antéro-postérieure plus importantes que celles mutées pour un seul gène suggérant une redondance d'activité. C'est le cas de l'inactivation de *HoxA4* ou *HoxB4* ou *HoxD4* qui a des effets mineurs alors que la co-inactivation de cet ensemble de gènes induit un déplacement rostral des vertèbres cervicales (Horan et al., 1995). Il peut donc aussi y avoir une redondance fonctionnelle entre les gènes *Hox* paralogues (Krumlauf, 1994; Mallo et al., 2009; Wellik, 2009).

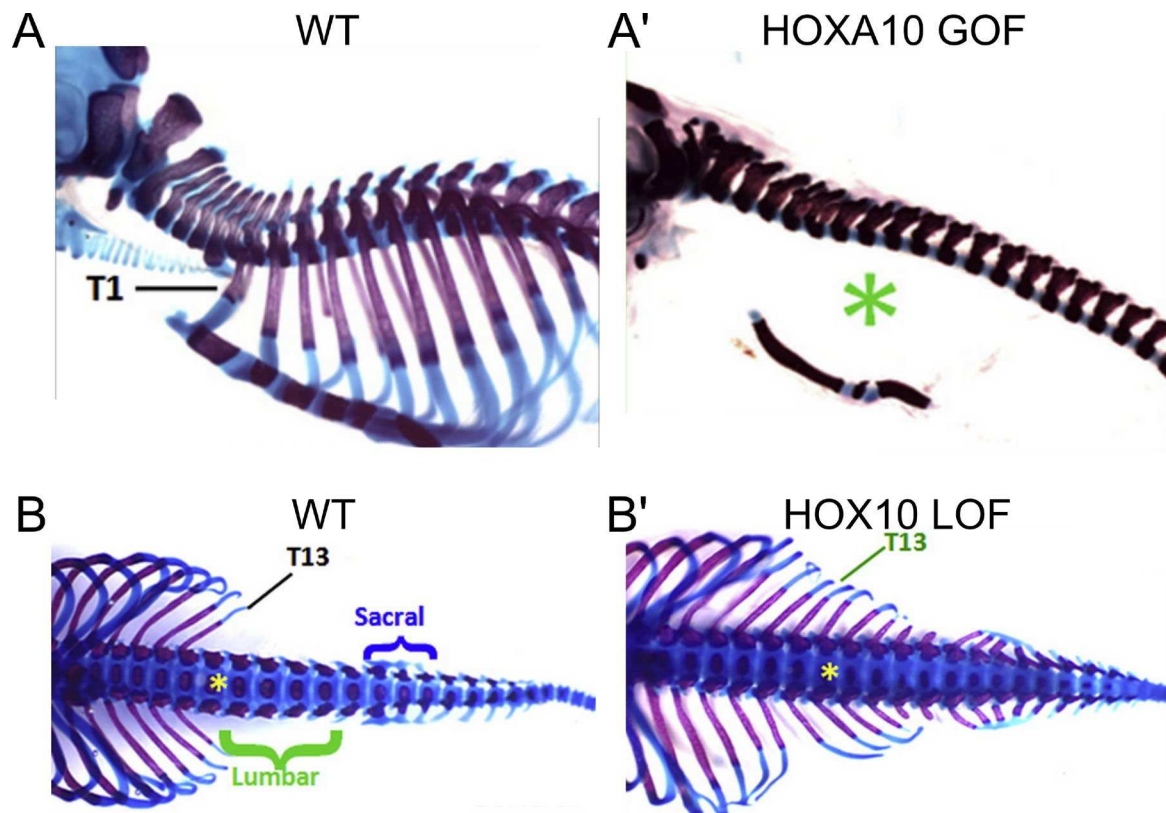


Figure 10. Les gènes *Hox* contrôlent l'organisation du plan vertébral. (a, a') Squelette axial chez des souris de type sauvage et après la surexpression de *HoxA10* dans le mésoderme paraxial. Le gain de fonction de *HoxA10* entraîne un défaut de formation des vertèbres thoraciques. (b, b') Squelette axial chez des souris de type sauvage et triple mutant *Hox10*. La perte de fonction de *Hox10* a pour conséquence la perte de vertèbres lombaires. (Adapté d'après Carapuço et al., 2005; Wellik and Capecchi, 2003)

b. La spécification des colonnes de motoneurones dépend des facteurs de transcription *Hox*

Les mutations des gènes *Hox* ont également un impact sur la spécification des sous-types de motoneurones le long de l'axe antéro-postérieur. Ainsi chez la souris, la mutation du gène *HoxC9* conduit à la transformation des motoneurones thoraciques en motoneurones brachiaux (Jung et al., 2010). En effet, chez ces mutants, on observe une absence de la colonne

motrice hypaxiale normalement présente au niveau thoracique et qui innervent les muscles intercostaux. Les motoneurones acquièrent alors une identité correspondant à la colonne motrice latérale qui est présente au niveau brachial et lombaire comme le montre l'expansion des facteurs de transcription *Foxp1* et *Scip* qui sont caractéristiques des motoneurones innervant les muscles des membres. En revanche, les domaines d'expression de *Foxp1* et *Scip* sont réduits chez les mutants *HoxC6* (Lacombe et al., 2013) ou *HoxC8* (Tiret et al., 1998; Vermot et al., 2005), et l'identité des motoneurones innervant les membres dans les régions brachiales où ces gènes sont normalement exprimés est altérée. De plus, la perte d'expression de *HoxC10/D10* conduit à une extension des motoneurones thoraciques dans les régions lombaires (Wu et al., 2008). Les facteurs Hox instruisent donc l'acquisition des identités des sous-types de motoneurones et l'organisation des colonnes motrices (**Fig. 11**).

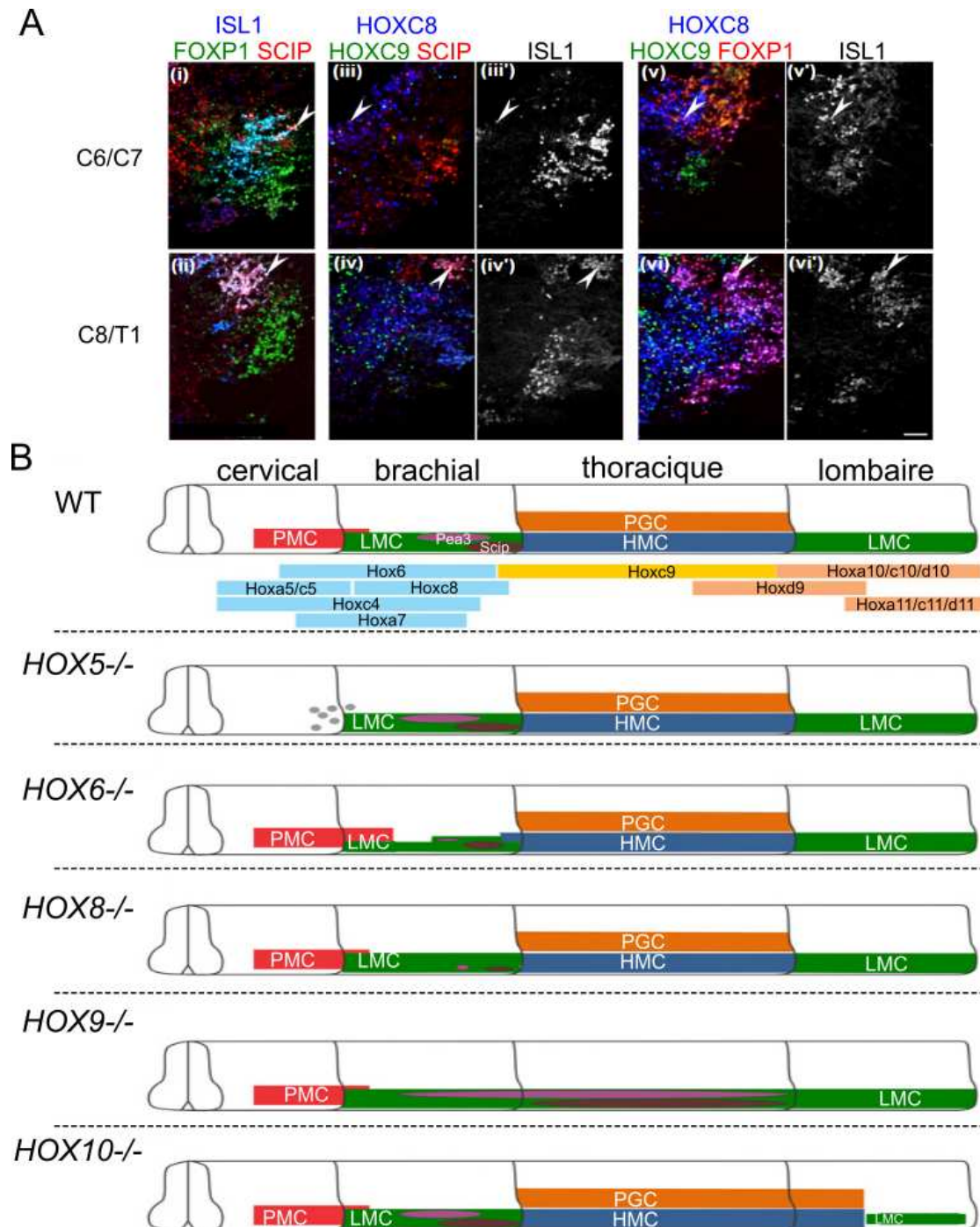


Figure 11. Effet des pertes de fonctions des gènes *Hox* sur la formation des populations de motoneurones. (a) Immunohistochimie de Isl1, Foxp1, Scip, HoxC8 et HoxC9. Les flèches en i) et ii) indiquent les motoneurones Foxp1/Scip/Is11-postitifs dans la moelle épinière brachiale caudale et thoracique antérieure. Les flèches en iii) et v) indiquent la présence de motoneurones Foxp1/HoxC8/Is11-positifs et HoxC9-négatifs (ISL1 montré en iii') et v'). En iv, iv' et vi, vi', les motoneurones Scip/Is11 ou Foxp1/Is11 autour et caudalement à la jonction brachiale caudale/ thoracique antérieure co-exprime HoxC8 et HoxC9. **(b)** La mutation de *HoxC5* réprime l'organisation de la colonne phrénique motrice tandis que les mutations de *HoxC6* et *HoxC8* vont avoir un effet au niveau brachial sur l'organisation des pools de motoneurones exprimant Scip et Pea3. La mutation de *HoxC9* conduit à une extension de la colonne motrice latérale des régions antérieures jusqu'au domaine lombaire alors que la mutation de *HoxC10* réduit le domaine cette colonne au niveau lombaire. (Adapté d'après Mouilleau et al., 2021; Philippidou and Dasen, 2013)

4. Induction temporelle des gènes *Hox* lors de l'élongation

Avant la gastrulation, les gènes *Hox* ne sont pas exprimés. L'initiation de la transcription des gènes *Hox* commence au stade HH4 chez le poulet et au stade E7.2 chez la souris dans les progéniteurs axiaux (Dasen et al., 2005; Neijts et al., 2016). Les gènes en 3' du complexe *HoxB* sont les premiers à être activés dans la zone postérieure de la ligne primitive, suivis par l'activation des gènes en 5' et ceux des autres complexes plus tardivement (Deschamps and Wijgerde, 1993; Forlani et al., 2003; Gaunt and Strachan, 1996). Les gènes en 3' des complexes sont exprimés les plus tôt, suivis progressivement de l'activation des gènes situés plus en 5' des complexes, un processus nommé colinéarité temporelle. Il y a donc un lien entre l'arrangement des gènes *Hox* au sein de leur complexe et leur ordre d'expression durant l'embryogenèse (Carpenter, 2002; Denans et al., 2015; Duboule and Dollé, 1989; Izpisua-Belmonte et al., 1991; Kmita and Duboule, 2003).

L'expression temporelle des gènes *Hox* dans la partie caudale de l'embryon est associée à la différenciation progressive de ces cellules pour former l'axe antéro-postérieur. En effet, lorsque les gènes *Hox* situés en 3' sont activés, leur domaine d'expression s'étend progressivement vers les régions antérieures par un processus qui n'implique pas de migration cellulaire (Deschamps and Wijgerde, 1993; Gaunt and Strachan, 1994). Par la suite, d'autres gènes *Hox* localisés en 5' sont induits et suivent la même dynamique. La coïncidence entre ces deux événements conduit à l'expression précoce des gènes *Hox* positionnés en 3' dans les régions antérieures de l'embryon, tandis que ceux en 5' sont exprimés plus caudalement et à des stades tardifs du développement (Nishimoto et al., 2014; Tschopp et al., 2012).

En résumé, ces résultats montrent que la mise en place des domaines d'expression des gènes *Hox* le long de l'axe antéro-postérieur est initiée au cours de l'élongation axiale dans les progéniteurs localisés dans le bourgeon caudal qui nourrissent aussi l'élongation axiale (Forlani et al., 2003). Les gènes *Hox* sont exprimés séquentiellement en fonction de leur position sur les complexes. A partir de ce constat, des études ont cherché à comprendre les mécanismes contrôlant la colinéarité temporelle des gènes *Hox*.

5. Mécanismes participant à la régulation transcriptionnelle des gènes *Hox*

Les gènes *Hox* sont regroupés dans des complexes génomiques de quelques kilobases et leur expression est régulée par des facteurs de transcription, des enhancers, la topologie des chromosomes et des marques épigénétiques. La coordination entre ces différents mécanismes de régulation est nécessaire pour permettre aux gènes *Hox* de s'exprimer au bon endroit et au

bon moment. Dans cette partie, nous allons décrire les différents mécanismes connus à ce jour qui participent à la régulation des gènes *Hox*.

a. Rôle des complexes polycomb et trithorax dans les complexes *Hox*

Avant, la gastrulation, les complexes *Hox* sont compactés et transcriptionnellement inactifs (Young et al., 2009). La répression des gènes *Hox* est assurée par des protéines de la famille des polycombs (PcG) (Eskeland et al., 2010; Schuettengruber and Cavalli, 2009). Les PcG s'associent pour former des complexes PRC1 et PRC2. Ces complexes induisent des modifications post-traductionnelles des histones affectant ainsi la structure chromatinienne. Les protéines Ezh1/2 appartenant à PRC2 initient ces modifications en ajoutant un groupement méthyle à la lysine 27 de l'histone H3 (H3K27) (Margueron et al., 2008). Puis des boucles de rétroaction positive gouvernées par des sous-unités Eed et Suz12 permettent à Ezh1/2 de continuer l'ajout de groupements méthyls jusqu'à la tri-méthylation (H3K27me3). Les marques H3K27me3 permettent le recrutement de PRC1 grâce à la protéine Cbx. Puis lorsque PRC1 est lié, les protéines Ring1A/B ubiquitinent la lysine 119 de l'histone H2A (H2K119ub) (Aranda et al., 2015). L'activité combinatoire des complexes PRCs conduit à la compaction de la chromatine. Par conséquent, l'ARN polymérase II ne peut plus être recruté au niveau des promoteurs et il y a donc une absence d'expression des gènes.

Chez la drosophile, les mutations de protéines appartenant au groupe des PcG (*Pc*, *esc*, *Scm*, *Pcl*...) se caractérisent par l'expression ectopique de gènes *Antennapedia* le long de l'axe antéro-postérieur (Jürgens, 1985; Simon et al., 1992). Ces différentes mutations provoquent toutes un développement segmentaire anormal ce qui est cohérent avec le fait que ces protéines du groupe PcG agissent via une voie unique ou un complexe moléculaire. Dans les cellules souches embryonnaires murines (CSEm), la mutation de *Eed* (composant de PRC2) provoque la perte des marques H3K27me3 dans les complexes *Hox* qui deviennent dé-réprimés (Boyer et al., 2006). Chez les embryons de souris, la perte fonctionnelle des protéines Bmi, M33, Mel18 ou Ring1A qui appartiennent toutes au complexe PRC1 entraîne des transformations homéotiques du squelette axial et la dérégulation des gènes *Hox* qui sont exprimés dans des régions plus antérieures que leur domaine d'expression normal (Akasaka et al., 1996; Core et al., 1997; Lorente et al., 2000; Van Der Lugt et al., 1994). L'inactivation simultanée de Bmi1 et Mel18 est caractérisée par des transformations homéotiques encore plus sévères (Akasaka et al., 2001) ce qui suggère qu'il y a une redondance fonctionnelle entre ces différents composants

du complexe PRC1. Les mutations simultanées de PRC1 et PRC2 sont quant à elles létales puisque les embryons murins meurent avant la gastrulation.

A l'inverse des PcG, le complexe Trithorax (TrxG) neutralise la répression en ajoutant des groupements méthylés (H3K4me3) (Byrd and Shearn, 2003) ou acétylés (H3K27ac) (Tie et al., 2014) qui correspondent à des marques transcriptionnellement permissives. Chez la drosophile, la mutation de *Trx*, qui code normalement pour une protéine associée au TrxG, entraîne des transformations homéotiques caractérisées par une extension des segments antérieurs (Breen and Harte, 1993). Ce phénotype est similaire à ceux observés lors de la perte de fonction des gènes *Hox*. Les conséquences phénotypiques causées par la mutation des protéines TrxG sont donc opposées à celles induites par la perte de fonction de PcG. Ces résultats suggèrent que ces deux complexes agissent en synergie pour contrôler l'expression des gènes *Hox*.

Les mécanismes de recrutement de PcG et TrxG dans les complexes *Hox* sont encore méconnus. Les complexes PRC1/2 répriment l'expression des gènes en se liant à des séquences 'polycomb response elements' (PREs). Bien que ces domaines aient été identifiés chez la drosophile, aucun motif consensus pouvant lier PcG n'est connu chez les Vertébrés. Pour le moment, seulement un élément semblant partager des caractéristiques similaires aux PREs est identifié entre le gène *HoxD11* et *HoxD12* chez l'homme (Woo et al., 2010). En effet, il est démontré que ce site peut lier PRC2 et réprimer l'expression d'un gène placé en aval. Cependant la délétion de cette séquence dans le complexe *HoxD* de la souris n'a pas d'effet détectable sur l'activation des gènes *HoxD11* et *HoxD12* pendant le développement embryonnaire. D'autres études ont alors suggéré que le recrutement des PcG et TrxG dépend de l'activité de lncRNA. Le lncRNA *HOTAIR* qui se trouve dans le complexe *HoxC* participe à la répression de *HoxD10* et *HoxD11* en interagissant avec PRC2 (Rinn et al., 2007). Chez la souris, la mutation de *HOTAIR* est caractérisée par une expansion antérieure du domaine d'expression de *HoxD10* et *HoxD11* dans le tronc (Li et al., 2013).

Etant donné que les complexes PcG et TrxG jouent un rôle crucial dans la régulation de l'expression des gènes *Hox*, l'évolution des marques d'histones associées à ces complexes protéiques a été étudiée au cours de l'élongation axiale. Chez la souris, une étude révèle que dans la région progénitrice caudale, il y a une perte progressive de H3K27me3 (une marque associée à une absence de transcription) dans le complexe *HoxD* qui est concomitante à un gain de H3K4me3 (une marque associée à une transcription active) (Soshnikova and Duboule, 2009). Ce changement épigénétique est progressif depuis les régions 3' vers les régions 5' des

complexes et est corrélé avec l'activation séquentielle des gènes *Hox*. Toutefois, ces observations ont été réalisées sur des tissus embryonnaires disséqués et reflètent donc probablement un état épigénétique moyen de types de cellules éventuellement distincts qui peuvent avoir été exposés à des signaux différents *in vivo*.

Contrairement à la dissection embryonnaire *in vivo*, la différenciation *in vitro* permet de contrôler étroitement la population cellulaire étudiée. En utilisant des progéniteurs spinaux dérivés de CSEm, Mazzoni et al. ont démontré que le remodelage saltatoire (par opposition au remodelage progressif) de la chromatine des complexes *Hox* était induit par l'exposition à des signaux extrinsèques (Mazzoni et al., 2013). Ils ont montré que RA induit rapidement la disparition des marques répressives des gènes *Hox1* à *Hox5* tandis que Fgf induit un changement d'état de la chromatine jusqu'à *Hox9*. Ces résultats démontrent que le remodelage de la chromatine des complexes *Hox* se produit à l'échelle du domaine et peut être induit par des séquences d'exposition à des facteurs extrinsèques *in vivo*. Cependant, ils ont observé que les gènes *Hox* situés dans les zones d'ouverture saltatoire restaient exprimés de manière progressive et colinéaire, ce qui suggère que l'ouverture de la chromatine n'est peut-être pas le seul mécanisme régulant le rythme des gènes *Hox*.

b. Les facteurs de transcription et les enhancers au sein des complexes *Hox*

Les enhancers sont des séquences génomiques caractérisées par des motifs consensus de 6 à 12 paires de bases qui peuvent lier des protéines pour stimuler ou réprimer la transcription des gènes. Ces séquences *cis*-régulatrices peuvent se trouver directement en amont de gènes ou bien à plusieurs mégabases de leurs cibles. Parfois même, les enhancers se trouvent dans des introns de gènes qu'ils ne vont pas réguler ou bien dans des séquences codantes. La position des enhancers par rapport aux gènes qu'ils régulent est donc très variable indiquant une grande flexibilité fonctionnelle qui est propre à ces séquences.

Des études génétiques ont démontré que plusieurs éléments *cis*-régulateurs RAREs flanquant les gènes paralogues *Hox1-5* étaient activés par RA (Langston et al., 1997; Langston and Gudas, 1992; Pöpperl and Featherstone, 1993). Lors d'une exposition à du RA, le recrutement de RAR est séquentiel et a lieu dans un premier temps sur les séquences RARE en 3' puis sur ceux plus en 5' (Gavalas et al., 1998). De plus, la délétion d'éléments RAREs a pour conséquence d'atténuer la réponse transcriptionnelle des gènes *Hox* antérieurs en réponse à une stimulation RA *in utero* (**Fig. 12A**) (Dupé et al., 1997). Ceci indique que le niveau d'expression

des gènes *Hox* n'est pas uniquement défini par leur séquence promotrice mais aussi par leur proximité avec des éléments *cis*-régulateurs.

D'autres expériences ont permis d'identifier des enhancers contrôlant la temporalité d'induction des gènes *Hox*. Par exemple, des expériences de transgénèse utilisant des gènes rapporteurs pour tester des éléments régulateurs ont isolé une séquence de 200 paires de bases en amont du promoteur de *HoxC8* qui récapitule la temporalité d'induction de ce gène (Shashikant et al., 1995). La délétion de cette région a pour conséquence de provoquer un délai dans l'expression de *HoxC8* (**Fig. 12C**). Si le patron d'expression de *HoxC8* est ensuite rétabli, cette induction plus tardive provoque néanmoins des défauts d'organisation du squelette axial (Juan and Ruddle, 2003; Shashikant et al., 1998). Des expériences de transgénèse chez la souris ont également révélé l'existence d'un enhancer entre *HoxD10* et *HoxD11* (Gérard et al., 1993). La mutation ciblée de cet enhancer a pour conséquence un déplacement caudal des domaines d'expression de *HoxD10* et *HoxD11* au 9^{ème} jour de gestation, et un déplacement postérieur de la jonction lombo-sacrée des vertèbres (**Fig. 12D**) (Zakany, 1997).

La difficulté à identifier et prédire les fonctions des enhancers réside en grande partie dans notre compréhension limitée des facteurs de transcription auxquels ils sont associés (Zeitlinger, 2020). Cependant, grâce à des techniques d'immuno-précipitation et séquençage, il est possible de déterminer les sites de liaison de facteurs le long de la chromatine. Ainsi en utilisant des cellules souches épiblastiques de souris, des études ont montré des domaines de liaison de Cdx2 entre *Hox1-9* (**Fig. 12B**) (Amin et al., 2016; Neijts et al., 2017). La présence de ces facteurs le long des complexes *Hox* est essentielle pour contrôler leur activation, puisque les cellules provenant de souris triple mutantes *Cdx1/2/3* ne peuvent pas induire l'expression du gène *Hox1-9* (Amin et al., 2016). L'étude de l'accessibilité à la chromatine chez ces mutants indique que les facteurs Cdx jouent un rôle important dans l'ouverture de la chromatine autour des gènes *HOX* participant à la spécification du tronc (Neijts et al., 2017). De plus, Cdx2 coopère avec la voie Fgf puisque la combinaison de Fgf et Cdx2 est nécessaire pour induire les gènes *HOX6-9* alors que Cdx2 seul n'est pas suffisant (Mazzoni et al., 2013).

En résumé, des enhancers et facteurs de transcription ont été trouvés au sein des complexes *Hox*. Les mutations de ces enhancers impactent l'expression des gènes *Hox* en générant souvent un délai d'activation, un déplacement des domaines d'expression et/ou une réduction de leur niveau transcriptionnel. Cependant, les gènes *Hox* régulés sont toujours exprimés. Cela suggère que les enhancers jouent un rôle crucial en induisant l'expression des

gènes *Hox* au bon moment, ce qui est essentiel pour déterminer l'identité positionnelle des tissus embryonnaires émergents.

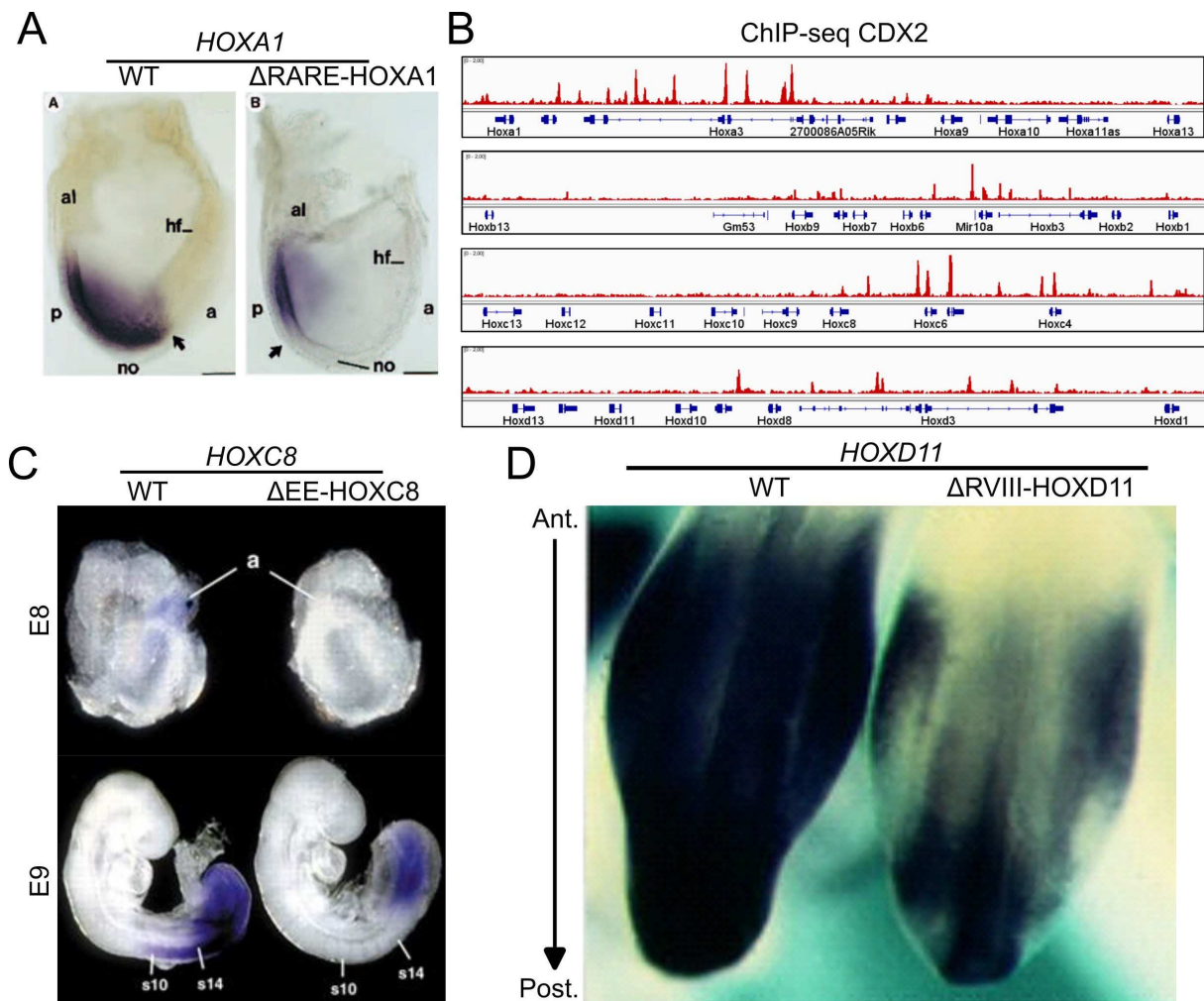


Figure 12. Rôle des facteurs de transcription et enhancers dans la régulation des gènes *Hox*. (a) Hybridation *in situ* de *HoxA1* dans embryon de souris E7.5 control et muté (délétion du site RARE en amont de *HoxA1*). Le domaine d'expression de *HoxA1* s'étend sur des domaines plus antérieurs par rapport au mutant. (b) Distribution de Cdx2 le long des complexes *HoxA/B/C/D* dans des cellules souches de l'épiblaste induites. (c) Hybridation *in situ* de *HoxC8* dans embryon de souris E9 control et muté (délétion de '*HoxC8* early enhancer'). Chez le mutant, l'expression de *HoxC8* est limitée au mésoderme paraxial. (d) Hybridation *in situ* de *HoxD11* dans embryon de souris control et muté (délétion de '*region VIII*' en amont de *HoxD11*). Les embryons mutants présentent une réduction des transcrits *HoxD10* dans le bourgeon de la queue. (Adapté d'après Amin et al., 2016; Dupé et al., 1997; Juan and Ruddle, 2003; Zakany, 1997)

c. Les enhancers en dehors des complexes *Hox*

Par ailleurs, des enhancers se trouvant en dehors des complexes peuvent également jouer un rôle clé dans la régulation des gènes *Hox*. C'est par exemple le cas durant la formation des doigts où *HoxD10-13* sont transcrits avec une efficacité progressivement plus faible, de sorte que *HoxD13* est exprimé à des niveaux élevés dans tous les doigts, tandis que l'expression de *HoxD12/D11/D10* est plus faible (**Fig. 13A**) (Montavon et al., 2008). Cette régulation colinéaire

inversée des gènes *HoxD10-13* lors de la formation des doigts dépend de l'existence d'un enhancer distal situé en amont de l'extrémité 5' du complexe *HoxD* (Spitz et al., 2003). Cet enhancer induit l'expression des gènes *Hox* en fonction d'une efficacité qui diminue à mesure que le nombre de gènes intermédiaires entre lui et le gène qu'il régule augmente, ce qui explique pourquoi *HoxD13* est plus fortement exprimé dans les membres que *HoxD10* (**Fig. 13B**) (Kmita et al., 2002).

Des enhancers distaux Ades1-6 qui participent à la formation de l'axe antéro-postérieur ont également été trouvés entre les gènes *HoxA1* et *Skap2* à environ 300kb du complexe *HoxA* (Zhang et al., 2013). La suppression de certains de ces enhancers réduit, mais n'élimine pas l'expression de *HoxA* en réponse à RA ou à la voie Wnt (Cao et al., 2017; Neijts et al., 2016). Récemment, le rôle de ces enhancers distaux et le contexte génomique dans lequel se trouve les complexes *Hox* a été étudié en synthétisant le complexe *HoxA* avec (enhancers+SynHoxA) ou sans ces enhancers distaux (SynHoxA) (Pingley et al., 2022). Ces séquences ont ensuite été intégrées dans le chromosome X de cellules embryonnaires de souris. Ces complexes ectopiques sont capables de reconstituer les mêmes événements transcriptionnels que le complexe endogène qui se trouve sur le chromosome 7. De plus, la séquence enhancers+SynHoxA induit des niveaux d'expression de *HoxA1-5* plus élevés que la séquence SynHoxA, en particulier aux temps les plus précoces après l'ajout de RA. Ces résultats indiquent donc que tout comme les enhancers se trouvant au sein des complexes *Hox*, les enhancers distaux sont importants pour contrôler le niveau d'expression des gènes *Hox*. Cependant, pour que ces enhancers distaux puissent réguler les gènes *Hox*, des changements de conformation de la chromatine sont nécessaires afin de les rapprocher de leurs cibles.

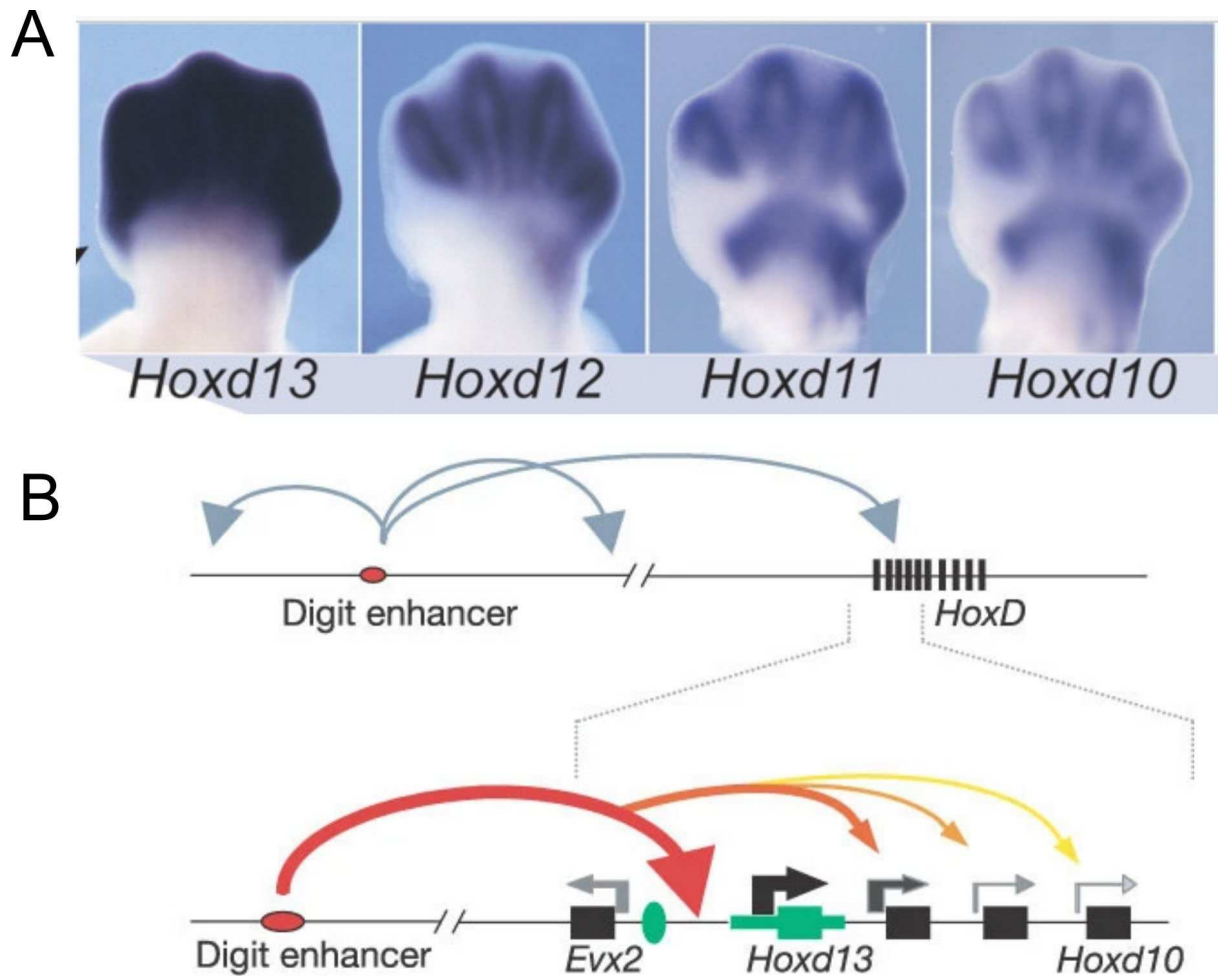


Figure 13. Rôle des enhancers distaux dans la régulation des gènes *Hox*. (a) Hybridation *in situ* de *HoxD10-13* dans membres de souris embryon de souris à E12.5. (b) Illustration montrant le rôle de « digit enhancer » dans la régulation des gènes *HoxD10-13*. (Adapté d'après Kmita et al., 2002; Montavon et al., 2008)

d. Topologie et dynamique génomique (TAD)

Dans les complexes *Hox*, des réarrangements de la chromatine ont été observés pour la première fois en visualisant la localisation de différents *Hox* à l'aide de l'hybridation fluorescente *in situ*. Lorsque les gènes *Hox* ne sont pas exprimés, les gènes *HoxB1* et *HoxB9*, situés à 90 kb l'un de l'autre sur le cluster *HoxB*, sont co-localisés dans les noyaux des cellules souches embryonnaires. Après ajout de RA, la distance entre ces deux loci augmente en même temps que l'activation transcriptionnelle de *HoxB1* puis *HoxB9* (Chambeyron and Bickmore, 2004). Cependant, la résolution des hybridations *in situ* fluorescentes ne permet pas de discriminer des loci séparés par de courtes distances génomiques. C'est pourquoi de nouvelles méthodes de séquençage ont émergé afin de capturer la conformation de la chromatine. Par exemple, la technique '3C-sequencing' permet de quantifier le nombre d'interactions entre des loci génomiques et de cartographier en tri-dimension (3D) la chromatine. Ces nouvelles

méthodes ont donc permis de révéler de nouvelles modulations chromatinienne associées avec l'expression des gènes *Hox*.

Lors de la formation des membres, *HoxD* est localisé entre deux grands domaines d'association topologique (TAD) de la chromatine qui s'étendent sur plusieurs centaines de kilobases en dehors des complexes (Andrey et al., 2013). Alors que le domaine télomérique (T-DOM) joue un rôle dans la phase précoce de formation des membres, le domaine centromérique (C-DOM), qui comprend l'enhancer distal induisant l'expression des gènes *HoxD10-13* (Spitz et al., 2003), intervient dans la phase finale de la formation des membres. Le réarrangement topographique de la chromatine permet donc de rapprocher les enhanceurs distaux impliqués dans la régulation de l'expression *Hox*.

Par ailleurs, il a aussi été démontré que la configuration 3D de la chromatine est connectée à l'accessibilité chromatinienne. Dans les CSEm, les gènes *Hox* qui ne sont pas exprimés sont organisés dans un unique TAD recouvert par des marques répressives H3K27me3 (Ferraiuolo et al., 2010; Montavon et al., 2011; Noordermeer et al., 2011). La topologie de la chromatine a été étudiée dans différentes régions antéro-postérieures de la souris (Noordermeer et al., 2011). Ces données montrent que les complexes *Hox* sont divisés en deux TAD distincts. Dans les cellules antérieures, les gènes transcrits comme *HoxD4* se trouvent dans un TAD caractérisé par des marques H3K4me3, une modification d'histone associée à des régions transcriptionnellement actives. A l'inverse dans les mêmes cellules, *HoxD13*, qui n'est pas exprimé, est associé à un autre TAD riche en H3K27me3. Cette marque correspond à une modification d'histone associée à des régions transcriptionnellement inactives. Dans les cellules postérieures, ces deux compartiments sont toujours présents mais la limite se trouvant entre eux n'est plus au même endroit sur les complexes *Hox*. Cette transition d'un domaine 3D inactif à un domaine 3D actif a particulièrement été illustré pour *HoxD9* qui est fortement exprimé dans les cellules du tronc postérieur mais faiblement dans les régions antérieures. Cette étude a ensuite été confortée dans une analyse temporelle chez la souris à partir de cellules du mésoderme pré-somitique à E8.5 et du bourgeon caudal à E9.5 et E10.5 (**Fig. 14A**) (Noordermeer et al., 2014). En effet, ces données indiquent qu'il y a un changement de conformation topologique progressif illustré une nouvelle fois par *HoxD9* qui est dans un compartiment chromatinien inactif à E8.5, puis dans un compartiment chromatinien actif à E9.5. De ces observations est apparu le modèle de Rosaire qui indique que les gènes *Hox* sont séquentiellement activés par un glissement linéaire de la chromatine allant d'un TAD transcriptionnellement réprimé, vers un TAD actif.

La formation des boucles intrachromosomiques est dépendante de l'association de facteurs CCCTC-binding (CTCFs) à l'ADN qui permettent alors le recrutement du complexe cohésine. Les CTCFs sont des protéines ubiquitaires qui agissent comme des facteurs isolateurs assurant l'organisation de la chromatine. Il existe un enrichissement particulièrement important de sites CTCF dans les complexes *Hox* (**Fig. 14B**). La cartographie de l'emplacement et de l'orientation des sites CTCFs au sein des complexes *Hox* a révélé trois sous-domaines principaux : un domaine antérieur où il n'y a pas de CTCFs, un domaine central où les CTCFs sont principalement orientés vers l'extrémité 3', et un domaine postérieur où les CTCFs sont orientés vers l'extrémité 5' (Amândio et al., 2021). Or les boucles chromatiniennes se forment de préférence entre des sites de liaison CTCFs orientés de manière convergente (Guo et al., 2015). La position et l'orientation de ces sites semble donc jouer un rôle dans l'activation séquentielle des gènes *Hox*.

En effet, durant la différenciation de CSEm en motoneurones, la stimulation par du RA active l'expression de *HoxA1-6* tandis que les *HoxA7-13* restent réprimés. Cependant, la mutation du site CTCF en amont de *HoxA7* conduit à l'expression des gènes *HOX* jusqu'à *HoxA9* en réponse au RA (Narendra et al., 2015). Ces changements transcriptionnels s'accompagnent également d'une réduction des marques H3K27me3 au niveau des gènes *HoxA7/A8/A9*. Le rôle des CTCFs a aussi été analysé durant la formation de l'axe antéro-postérieur. Chez la souris, la mutation du site CTCF en aval de *HoxD9* provoque une antériorisation très forte du domaine d'expression de *HoxD9* (**Fig. 14C**). Dans ce mutant le domaine d'expression de *HoxD9* devient alors similaire à des gènes plus antérieurs comme *HoxD1* et *HoxD3* (Amândio et al., 2021; Narendra et al., 2016). Dans une étude plus récente, des CSEm ont été différenciés en gastruloïde, un modèle *in vitro* 3D qui mime l'élongation axial, l'activation séquentielle des gènes *Hox* et la formation de patrons spatiaux le long de l'axe antéro-postérieur (voir aussi plus bas). Chez ces gastruloïdes où les sites CTCFs entre *HoxD8* et *HoxD13* ont été mutés, l'expression des gènes postérieurs *HoxD10/D11/D13* est fortement augmentée et il n'y a plus de séparation topographique de la chromatine entre la partie 3' et la partie 5' du complexe *HoxD* (Rekaik et al., 2023). Ces résultats indiquent donc que les protéines CTCF sont essentielles pour contrôler le rythme d'activation des gènes *Hox* en modulant l'organisation chromatinienne tridimensionnelle.

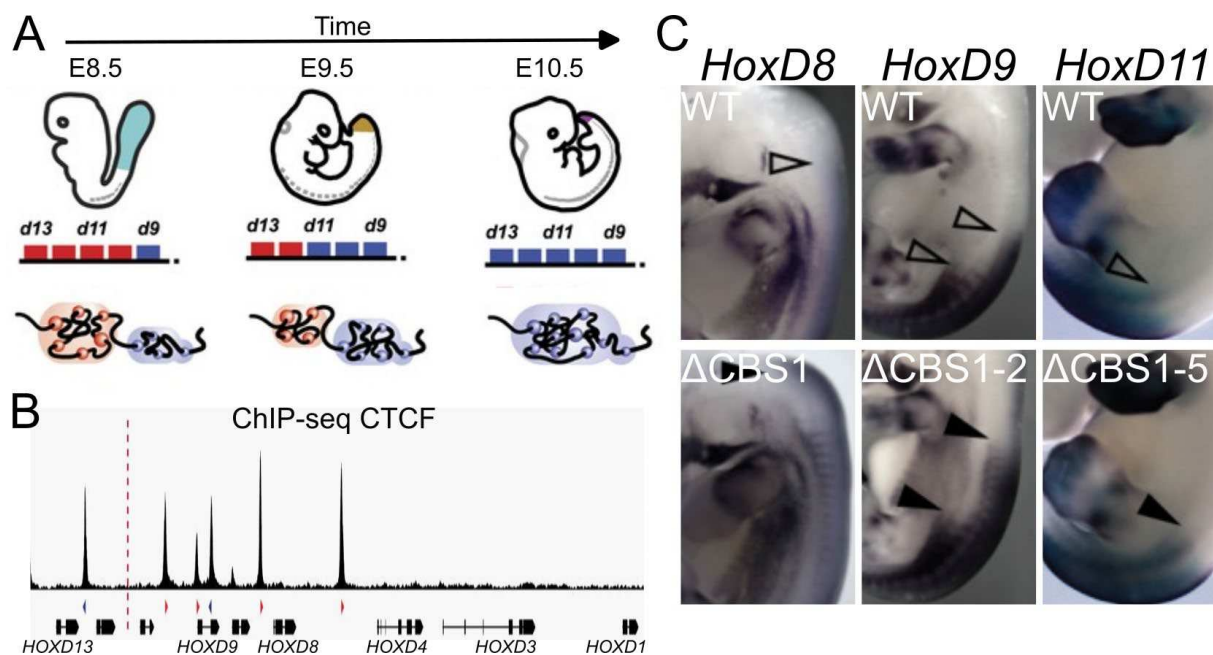


Figure 14. Organisation 3D des complexes *Hox* au cours de l'élongation axiale. (a) Modèle illustrant l'organisation topographique des complexes *Hox* durant la formation de l'axe. Les gènes transcriptionnellement inactifs sont représentés en rouge et les gènes actifs en bleu. Dans les complexes *Hox*, l'activation progressive des gènes s'accompagne d'une transition d'un domaine transcriptionnellement inactif marqué par H3K27me3 (chromatine rouge) à un autre domaine transcriptionnellement actif marqué par H3K4me3 (chromatine bleue). (b) Distribution de CTCF le long du complexe *HoxD* dans des cellules du tronc de souris à E9.5. Les flèches rouges indiquent des CTCF orientés vers 3' et les flèches bleues des CTCF orientés vers 5'. Les boucles chromatiniennes se forment lorsque deux CTCF ont des orientations opposées. (c) Hybridation *in situ* de *HoxD8/D9/D11* dans des embryons control ou mutés avec des délétions des sites de liaison CTCF en amont de *HoxD4* (CBS1), *HoxD8* (CBS2) et *HoxD10* (CBS5). Une antériorité de l'expression dans le tronc est observable pour le gène positionné immédiatement en 5' du site de liaison CTCF supprimé. (Adapté d'après Amândio et al., 2021; Noordermeer et al., 2014)

6. Régulation des gènes *Hox* par les facteurs extrinsèques

L'activation séquentielle des gènes *Hox* de 3' vers 5' corrèle avec l'expression séquentielle de signaux environnementaux contrôlant l'élongation axiale dans la zone proliférative caudale. Les voies Wnt, RA, Fgf et Gdf qui régulent la morphogenèse de l'axe antéro-postérieur jouent aussi un rôle primordial dans la régulation de l'expression des gènes *Hox*.

a. La voie Wnt régule l'induction des gènes *Hox* situés dans la partie 3' des complexes

Dans les progéniteurs axiaux localisés dans la région postérieure de la ligne primitive, l'expression des gènes *Hox* en 3' des complexes est initiée en réponse à l'activation de la voie Wnt. En effet, l'expression de *Wnt3* dans l'épiblaste postérieur précède la formation de la ligne primitive et l'absence de *Wnt3* bloque l'expression des gènes *Hox* (Fig. 15B) (Liu et al., 1999). Par ailleurs, le rôle de Wnt dans l'induction de l'expression des gènes *Hox* a été confirmé dans

les embryons pré-gastrulation. L'exposition de ces embryons à un agoniste de la voie Wnt induit précocement l'expression des gènes *HoxA1* et *HoxB1* (**Fig. 15A**) (Neijts et al., 2016).

Après cette étape d'initiation, l'expansion rostrale du domaine *Wnt3* lors de la gastrulation est suivie d'une expansion similaire des transcrits *Hox* localisés en 3' (Deschamps and Duboule, 2017). Les gènes *Hox* situés un peu plus en 5' sur le complexe sont ensuite induits en suivant la même dynamique que les *Hox* en 3' jusqu'aux paralogues 4. A noter que l'expansion des domaines d'expression vers les régions antérieures est indépendante de la migration cellulaire (Deschamps and Wijgerde, 1993).

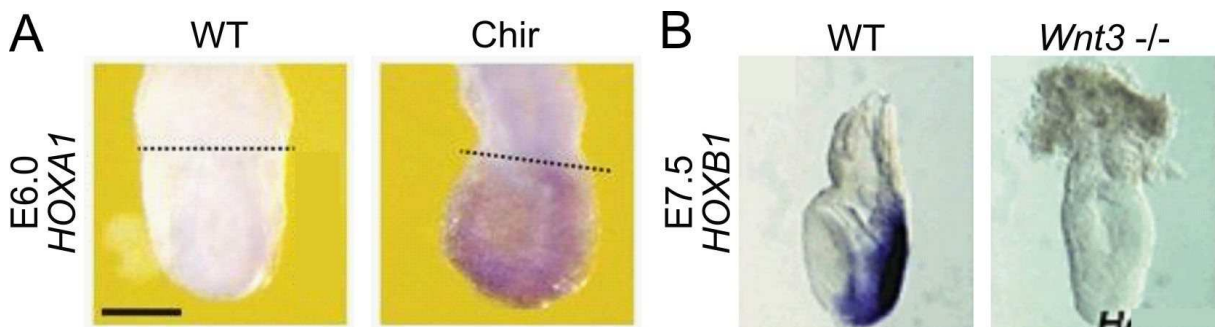


Figure 15. L'activation transcriptionnelle des complexes *Hox* est initiée par la voie Wnt. (a) Embryon de souris à E6 stimulé pendant 10h par du Chir (activateur de la voie Wnt). Il y a une induction précoce de l'expression de *HoxA1* lorsque la voie Wnt est activée à ce stade. (b) Hybridation *in situ* de *HoxB1* dans embryon de souris de type sauvage et mutant *Wnt3*^{-/-}. *Hoxb1* est exprimé dans la région postérieure de l'embryon control, mais absent chez le mutant. (Adapté d'après Liu et al., 1999; Neijts et al., 2016)

b. L'acide rétinoïque module l'expression des gènes *Hox1-5*

Le RA a également été identifié comme jouant un rôle dans la régulation de l'expression des gènes *Hox* dans le tronc cérébral et la partie la plus rostrale de la moelle épinière. En effet, l'exposition d'explants de tube neural de poulet à de fortes concentrations de RA induit l'augmentation de l'expression *HoxC5* (Liu et al., 2001). Au niveau cervical, l'expression de *HoxC5* peut être bloquée par l'utilisation d'antagonistes aux RARs, suggérant un rôle du RA dans la spécification des motoneurones cervicaux et brachiaux antérieurs (Liu et al., 2001). De plus des explants provenant de la moelle thoracique qui donnent normalement des motoneurones *HoxC9*, une fois exposés au RA produisent des motoneurones exprimant *HoxC6* ce qui entraîne la perte de *HoxC9*.

Malgré la capacité de cette voie à induire les gènes *Hox*, le rôle du RA endogène dans l'activation initiale des complexes *Hox* reste incertain. En effet, il n'y a pas d'activité du RA dans la partie caudale de la ligne primitive où l'expression des gènes *Hox* est initiée. De plus, malgré l'absence de RA chez les souris mutantes *Raldh2* (enzyme synthétisant le RA), l'expression de *HoxB1* est maintenue (Roelen et al., 2002).

En résumé, même si le RA semble pouvoir contrôler l'expression des gènes *Hox* dans les parties antérieures et que le RA exogène peut moduler l'expression des gènes *Hox* dans des modèles en culture (Simeone et al., 1991), le RA ne semble pas contribuer directement à l'initiation de l'expression des *Hox* dans les progéniteurs axiaux.

c. Fgf coopèrent avec les facteurs Cdx pour activer l'expression des gènes *Hox* se trouvant dans la partie centrale des complexes

La voie Fgf joue également un rôle dans la régulation de l'expression de certains gènes *Hox*. Par exemple, la surexpression de *Fgf4* ou *Fgf8* par électroporation dans les segments antérieurs du tube neural chez le poulet induit une expression ectopique, dans ces régions, de gènes se trouvant dans la partie centrale des complexes *Hox* (**Fig. 16A/B/C**) (e.g. *HoxC6/A9/B9/C9*) (Bel-Vialar et al., 2002; Dasen et al., 2003). Pour caractériser l'influence de ces signaux sur l'expression des gènes *Hox*, des explants de la partie antérieure du nœud primitif chez le poulet au stade HH8, composés de cellules normalement destinées à donner naissance aux parties antérieures du domaine cervical, ont été exposés à différentes concentrations de Fgf2. Lorsque ces explants sont exposés à des faibles concentrations de Fgf2 (5 ng/ml), ils donnent naissance à des motoneurones localisés dans les régions brachiales (*HoxC6+*), alors que ceux exposés à des concentrations plus fortes (625 ng/ml) donnent naissance à des motoneurones se trouvant normalement dans les segments lombaires (*HoxC10+*) (Liu et al., 2001). Ce résultat suggère un effet concentration dépendant du FGF sur l'expression des gènes *Hox*. Cependant, ces expériences ont été réalisées sur des explants de cellules de plaque neurale et ne permettent pas de conclure sur le rôle de Fgf dans l'initiation de l'expression *Hox* dans les progéniteurs axiaux.

L'activité de la voie Fgf sur les gènes *Hox* est en partie médiée par les facteurs de transcription de la famille des *Cdx*. En effet, lorsqu'un embryon est électroporé avec une forme dominante négative de *Cdx* et cultivé pendant une nuit en présence de Fgf2, *HoxB9* n'est plus exprimé dans les parties antérieures de l'axe (**Fig. 16D**) (Bel-Vialar et al., 2002). A l'inverse, l'expression ectopique d'une forme activatrice de *Cdx* dans le cerveau postérieur a pour conséquence d'étendre l'expression de *HoxB4*. Une autre étude a caractérisé le lien entre Fgf et *Cdx* dans des progéniteurs dérivés de cellules souches se différenciant en motoneurones *in vitro*. Dans ce contexte, la surexpression de *Cdx2* en présence de Fgf permet l'induction *HoxC6/C8/C9* et la spécification de motoneurones thoraciques exprimant *HoxC9* (Mazzoni et

al., 2013). En revanche, Fgf seul ou Cdx2 seul ne permettent pas la spécification des motoneurons HoxC9. L'activation de *Hox6-9* dépend donc de la coopération entre la voie Fgf et les facteurs Cdx. Toutefois, les mécanismes de l'induction séquentielle des gènes *HOXC6* à *HOXC9* restent encore inconnus.

En résumé, la signalisation Fgf joue un double rôle au cours de la formation de l'axe antéro-postérieur : dans un premier temps, les voies Fgf/Wnt induisent l'expression de *Cdx2*, puis la voie Fgf pourrait agir en synergie avec *Cdx2* pour activer l'expression des gènes *Hox* se trouvant dans la partie centrale des complexes (*Hox6-9*). Toutefois, ces études ont principalement été réalisées par électroporation, explant et différenciation *in vitro* sans cibler spécifiquement les progéniteurs axiaux, mais plutôt des cellules plus différenciées du PNT ou du tube neural (e.g. progéniteurs neuraux). Cela implique donc que le rôle du Fgf dans l'initiation de l'expression des gènes *Hox* dans les progéniteurs axiaux n'a pas été précisément étudié en détail. De plus, jusqu'à présent, aucun composant de la voie Fgf en aval des récepteurs n'a été démontré comme étant impliqué dans la régulation des gènes *Hox*.

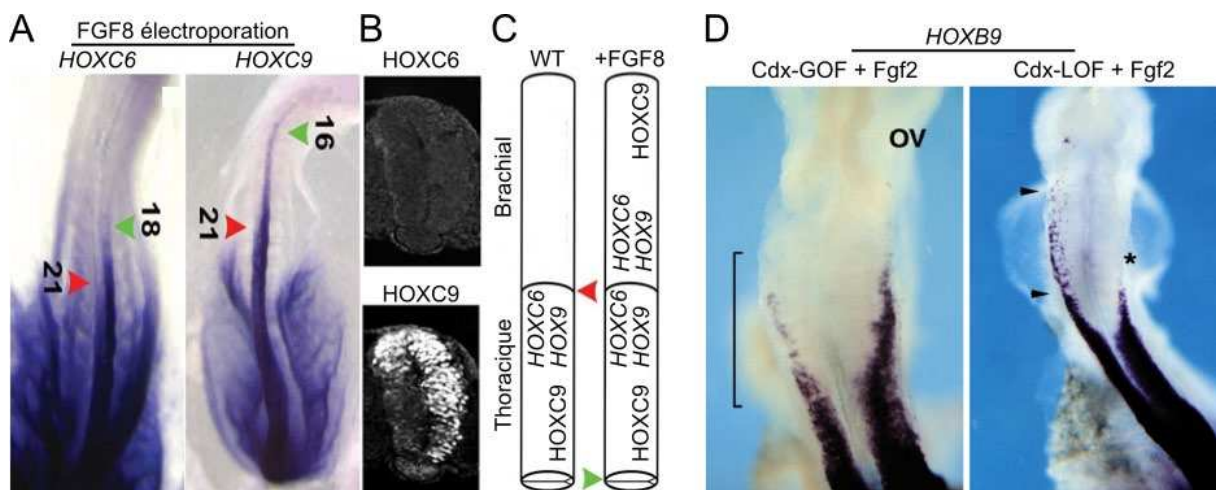


Figure 16. L'initiation des gènes *Hox* centraux dépend de la voie Fgf-Cdx (a) Hybridation *in situ* de *HoxC6* et *HoxC9* 8h après électroporation de Fgf8 chez l'embryon de poulet HH15. La pointe de flèche rouge indique la limite d'expression rostrale normale et la pointe de flèche verte indique la limite rostrale après électroporation de Fgf8. L'expression ectopique de Fgf8 conduit à une expansion rostrale des domaines d'expression de *HoxC6* et *HoxC9*. (b) Immunohistochimie révélant l'expression ectopique de *HoxC9* et la perte de *HoxC6* dans le domaine brachial après électroporation de *Fgf8*. (c) Schéma de l'ARNm et des protéines *Hox6* et *Hox9* dans le tube neural brachial et thoracique chez les embryons normaux et exposés au Fgf8. (d) Hybridation *in situ* de *HoxB9* chez embryons de poulet électroporés unilatéralement sur le côté gauche du tube neural avec une forme dominante active ou négative de Cdx. Les embryons ont ensuite été cultivés toute la nuit en présence de Fgf2 exogène. L'induction de *HoxB9* médiée par le Fgf est réduite chez les embryons Cdx-négatif. A l'inverse, il y a une expansion antérieure de *HoxB9* chez les embryons Cdx-positif. (Adapté d'après Bel-Vialar et al., 2002; Dasen et al., 2003)

d. La voie Gdf participe à l'induction des gènes *Hox* situés dans la partie 5' des complexes

L'induction des gènes *Hox10-13* exprimés dans les régions lombaires et sacrées dépend de la voie de signalisation Gdf. En effet, les souris mutantes *Gdf11* sont caractérisées par des

transformations homéotiques avec une augmentation du nombre de segments thoraciques et un décalage des membres postérieurs de plusieurs vertèbres (McPherron et al., 1999; Suh et al., 2019). Ces anomalies dans le plan d'organisation sont liées à un décalage vers la partie caudale des domaines d'expression des protéines Hox notamment une expansion postérieure de HoxC6 et HoxC8 et un déplacement postérieur de la limite rostrale des domaines d'expression de HoxC10 et HoxC11 (Liu, 2006; McPherron et al., 1999).

D'autres expériences ont induit l'expression ectopique de *Gdf11* dans le tube neural de poulet, conduisant à une antériorisation des domaines d'expression des gènes *HoxC6-C10* (**Fig. 17A/B/C**) (Liu, 2006). A l'inverse, l'électroporation de la follistatine, un antagoniste des voies TGF- β /GDF, provoque le déplacement caudal d'un à deux segments des domaines d'expression des facteurs HoxC6-C10 (Liu, 2006).

La voie de signalisation par laquelle *Gdf11* agit pour induire l'expression des gènes *Hox* a fait l'objet d'études approfondies. Chez les souris mutantes *Alk5*, il n'y a pas d'expression de *HoxC10* dans le bourgeon caudal (Andersson et al., 2006). Lorsque des explants du bourgeon caudal, prélevés à E10.5, sont cultivés dans un milieu contenant du SB431542 (inhibiteur de *Alk5*), les facteurs HoxB13/C13 ne sont plus exprimés (Aires et al., 2019). En revanche, l'expression ectopique d'une forme constitutivement active du récepteur *Alk5* dans la ligne primitive et épiblaste postérieur conduit à l'expression des gènes *HoxA9* et *HoxC10* dans des régions plus antérieures (Jurberg et al., 2013). De même, l'électroporation d'une construction exprimant une forme constitutivement active de *Smad2* (*Smad2-2E*) dans le tube neural de poulet induit l'extension des domaines d'expression des facteurs HoxC les plus caudaux (*HoxC9* et *HoxC10*) aux dépens des facteurs HoxC rostraux (*HoxC6* et *HoxC8*) (Liu, 2006). Ces différents phénotypes observés suite au gain de fonction de *Alk5* et *Smad2* sont cohérents avec les expériences précédentes de gain de fonction de *Gdf11* (Liu, 2006).

Le rôle direct de la voie de signalisation *Smad* dans la régulation des gènes *Hox* a également été mis en évidence par la présence de motifs de liaison pour les facteurs de transcription *Smad* dans les enhancers en aval de *HoxC8* et *HoxD11* (Gaunt, 2017; Gaunt et al., 2013). En effet dans des cellules HepG2, lorsque le site de liaison des *Smad* pour l'enhancer en aval de *HoxC8* ou *HoxD11* sont mutés, les rapporteurs associés à ces séquences ne sont plus exprimés en réponse à *Gdf11*. Par ailleurs, un transgène *HoxD11* portant une mutation sur le motif de liaison *Smad* n'est plus exprimé dans le bourgeon de queue de l'embryon de souris (Gaunt et al., 2013).

En résumé, la voie Gdf signal via le récepteur Alk5 et les facteurs Smads afin de réguler l'expression des gènes *Hox8-13* en fonction du moment du développement où la voie est activée. Cependant, tout comme les expériences qui ont étudié le rôle du Fgf, ces différentes approches (électroporations, explants, mutants) ne permettent pas de contrôler précisément les types de cellules affectées, et par conséquent, ne permettent pas de conclure de manière définitive sur le rôle de la voie Gdf dans l'initiation de l'expression des gènes *Hox* dans les progéniteurs axiaux. De plus, la plupart des interprétations concernant le rôle des facteurs extrinsèques sont basées sur l'observation des domaines d'expression des gènes *Hox* le long de l'axe antéro-postérieur à un moment spécifique. Ce sont donc des études spatiales qui ne permettent pas de tirer de conclusions directes sur le rôle temporel des facteurs extrinsèques dans la régulation de l'expression des gènes *Hox*.

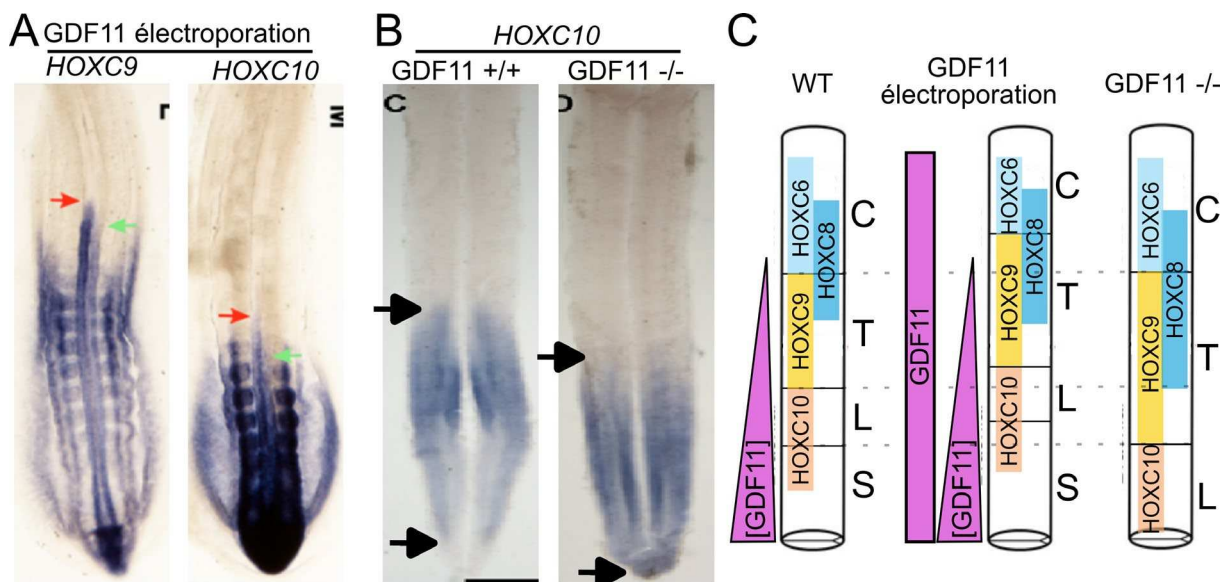


Figure 17. L'initiation des gènes *Hox* les plus caudaux dépend de la voie Gdf (a) Hybridation *in situ* de *HoxC9* et *HoxC10* 20-24h après électroporation de *Gdf11* chez l'embryon de poulet à HH15-17. La pointe de flèche verte indique la limite d'expression rostrale normale et la pointe de flèche rouge indique la limite rostrale après électroporation de *Gdf11*. L'expression ectopique de *Gdf11* conduit à une expansion rostrale des domaines d'expression de *HoxC9* et *HoxC10*. (b) Hybridation *in situ* de *HoxC10* dans embryon de souris de type sauvage et mutant *Gdf11*^{-/-}. L'expression de *HoxC10* est plus restreinte chez le mutant par rapport au contrôle. (c) Schéma récapitulant les résultats présentés en (a, b) mettant en évidence le rôle de *Gdf11* dans l'expression des gènes *Hox* caractérisant les segments postérieurs. (Adapté d'après Liu, 2006)

7. Contribution des signaux intrinsèques et extrinsèques dans la régulation des gènes *Hox*

Plusieurs signaux participant à la régulation de l'expression de *Hox* ont été présentés jusqu'à présent. Cependant, l'importance relative des signaux intrinsèques et extrinsèques dans la régulation des gènes *Hox* fait toujours l'objet de débats. En effet, des observations contradictoires ont été faites concernant la contribution des signaux environnementaux dans la régulation des gènes *Hox*, notamment lors d'expériences de greffes hétérochroniques où des

progéniteurs sont transplantés dans une région homologue mais à un stade de développement différent.

Chez l'embryon aviaire, une sous-population de progéniteurs se trouvant dans la ligne primitive a été décrite et participe à la production de la partie médiane des somites (Imura et al., 2007; Psychoyos and Stern, 1996). Ce sont les progéniteurs du somite médian (MSPs) qui expriment *HoxA2/A3/A6* à partir de HH8. Une récente étude indique que l'induction des *Hox* dépend de mécanismes intrinsèques dans ces cellules car 1) l'expression de *HoxA2/A3/A6* est maintenue après la greffe hétérochronique de MSPs « vieux » dans un environnement plus « jeune », 2) l'expression de *HoxA2/A3/A6* continue d'augmenter dans des explants de MSP cultivés *ex vivo* et 3) la greffe hétérochronique de MSPs « jeunes » dans un environnement plus « vieux » ne permet pas l'activation de *HoxA2* (Busby et al., 2023). Le maintien de l'expression des gènes *Hox* après greffe dans un environnement plus « jeune » a également été démontré par la persistance de l'expression de *HoxB9* après la greffe hétérochronique d'un fragment de la ligne primitive donnant naissance à des cellules mésodermiques (Imura and Pourquié, 2006).

Toutefois, ces résultats diffèrent de ceux rapportés par une autre étude indiquant que l'environnement extrinsèque dans lequel se trouve les progéniteurs axiaux influence la dynamique d'expression des gènes *Hox* tout comme le potentiel de différenciation neuro-mésodermique. Dans cette étude, des progéniteurs axiaux exprimant *HoxA10* au stade HH15 et se trouvant dans le CNH ont été greffés dans la région caudale de poulet à HH8. 8h après la greffe hétérochronique de ces progéniteurs « vieux » dans une niche « jeune », une perte de l'expression de *HoxA10* est observée (McGrew et al., 2008). Ce résultat suggère que l'expression des gènes *Hox*, du moins leur maintien, pourrait être contrôlée par des signaux environnementaux et que l'activation des *Hox* serait donc possiblement réversible (**Fig. 18**).

Les différences observées entre ces deux études (Busby et al., 2023; McGrew et al., 2008) peuvent s'expliquer par 1) des différences dans l'origine des populations cellulaires transplantées (MSP vs CNH) et 2) la variation des contributions extrinsèques et intrinsèques pour réguler les gènes *Hox* en fonction de leur position génomique. En effet, il est possible que les gènes *Hox* antérieurs se trouvant en 3' des complexes soient régulés par un mode intrinsèque, puis plus tard dans le développement, un mode extrinsèque permettrait de réguler les *Hox* plus en 5'.

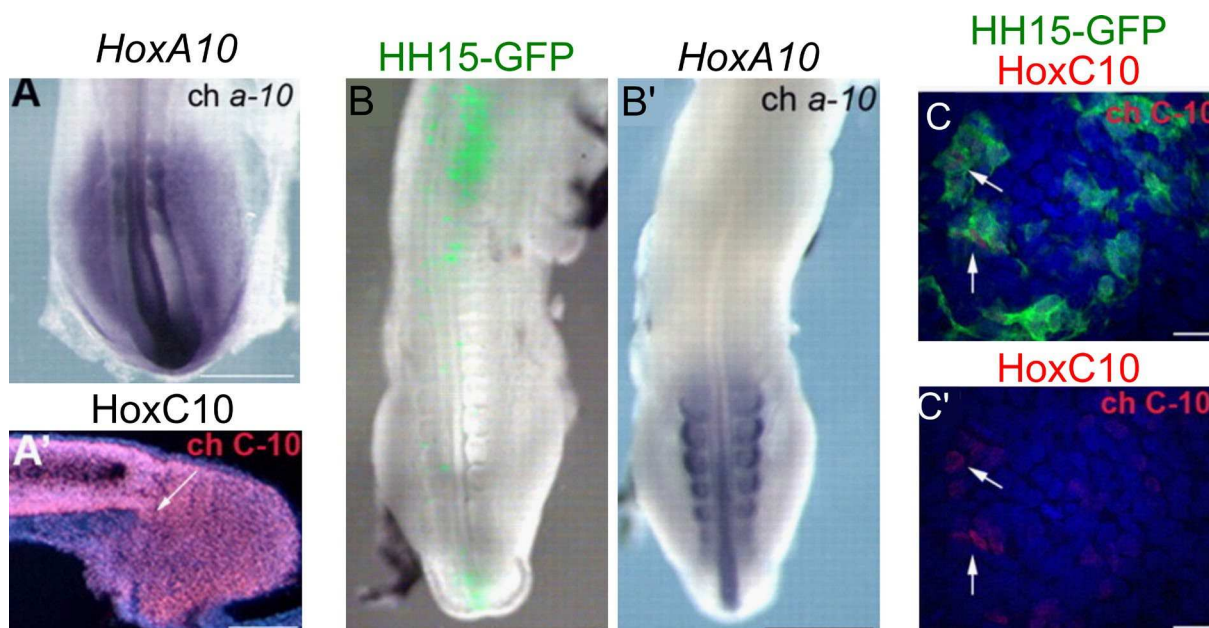


Figure 18. Les facteurs environnementaux régulent l'expression des gènes *Hox*. (a, a') Dans le bourgeon caudal du poulet à HH15, hybridation *in situ* de *HoxA10* (A) et immunomarquage de HoxC10 (A'). (b, b') Hybridation *in situ* de *HoxA10* après greffe hétérotopique de HH15 à HH8. Les cellules GFP+ proviennent du bourgeon caudal de poulet HH15 (B) et le tissu antérieur GFP+ est négatif pour *HoxA10* (B'). (c, c') Immunomarquage de HoxC10 dans coupe transversale antérieure d'un embryon 48 heures après greffe hétérotopique de HH15 à HH8. La majorité des cellules GFP+ ne sont plus positives pour HoxC10. (Adapté d'après McGrew et al., 2008)

En résumé, les signaux environnementaux participant à la régulation des gènes *Hox* ont été caractérisés dans des contextes expérimentaux limités que nous avons détaillés au cours de ce chapitre. Le rôle exact des signaux extrinsèques dans l'initiation de l'expression des gènes *Hox* dans les progéniteurs axiaux doit donc encore être caractérisé. Par ailleurs, il n'est pas clair si l'activité des voies de signalisation passe par 1) des changements de l'état chromatinien précédant ou suivant la liaison de facteurs de signalisation, et 2) si ces changements se passent selon un mode saltatoire (i.e. ouverture complète des régions activées indépendamment de l'activation séquentielle des gènes) ou bien progressif (i.e. changement chromatinien progressif en lien avec l'activité transcriptionnelle d'un gène donné) (Mazzoni et al., 2013; Soshnikova and Duboule, 2009).

La génération *in vitro* d'un grand nombre de progéniteurs axiaux à partir de cellules souches pluripotentes offre une alternative prometteuse pour répondre à ces questions. En effet, grâce à cette approche, il est possible 1) d'avoir un accès illimité à des populations cellulaires rares et difficilement accessibles, en particulier chez l'homme, 2) de moduler précisément l'activité d'un facteur afin d'étudier son rôle spécifique, et 3) de caractériser des processus de signalisation et de transcription rapide.

III. Cellules souches pluripotentes

Les cellules souches pluripotentes (CSPs) sont des cellules indifférenciées qui sont à l'origine, au cours du développement, de toutes les cellules de l'organisme. Ces cellules peuvent être isolées, amplifiées et différenciées *in vitro*. L'un des principaux avantages de l'utilisation de ces cellules est leur capacité à générer tous les types de cellules dérivées de l'endoderme, de l'ectoderme et du mésoderme. Cela en fait donc un outil pertinent afin d'étudier les processus contrôlant le développement et orchestrant la diversification cellulaire au cours de l'embryogenèse.

1. Généralités

Les premières CSEm ont été générées à partir d'embryons de souris. Les méthodes de dérivation des CSEm consistent soit à cultiver des embryons entiers, soit à effectuer une dissection des blastocystes pour isoler les CSE du reste de l'embryon. Ces CSEm sont ensuite cultivées sur des cellules nourricières telles que des fibroblastes embryonnaires murins, qui ont été rendus non proliférants et qui sécrètent des facteurs essentiels au maintien de la pluripotence. Dans ces conditions, les CSEm forment des colonies compactes qui peuvent être dissociées en petits morceaux par des méthodes enzymatiques ou mécaniques, générant ainsi de nouvelles colonies (Evans and Kaufman, 1981). Quelques années plus tard, ces cellules ont été isolées depuis des tissus humains, à partir de blastocystes issus de fécondations *in vitro* offrant de nouvelles opportunités pour aborder des questions liées au développement humain, l'étude des maladies et la recherche thérapeutique (Thomson et al., 1998).

En 2006, l'équipe du Pr. Yamanaka a découvert la possibilité de convertir génétiquement des cellules somatiques adultes en cellules présentant les caractéristiques fondamentales des cellules souches embryonnaires, désignées par la suite sous le nom de "cellules souches pluripotentes induites" (CSPi) (Takahashi and Yamanaka, 2006). Ces recherches, initialement menées sur des cellules murines, ont rapidement été étendues aux cellules humaines (Takahashi et al., 2007). En étudiant 24 gènes connus pour être impliqués dans le maintien de la pluripotence, l'équipe du Pr. Yamanaka a pu identifier une combinaison de 4 gènes (*Oct4*, *Sox2*, *cMyc* et *Klf4*) qui sont nécessaires et suffisants pour induire la reprogrammation (**Fig. 19**).

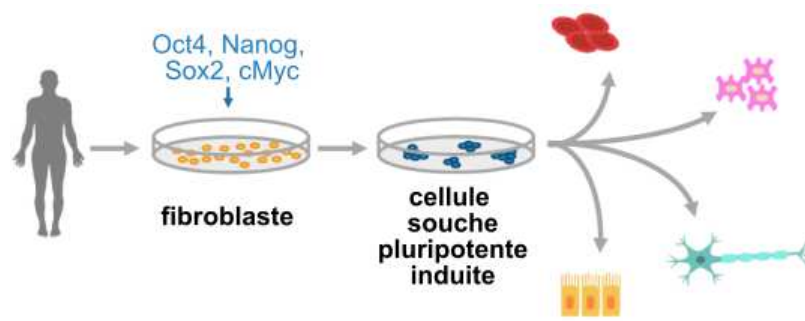


Figure 19. Génération de cellules souches pluripotentes induites. Des cellules différenciées (fibroblaste, sang...) sont prélevées, mises en culture puis infectées à l'aide de virus exprimant les facteurs de pluripotence (Oct4, Nanog, Sox2, cMyc). Les colonies de CSPi sont ensuite isolées puis peuvent être utilisées dans des protocoles de différenciation et donner naissance à n'importe quelles cellules d'un des trois feuillet embryonnaires (ectoderme, mésoderme, endoderme).

2. Modéliser le développement de l'axe antéro-postérieur *in vitro*

L'utilisation des CSE et la découverte de CSPi ont ouvert un nouveau champ afin d'étudier des populations cellulaires rares de l'organisme et des processus impliqués dans leur développement. Des protocoles permettant la différenciation de progéniteurs axiaux *in vitro* ont donc été développés en reproduisant les événements qui conduisent à leur genèse dans l'embryon.

a. Progéniteurs axiaux *in vitro*

Le premier protocole permettant de générer des progéniteurs axiaux *in vitro* a été mis au point en traitant des cellules souches d'épiblaste de souris avec du CHIR99021 (agoniste de la voie Wnt) en combinaison avec du FGF2 et de l'activine A (Tsakiridis et al., 2014). Les populations de cellules dérivées de CSP ont des caractéristiques similaires aux NMPs *in vitro* puisqu'ils 1) coexpriment Sox2 et Brachyury, 2) génèrent des types de cellules neurales et mésodermiques et 3) contribuent à la fois au tube neural et au mésoderme paraxial après leur greffe dans des embryons hôtes (Edri et al., 2019; Frith et al., 2018; Gouti et al., 2017, 2014; Tsakiridis et al., 2014; Turner et al., 2014; Verrier et al., 2018). Par ailleurs, les analyses transcriptomiques ont également révélé que ces cellules dérivées des CSP ont des profils d'expression similaires à ceux des progéniteurs axiaux embryonnaires puisqu'ils expriment également des marqueurs caudaux (e.g. *Cdx2*, *Fgf8*) (Frith et al., 2018; Gouti et al., 2017; Verrier et al., 2018; Wymeersch et al., 2018).

De nombreux protocoles ont été décrits et améliorés, utilisant parfois d'autres types de cellules pluripotentes comme des cellules humaines, pour générer des cellules présentant certaines caractéristiques moléculaires et fonctionnelles des progéniteurs axiaux (Amin et al.,

2016; Edri et al., 2019; Frith et al., 2018; Gouti et al., 2014; Lippmann et al., 2015; Maury et al., 2015; Verrier et al., 2018). Malgré des variations concernant la durée de traitement, de composition des milieux et de choix des molécules de signalisation, tous ces protocoles de génération *in vitro* de progéniteurs axiaux reposent sur la stimulation de la voie de signalisation Wnt, et souvent de la voie Fgf, récapitulant ainsi l'environnement de signalisation *in vivo* dans lequel se trouve ces progéniteurs dans l'embryon. L'ajout de Fgf exogène distingue souvent les protocoles adhérents des protocoles en 3D dans lesquels le Fgf produit par les progéniteurs eux-mêmes semble être suffisant (Maury et al., 2015; Turner et al., 2014).

b. Spécification des sous-types cellulaires de l'axe antéro-postérieur *in vitro*

La différenciation ciblée des progéniteurs axiaux en cellules présentant à la fois un lignage et une identité axiale spécifique constitue un défi majeur pour la recherche fondamentale et appliquée. En effet, l'identité antéro-postérieure des cellules joue un rôle essentiel dans leur fonction. En lien, certaines maladies affectent des types de cellules spécifiques et la compréhension des mécanismes sous-jacents qui conduisent à leur dysfonctionnement nécessite l'amélioration de nos capacités à pouvoir les générer.

Par exemple, les premiers motoneurones générés *in vitro* à partir de cellules souches murines ou humaines ont été obtenus en utilisant uniquement du RA et un agoniste de la voie Sonic Hedgehog, sans passer par une étape de différenciation en progéniteurs axiaux (Shin et al., 2005; Wichterle et al., 2002). En accord ces motoneurones expriment des gènes *Hox* qui sont associés aux régions du tronc cérébral caudal et de la moelle épinière antérieure. La génération de sous-types de motoneurones plus caudaux n'a été permise que plus récemment après la spécification de NMPs à partir de CSP murines ou humaines (**Fig. 20**) (Gouti et al., 2014; Lippmann et al., 2015). La génération de NMPs *in vitro* déclenchée par les voies Wnt et Fgf s'accompagne d'une régulation séquentielle à la hausse des gènes *Hox* jusqu'aux paralogues 9, reflétant ainsi un caractère brachial et thoracique postérieur prédominant chez ces progéniteurs (Lippmann et al., 2015). Le passage à des gènes *Hox* paralogues plus postérieurs caractérisant les régions lombo-sacrés est ensuite possible en augmentant l'activité de signalisation Fgf et/ou en traitant avec le Gdf11 (Lippmann et al., 2015). Cependant, les taux de motoneurones produits à partir de ces protocoles est faible et les autres types cellulaires mal définis.

Des recherches supplémentaires sont donc nécessaires pour comprendre les conditions optimales permettant la génération de l'ensemble des tissus composant l'axe antéro-postérieur. L'amélioration de ces méthodes permettra d'isoler des populations de cellules rares et difficiles d'accès, afin de mieux comprendre les mécanismes qui conduisent à leur formation ou leur dégénérescence.

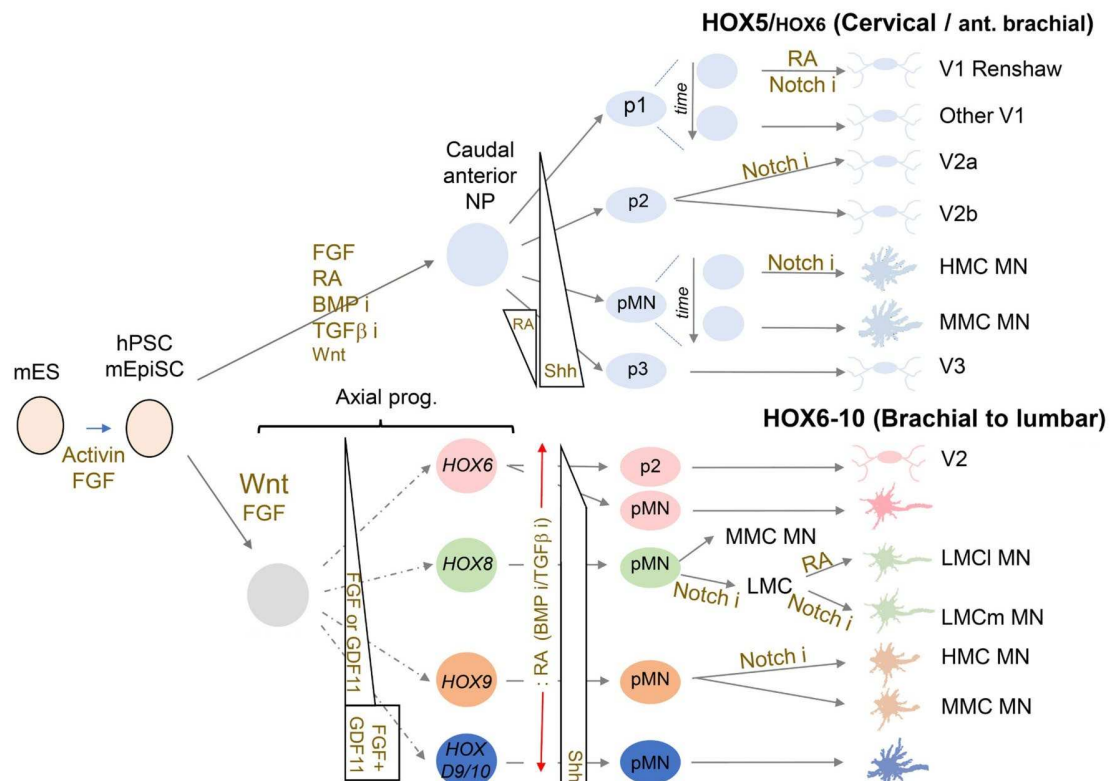


Figure 20. Résumé de la différenciation ciblée *in vitro* en sous-types de neurones spinaux. Le schéma résume différentes études où des CSP ont été différenciées en différentes cellules impliquées dans la formation de l'axe antéro-postérieur (de la région cervicale à la région lombaire). (Adapté d'après Nedelec and Martinez-Arias, 2021)

c. Modèle *in vitro* de la morphogenèse antéro-postérieure basé sur l'auto-organisation des cellules souches pluripotentes

Suite à la découverte des CSE, leur capacité à imiter l'embryogenèse a été testée par la formation d'agrégats tridimensionnels appelés corps embryoides (Doetschman et al., 1985). Les corps embryoides sont devenus un outil largement utilisé pour explorer les mécanismes de spécification du destin cellulaire et de différenciation (Murry and Keller, 2008). Cependant, la plupart des modèles de différenciation utilisant des corps embryonnaires présentent des limites dans l'étude de la morphogenèse axiale, car ces agrégats cellulaires génèrent généralement des types de cellules qui ne sont pas organisés spatialement.

Les travaux pionniers de ten Berge ont été les premiers à montrer que l'activation locale de la voie Wnt dans des CSEm permet d'établir des gradients de morphogènes auto-organisés et une spécification locale des cellules mésodermiques (Ten Berge et al., 2008). Suite à cette découverte, des conditions expérimentales appropriées ont montré que la courte stimulation de la voie Wnt dans des agrégats de CSEm permet 1) de rompre la symétrie (déclenchée par l'émergence d'un pôle de cellules exprimant Brachyury), 2) l'émergence des trois couches germinales et 3) la spécification axiale (caractérisée par la mise en place de patrons d'expression spatiale des gènes *Hox*) (Beccari et al., 2018; Turner et al., 2017; Van Den Brink et al., 2014). Ce modèle *in vitro*, appelé gastruloïde, est donc capable de donner naissance aux progéniteurs des différents organes du tronc à la bonne position. Des gastruloïdes humains ont ensuite été générés à partir de CSE humaines (Moris et al., 2020). L'analyse transcriptomique des gastruloïdes humains a démontré que ce modèle *in vitro* a des similarités avec l'embryon humain à 21 jours post-fertilisation. A ce stade, des somites commencent à émerger.

Cependant, malgré la possibilité d'avoir des domaines d'expression génétique organisés, les gastruloïdes restent dépourvus d'une morphogénèse ressemblant à celle de l'embryon. Une étude a démontré que l'inclusion de gastruloïdes de souris dans du matrigel, un substitut de la matrice extracellulaire, conduit à la formation d'un tube neural entouré de somites bilatéraux mimant ainsi des aspects de la morphogénèse du tronc murin (Veenvliet et al., 2020).

Plus récemment, d'autres modèles 3D, en particulier humain, permettant de mimer la morphogénèse des somites et/ou de la moelle épinière ont été décrits (Libby et al., 2021; Miao et al., 2023; Sanaki-Matsumiya et al., 2022; Yaman and Ramanathan, 2023; Yamanaka et al., 2023). Ces différents modèles sont tous basés sur la spécification de progéniteurs axiaux bipotents en exposant des CSP humains à un activateur de la voie Wnt, et en présence ou en l'absence de Fgf exogène.

L'un de ces modèles, appelé "axioloïdes", recrée le processus de somitogénèse et est caractérisé par la régulation spatiale et temporelle des gènes *HOX* (Yamanaka et al., 2023). Ainsi, *HOXC4* est exprimé tout au long de la structure antéro-postérieure, tandis que l'expression de *HOXC10* est plus tardive et se trouve limitée à la partie caudale des "axioloïdes" (**Fig. 21**). Cependant, aucun de ces modèles intégrés ne parvient à récapituler la comorphogénèse des somites et du tube neural humain, et parallèlement à reproduire la mise en place de patrons d'expression spatiaux des gènes *HOX* puis des facteurs HOX le long de l'axe antéro-postérieur. Cela indique l'importance de développer de nouveaux modèles

expérimentaux pour mieux comprendre la complexité de la régulation des gènes *HOX* dans le contexte du développement humain.

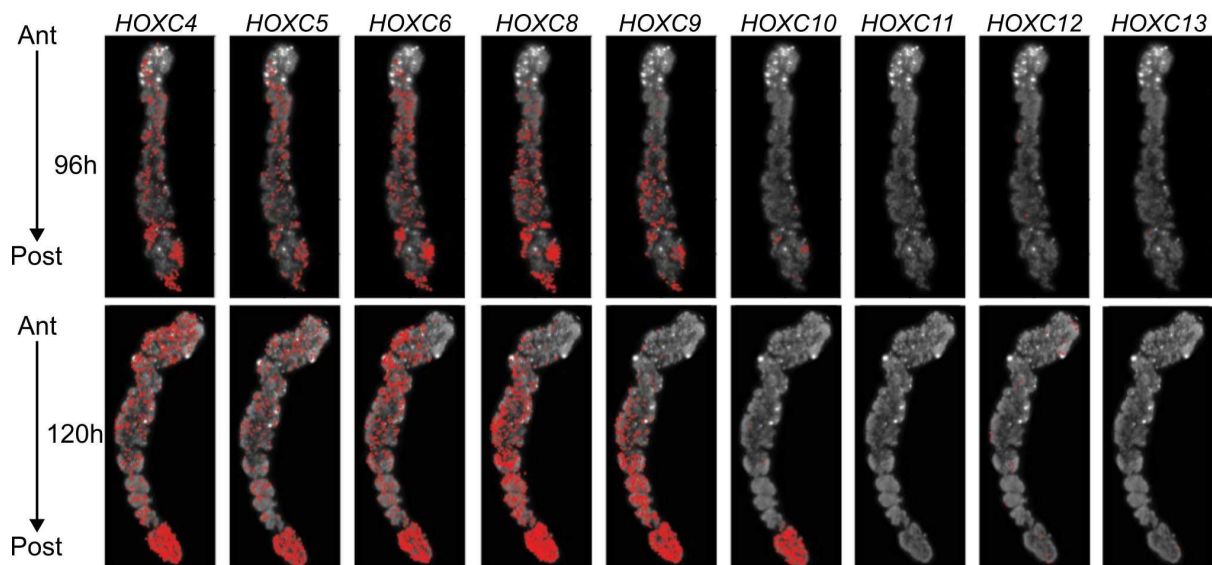


Figure 21. Organisation cellulaire le long de l'axe antéro-postérieur des axioloïdes. Distribution spatiale des transcrits *HOXC* dans les axioloïdes à 96h et 120h. (D'après Yamanaka et al., 2023)

Ces nouveaux modèles constituent des plateformes accessibles qui peuvent être caractérisées avec une haute résolution spatio-temporelle afin d'étudier la coordination entre les décisions relatives à la lignée et les processus morphogénétiques qui façonnent l'embryon au milieu de la gestation. Cependant, aucun de ces modèles intégrés ne parvient à reproduire la comorphogenèse et la structuration antéro-postérieure des lignées neuronales et mésodermiques. De plus, ces modèles ne permettent pas de reproduire des stades embryonnaires avancés.

Résultats

Partie 1

Rôle des facteurs extrinsèques dans le contrôle de l'induction des gènes *HOX*

La régulation par les facteurs extrinsèques de l'horloge *HOX* dans les progéniteurs axiaux humains contrôle la diversification des motoneurones

Article publié dans le journal Development

a. Contexte

Au cours de l'embryogenèse, la mise en place de la diversité cellulaire le long de l'axe antéro-postérieur dépend de l'expression des facteurs de transcription *HOX* le long de l'axe (Carpenter, 2002; Philippidou and Dasen, 2013). Ces facteurs sont produits par des gènes organisés dans quatre complexes génomiques (*HOXA/B/C/D*) et leur expression est initiée dans les progéniteurs axiaux, une population cellulaire localisée dans le bourgeon caudal de l'embryon qui donne naissance aux différents tissus postérieurs à partir du cou (Forlani et al., 2003; Neijts et al., 2016). L'activation transcriptionnelle des gènes *HOX* est progressive et colinéaire puisqu'ils sont différentiellement exprimés le long de l'axe antéro-postérieur en accord avec leur position dans le génome. Ainsi, les gènes *HOX* situés en 3' des complexes sont les premiers à être exprimés et participeront à la spécification des régions antérieures. A l'inverse, les gènes situés en 5' des complexes *HOX* sont exprimés plus tardivement et participeront à la spécification des régions postérieures. (Dasen et al., 2003; Duboule and Dollé, 1989). Chez la souris et le poulet, les domaines d'expression *HOX* le long de l'axe antéro-postérieur ont été bien caractérisés dans la moelle épinière. Cependant, les mécanismes contrôlant le rythme d'activation des gènes *HOX* dans les progéniteurs axiaux afin de permettre en place la spécification cellulaire de l'axe restent largement inconnus.

Des études ont démontré le rôle de facteurs extrinsèques (RA, FGF et GDF11) dans le contrôle de l'expression des gènes *Hox* (Bel-Vialar et al., 2002; Liu, 2006; Stifani, 2014). D'autre part, des études de génomique fonctionnelle ont démontré une ouverture progressive de la chromatine en lien avec l'activation séquentielle des gènes *Hox* au cours du temps

(Soshnikova and Duboule, 2009). Restait à savoir si la temporalité de l'ouverture progressive de la chromatine initiée et terminée par ces facteurs extrinsèques jouait un rôle d'horloge intrinsèque dans ces cellules, ou bien si la temporalité d'induction des gènes *HOX* était déterminée par l'exposition séquentielle à ces différents facteurs extrinsèques. Cette question était difficile à aborder dans un contexte *in vivo*, car elle nécessitait l'accès à des populations de cellules rares dans des états très dynamiques, que l'on trouve dans des tissus complexes composés de plusieurs types de cellules.

Puisqu'il était difficile d'étudier la dynamique de l'induction des gènes *HOX* et le rôle des facteurs intrinsèques ou extrinsèques dans ce contexte, nous avons utilisé un modèle de différenciation *in vitro* de cellules souches pluripotentes humaines en progéniteurs axiaux puis en motoneurones dont la diversité est instruite par l'expression différentielle des gènes *HOX*. Grâce à ce modèle nous avons pu étudier comment la temporalité d'induction des *HOX* est régulée au cours du développement de la moelle épinière.

b. Résultats

Tout d'abord, nous avons identifié les domaines d'expression des protéines HOXC dans la moelle épinière de l'embryon humain. Ces résultats sont très similaires à ceux déjà observés chez les souris et les poulets, ce qui renforce l'idée que l'expression des gènes *HOX* est conservée au cours de l'évolution. De plus, la caractérisation des domaines d'expression des facteurs HOX dans l'embryon humain nous a permis d'identifier le code HOX des motoneurones humains. Par la suite, ces résultats ont permis de caractériser la différenciation des motoneurones *in vitro*.

En nous appuyant sur les travaux antérieurs de l'équipe (Maury et al., 2015), nous avons différencié des cellules souches pluripotentes induites humaines en progéniteurs ressemblant aux progéniteurs axiaux et capables de donner naissance à des motoneurones spinaux d'identité cervicale (Cambray and Wilson, 2007, 2002; Forlani et al., 2003). En décalant progressivement le jour où ces progéniteurs sont exposés à RA, nous avons mis en évidence la genèse de motoneurones exprimant des facteurs HOX présent de plus en plus caudalement dans l'embryon jusqu'aux régions thoraciques. Ce résultat indique que le moment d'exposition de ces progéniteurs au RA détermine l'identité des progéniteurs le long de l'axe antéro-postérieur.

Nous avons ensuite cherché à définir les changements moléculaires qui accompagnent ce changement de potentiel des progéniteurs au cours du temps. L'analyse transcriptomique, a montré que ces progéniteurs expriment les marqueurs typiques des progéniteurs axiaux et de

plus induisent progressivement les gènes *HOX* de 3' vers 5' au cours du temps. Cette induction progressive des gènes *HOX* est corrélée avec une augmentation de l'expression de gènes de la famille des ligands FGFs et des gènes cibles de la voie MEK-ERK1/2 des MAPKs, une voie de signalisation connue pour transduire les signaux FGFs.

Pour déterminer si cette voie avait un rôle dans l'activation progressif des gènes *HOX*, nous avons inhibé les récepteurs aux FGF ou les effecteurs MEK1/2. Dans les deux cas, l'activation progressive des gènes *HOX* est bloquée et les progéniteurs axiaux donnent naissance à des motoneurones ayant une identité brachiale antérieure au lieu de thoracique. A l'inverse, l'exposition des progéniteurs axiaux à du FGF2 exogène accélère l'induction des gènes *HOX* et conduit à la spécification de motoneurones ayant une identité plus caudale.

De plus, l'utilisation combinée du FGF2 avec le facteur extrinsèque GDF11 entraîne une accélération encore plus importante du rythme d'activation des gènes *HOX*, avec l'initiation de nouveaux gènes tels que *HOXC10* ainsi que la génération de motoneurones ayant une identité lombaire.

Dans l'ensemble, ces travaux montrent que le rythme d'activation des gènes *HOX* est régulée de manière critique par des facteurs extrinsèques. Notre travail a non seulement permis d'élucider certains mécanismes participant à la régulation de l'induction des gènes *HOX*, mais également de développer des protocoles efficaces et rapides permettant de produire efficacement différents sous-types de motoneurones. Certains de ces protocoles ont été brevetés.

c. Contribution

Pour cette étude, j'ai analysé l'ensemble des données transcriptomiques qui ont permis de caractériser 1) l'identité des progéniteurs *in vitro*, 2) la signalisation endogène de la voie FGF et 3) l'activation progressive de *HOX* au cours du temps. J'ai également contribué à la caractérisation de l'évolution des progéniteurs axiaux au cours de la différenciation, révélant leur engagement progressif et spontané vers un destin pré-neural.

Dynamic extrinsic pacing of the HOX clock in human axial progenitors controls motor neuron subtype specification

Vincent Mouilleau^{1,2,3,4,*}, Célia Vaslin^{1,2,3,*}, Rémi Robert^{1,2,3}, Simona Gribaudo^{1,2,3}, Nour Nicolas⁵, Margot Jarrige⁴, Angélique Terray^{1,2,3}, Léa Lesueur⁴, Mackenzie W. Mathis⁶, Gist Croft⁶, Mathieu Daynac^{1,2,3}, Virginie Rouiller-Fabre⁵, Hynek Wichterle⁶, Vanessa Ribes⁷, Cécile Martinat⁴ and Stéphane Nedelec^{1,2,3,†}

ABSTRACT

Rostro-caudal patterning of vertebrates depends on the temporally progressive activation of HOX genes within axial stem cells that fuel axial embryo elongation. Whether the pace of sequential activation of HOX genes, the 'HOX clock', is controlled by intrinsic chromatin-based timing mechanisms or by temporal changes in extrinsic cues remains unclear. Here, we studied HOX clock pacing in human pluripotent stem cell-derived axial progenitors differentiating into diverse spinal cord motor neuron subtypes. We show that the progressive activation of caudal HOX genes is controlled by a dynamic increase in FGF signaling. Blocking the FGF pathway stalled induction of HOX genes, while a precocious increase of FGF, alone or with GDF11 ligand, accelerated the HOX clock. Cells differentiated under accelerated HOX induction generated appropriate posterior motor neuron subtypes found along the human embryonic spinal cord. The pacing of the HOX clock is thus dynamically regulated by exposure to secreted cues. Its manipulation by extrinsic factors provides synchronized access to multiple human neuronal subtypes of distinct rostro-caudal identities for basic and translational applications.

This article has an associated 'The people behind the papers' interview.

KEY WORDS: Axial, HOX genes, Human, Motor neurons, Pluripotent stem cells, Spinal cord

INTRODUCTION


The patterning of bilaterian body is orchestrated by the differential expression of HOX transcription factors along their rostro-caudal axis. In vertebrates, HOX proteins are encoded by 39 genes,

organized into four genomic clusters (*HOXA*, *HOXB*, *HOXC* and *HOXD*), that display anterior boundaries of expression co-linear to their chromosomal location. Genes located in the 3' region of the clusters are expressed more anteriorly than their 5' neighbors. The regionalized expression of HOX genes in post-occipital regions is initiated by their sequential, collinear, activation in axial progenitors, a dynamic cell population that fuels rostro-caudal extension of the embryo (Cambray and Wilson, 2002, 2007; Deschamps and Duboule, 2017; Forlani et al., 2003; Gouti et al., 2017; Henrique et al., 2015; Kondoh et al., 2016; Wymeersch et al., 2019). As axial progenitors contribute to progressively more caudal mesodermal and neuroectodermal structures, the 3' to 5' sequence of HOX gene activation is translated into a co-linear spatial pattern of expression in the progenies (Deschamps and Duboule, 2017; Henrique et al., 2015). Hence, the coupling between rostro-caudal extension of the body axis and the sequential activation of Hox genes over time, the HOX clock, is a central element of axial patterning. However, the mechanisms pacing the clock in axial progenitors are still unclear. Although extrinsic factors such as retinoic acid (RA), Wnt, fibroblast growth factors (FGFs) or growth differentiation factors (GDFs) control HOX patterns of expression, cell-intrinsic changes from a transcriptionally repressive to a permissive chromatin state occur within HOX complexes and correlate with the HOX temporal sequence of induction (Bel-Vialar et al., 2002; Deschamps and Duboule, 2017; Liu et al., 2001; Mazzoni et al., 2013; Narendra et al., 2015; Neijts et al., 2017; Noordermeer et al., 2014; Philippidou and Dasen, 2013; Soshnikova and Duboule, 2009). Whether the progressive opening of chromatin along the complexes serves as an internal timer actuated by extrinsic cues or whether sequences of secreted factors that activate progressively more caudal HOX genes largely defines the tempo of the induction remains unclear (Bel-Vialar et al., 2002; Del Corral and Storey, 2004; Deschamps and Duboule, 2017; Ebisuya and Briscoe, 2018; Lippmann et al., 2015; Mazzoni et al., 2013; Wymeersch et al., 2019). Inability to distinguish between these two hypotheses might underlie our limited ability to finely tune the rostro-caudal identity of caudal cell types derived from pluripotent stem cells [PSCs: embryonic stem cells (ESCs) and induced pluripotent stem cells (iPSCs); Diaz-Cuadros et al., 2020; Du et al., 2015; Duval et al., 2019; Faustino Martins et al., 2020; Frith et al., 2018; Gouti et al., 2014; Li et al., 2005; Lippmann et al., 2015; Matsuda et al., 2020; Maury et al., 2015; Ogura et al., 2018; Peljto et al., 2010; Sasai et al., 2014; Verrier et al., 2018]. The two models imply distinct strategies for *in vitro* differentiation: the 'intrinsic model' predicts that the specification of posterior identities will require a precise synchronization between the timing of differentiation and the internal HOX timer. Alternatively, the 'extrinsic model' implies that exposing axial progenitors to the relevant extrinsic cues should control the HOX clock to generate progenies of defined rostro-caudal identities. To approach this

¹Institut du Fer à Moulin, 75005 Paris, France. ²Inserm, UMR-S 1270, 75005 Paris, France. ³Sorbonne Université, Science and Engineering Faculty, 75005 Paris, France. ⁴I-STEM, UMR 861, Inserm, UEPS, 91100 Corbeil-Essonnes, France. ⁵Laboratory of Development of the Gonads, Unit of Genetic Stability, Stem Cells and Radiation, UMR 967, INSERM, CEA/DSV/IRCM/SCSR, Université Paris Diderot, Sorbonne Paris Cité, Université Paris-Sud, Université Paris-Saclay, Fontenay aux Roses F-92265, France. ⁶Departments of Pathology and Cell Biology, Neuroscience, and Neurology, Center for Motor Neuron Biology and Disease, Columbia Stem Cell Initiative, Columbia University Medical Center, New York, NY 10032, USA. ⁷Université de Paris, CNRS, Institut Jacques Monod, 15 rue Héliène Brion, 75013 Paris, France.

*These authors contributed equally to this work

†Author for correspondence (stephane.nedelec@inserm.fr)

 C.V., 0000-0002-0047-5132; R.R., 0000-0002-9361-3855; N.N., 0000-0001-8266-8036; M.W.M., 0000-0001-7368-4456; G.C., 0000-0002-9721-0148; M.D., 0000-0002-4956-6289; V.R., 0000-0002-3927-5701; H.W., 0000-0002-7817-0080; V.R., 0000-0001-7016-9192; C.M., 0000-0002-5234-1064; S.N., 0000-0001-8044-2498

question and its consequences for cell engineering, we aimed to generate axial progenitors from human pluripotent stem cells (hPSCs) to study the mechanisms pacing the HOX clock during their differentiation into spinal motor neurons (MNs), a cell type that relies on a precise HOX code to acquire appropriate rostro-caudal subtype identities. Here, we report the generation of functional axial progenitors from hPSCs demonstrated by their ability to generate spinal MN subtypes located along the rostro-caudal axis of human embryos. As they become older, hPSC-derived axial progenitors generated MN subtypes born later in development and located more caudally in embryos, showing that they undergo a temporal shift in their rostro-caudal potential. Transcriptomic analysis indicated that this shift is linked to a temporal activation of HOX genes paralleled by a graded increase in FGF signaling. Blocking the FGF pathway stalled HOX temporal activation which resulted in the specification of early-born anterior MNs from late progenitors. In contrast, FGF precocious increase accelerated the temporal sequence of HOX activation and promoted the generation of more caudal, thoracic, MN subtypes. Combining FGFs with GDF11, another extrinsic factor, further accelerated the clock, favoring the specification of late-born lumbar MNs. Overall, our results indicate that extrinsic factors function as key pacers of the HOX clock tempo. Manipulations of the pacing pathways allow the efficient and synchronous engineering of human cell types of defined rostro-caudal identities for basic and translation applications.

RESULTS

HOX expression profiles and motor neuron subtypes in human embryonic spinal cord

The differential expression of Hox transcription factors along the rostro-caudal axis of the vertebrate spinal cord controls the formation of molecularly and functionally distinct neuronal populations, which therefore represent a functional product of Hox gene regulation. In particular, the Hox code orchestrates the acquisition of subtype-specific features in mouse and chick spinal motor neuron (MNs), controlling the formation of the distinct locomotor circuits (Dasen, 2017; Philippidou and Dasen, 2013). Whether this spinal HOX code and associated MN subtypes is conserved in human remained unknown, thus preventing faithful assessment of HOX regulation and its link with cell 'rostro-caudal' identity during hPSC differentiation. We therefore mapped on human embryos, at 6.3 and 7.5 weeks of development, seven HOX transcription factors that display spinal collinear expression patterns and instruct MN subtype specification in mouse and chick (Philippidou and Dasen, 2013) (Fig. 1A, Fig. S1). As in animal models, human MNs expressed ISL1 or HB9 all along the spinal cord (Fig. 1A, circles in Fig. S1) (Amoroso et al., 2013). Within MNs, HOX displayed rostro-caudal patterns resembling those of mouse and chick (Dasen et al., 2003; Liu et al., 2001): cervical MNs expressed HOXA/C5, while brachial MNs expressed HOXC6, thoracic MNs expressed HOXC9 and lumbar MNs expressed HOXC10. Caudal brachial MNs co-expressed HOXC6 and HOXC8, and anterior thoracic MNs expressed HOXC8 and HOXC9. HOXD9 labeled caudal thoracic MNs as well as anterior lumbar MNs, which also expressed HOXC10 (Fig. 1A, Fig. S1). In amniotes, this Hox code instructs the formation of distinct motor columns that innervate common muscle groups, and motor pools that innervate a single muscle (Philippidou and Dasen, 2013). To be able to assess *in vitro* whether changes in HOX expression have functional consequences, we mapped columnar and pool markers within HOX expression domains (Fig. 1A, Figs S1 and S2). As in mouse and chick, MNs expressing high levels of FOXP1 (FOXP1^{high}) were observed at brachio-thoracic (HOXC6 and HOXC8/HOXC9) and

lumbar (HOXC10) levels, where they formed a lateral motor column (LMC) matching the location of limb-innervating MNs (Amoroso et al., 2013; Routal and Pal, 1999) (Fig. 1A, Figs S1 and S2). Within the FOXP1^{high} LMC, MNs expressing the transcription factor SCIP (POU3F1) together with HOXC8 were observed in the caudal brachial spinal cord (Helmbacher et al., 2003). In addition, SCIP/HOXC8/HOXC9-positive MNs were detected in the anterior thoracic region, which might correspond to hand-innervating MNs (Fig. S2) (Bell et al., 2017; Mendelsohn et al., 2017).

Overall, HOX transcription factors are regionally expressed along the rostro-caudal axis of the human spinal cord. Within the HOX domains, distinct MN subtypes, identifiable by specific combination of transcription factors, are generated at stereotyped positions. These data provided readouts to assess the mechanisms regulating HOX expression in axial progenitors and their impact on cell type specification during hPSC differentiation.

Temporal regulation of HOX genes and FGF pathway in hPSC-derived axial progenitors

We first sought to test whether axial progenitors (i.e. progenitors producing derivatives along the anteroposterior axis; Cambrey and Wilson, 2002, 2007; Forlani et al., 2003; Kondoh and Takemoto, 2012) could be generated from hPSCs using MN subtype specification as a readout. We have previously reported sequences of extrinsic cues leading to the efficient generation of spinal MNs from hPSCs through a putative axial progenitor stage (Maury et al., 2015). Exposure of free-floating embryoid bodies (EBs) to a WNT pathway agonist (CHIR99021, CHIR) and inhibitors of BMP and TGF β pathways (LDN193189 and SB431542) generated progenitors expressing *CDX2*, which encodes a transcription factor expressed in axial progenitors and a regulator of caudal HOX gene induction (Bel-Vialar et al., 2002; Bialecka et al., 2010; Gouti et al., 2014; Maury et al., 2015; Mazzoni et al., 2013; Neijts et al., 2017). Then, human spinal MN progenitors (83% of the cells on average) were generated upon exposure of these progenitors to: (1) retinoic acid (RA), which was previously shown to guide axial progenitors toward a neural fate; and (2) an agonist of the sonic hedgehog pathway (SAG), which induces a ventral MN progenitor fate (Briscoe and Novitsch, 2008; Maury et al., 2015; Ribes et al., 2009). Finally, inhibition of Notch signaling on day 9 (D9) promoted the differentiation of MN progenitors in post-mitotic MNs (Maury et al., 2015). We thus used MN subtype markers, including HOX transcription factors, to assess the rostro-caudal identity of MNs produced under the previously reported conditions (Fig. 1B) (Maury et al., 2015). First, staining for HB9, ISL1 and the pan-neuronal marker neurofilament light chain (NEFL), together with quantification of ISL1⁺ cells, confirmed the efficient generation of spinal MNs as previously shown (Fig. 1C, Fig. S3A). Analysis of HOX expression showed that exposing cells to RA from D2 followed by SAG on D4 gave rise to some HOXC6 MNs corresponding to anterior brachial location *in vivo*. This identity was acquired by most MNs when RA was added 1 day later, on D3 (70.1% of HOXC6⁺ and 13.9% HOXC8⁺ MNs) (Fig. 1D,E). Addition of RA/SAG from D4 generated MNs with a caudal brachial identity (44.6% HOXC6/C8⁺) from which 37.1% expressed a high level of FOXP1, which identifies limb-innervating MNs in the spinal cord (Fig. 1D-F, Fig. S3B). Overall, these results suggested that the specification of more caudal MN subtypes was dependent on either: (1) the duration of exposure to RA, with a shorter RA exposure promoting caudalization; or (2) the time-point at which the progenitors received RA, as previously suggested (Bel-Vialar et al., 2002; Del Corral and Storey, 2004; Lippmann et al., 2015). To distinguish between these two possibilities, D3 progenitors

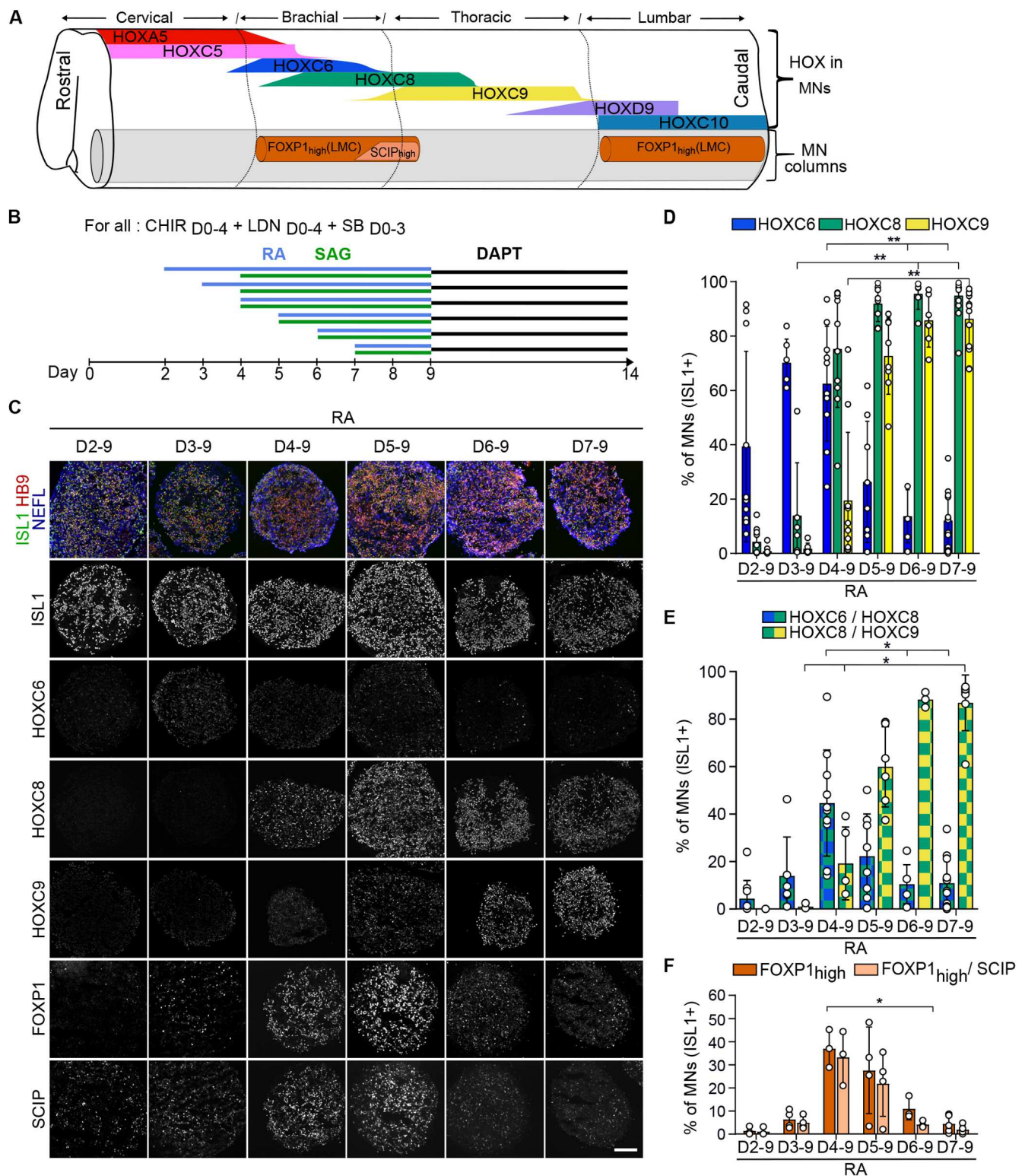


Fig. 1. Human axial progenitors generate progressively more caudal motor neuron subtypes. (A) Schematic summary of data in Figs S1, S2. MNs are defined by the expression of ISL1 or HB9 (gray) are organized in motor columns in spinal ventral horns. HOX expression profiles within MNs and localization of MNs expressing high levels of FOXP1 and SCIP are represented as observed in the human embryonic spinal cord at 6.3 and 7.5 weeks of gestation. Changes in shapes indicate an increase or decrease in the number of MNs expressing a given marker. FOXP1^{high} MNs are observed selectively in lateral motor columns (LMCs) of the brachio-thoracic and lumbar spinal cord. SCIP^{high} MNs are a subset of LMC MNs in the caudal brachial spinal cord. (B) Differentiation conditions used in C to F in which the time of exposure to retinoic acid (RA)/SAG (an agonist of the sonic hedgehog pathway) is modulated. (C) Immunostaining for ISL1/HB9 (MNs), NEFL (neurons), HOX transcription factors, FOXP1 and SCIP on cryostat sections of hESC-derived EBs on day 14 of differentiation. The later RA is applied, the more caudal the MNs are. FOXP1 and SCIP MNs are mostly generated when RA is applied at day 4 and day 5, further defining the rostro-caudal identity of the MNs within HOXC8⁺ conditions. Scale bar: 100 μ m. (D-F) Proportion of MNs (ISL1⁺ cells) expressing the indicated markers. Data are mean \pm s.d. Each circle is an independent biological replicate. (D) $n=6-13$, (E) $n=3-13$, (F) $n=3$ or 4. * $P \leq 0.05$, ** $P \leq 0.01$ (ANOVA with Kruskal-Wallis post-hoc test).

that gave rise mostly to HOXC6 MNs received shorter RA treatments (D3-5 or D3-8 versus D3-9). These two conditions failed to promote more caudal identities (Fig. S3C-E). Instead, HOXC6 expression was lost in the shortest RA treatment (48 h, D3-5, Fig. S3D), indicating a potential role for RA in HOXC6 MNs specification as previously demonstrated in chick (Liu et al., 2001). These results indicate that the day at which progenitors are exposed to RA/SAG is the main trigger of caudalization (Fig. S3C-E). Further delaying RA/SAG then induced even more caudal MN subtypes. Treatment of cells with RA/SAG on D5 generated MNs with an anterior thoracic identity (59.8% HOXC8/C9⁺) (Fig. 1B-E) from which 27.6% MNs acquired a FOXP1⁺ limb-innervating identity (Fig. 1F). Accordingly, HOXC9/FOXP1/SCIP-positive MNs, which are located in the human anterior thoracic spinal cord and might correspond to hand-innervating MNs, were observed almost exclusively in this condition (Fig. 1F, Fig. S3G). Even later treatment, on D6 or D7, specified MNs acquiring a mid-thoracic identity, as demonstrated by the expression of HOXC9 and the loss of FOXP1^{high} MNs (Fig. 1B-F, Fig. S3F,G). The progressive caudalization of MN identity upon incremental delays of RA/SAG addition was confirmed with another hESC line and two iPSC lines (Fig. S4A,B). Interestingly, for one iPSC line, we had to increase CHIR concentration from 3 to 4 μ M to observe this caudalization, suggesting line-to-line differences in WNT pathway integration (Fig. S4A-E). Indeed, and in agreement with the lack of caudalization, upon 3 μ M of CHIR, this iPSC line exhibited a decreased induction of CDX2 at D3, a transcription factor induced in axial progenitors by WNT and FGF signaling pathways, and required for caudal HOX gene induction (Fig. S4F) (Amin et al., 2016; Bel-Vialar et al., 2002; Mazzoni et al., 2013; Nordström et al., 2006; Young et al., 2009).

Overall, WNT activation combined with TGF- β /BMP pathway inhibition converts hPSC into cells that can generate progenies of distinct rostro-caudal identities, a hallmark of axial progenitors. The time window between the initial exposure to WNT agonist and later exposure to RA defines HOX expression patterns in the generated progenies.

Temporal induction of caudal HOX genes and FGF target genes in axial progenitors

We then aimed at defining more precisely the identity of these hPSC derived-axial progenitors and the molecular changes paralleling their temporal switch in rostro-caudal potential, including HOX gene induction patterns. We first wondered whether the progenitors generated between D2 and D4 share a molecular identity with other vertebrate axial progenitors. In mouse, axial progenitor activity is carried in part by cells located in the caudal lateral epiblast (CLE) and the node-primitive streak border, which express *Sox2*, *Nkx1.2*, *Cdx2* and *TbxT* (Brachyury) (Albors et al., 2018; Amin et al., 2016; Cambray and Wilson, 2002, 2007; Edri et al., 2019; Forlani et al., 2003; Gouti et al., 2014, 2017; Henrique et al., 2015; Koch et al., 2017; Kondoh et al., 2016; Mathis and Nicolas, 2003; Tzouanacou et al., 2009; Wymeersch et al., 2016, 2019). Single-cell lineage-tracing experiments showed that some of these axial progenitors are dual-fated neuromesodermal progenitors (NMPs) that generate progenies both in the spinal cord and paraxial mesoderm (Tzouanacou et al., 2009). The relative levels of Sox2 and TbxT transcription factors correlate with a preponderant differentiation of the cells toward the mesodermal (higher TbxT) or neural lineages (higher Sox2). We thus monitored the relative level of expression of *SOX2*, *TBXT* and *NKX1.2* transcripts in D2, D3 and D4 progenitors in comparison with both hESCs and hiPSCs (Fig. S5A,B). The three markers were present from D2 to D4. *SOX2* displayed levels similar to those in hPSCs, while *NKX1.2* and *TBXT* were highly induced

(Fig. S5A,B). We then stained for *SOX2*, *TBXT* and *CDX2* (Fig. 2A, Fig. S5C). They were all co-expressed during the timeframe, yet *TBXT* progressively decreased over time to be weakly expressed by D4. This decrease in *TBXT* expression likely indicates a progressive transition towards preneural plate progenitors, which lose *TbxT* expression and are committed to the neural lineage (Del Corral and Storey, 2004; Delfino-Machín et al., 2005; Olivera-Martinez et al., 2014; Wymeersch et al., 2016).

Whole-transcriptome analysis confirmed that, beside these markers, D2, D3 and D4 hPSC-derived progenitors shared many similarities with mouse axial progenitors found in the caudal epiblast (Fig. 2, Fig. S5 and Tables S1, S2). Functional enrichment analysis of genes increased in progenitors in comparison with hESCs indicated an activation of HOX genes and the WNT pathway in D3 and D4 progenitors, accompanied by a decrease of key pluripotency markers, such as *NANOG* or *KLF2* and *KLF4*, as observed in CLE cells (Fig. 2C, Fig. S5D-F) (Del Corral and Storey, 2004; Deschamps and Duboule, 2017; Olivera-Martinez and Storey, 2007; Olivera-Martinez et al., 2014; Wymeersch et al., 2019). More strikingly, the most common genes defining mouse CLE cells (Edri et al., 2019; Gouti et al., 2017; Olivera-Martinez et al., 2014; Wymeersch et al., 2019) were present from D2 to D4 (Fig. 2C, Table S1), and D2 and D3 progenitors expressed 122 out of 142 genes defining mouse E8.5 and E9.5 NMPs in comparison with mesodermal progenitors (RPM \geq 1.0, Fig. S5G, Table S4). Out of these 122 genes, the CLE genes *CDX1*, *FGF17*, *RXRG*, *SP8*, *WNT5A*, *WNT8A* and *FST* all belonged to the 50 most enriched genes in D2 and D3 progenitors compared with hESCs (Fig. S5H, Tables S1 and S4). There was also a strong overlap with genes found in recently reported hPSC-derived axial progenitors (Fig. S5J, Table S3) (Frith et al., 2018). Overall, the transcriptomic signature of the progenitors, together with the co-expression of *SOX2*, *TBXT* and *CDX2*, demonstrate the WNT-induced specification of human axial progenitors resembling mouse CLE progenitors (in which NMPs reside) (Edri et al., 2019; Gouti et al., 2014, 2017; Henrique et al., 2015; Koch et al., 2017; Liu et al., 2001; Wymeersch et al., 2016). However, the degree of similarity with the different states of mouse axial progenitors, including early and late NMPs and whether individual cells are bipotent, would require additional analysis and experiments. Furthermore, the progressive decrease of *TBXT* indicates a transition toward more neural potent progenitors, as observed in the mouse anterior CLE and in regions anterior to the node where pre-neural axial progenitors are located (Del Corral et al., 2003; Diez del Corral et al., 2002; Forlani et al., 2003; Wymeersch et al., 2016).

Next, we aimed to define the molecular changes paralleling the temporal change in rostro-caudal potential of the progenitors. A comparative analysis of axial progenitor transcriptomes indicated that D2 progenitors already activated most HOXB genes (except HOXB13), which are also activated early in development in mouse and chick and are not differentially expressed along the spinal cord (Dasen et al., 2003; Denans et al., 2015; Deschamps and van Nes, 2005) (Fig. 2D). Similarly, members of the *HOXA*, *HOXC* and *HOXD* complexes expressed early in development with anterior borders of expression in the hindbrain or anterior spinal cord were already activated at D2: *HOXA4* and *HOXA5* were expressed at low levels (Fig. 2D, Table S1) whereas *HOXC4* and *HOXC5* were below the detection level but detectable using real time PCR (see Fig. 3B,C). HOX genes regionally expressed along the brachial and thoracic spinal cord (Fig. 1A, Fig. S1) then displayed a temporal co-linear activation (Fig. 2D). At D2, the progenitors expressed low levels of *HOXC6* that increased by D3; *HOXC8* was expressed at higher levels than *HOXC9* on D3 (Fig. 2D). The expression of *HOXC8* and *HOXC9* further

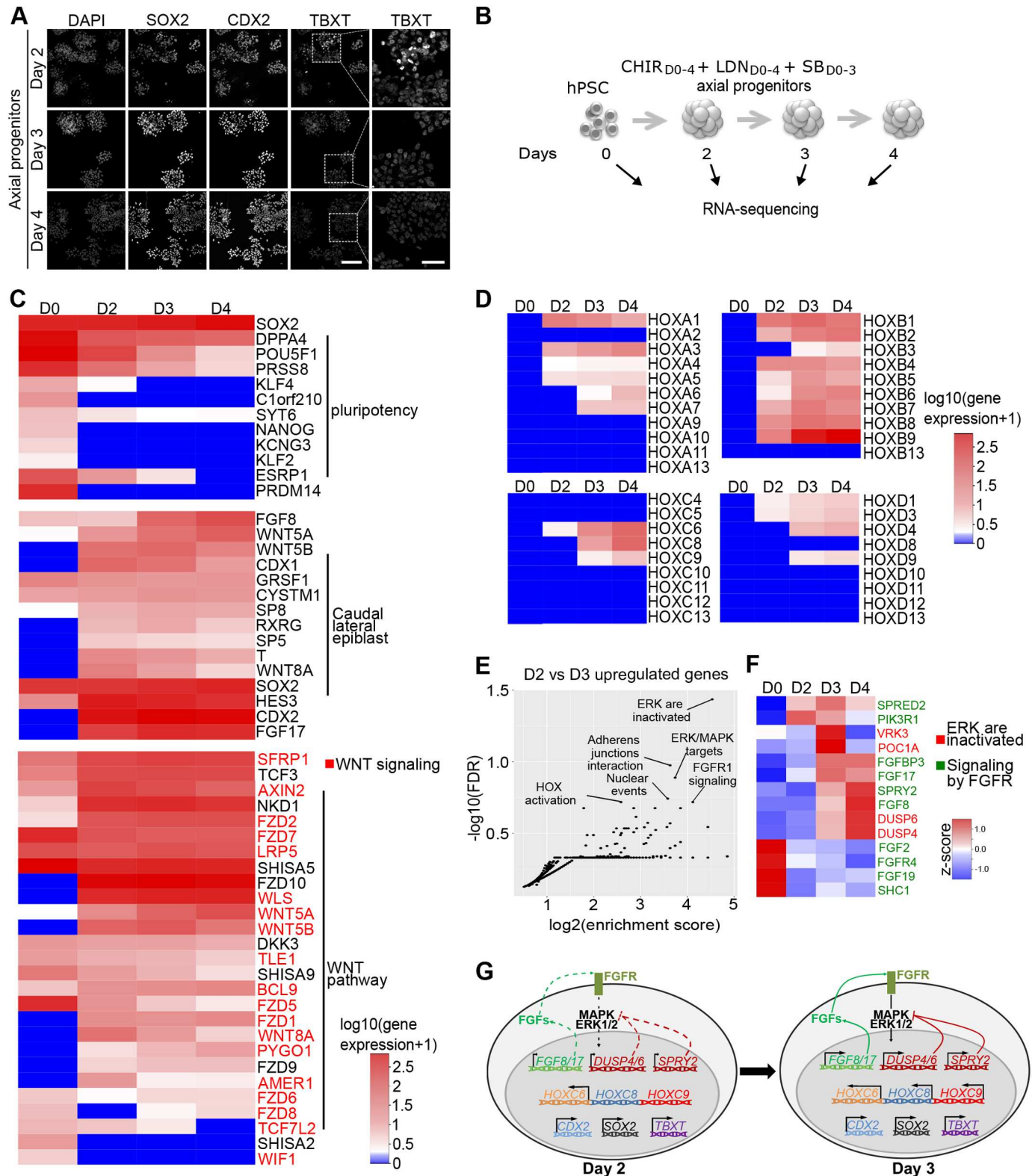


Fig. 2. Temporal transcriptomic analysis of hPSC-derived axial progenitors. (A) Immunostaining for axial progenitor markers on hESC-derived progenitors at day 2, day 3 and day 4 of differentiation. Scale bars: 40 μ m. (B) Experimental design of RNA-seq experiment to profile the transcriptome of hESC (SA001)-derived axial progenitors. *n*=2 per sample. (C) Heatmap of gene expression [$\log_{10}(\text{gene expression} + 1)$] for pluripotency genes, most common mouse NMPs markers and Wnt pathway-related genes. Genes in red are part of the 'Formation β -catenin:TCF transactivating complex' annotation found enriched in reactome pathway analysis in Fig. S3E,F. (D) Heatmap showing temporal transcriptional changes [$\log_{10}(\text{gene expression} + 1)$] of all HOX genes. (E) Functional enrichment analysis (Reactome pathway) of the 232 genes upregulated twofold ($P < 0.05$) between D3 and D2; y axis, FDR (false discovery rate); x-axis, enrichment score calculated for a given Reactome pathway. (F) Heatmap based on z-score of the genes associated with the annotations 'ERK/MAPK target' (red labels) or 'signaling by FGFR' (green labels) in reactome analysis in 2E. (G) Schematic representation of the transcriptional and immunostaining analysis of day 2 and day 3 progenitors.

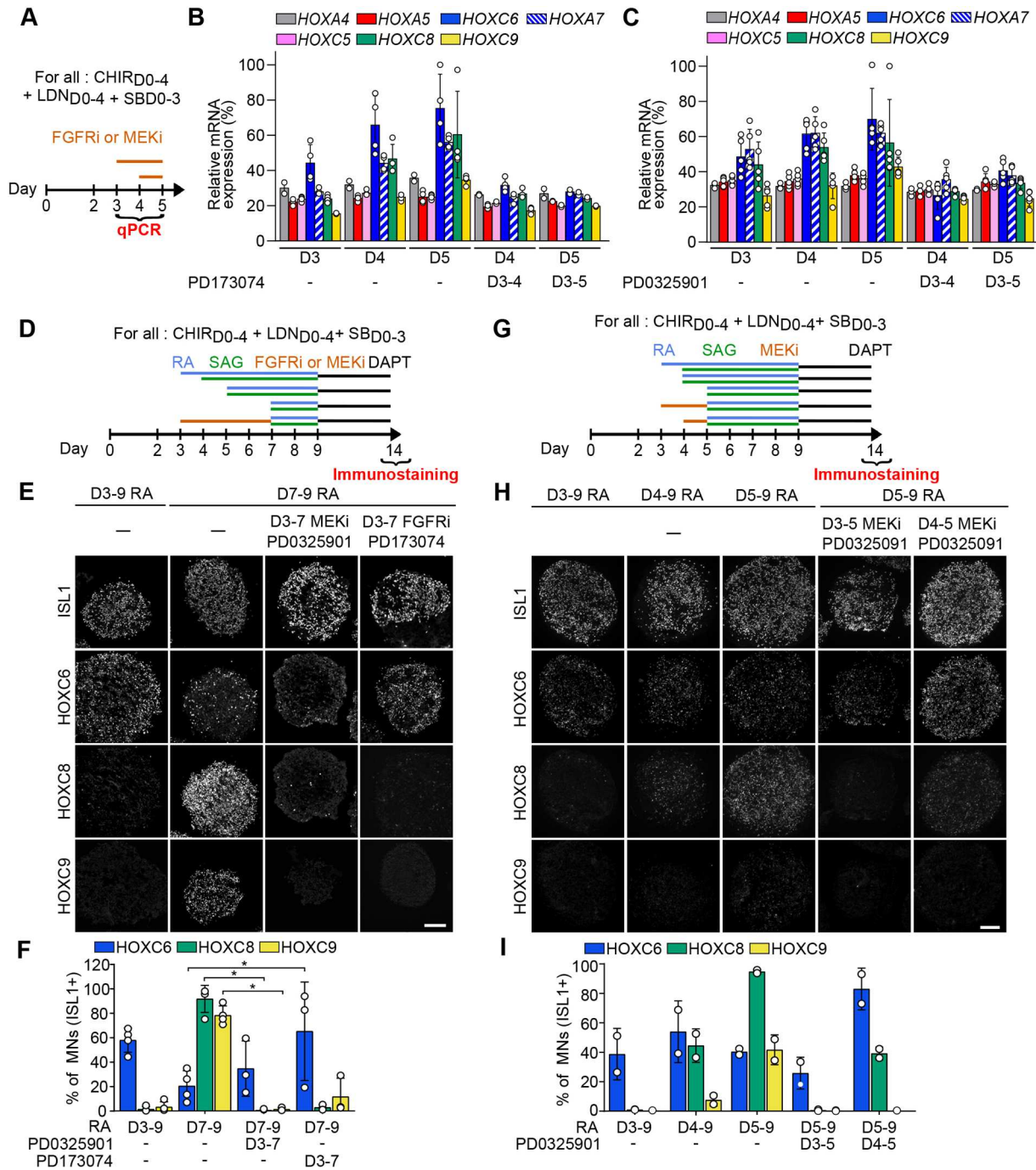


Fig. 3. FGF and MEK pathway inhibitors stall temporal induction of HOX genes in axial progenitors and prevent specification of caudal motor neuron subtypes. (A) Differentiation conditions. PD173074 (FGFR1-3 inhibitor) or PD032590 (MEK1/2 inhibitor) were added on day 3 up to day 5 and hESC-derived progenitors were collected from day 3 to 5 for qPCR analysis. (B,C) Real-time PCR analysis of HOX mRNAs in day 3, 4 and 5 progenitors. Data represent the expression of the different genes relative to the highest expressed gene for all time points and conditions (*HOXC6*). MEK1/2 and FGFR inhibitors, applied from day 3, prevent the temporal increase in caudal HOX expression. (D,G) Differentiation conditions. PD173074 (FGFR1-3 inhibitor; FGFRi) or PD032590 (MEK1/2 inhibitor; MEKi) was added on day 3 up to day 7 (D) or from day 3 or day 4 until day 5 (G) and hESC-derived MNs were collected at day 14 for immunostaining analysis. Retinoic acid (RA) and SAG (an agonist of the sonic hedgehog pathway) were added at the indicated time points, between day 3 and day 7. (E,H) Immunostaining for ISL1 (MNs) and HOX transcription factors on cryostat sections of embryoid bodies on day 14 of differentiation, according to conditions presented in D for E and in G for H. MEK and FGFR inhibitors prevent the specification of *HOXC8*⁺ and *HOXC9*⁺ MNs. Instead, *HOXC6*⁺ MNs are generated. Scale bars: 100 μm. (F,I) Quantification of *HOXC6*, *HOXC8* and *HOXC9* MNs on day 14 of differentiation, according to conditions presented in D for F and in G for I. Data are mean±s.d. Each circle is an independent biological replicate: (B) *n*=4, (C) *n*=5, (F) *n*=3 or 4 and (I) *n*=2. **P*≤0.05 (ANOVA with Kruskal-Wallis post-hoc test).

increased by D4 (by 3.8- and 3.4-fold, respectively). In agreement, this progenitor state will give rise to *HOXC8* and some *HOXC9* MNs when exposed to RA/SAG (Fig. 1C,D). Non-expressed HOX genes

corresponded to the 10 to 13 paralog groups – the latest and most caudally expressed genes (Gaut, 1991; Izpisua-Belmonte et al., 1991; Philippidou and Dasen, 2013; Fig. 1A, Fig. S1). Hence, WNT

activation induced a temporal co-linear activation of HOX genes, which prefigured the generation of MNs of progressively more caudal identities upon incremental delays of RA/SAG treatment (Fig. 1).

To define the pathways activated in parallel to the sequential induction of HOX genes, we performed functional enrichment analysis on the upregulated genes between D2 and D3 when brachial and thoracic spinal HOX genes start to be induced [$\log_2(\text{FC}) \geq 1.0$, $P\text{-value} \leq 0.05$, 232 genes]. We detected a significant enrichment for target genes associated with an activation of the mitogen-activated protein kinase (MAPK) pathways (Fig. 2E). Indeed, typical target genes of the MAPK ERK1/2 pathway (such as *DUSP4*, *DUSP6*, *SPRED2* and *SPRY2*) gradually increased from D2 to D4 (Fig. 2E,F), which was confirmed by real-time PCR in hiPSC-derived progenitors (Fig. S5K). MAPKs are classical mediators of FGF signaling (Lunn et al., 2007). In agreement, we observed a temporal increase of *FGF8* and *FGF17* expression, two secreted FGF ligands, previously described to increase over time in the caudal epiblast of chick and mouse embryos (Fig. 2F,G) (Liu et al., 2001; Wymeersch et al., 2019).

Hence, the sequential co-linear activation of HOX genes within axial progenitors is paralleled by an increase in FGF ligands and MAPK target genes, suggesting a temporal increase in FGF pathway activity. As FGFs modulate the expression of caudal Hox genes in other systems (Bel-Vialar et al., 2002; Dasen et al., 2003; Liu et al., 2001), we hypothesized that graded paracrine or autocrine FGF signaling might be triggering the sequential induction of HOX genes.

FGF signaling is necessary for HOX sequential activation and caudal MN specification

To test whether endogenous FGF signaling was necessary for the temporal sequence of HOX gene induction and caudal MN subtype specification, we first exposed progenitors to PD173074, a selective inhibitor of FGFR1/3, and monitored the impact on *HOXA* and *HOXC* genes, which undergo temporal activation in axial progenitors (Fig. 3A,B, Fig. S6A). Inhibition of FGFR1/3 for 24 h blocked the temporal increase of *HOXC6*, *HOXA7*, *HOXC8* and *HOXC9* but did not impact the expression of anteriorly expressed *HOXA4*, *HOXA5* and *HOXC5*, which are not temporally activated in this time-frame (Fig. 3A,B). We then tested whether this effect was mediated by MEK1/2, a downstream effector of the FGFR pathway. Exposing D3 progenitors to PD0325901, a selective MEK1/2 inhibitor, for 24 h or 48 h also blocked the temporal increase of *HOXC6*, *HOXA7*, *HOXC8* and *HOXC9* (Fig. 3C, Fig. S6A). MEK inhibition even induced a drop in expression of all HOX genes, including *HOXA4*, *HOXA5* and *HOXC5*, that was less pronounced at 48 h than at 24 h (Fig. 3C, Fig. S6B). This result suggests a stronger impact of MEK inhibition on HOX genes, potentially reflecting a more pronounced and rapid inhibition of pathway activity when targeting downstream effectors.

To determine the functional consequences of the stalled HOX temporal induction, we monitored MN subtype specification after FGFR1/3 and MEK1/2 inhibition (Fig. 3D-I, Fig. S6). First, the efficiency of MN generation was not impacted by the two inhibitors in the different tested conditions, even though we collected a reduced number of EBs in the D3-D7 PD173074 condition (Fig. S6D,F). Exposure to RA/SAG on D7 normally leads to the specification of HOXC8 and C9 MNs (Figs 1B-D and 3D,F,H). Addition of PD173074 or PD0325901 from D3 to D7 followed by RA/SAG at D7 fully prevented the generation of HOXC8/9 MNs (Fig. 3E,F). Instead, HOXC6 MNs were specified, an identity normally obtained when D3 progenitors are exposed to RA (Figs 1B,C and 3D-F). A similar result was obtained with only a 48 h inhibition from D3 to D5 followed by RA on D5, which also gave rise to HOXC8/9 MNs in

absence of inhibitor (Fig. 3E,G-I, Fig. S6B,C). These results agreed with the view that FGF pathway inhibition stalled the progression of the HOX clock, resulting in the specification of MNs normally obtained from younger progenitors. To further verify this hypothesis, we inhibited the pathway 1 day later, at D4, followed by RA on D5. In contrast to D3 inhibition, this later inhibition did not prevent HOXC8 MN specification but blocked selectively HOXC9 MNs specification (Figs 1B and 3G-I). Altogether, these results demonstrate that autocrine and/or paracrine FGF signaling is necessary for HOX temporal induction within axial progenitors and the subsequent specification of caudal brachial and anterior thoracic MNs.

FGF level regulates the pace of the HOX clock and MN subtype specification

Our transcriptomic analysis indicated a temporal increase in FGF ligand gene expression, suggesting that an increase in FGF concentration and/or duration of exposure might be pacing the HOX clock. To test whether modulating the level of environmental FGF was sufficient to modify HOX induction, we exposed early D3 progenitors to RA/SAG, together with FGF2 or FGF8 at different concentrations and for different durations. In all conditions that received FGF2, more caudal MN identities were induced without the need for delaying RA addition (Fig. 4A-D, Fig. S7). In addition, the extent of caudalization varied with FGF2 concentration, with increasing concentrations promoting more caudal identities: caudal brachial HOXC8⁺ MNs were induced at 15 ng/ml (68.2%, 8.6-fold increase) and HOXC9⁺/HOXC6⁻ thoracic MNs at 60 ng/ml (58.9%, 825-fold increase), while 120 ng/ml further reduced HOXC6⁺ MNs (Fig. 4A-D, Fig. S7A-G). Similar results were obtained with FGF8 (Fig. S7B). FGFs acted directly and rapidly on axial progenitors, as addition of 120 ng/ml FGF2 for 24 h or 48 h induced MNs of a caudal brachial or mid-thoracic identity (after 24h, 49.7% of the MNs were HOXC6⁺, 79.4% HOXC8⁺, 14.7% HOXC9⁺ and 47.9% FOXP1^{high}/SCIP⁺. After 48 h, 7.5% were HOXC6⁺, 76.4% HOXC9⁺ and 13.2% FOXP1^{high}/SCIP⁺; Fig. 4A-C, S7C-I).

To determine whether this caudalizing effect was linked to an accelerated induction of brachial and thoracic HOX genes, we performed real-time PCR analysis 24 h and 48 h post-FGF2 treatment. We indeed observed an earlier increase of *HOXC8*, *HOXC9* and *HOXD9* mRNA expression compared with RA/SAG controls (Fig. 4F).

Hence, a precocious rise in FGF signaling accelerates the induction of caudal HOX genes in early axial progenitors, resulting in the specification of more caudal cell types within the same differentiation timeline (14 days). These results demonstrate that the levels or duration of FGF signaling dynamically regulate the pace of HOX co-linear activation in axial progenitors.

FGF and GDF11 synergize to further accelerate the HOX clock

FGF2 accelerated HOX induction to generate MNs up to the mid-thoracic level, suggesting that early axial progenitors might lack competency to generate the most caudal segments. Alternatively, other extrinsic factors might be required to further accelerate the induction of HOX genes and promote more caudal identities. GDF11 is a member of the TGF β family implicated in the control of axial elongation and MN subtype specification, and is required for the expression of the most 5' HOX gene starting from the 10 paralogs *in vivo* and *in vitro* in late NMP-like cells (Aires et al., 2019; Gaunt et al., 2013; Lippmann et al., 2015; Liu, 2006; Liu et al., 2001; McPherron et al., 1999; Peljto et al., 2010). We exposed D3

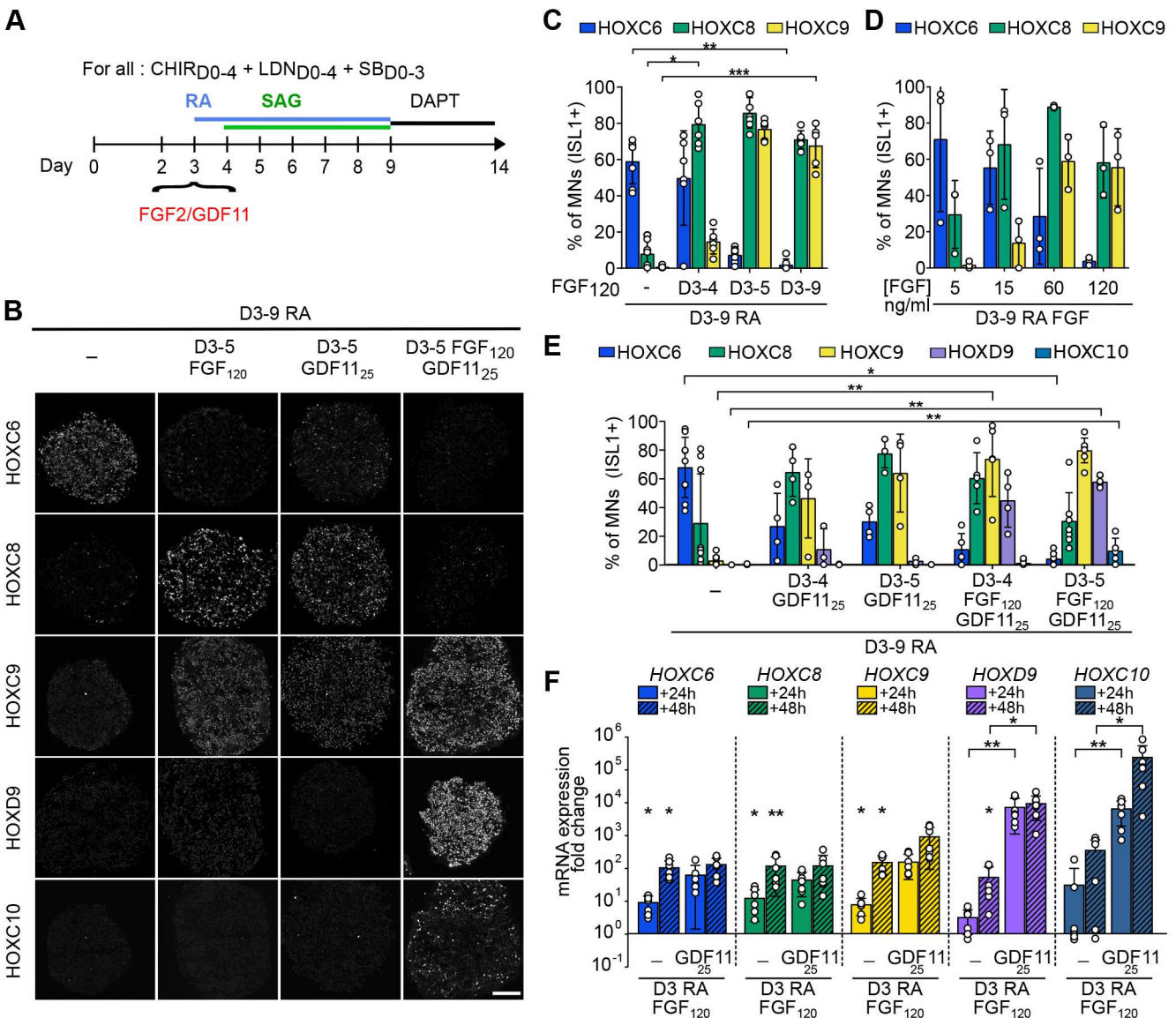


Fig. 4. Dynamic pacing of HOX induction in axial progenitors by changes in extrinsic FGF2 and GDF11 levels. (A) Differentiation conditions. Extrinsic cues, FGF2, GDF11 or FGF2+GDF11 were added on day 3 of differentiation at various concentrations or for different durations. (B) Immunostaining for HOX proteins on cryostat sections of hESC-derived EBs on day 14 of differentiation. FGF2, GDF11 and FGF2/GDF11 induce more caudal MN subtypes. Scale bar: 100 μ m. (C-E) Proportion of MNs (ISL1⁺ cells) expressing the indicated markers. The effect of the duration of FGF2 treatment (C), FGF2 concentration (D) and duration of GDF11 or FGF2+GDF11 (E) were monitored. (F) Real-time quantitative PCR analysis of the expression of HOX genes regionally expressed in human MNs *in vivo*. HOX mRNAs were monitored at day 4 (24 h post-treatment) and day 5 (48 h post-treatment) upon addition of FGF2 or GDF11, or a combination of FGF2 (120 ng/ml) and GDF11 (25 ng/ml). Data are expressed as fold changes to their respective control [day 4 retinoic acid (RA) 24 h or day 5 RA 48 h]. Asterisks above the horizontal lines indicate significance of statistical comparison between the indicated conditions; asterisks above the histogram bars are for statistical comparison with the control. Data are mean \pm s.d. Each circle is an independent biological replicate: (C) $n=6$ or 7, (D) $n=3$, (E) $n=3-9$, (F) $n=6$. * $P\leq 0.05$, ** $P\leq 0.01$, *** $P\leq 0.001$ (ANOVA with Kruskal-Wallis post-hoc test).

progenitors to a combination of RA and GDF11 (25 ng/ml; Liu et al., 2001) for 24 h or 48 h. A 24 h GDF11 treatment induced caudal brachial and anterior thoracic MNs (64.7% HOXC8⁺ and 46.4% HOXC9⁺), the latter increasing after a 48 h treatment (Fig. 4A,B,E, Fig. S8A,B). However, as with FGF2, none of the more caudal identities was observed. In chick, exposure of spinal cord explants to combination of FGF2 and GDF11 promoted more caudal MNs than the two factors separately (Liu et al., 2001). We thus tested whether this combination might accelerate the HOX clock and promote caudal thoracic or lumbar identities. 24 h after exposure of D3 axial progenitors to FGF2 and GDF11, HOXC9, HOXD9 and HOXC10 mRNAs were strongly induced (respectively 65.4-, 2774.77- and

3329.44-fold compared with controls) and further increased after 48 h (Fig. 4F). In agreement, a 24 h (D3-D4) treatment combined with RA led to the specification of caudal thoracic MNs (73.9% HOXC9⁺ and 45% HOXD9⁺) whereas lumbar HOXC10⁺ MNs were observed after a 48 h FGF/GDF11 treatment (10.1% HOXC10⁺, 58% HOXD9⁺) (Fig. 4B,E, Fig. S8B-F). Importantly, adding FGF2/GDF11 prior RA (FGF/GDF11 at D3 followed by RA D5) largely increased the generation of lumbar MNs (81% HOXC10⁺, 89% HOXD9⁺) (Fig. S8G-I). This result suggests that RA antagonized the caudalizing activity of FGF2/GDF11, potentially by repressing FGF8 gene expression as previously shown (Del Corral and Storey, 2004; Del Corral et al., 2003) and/or by promoting a differentiation stage at

which HOX expression is stalled. Hence, a combination of FGF2 and GDF11 further accelerates the HOX clock in young D3 progenitors to induce late and caudally expressed *HOX10* genes, resulting in the generation of lumbar MNs.

DISCUSSION

In this study, using axial progenitors derived from hPSCs, we demonstrate that the parameters of exposure to two extrinsic factors, FGFs and GDF11, dynamically regulate the speed at which the HOX clock proceeds. Extrinsic control of the clock results in the synchronous specification of progenies of distinct rostro-caudal identities, which were born at different stages of embryogenesis. These results support the view that the pace of the HOX clock is largely regulated by extrinsic factors rather than by an intrinsic timer.

To reach these conclusions, we used 3D, embryoid body-based differentiation, of hPSC to selectively generate axial progenitors, an identity demonstrated by the co-expression of SOX2, TBXT and CDX2, as observed in mouse, chick and human embryos, complemented by detailed transcriptomic analysis and, more importantly, by functional analyses showing their ability to generate cell types found along the rostro-caudal axis of the human embryonic spinal cord (Figs 1, 2 and Figs S1-S5). Axial progenitors feed axial elongation in the embryo by generating diverse neural and mesodermal derivatives. They appear to form a heterogeneous dynamic population of progenitors comprising progenitors differentiating into diverse neural and mesodermal progenies (lateral, axial and paraxial), bi-fated NMPs that generate both spinal and paraxial mesoderm, as well as mono-fated progenitors that give rise either to mesodermal or spinal progenies (Albors et al., 2018; Forlani et al., 2003; Henrique et al., 2015; Tzouanacou et al., 2009; Wymeersch et al., 2016, 2019). Defining the precise degree of similarity between the reported hPSC-derived axial progenitors with these distinct states or with previously reported *in vitro*-generated axial progenitors and, in particular, whether individual progenitors are mono-fated or bi-fated NMPs will require further investigations (Wymeersch et al., 2021).

With efficient access to axial progenitors, we demonstrated that modulating the concentration, duration and combination of the caudalizing factors FGFs and GDF11 controls the speed at which the temporal activation of HOX genes occurs. The pace of the HOX clock is thus dynamically controlled by exposure parameters to extrinsic cues. Hence, the sequential changes in chromatin structure occurring along HOX complexes during their activation might be a consequence of changes in extrinsic signals rather than in the main mechanism intrinsically controlling the timing of HOX induction. (Bel-Vialar et al., 2002; Del Corral and Storey, 2004; Kimelman and Martin, 2012; Kmita and Duboule, 2003; Lippmann et al., 2015; Mazzoni et al., 2013; Narendra et al., 2015; Neijts and Deschamps, 2017; Noordermeer et al., 2011, 2014; Soshnikova and Duboule, 2009; Tschopp et al., 2009). As the sequence of activation from 3' to 5' genes was maintained under the different experimental conditions, the extrinsic cues seem to modulate the speed but not the directionality at which a repressive chromatin state is cleared from HOX clusters. This clearance could be temporally progressive or, alternatively, individual extrinsic factors might induce domain-wide, saltatory remodeling of repressive chromatin marks (up to *HOXC9* in presence of FGFs, up to *HOXD9/HOXC10* or more caudally in presence of FGF and GDF11) (Kmita and Duboule, 2003; Mazzoni et al., 2013; Narendra et al., 2016; Soshnikova and Duboule, 2009). Individual HOX genes within these transcriptionally permissive domains might be activated in a 3' to 5' gene direction due to their reliance on different combinations of transcription factors or

differential strength of enhancer-promoter interactions (Kmita and Duboule, 2003; Mazzoni et al., 2013; Neijts and Deschamps, 2017). Determining the changes in histones marks, chromatin structure and the binding of signaling effectors within HOX complexes should help determine how extrinsic cues signal to the genome to control rostro-caudal patterning.

Dynamic environmental control of Hox expression was already suggested by grafting experiments in chick (Ensigni et al., 1998; McGrew et al., 2008). In particular, it was demonstrated that heterochronic grafting of 'old' axial progenitors to a 'younger' caudal stem zone reverted their HOX profile to the 'young' one (McGrew et al., 2008). Although our results might provide a potential molecular basis for environmental changes in HOX expression in caudal progenitors, we did not observe a reversion of the HOX profile upon FGF pathway inhibition but a stalled activation (Fig. 3E-I). This observation suggests that the FGF pathway regulates the tempo at which HOX activation proceeds rather than the maintenance of the HOX profile.

Importantly, the pacing of the HOX clock by secreted factors might ensure a community effect to synchronize HOX expression between neighboring progenitors, allowing the emergence of expression domains at tissue level (Durstion, 2019). Furthermore, FGFs and GDFs are also implicated in embryo axial elongation (Aires et al., 2019; Boulet and Capecchi, 2012; Jurberg et al., 2013; Mallo et al., 2009; McPherron et al., 1999). A common mechanism to control rostro-caudal extension of the body axis together with Hox gene induction would be a parsimonious way to couple morphogenesis and patterning (Denans et al., 2015; Young et al., 2009). Hence, determining the mechanisms controlling the temporality of FGF and GDF expression onset, levels and duration should provide a better understanding of post-occipital tissues development and evolution.

In a bioengineering perspective, the extrinsic control of the HOX clock provides a means with which to manipulate HOX expression for the synchronous and efficient engineering of neuronal subtypes found at distinct rostro-caudal identities. In particular, we provide evidence for the controlled generation of putative hand- and leg-controlling MNs (Mendelsohn et al., 2017). Importantly, MN subtypes display differential vulnerabilities in MN diseases. The targeted differentiations in MN subtypes reported here should thus help in the more accurate study of these incurable diseases or allow the access of more controlled sources of cells for putative cell therapy approaches (Abati et al., 2019; An et al., 2019; Baloh et al., 2018; Nijssen et al., 2017; Ragagnin et al., 2019; Sances et al., 2016; Steinbeck and Studer, 2015; Tung et al., 2019). Considering the role of HOX transcription factors in instructing cell diversity in other derivatives of axial progenitors, such as control over their expression, might have broader applications for cell engineering beside spinal cell types (Deschamps and Duboule, 2017; Frith et al., 2018; Helmbacher et al., 2003; Iimura et al., 2009). More broadly, the temporal generation of distinct types of neurons or glia from the same progenitor domain is a widely used strategy to increase cell diversity in the nervous system (Dias et al., 2014; Kohwi and Doe, 2013; Oberst et al., 2019a; Rossi et al., 2017). Extrinsic cues play important roles in the unfolding of these temporal sequences (Kawaguchi, 2019; Oberst et al., 2019a,b; Syed et al., 2017; Tiberi et al., 2012). Extrinsic manipulation of the temporality of these lineages might help improve the generation of early and late-born cells for both basic research, disease modeling and cell therapy.

MATERIALS AND METHODS

Human embryonic spinal cord histology

Human fetal embryos at 6.3 weeks of gestation were obtained from pregnant women referred to the Department of Gynecology and Obstetrics at the

Antoine B  cl  re hospital (Clamart, France) for legally induced abortions in the first trimester of pregnancy, as previously described (Lambrot et al., 2006). All women provided written informed consent for scientific use of the fetal tissues. None of the abortions was due to fetal abnormality. The fetal age was determined by measuring the length of limbs and feet (Evtouchenko et al., 1996). The project was approved by the local Medical Ethics Committee and by the French Biomedicine Agency (reference number PFS 12-002). Alternatively, human embryonic spinal cords ($n=2$) at stage 7.5 weeks of gestation were collected in accordance with the national guidelines of the USA (National Institutes of Health, US Food and Drug Administration) and the State of New York and under Columbia University institutionally approved ethical guidelines relating to anonymous tissue. The material was obtained after elective abortions and was classified on the basis of external morphology according to the Carnegie stages. Gestational age was determined by the date of the last menstrual period of the patient or by ultrasound, if the ultrasound estimated date differed by 1 week (as indicated by the obstetrician). In all cases, the spinal cord was removed as intactly as possible before fixation with fresh, ice-cold 4% PFA for 1.5 h, washed abundantly with PBS and then cryoprotected overnight in 30% sucrose. Post-fixation, the cord was measured and cut into anatomical sections to accommodate embedding in OCT compound (Leica) and stored at -80°C before cutting on a cryostat. Sections ($16\ \mu\text{m}$) were cut along the full length of the cord.

Human pluripotent stem cell lines

Human SA001 embryonic stem cell (ESC) line (male, RRID: CVCL_B347) was obtained from Collectis and H9 ESC line (female, RRID: CVCL_9773) was obtained from Wicell. Both cell lines were used according to the French current legislation for hESC (Agency of Biomedicine, authorization number AFSB1530532S). The induced pluripotent stem cell (iPSC) line WTSli002 (male, RRID: CVCL_AH30, alternative name HPSi0913i-eika_2) was obtained from the European Bank for Pluripotent Stem Cells (EBISC). WTC-mEGFP-Safe harbor locus (AAVS1)-cl6 produced by the Allen Cell Institute was obtained from Coriell (AICS-0036-006, male, RRID: CVCL_JM19). Experiments with iPSCs were approved by relevant ethic committees (declaration DC-2015-2559). All PSC lines were cultured at 37°C on Matrigel (Corning) in mTSE1 medium (Stem Cell Technologies) and amplified using EDTA (Life Technologies) clump-based passaging. They were tested for potential mycoplasma contamination every other week (MycAlertTM Mycoplasma Detection Kit, Lonza, LT07-118). No contamination was detected during the study. PSCs were thawed in presence of Y-27632 ($10\ \mu\text{M}$, Stemgent or Stem Cell Technologies) and the culture medium was changed every day.

Human pluripotent stem cell differentiation

Human PSC embryoid body-based differentiation was performed as previously described (Maury et al., 2015). hPSC were dissociated with cold Accutase (Life Technologies) for 3-5 min at 37°C and resuspended in differentiation medium N2B27 [Advanced DMEM F12, Neurobasal vol:vol (Life Technologies)], supplemented with N2 (Life Technologies), B27 without Vitamin A (Life Technologies), penicillin/streptomycin 1%, β -mercaptoethanol 0.1% (Life Technologies), Y-27632 ($10\ \mu\text{M}$, Stemgent or Stem Cell Technologies), CHIR-99021 ($3\ \mu\text{M}$ or $4\ \mu\text{M}$ Selleckchem), SB431542 ($20\ \mu\text{M}$, Selleckchem) and LDN 193189 ($0.1\ \mu\text{M}$, Selleckchem). Cells were seeded in ultra-low attachment 6 well plates (Corning) (2×10^5 cells ml^{-1}) to form embryoid bodies (EBs). All conditions of differentiation received the same medium, without Y-27632, at day 2 and at day 3, but SB431542 was removed at day 3. The differentiation then proceeded according to the schematics presented in Fig. 1B, Fig. 3A,D,G, Fig. 4A. SAG (smoothed agonist, Merck millipore), FGF2 (recombinant human FGF basic, Peprotech), RA (Sigma-Aldrich), GDF11 (recombinant human/murine/rat GDF11, Peprotech), PD0325901 (Selleckchem), PD173074 (Selleckchem) and DAPT (Stemgent or Stem Cell Technologies) were added at indicated time points. For concentrations used, see Table S5. Media were changed every other day unless specified.

Embryoid body processing for immunostaining

EBs were collected, rinsed with PBS then fixed with 4% PFA for 5 min at 4°C and rinsed with PBS three times for 5 min. EBs were cryoprotected with

30% sucrose and embedded in OCT (Leica) prior to sectioning with a cryostat. Alternatively, day 2, 3 and 4 progenitors were plated on Matrigel (Corning, diluted according to the manufacturer's recommendation)-coated coverslips and allowed to adhere between 30 and 60 min prior to fixation with 4% PFA for 5 min at 4°C .

Immunostaining

All immunostainings were performed as follows: cells or sections were incubated with a saturation solution (PBS, 10% FBS, 0.2% Triton) for 10 min. Primary antibodies (Table S5) were diluted in a staining solution (PBS, 2% FBS, 0.2% Triton) and incubated overnight at 4°C in a humidified chamber. After four PBS washes (10 min each), secondary antibodies (Alexa488, Alexa555 and Alexa647, Life Technologies, 1:1000) were added for 1 h at room temperature. After three PBS washes, DAPI was added to cells (Invitrogen, 1:3000) for 5 min. Cells or slices were then mounted in Fluoromount (Sigma-Aldrich or Cliniscience).

Image acquisition

Samples were visualized and imaged using either a Zeiss LSM 880 Confocal Laser Scanning Microscope controlled by Zen black software (Zeiss), a confocal microscope TCS SP5 II (Leica) or a DM6000 microscope (Leica) equipped with CoolSNAP EZ CDD camera, controlled by MetaMorph software. Alternatively, images were acquired using the automated microscope Cell Discoverer 7 (Zeiss), equipped with an AxioCam 506m camera, with Zen black software (Zeiss).

Quantitative RT-PCR analysis

Total RNA was extracted (RNeasy Plus Mini Kit, Qiagen) and cDNA synthesized using SuperScript III (Invitrogen). Quantitative real-time PCR was performed using a 7900HT fast real-time PCR system (Applied Biosystems) with Sybr Green PCR Master Mix (Applied Biosystems) or performed using QuantStudio 5 Real-Time PCR System (ThermoFisher Scientific) and a mix with qPCR Brilliant II SYBR MM with low ROX (Agilent). Primers are listed in Table S5. All expression data were normalized to cyclophilin A mRNA (ΔCt). All analyses were performed with three technical replicates per plate. In Fig. 3 and Fig. S5, mRNA expression levels are represented as a percentage of the highest expressed gene ($\Delta\text{Ct}_{\text{max}}$) among all the conditions [% relative to max gene expression = $(\Delta\text{Ct}_{\text{max}}/\Delta\text{Ct}) \times 100$]. In Fig. 4F, expression levels are represented as a percentage relative to the maximal expression level of each gene ($\Delta\text{Ct}_{\text{geneMax}}$) among all conditions [% relative to max gene expression = $(\Delta\text{Ct}_{\text{geneMax}}/\Delta\text{Ct}) \times 100$]. In Fig. S4, data are presented as relative expression levels determined by calculating $2^{-\Delta\Delta\text{Ct}}$.

Transcriptomic analysis

hESCs were exposed to CHIR-99021 together with LDN193189 and SB431542 while forming EBs. Progenitors were collected on days 2, 3 and 4 of differentiation, and processed for total RNA preparation. For each of the eight samples, total RNA was reverse transcribed using the Ion AmpliSeq Transcriptome Human Gene Expression kit (ThermoFisher Scientific). The cDNA libraries were amplified and barcoded using the Ion AmpliSeq Transcriptome Human Gene Expression core panel and Ion Xpress Barcode Adapter (ThermoFisher Scientific). The amplicons were quantified using Agilent High Sensitivity DNA kit before the samples were pooled in sets of eight. Emulsion PCR and enrichment was performed on the Ion OT2 system Instrument using the Ion PI Hi-Q OT2 200 kit (ThermoFisher Scientific). Samples were loaded on an Ion PI v3 Chip and sequenced on the Ion Proton System using Ion PI Hi-Q sequencing 200 kit chemistry (200 bp read length; ThermoFisher Scientific). The Ion Proton reads (FASTQ files) were imported into the RNA-seq pipeline of Partek Flow software (v6 Partek) using hg19 as a reference genome. The number of reads per sample ranged from 7.5 million to 12 million reads. To determine genes that are differentially expressed between groups, mapped reads were quantified using Partek E/M algorithm normalized by the Total count/sample (the resulting counts represent the gene expression levels on reads/millions for over 20,800 different genes present in the AmpliSeq Human Gene Expression panel). Genes with an average of reads ≤ 1.0 in all the time points were excluded from the analysis. The evaluation of the differential

expression between two conditions was performed using the EdgeR package under R. Pathway enrichment analyses were performed on upregulated genes ($FC \geq 2.0$, P -value < 0.05) between two time points by interrogating Reactome database. Hierarchical clustering was performed using Ward's method with the ward.d2 algorithm. The Euclidean distance is represented. Significant enrichments were calculated using a hypergeometrical test and Benjamini-Hochberg correction for multiple comparisons. The enrichment score was calculated as described by Wang et al. (2017).

The normalized transcriptomic data are provided in Table S1. The GEO accession number for the raw data is GSE153519. Lists of genes differentially expressed between time points are available in Table S2. Lists of common genes between D3 enriched genes and NMPs-like genes have been identified by Frith et al. (2018) and are available in Table S3. Expression of the genes in our model that have been identified in NMP cells at E8.5 and E9.5 by Gouti et al. (2017) are listed in Table S4.

Quantification and statistical analysis

All statistics were computed using Graphpad Prism software. One-way analysis of variance (ANOVA) with a Kruskal-Wallis post-hoc analysis was performed following normality tests provided by Prism. Number of samples (n), dispersion measures and P -values are indicated in figure legends. In all figures, n are independent differentiations started from independent newly thawed hPSCs vials. For each condition at least four independent EB sections were imaged in which all the cells were quantified by automated image analysis (four EBs were quantified in one replicate of the FGFRi 3-7 conditions; in all the other experiments 6 to 9 EBs were quantified). The images were exported and saved as TIFF with Fiji if needed (Schindelin et al., 2012). Quantitative analyses on images were performed using the CellProfiler software (Carpenter et al., 2006) (Broad Institute open source at www.cellprofiler.org). DAPI-stained nuclei were segmented into primary objects using the CellProfiler segmentation pipeline and the nuclear mask was used to define objects on the target channels. The threshold for defining positive nuclei for a given target was obtained using EB sections negative for the target of interest. All images across conditions were then automatically analyzed in batch to ensure unbiased analysis. The percentage of colocalization between markers (ISL1/HOXC6/C8, ISL1/C8/C9, ISL1/C9/C10) was obtained after co-immunostaining for the markers of interest followed by quantification in Cell profiler. The analyses of the FOXP1 and SCIP immunostaining intensities were performed by combining CellProfiler with the software FCS express 7 (DeNovo Software). Nuclei were segmented into primary objects as described above, and FOXP1 and SCIP fluorescence intensities were calculated for each primary object. Fluorescence intensity plots for FOXP1 and SCIP were then generated by FCS express 7 software to visualize the intensity levels of the different markers for each individual cell and determine the percentage of cells above a given threshold. Cell profiler pipelines for quantification are available upon request. For real-time PCR analysis, one-way analysis of variance (ANOVA) with a Kruskal-Wallis post-hoc analysis and correction of multiple comparisons by controlling the FDR (Benjamini and Hochberg) was performed following normality tests provided by Prism. Number of samples (n), dispersion measures and P -values are indicated in figure legends.

Acknowledgements

We thank Frédéric Causeret for providing sections of mouse embryos; Susan Morton, Thomas Jessell and Jeremy Dasen for the generous gift of antibodies; and F. Causeret, P. Gilardi, A. Rebsam, F. Giudicelli and E. Mazzoni for critical reading of the manuscript. Differentiation were performed at the 'Cell engineering facility' of the Institut du Fer à Moulin (IFM) and at I-STEM. Imaging experiments were carried out at the Imaging facility of the Institut du Fer à Moulin (IFM) and at I-STEM.

Competing interests

The authors declare no competing or financial interests.

Author contributions

Conceptualization: C.M., S.N.; Methodology: V.M., C.V., R.R., S.G., V.R., S.N.; Software: S.G.; Validation: V.M., C.V., R.R., A.T., M.D., C.M., S.N.; Formal analysis: V.M., C.V., R.R., S.G., M.J., M.D., S.N.; Investigation: V.M., C.V., R.R., M.J., A.T.,

L.L., M.W.M., G.C.; Resources: N.N., V.R.-F., S.N.; Data curation: R.R., M.J.; Writing - original draft: V.M., S.N.; Writing - review & editing: V.M., C.V., R.R., S.G., N.N., M.J., A.T., L.L., M.W.M., G.C., M.D., V.R.-F., H.W., V.R., C.M., S.N.; Visualization: V.M., C.V., R.R., S.G., M.W.M., G.C., M.D., S.N.; Supervision: H.W., C.M., S.N.; Project administration: H.W., C.M., S.N.; Funding acquisition: S.N., C.M., H.W.

Funding

This project was supported by grants from AVENIR/ATIP, from the Association Française contre les Myopathies (AFM) and from Laboratoire d'Excellence (Labex) Biopsy and Revive (ANR-10-LABX-73 and 11-LABX-0035, respectively), and from the Agence Nationale de la Recherche SYMASYM (ANR-18-CE16-0021-03) and Atomy (ANR-19-CE16-0002-02). V.M. received fellowships from Domaine d'Investissement Majeur (DIM) 'Biothérapies' of the region Ile-de-France and from the Fondation pour la Recherche Médicale (FRM). C.V. was supported by a PhD fellowship of the Ministère de la Recherche and the Fondation pour la Recherche Médicale (FRM). V.R. and S.N.'s salaries are funded by the Institut National de la Santé et de la Recherche Médicale (INSERM). H.W., M.W.M. and G.C. were supported by Project ALS and by the National Institutes of Health (R01NS109217). Deposited in PMC for release after 12 months.

Data availability

The normalized transcriptomic data are provided in Table S1. FASTQ files are available in GEO under accession number GSE153519.

Supplementary information

Supplementary information available online at <https://dev.biologists.org/lookup/doi/10.1242/dev.194514.supplemental>

Peer review history

The peer review history is available online at <https://dev.biologists.org/lookup/doi/10.1242/dev.194514.reviewer-comments.pdf>

References

- Abati, E., Bresolin, N., Comi, G. and Corti, S. (2019). Advances, challenges, and perspectives in translational stem cell therapy for amyotrophic lateral sclerosis. *Mol. Neurobiol.* **56**, 6703-6715. doi:10.1007/s12035-019-1554-x
- Aires, R., de Lemos, L., N6voa, A., Jurberg, A. D., Mascrez, B., Duboule, D. and Mallo, M. (2019). Tail bud progenitor activity relies on a network comprising Gdf11, Lin28, and Hox13 genes. *Dev. Cell* **48**, 383-395.e8. doi:10.1016/j.devcel.2018.12.004
- Albors, A., Halley, P. A. and Storey, K. G. (2018). Lineage tracing of axial progenitors using Nkx1-2CreER T2 mice defines their trunk and tail contributions. *Development* **145**, dev164319. doi:10.1242/dev.164319
- Amin, S., Neijts, R., Simmini, S., van Rooijen, C., Tan, S. C., Kester, L., van Oudenaarden, A., Creighton, M. P. and Deschamps, J. (2016). Cdx and T Brachyury co-activate growth signaling in the embryonic axial progenitor niche. *Cell Rep.* **17**, 3165-3177. doi:10.1016/j.celrep.2016.11.069
- Amoroso, M. W., Croft, G. F., Williams, D. J., O'Keeffe, S., Carrasco, M. A., Davis, A. R., Roybon, L., Oakley, D. H., Maniatis, T., Henderson, C. E. et al. (2013). Accelerated high-yield generation of limb-innervating motor neurons from human stem cells. *J. Neurosci.* **33**, 574-586. doi:10.1523/JNEUROSCI.0906-12.2013
- An, D., Fujiki, R., Iannitelli, D. E., Smerdon, J. W., Maity, S., Rose, M. F., Gelber, A., Wanaseleja, E. K., Yagudayeva, I., Lee, J. Y. et al. (2019). Stem cell-derived cranial and spinal motor neurons reveal proteostatic differences between ALS resistant and sensitive motor neurons. *eLife* **8**, e44423. doi:10.7554/eLife.44423
- Baloh, R. H., Glass, J. D. and Svendsen, C. N. (2018). Stem cell transplantation for amyotrophic lateral sclerosis. *Curr. Opin. Neurol.* **31**, 655-661. doi:10.1097/WCO.0000000000000598
- Bel-Vialar, S., Itasaki, N. and 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**, 5103-5115.
- Bell, S. W., Brown, M. J. C. and Hems, T. J. (2017). Refinement of myotome values in the upper limb: evidence from brachial plexus injuries. *Surgeon* **15**, 1-6. doi:10.1016/j.surge.2015.08.004
- Bialecka, M., Wilson, V. and Deschamps, J. (2010). Cdx mutant axial progenitor cells are rescued by grafting to a wild type environment. *Dev. Biol.* **347**, 228-234. doi:10.1016/j.ydbio.2010.08.032
- Boulet, A. M. and Capocchi, M. R. (2012). Signaling by FGF4 and FGF8 is required for axial elongation of the mouse embryo. *Dev. Biol.* **371**, 235-245. doi:10.1016/j.ydbio.2012.08.017
- Briscoe, J. and Novitch, B. G. (2008). Regulatory pathways linking progenitor patterning, cell fates and neurogenesis in the ventral neural tube. *Philos. Trans. R. Soc. B Biol. Sci.* **363**, 57-70. doi:10.1098/rstb.2006.2012
- Cambay, N. and Wilson, V. (2002). Axial progenitors with extensive potency are localised to the mouse chordoneural hinge. *Development* **129**, 4855-4866.

- Cambray, N. and Wilson, V.** (2007). Two distinct sources for a population of maturing axial progenitors. *Development* **134**, 2829-2840. doi:10.1242/dev.02877
- Carpenter, A. E., Jones, T. R., Lamprecht, M. R., Clarke, C., Kang, I. H., Friman, O., Guertin, D. A., Chang, J. H., Lindquist, R. A., Moffat, J. et al.** (2006). CellProfiler: Image analysis software for identifying and quantifying cell phenotypes. *Genome Biol.* **7**, R100. doi:10.1186/gb-2006-7-10-r100
- Dasen, J. S.** (2017). Master or servant? emerging roles for motor neuron subtypes in the construction and evolution of locomotor circuits. *Curr. Opin. Neurobiol.* **42**, 25-32. doi:10.1016/j.conb.2016.11.005
- Dasen, J. S., Liu, J.-P. and Jessell, T. M.** (2003). Motor neuron columnar fate imposed by sequential phases of Hox-c activity. *Nature* **425**, 926-933. doi:10.1038/nature02051
- Del Corral, R. D. and 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**, 857-869. doi:10.1002/bies.20080
- Del Corral, R. D., Olivera-Martinez, I., Goriely, A., Gale, E., Maden, M. and Storey, K.** (2003). Opposing FGF and retinoid pathways control ventral neural pattern, neuronal differentiation, and segmentation during body axis extension. *Neuron* **40**, 65-79. doi:10.1016/S0896-6273(03)00565-8
- Delfino-Machín, M., Lunn, J. S., Breikreuz, D. N., Akai, J. and Storey, K. G.** (2005). Specification and maintenance of the spinal cord stem zone. *Development* **132**, 4273-4283. doi:10.1242/dev.02009
- Denans, N., Iimura, T. and Pourquié, O.** (2015). Hox genes control vertebrate body elongation by collinear Wnt repression. *eLife* **4**, e04379.
- Deschamps, J. and Duboule, D.** (2017). Embryonic timing, axial stem cells, chromatin dynamics, and the Hox clock. *Genes Dev.* **31**, 1406-1416. doi:10.1101/gad.303123.117
- Deschamps, J. and van Nes, J.** (2005). Developmental regulation of the Hox genes during axial morphogenesis in the mouse. *Development* **132**, 2931-2942. doi:10.1242/dev.01897
- Dias, J. M., Alekseenko, Z., Applequist, J. M. and Ericson, J.** (2014). Tgfb signaling regulates temporal neurogenesis and potency of neural stem cells in the CNS. *Neuron* **84**, 927-939. doi:10.1016/j.neuron.2014.10.033
- Diaz-Cuadros, M., Wagner, D. E., Budjan, C., Hubaud, A., Tarazona, O. A., Donnelly, S., Michaut, A., Al Tanoury, Z., Yoshioka-Kobayashi, K., Niino, Y. et al.** (2020). In vitro characterization of the human segmentation clock. *Nature* **580**, 113-118. doi:10.1038/s41586-019-1885-9
- Diez del Corral, R., Breikreuz, D. N. and Storey, K. G.** (2002). Onset of neuronal differentiation is regulated by paraxial mesoderm and requires attenuation of FGF signalling. *Development* **129**, 1681-1691.
- Du, Z.-W., Chen, H., Liu, H., Lu, J., Qian, K., Huang, C. T.-L., Zhong, X., Fan, F. and Zhang, S.-C.** (2015). Generation and expansion of highly pure motor neuron progenitors from human pluripotent stem cells. *Nat. Commun.* **6**, 6626. doi:10.1038/ncomms7626
- Durston, A. J.** (2019). Some questions and answers about the role of Hox temporal collinearity in vertebrate axial patterning. *Front. Cell Dev. Biol.* **7**. doi:10.3389/fcell.2019.00257
- Duval, N., Vaslin, C., Barata, T. C., Frarma, Y., Contremoulins, V., Baudin, X., Nedelec, S. and Ribes, V. C.** (2019). Bmp4 patterns smad activity and generates stereotyped cell fate organization in spinal organoids. *Development* **146**, dev175430. doi:10.1242/dev.175430
- Ebisuya, M. and Briscoe, J.** (2018). What does time mean in development? *Development* **145**, dev164368. doi:10.1242/dev.164368
- Edri, S., Hayward, P., Jawaid, W. and Arias, A. M.** (2019). Neuro-mesodermal progenitors (NMPs): a comparative study between pluripotent stem cells and embryo-derived populations. *Development* **146**, dev180190. doi:10.1242/dev.180190
- Ensign, M., Tsuchida, T. N., Belting, H. G. and Jessell, T. M.** (1998). The control of rostrocaudal pattern in the developing spinal cord: specification of motor neuron subtype identity is initiated by signals from paraxial mesoderm. *Development* **125**, 969-982.
- Evtouchenko, L., Studer, L., Spenger, C., Dreher, E. and Seiler, R. W.** (1996). A mathematical model for the estimation of human embryonic and fetal age. *Cell Transplant.* **5**, 453-464. doi:10.1177/096368979600500404
- Faustino Martins, J.-M., Fischer, C., Urzi, A., Vidal, R., Kunz, S., Ruffault, P.-L., Kabuss, L., Hube, I., Gazzo, E., Birchmeier, C. et al.** (2020). Self-organizing 3D human trunk neuromuscular organoids. *Cell Stem Cell* **26**, 172-186.e6. doi:10.1016/j.stem.2019.12.007
- Forlani, S., Lawson, K. A. and Deschamps, J.** (2003). Acquisition of Hox codes during gastrulation and axial elongation in the mouse embryo. *Development* **130**, 3807-3819. doi:10.1242/dev.00573
- Frith, T. J. R., Granata, I., Wind, M., Stout, E., Thompson, O., Neumann, K., Stavish, D., Heath, P. R., Ortmann, D., Hackland, J. O. S. et al.** (2018). Human axial progenitors generate trunk neural crest cells in vitro. *eLife* **7**, e35786. doi:10.7554/eLife.35786
- Gaunt, S. J.** (1991). Expression patterns of mouse hox genes: clues to an understanding of developmental and evolutionary strategies. *BioEssays* **13**, 505-513. doi:10.1002/bies.950131004
- Gaunt, S. J., George, M. and Paul, Y.-L.** (2013). Direct activation of a mouse Hoxd11 axial expression enhancer by Gdf11/Smad signalling. *Dev. Biol.* **383**, 52-60. doi:10.1016/j.ydbio.2013.08.025
- Gouti, M., Tsakiridis, A., Wymeersch, F. J., Huang, Y., Kleinjung, J., Wilson, V. and Briscoe, J.** (2014). In vitro generation of neuromesodermal progenitors reveals distinct roles for wnt signalling in the specification of spinal cord and paraxial mesoderm identity. *PLoS Biol.* **12**, e1001937. doi:10.1371/journal.pbio.1001937
- Gouti, M., Delile, J., Stamataki, D., Wymeersch, F. J., Huang, Y., Kleinjung, J., Wilson, V. and Briscoe, J.** (2017). A gene regulatory network balances neural and mesoderm specification during vertebrate trunk development. *Dev. Cell* **41**, 243-261.e7. doi:10.1016/j.devcel.2017.04.002
- Helmbacher, F., Dessaud, E., Arber, S., DeLapeyrière, O., Henderson, C. E., Klein, R. and Maina, F.** (2003). Met signaling is required for recruitment of motor neurons to PEA3-positive motor pools. *Neuron* **39**, 767-777. doi:10.1016/S0896-6273(03)00493-8
- Henrique, D., Abranches, E., Verrier, L. and Storey, K. G.** (2015). Neuromesodermal progenitors and the making of the spinal cord. *Development* **142**, 2864-2875. doi:10.1242/dev.119768
- Iimura, T., Denans, N. and Pourquié, O.** (2009). Chapter 7 establishment of Hox vertebral identities in the embryonic spine precursors. *Curr. Top. Dev. Biol.* **88**, 201-234. doi:10.1016/S0070-2153(09)88007-1
- Izpisua-Belmonte, J. C., Falkenstein, H., Dollé, P., Renucci, A. and 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. doi:10.1002/j.1460-2075.1991.tb07764.x
- Jurberg, A. D., Aires, R., Varela-Lasheras, I., Nóvoa, A. and Mallo, M.** (2013). Switching axial progenitors from producing trunk to tail tissues in vertebrate embryos. *Dev. Cell* **25**, 451-462. doi:10.1016/j.devcel.2013.05.009
- Kawaguchi, A.** (2019). Temporal patterning of neocortical progenitor cells: how do they know the right time? *Neurosci. Res.* **138**, 3-11. doi:10.1016/j.neures.2018.09.004
- Kimelman, D. and Martin, B. L.** (2012). Anterior-posterior patterning in early development: three strategies. *Wiley Interdiscip. Rev. Dev. Biol.* **1**, 253-266. doi:10.1002/wdev.25
- Kmita, M. and Duboule, D.** (2003). Organizing axes in time and space; 25 years of colinear tinkering. *Science* **301**, 331-333. doi:10.1126/science.1085753
- Koch, F., Scholze, M., Wittler, L., Schifferl, D., Sudheer, S., Grote, P., Timmermann, B., Macura, K. and Herrmann, B. G.** (2017). Antagonistic activities of Sox2 and brachyury control the fate choice of neuro-mesodermal progenitors. *Dev. Cell* **42**, 514-526.e7. doi:10.1016/j.devcel.2017.07.021
- Kohwi, M. and Doe, C. Q.** (2013). Temporal fate specification and neural progenitor competence during development. *Nat. Rev. Neurosci.* **14**, 823-838. doi:10.1038/nrn3618
- Kondoh, H. and Takemoto, T.** (2012). Axial stem cells deriving both posterior neural and mesodermal tissues during gastrulation. *Curr. Opin. Genet. Dev.* **22**, 374-380. doi:10.1016/j.gde.2012.03.006
- Kondoh, H., Takada, S. and Takemoto, T.** (2016). Axial level-dependent molecular and cellular mechanisms underlying the genesis of the embryonic neural plate. *Dev. Growth Differ.* **58**, 427-436. doi:10.1111/dgd.12295
- Lambrot, R., Coffigny, H., Pairault, C., Donnadiou, A.-C., Frydman, R., Habert, R. and Rouiller-Fabre, V.** (2006). Use of organ culture to study the human fetal testis development: Effect of retinoic acid. *J. Clin. Endocrinol. Metab.* **91**, 2696-2703. doi:10.1210/jc.2005-2113
- Li, X.-J., Du, Z.-W., Zarnowska, E. D., Pankratz, M., Hansen, L. O., Pearce, R. A. and Zhang, S.-C.** (2005). Specification of motoneurons from human embryonic stem cells. *Nat. Biotechnol.* **23**, 215-221. doi:10.1038/nbt1063
- Lippmann, E. S., E. Williams, C., Ruhl, D. A., Estevez-Silva, M. C., Chapman, E. R., Coon, J. J. and Ashton, R. S.** (2015). Deterministic HOX patterning in human pluripotent stem cell-derived neuroectoderm. *Stem Cell Rep.* **4**, 632-644. doi:10.1016/j.stemcr.2015.02.018
- Liu, J.-P.** (2006). The function of growth/differentiation factor 11 (Gdf11) in rostrocaudal patterning of the developing spinal cord. *Development* **133**, 2865-2874. doi:10.1242/dev.02478
- Liu, J.-P., Laufer, E. and Jessell, T. M.** (2001). Assigning the positional identity of spinal motor neurons: Rostrocaudal patterning of Hox-c expression by FGFs, Gdf11, and retinoids. *Neuron* **32**, 997-1012. doi:10.1016/S0896-6273(01)00544-X
- Lunn, J. S., Fishwick, K. J., Halley, P. A. and Storey, K. G.** (2007). A spatial and temporal map of FGF/Erk1/2 activity and response repertoires in the early chick embryo. *Dev. Biol.* **302**, 536-552. doi:10.1016/j.ydbio.2006.10.014
- Mallo, M., Vinagre, T. and Carapuço, M.** (2009). The road to the vertebral formula. *Int. J. Dev. Biol.* **53**, 1469-1481. doi:10.1387/jdb.072276mm
- Mathis, L. and Nicolas, J.-F.** (2003). Progressive restriction of cell fates in relation to neuroepithelial cell mingling in the mouse cerebellum. *Dev. Biol.* **258**, 20-31. doi:10.1016/S0012-1606(03)00098-8
- Matsuda, M., Yamanaka, Y., Uemura, M., Osawa, M., Saito, M. K., Nagahashi, A., Nishio, M., Guo, L., Ikegawa, S., Sakurai, S. et al.** (2020). Recapitulating the human segmentation clock with pluripotent stem cells. *Nature* **580**, 124-129. doi:10.1038/s41586-020-2144-9

- Maury, Y., Côme, J., Piskorowski, R. A., Salah-Mohellibi, N., Chevalleyre, V., Peschanski, M., Martinat, C. and Nedelec, S. (2015). Combinatorial analysis of developmental cues efficiently converts human pluripotent stem cells into multiple neuronal subtypes. *Nat. Biotechnol.* **33**, 89-96. doi:10.1038/nbt.3049
- Mazzoni, E. O., Mahony, S., Peljto, M., Patel, T., Thornton, S. R., McCuine, S., Reeder, C., Boyer, L. A., Young, R. A., Gifford, D. K. et al. (2013). Saltatory remodeling of Hox chromatin in response to rostrocaudal patterning signals. *Nat. Neurosci.* **16**, 1191-1198. doi:10.1038/nn.3490
- McGrew, M. J., Sherman, A., Lillico, S. G., Ellard, F. M., Radcliffe, P. A., Gilhooley, H. J., Mitrophanous, K. A., Cambay, N., Wilson, V. and Sang, H. (2008). Localised axial progenitor cell populations in the avian tail bud are not committed to a posterior Hox identity. *Development* **135**, 2289-2299. doi:10.1242/dev.022020
- McPherron, A. C., Lawler, A. M. and Lee, S.-J. (1999). Regulation of anterior/posterior patterning of the axial skeleton by growth/differentiation factor 11. *Nat. Genet.* **22**, 260-264. doi:10.1038/10320
- Mendelsohn, A. I., Dasen, J. S. and Jessell, T. M. (2017). Divergent Hox coding and evasion of retinoid signaling specifies motor neurons innervating digit muscles. *Neuron* **93**, 792-805.e4. doi:10.1016/j.neuron.2017.01.017
- Narendra, V., Rocha, P. P., An, D., Raviram, R., Skok, J. A., Mazzoni, E. O. and Reinberg, D. (2015). CTCF establishes discrete functional chromatin domains at the Hox clusters during differentiation. *Science* **347**, 1017-1021. doi:10.1126/science.1262088
- Narendra, V., Bulajić, M., Dekker, J., Mazzoni, E. O. and Reinberg, D. (2016). CTCF-mediated topological boundaries during development foster appropriate gene regulation. *Genes Dev.* **30**, 2657-2662. doi:10.1101/gad.288324.116
- Neijts, R. and Deschamps, J. (2017). At the base of colinear Hox gene expression: cis-features and trans-factors orchestrating the initial phase of Hox cluster activation. *Dev. Biol.* **428**, 293-299. doi:10.1016/j.ydbio.2017.02.009
- Neijts, R., Amin, S., van Rooijen, C. and Deschamps, J. (2017). Cdx is crucial for the timing mechanism driving colinear Hox activation and defines a trunk segment in the Hox cluster topology. *Dev. Biol.* **422**, 146-154. doi:10.1016/j.ydbio.2016.12.024
- Nijssen, J., Comley, L. H. and Hedlund, E. (2017). Motor neuron vulnerability and resistance in amyotrophic lateral sclerosis. *Acta Neuropathol.* **133**, 863-885. doi:10.1007/s00401-017-1708-8
- Noordermeer, D., Leleu, M., Splinter, E., Rougemont, J., De Laat, W. and Duboule, D. (2011). The dynamic architecture of Hox gene clusters. *Science* **334**, 222-225. doi:10.1126/science.1207194
- Noordermeer, D., Leleu, M., Schorderet, P., Joye, E., Chabaud, F. and Duboule, D. (2014). Temporal dynamics and developmental memory of 3D chromatin architecture at Hox gene loci. *eLife* **3**, e02557. doi:10.7554/eLife.02557
- Nordström, U., Maier, E., Jessell, T. M. and 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**, 1438-1452. doi:10.1371/journal.pbio.0040252
- Oberst, P., Fièvre, S., Baumann, N., Concetti, C., Bartolini, G. and Jabaudon, D. (2019a). Temporal plasticity of apical progenitors in the developing mouse neocortex. *Nature* **573**, 370-374. doi:10.1038/s41586-019-1515-6
- Oberst, P., Agirman, G. and Jabaudon, D. (2019b). Principles of progenitor temporal patterning in the developing invertebrate and vertebrate nervous system. *Curr. Opin. Neurobiol.* **56**, 185-193. doi:10.1016/j.conb.2019.03.004
- Ogura, T., Sakaguchi, H., Miyamoto, S. and Takahashi, J. (2018). Three-dimensional induction of dorsal, intermediate and ventral spinal cord tissues from human pluripotent stem cells. *Development* **145**, dev162214. doi:10.1242/dev.162214
- Olivera-Martinez, I. and Storey, K. G. (2007). Wnt signals provide a timing mechanism for the FGF-retinoid differentiation switch during vertebrate body axis extension. *Development* **134**, 2125-2135. doi:10.1242/dev.000216
- Olivera-Martinez, I., Schurch, N., Li, R. A., Song, J., Halley, P. A., Das, R. M., Burt, D. W., Barton, G. J. and Storey, K. G. (2014). Major transcriptome re-organisation and abrupt changes in signalling, cell cycle and chromatin regulation at neural differentiation in vivo. *Development* **141**, 3266-3276. doi:10.1242/dev.112623
- Peljto, M., Dasen, J. S., Mazzoni, E. O., Jessell, T. M. and Wichterle, H. (2010). Functional diversity of ESC-derived motor neuron subtypes revealed through intraspinal transplantation. *Cell Stem Cell* **7**, 355-366. doi:10.1016/j.stem.2010.07.013
- Philippidou, P. and Dasen, J. S. (2013). Hox genes: choreographers in neural development, architects of circuit organization. *Neuron* **80**, 12-34. doi:10.1016/j.neuron.2013.09.020
- Ragagnin, A. M. G., Shadfar, S., Vidal, M., Jamali, M. S. and Atkin, J. D. (2019). Motor neuron susceptibility in ALS/FTD. *Front. Neurosci.* **13**, 532. doi:10.3389/fnins.2019.00532
- Ribes, V., Le Roux, I., Rhinn, M., Schuhbaur, B. and Dolle, P. (2009). Early mouse caudal development relies on crosstalk between retinoic acid, Shh and Fgf signalling pathways. *Development* **136**, 665-676. doi:10.1242/dev.016204
- Rossi, A. M., Fernandes, V. M. and Desplan, C. (2017). Timing temporal transitions during brain development. *Curr. Opin. Neurobiol.* **42**, 84-92. doi:10.1016/j.conb.2016.11.010
- Routal, R. V. and Pal, G. P. (1999). A study of motoneuron groups and motor columns of the human spinal cord. *J. Anat.* **195**, 211-224. doi:10.1046/j.1469-7580.1999.19520211.x
- Sances, S., Bruijn, L. I., Chandran, S., Eggen, K., Ho, R., Klim, J. R., Livesey, M. R., Lowry, E., Macklis, J. D., Rushton, D. et al. (2016). Modeling ALS with motor neurons derived from human induced pluripotent stem cells. *Nat. Neurosci.* **19**, 542-553. doi:10.1038/nn.4273
- Sasai, N., Kutejova, E. and Briscoe, J. (2014). Integration of signals along orthogonal axes of the vertebrate neural tube controls progenitor competence and increases cell diversity. *PLoS Biol.* **12**, e1001907. doi:10.1371/journal.pbio.1001907
- Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B. et al. (2012). Fiji: An open-source platform for biological-image analysis. *Nat. Methods* **9**, 676-682. doi:10.1038/nmeth.2019
- Soshnikova, N. and Duboule, D. (2009). Epigenetic temporal control of mouse hox genes in vivo. *Science* **324**, 1321-1323. doi:10.1126/science.1171468
- Steinbeck, J. A. and Studer, L. (2015). Moving stem cells to the clinic: potential and limitations for brain repair. *Neuron* **86**, 187-206. doi:10.1016/j.neuron.2015.03.002
- Syed, M. H., Mark, B. and Doe, C. Q. (2017). Playing well with others: extrinsic cues regulate neural progenitor temporal identity to generate neuronal diversity. *Trends Genet.* **33**, 933-942. doi:10.1016/j.tig.2017.08.005
- Tiberi, L., Vanderhaeghen, P. and van den Aemele, J. (2012). Cortical neurogenesis and morphogens: diversity of cues, sources and functions. *Curr. Opin. Cell Biol.* **24**, 269-276. doi:10.1016/j.ceb.2012.01.010
- Tschopp, P., Tarchini, B., Spitz, F., Zakany, J. and Duboule, D. (2009). Uncoupling time and space in the collinear regulation of Hox genes. *PLoS Genet.* **5**, e1000398. doi:10.1371/journal.pgen.1000398
- Tung, Y.-T., Peng, K.-C., Chen, Y.-C., Yen, Y.-P., Chang, M., Thams, S. and Chen, J.-A. (2019). Mir-17~92 confers motor neuron subtype differential resistance to ALS-associated degeneration. *Cell Stem Cell* **25**, 193-209.e7. doi:10.1016/j.stem.2019.04.016
- Tzouanacou, E., Wegener, A., Wymeersch, F. J., Wilson, V. and Nicolas, J.-F. (2009). Redefining the progression of lineage segregations during mammalian embryogenesis by clonal analysis. *Dev. Cell* **17**, 365-376. doi:10.1016/j.devcel.2009.08.002
- Verrier, L., Davidson, L., Gierlinski, M., Dady, A. and Storey, K. G. (2018). Neural differentiation, selection and transcriptomic profiling of human neuromesodermal progenitor-like cells in vitro. *Development* **145**, dev166215. doi:10.1242/dev.166215
- Wang, J., Vasaikar, S., Shi, Z., Greer, M. and Zhang, B. (2017). WebGestalt 2017: A more comprehensive, powerful, flexible and interactive gene set enrichment analysis toolkit. *Nucleic Acids Res.* **45**, W130-W137. doi:10.1093/nar/gkx356
- Wymeersch, F. J., Huang, Y., Blin, G., Cambay, N., Wilkie, R., Wong, F. C. K. and Wilson, V. (2016). Position-dependent plasticity of distinct progenitor types in the primitive streak. *eLife* **5**, e10042. doi:10.7554/eLife.10042
- Wymeersch, F. J., Skylaki, S., Huang, Y., Watson, J. A., Economou, C., Marek-Johnston, C., Tomlinson, S. R. and Wilson, V. (2019). Transcriptionally dynamic progenitor populations organised around a stable niche drive axial patterning. *Development* **146**, dev168161. doi:10.1242/dev.168161
- Wymeersch, F. J., Wilson, V. and Tsakiridis, A. (2021). Understanding axial progenitor biology in vivo and in vitro. *Development* **148**, dev180612. doi:10.1242/dev.180612
- Young, T., Rowland, J. E., van de Ven, C., Bialecka, M., Novoa, A., Carapuco, M., van Nes, J., de Graaff, W., Duluc, I., Freund, J.-N. et al. (2009). Cdx and Hox genes differentially regulate posterior axial growth in mammalian embryos. *Dev. Cell* **17**, 516-526. doi:10.1016/j.devcel.2009.08.010

Figure S1

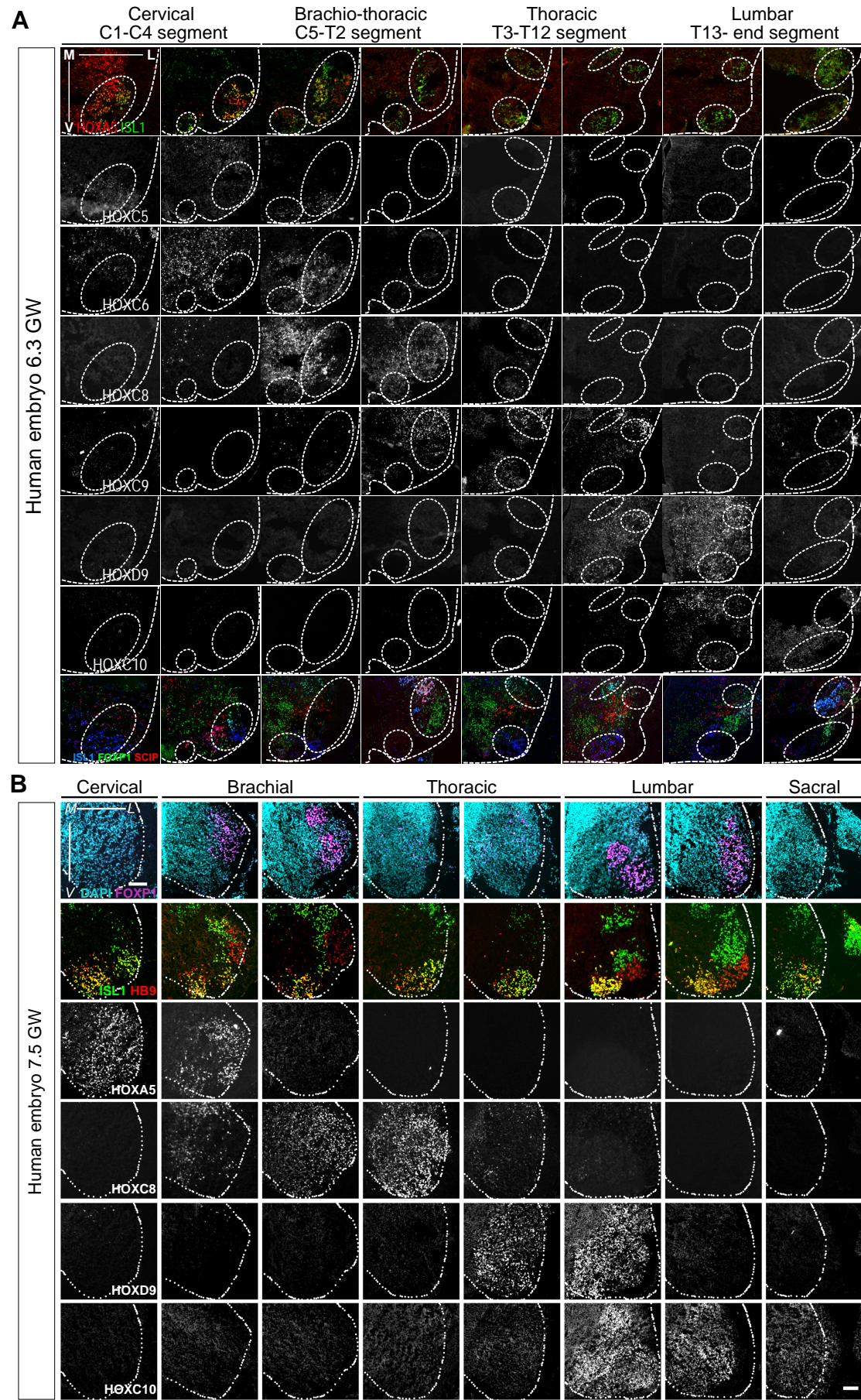


Figure S1 - Rostro-caudal distribution of motor neuron subtypes in human embryonic spinal cords. Related to Figure 1A

Immunostaining on transverse sections of human embryonic spinal cords at gestation weeks (GW) 6.3 **(A)** and 7.5 **(B)** to define HOX and MN subtype markers expression profiles in spinal MNs. MNs organized in motor columns that innervate distinct muscle groups were identified by their localization in the ventral horn and the expression of ISL1 or HB9. HOXA/C5 MNs are localized in the cervical regions, while HOXC6, HOXC9 and HOXC10 MNs are respectively at brachial, thoracic and lumbar levels. HOXC8 MNs are in the caudal brachial region and the anterior thoracic domain where some MNs co-express HOXC9. HOXD9 MNs are in the caudal thoracic and the anterior lumbar regions. FOXP1^{high} MNs form lateral motor columns at brachial, anterior thoracic and lumbar levels which correspond to limb innervating MNs in mouse (Dasen et al., 2008; , Rousso et al., 2008). FOXP1^{high}/SCIP^{high} MNs are present in the caudal brachial and anterior thoracic regions which respectively characterize forearm and forepaw innervating MNs in mouse (Dasen et al., 2008, Mendelsohn et al., 2017). External dotted lines in A and B mark the side of the ventral horn. M: medial L: lateral V: Ventral. Scale bars : 100µm.

Figure S2

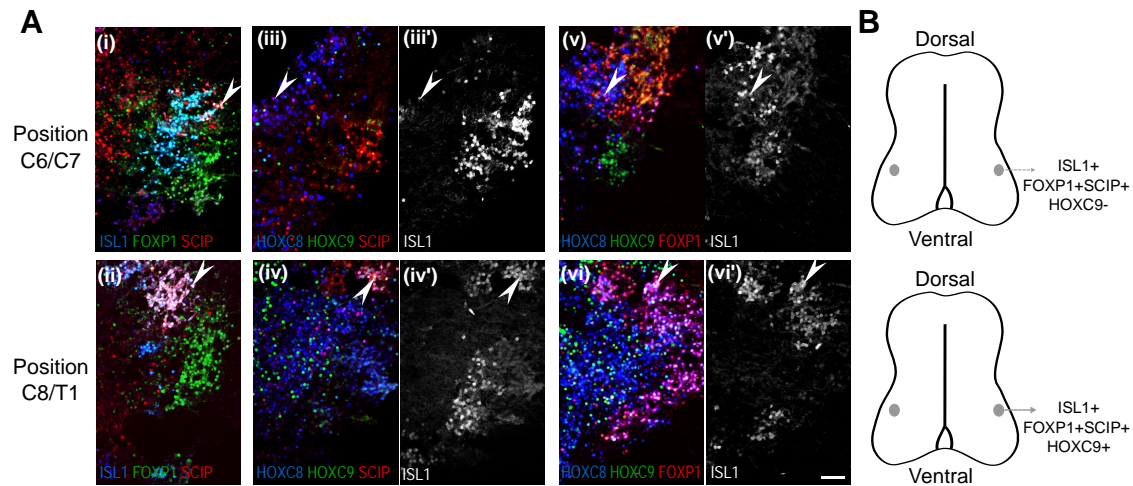


Figure S2 – Characterization of brachial and anterior thoracic FOXP1/SCIP MN subtypes

(A) Arrows in i) and ii) indicate FOXP1/SCIP/ISL1 MNs in caudal brachial and anterior thoracic spinal cord. Arrows in iii) and v) point to the presence of HOXC9-negative FOXP1/HOXC8/ISL1-positive MNs (ISL1 shown in iii' and v'). In iv, iv' and vi, vi', SCIP/ISL1 or FOXP1/ISL1 MNs around and caudal to the caudal brachial/anterior thoracic border co-expressed HOXC8 and HOXC9, which correspond to digit innervating MNs in mouse (Mendelsohn et al., 2017). **(B)** Schematic representation of the FOXP1/SCIP/ISL1/HOXC8 and FOXP1/SCIP/ISL1/HOXC9 MNs in brachial and thoracic regions. Scale bar: 100µm.

Figure S3

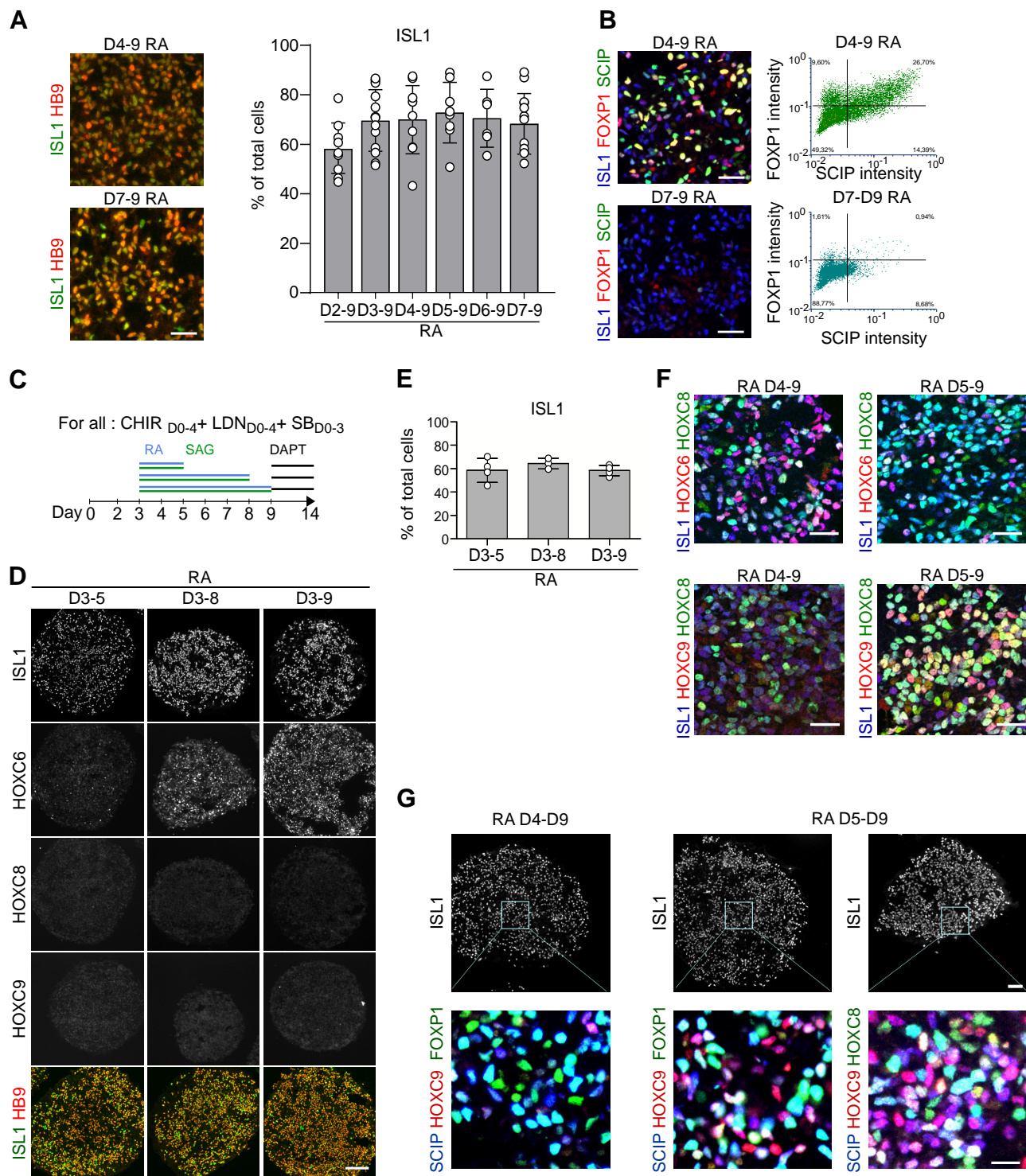


Figure S3 - The timing of exposure to RA and not the exposure duration controls the specification of caudal motor neurons from hESC and hiPSC. Related to Figure 1

(A) Proportion of MNs generated on Day 14 of differentiation when modulating the Day at which RA/SAG is added. Scale bar: 30 μ m. **(B)** Quantification of FOXP1^{high} and FOXP1^{high}/SCIP MNs. Intensity of FOXP1 and SCIP staining in all ISL1⁺ cells were extracted using Cell Profiler software. Intensity plots were generated for all conditions. FOXP1 and SCIP weak or negative cells were defined on the D7-D9 RA condition. Scale bar: 30 μ m. **(C-E)** Characterization on Day 14 of differentiation of the effect of varying RA/SAG duration of exposure on MN subtypes identity. **(C)** Differentiation conditions. **(D)** Immunostaining on cryostat sections of EBs for ISL1 and HOXs showing the absence of caudal HOXs when RA is added on Day 3 and the duration of exposure modulated. Scale bar: 100 μ m. **(E)** Proportion of MNs (ISL1⁺ cells, mean \pm SD) generated on Day 14 in the three tested conditions. **(F)** Immunostaining on cryostat sections of Day 14 EBs for the indicated markers quantified in Fig. 1D-F. Scale bar : 30 μ m. **(G)** Immunostaining on cryostat sections of EBs showing the presence of HOXC8/HOXC9/SCIP/ISL1 MNs and HOXC9/SCIP/FOXP1/ISL1 MNs that correspond to forepaw innervating-MNs in mouse and might represent hand innervating MNs in human. Scale bar: 40 μ m (top) and 15 μ m (bottom). Data are shown as mean \pm SD. Each circle is an independent biological replicate: n=6 to 13 (A), n=3 to 4 (E). * if $P \leq 0.05$, ** if $P \leq 0.01$ and *** if $P \leq 0.001$. ANOVA with KW post hoc.

Figure S4

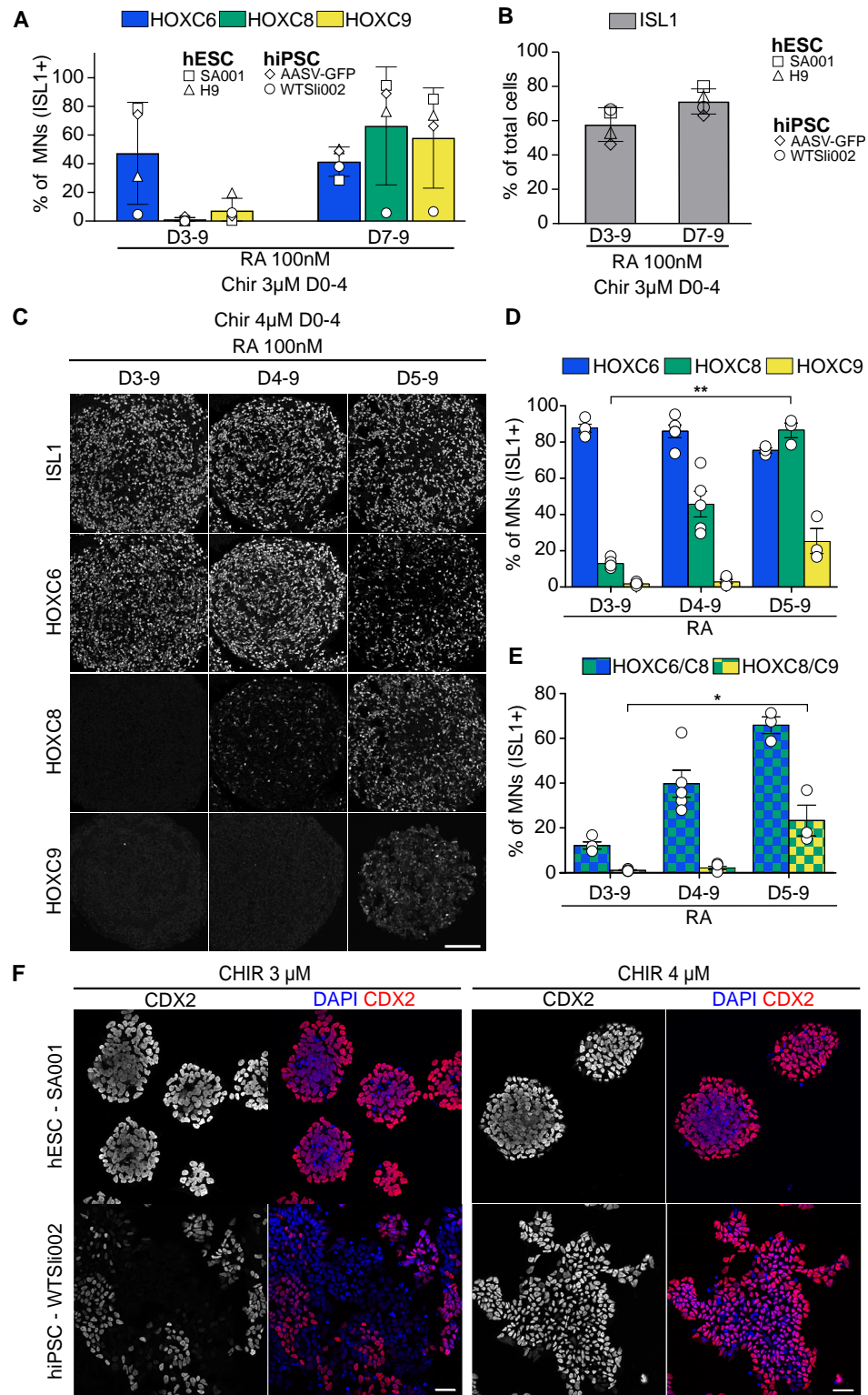
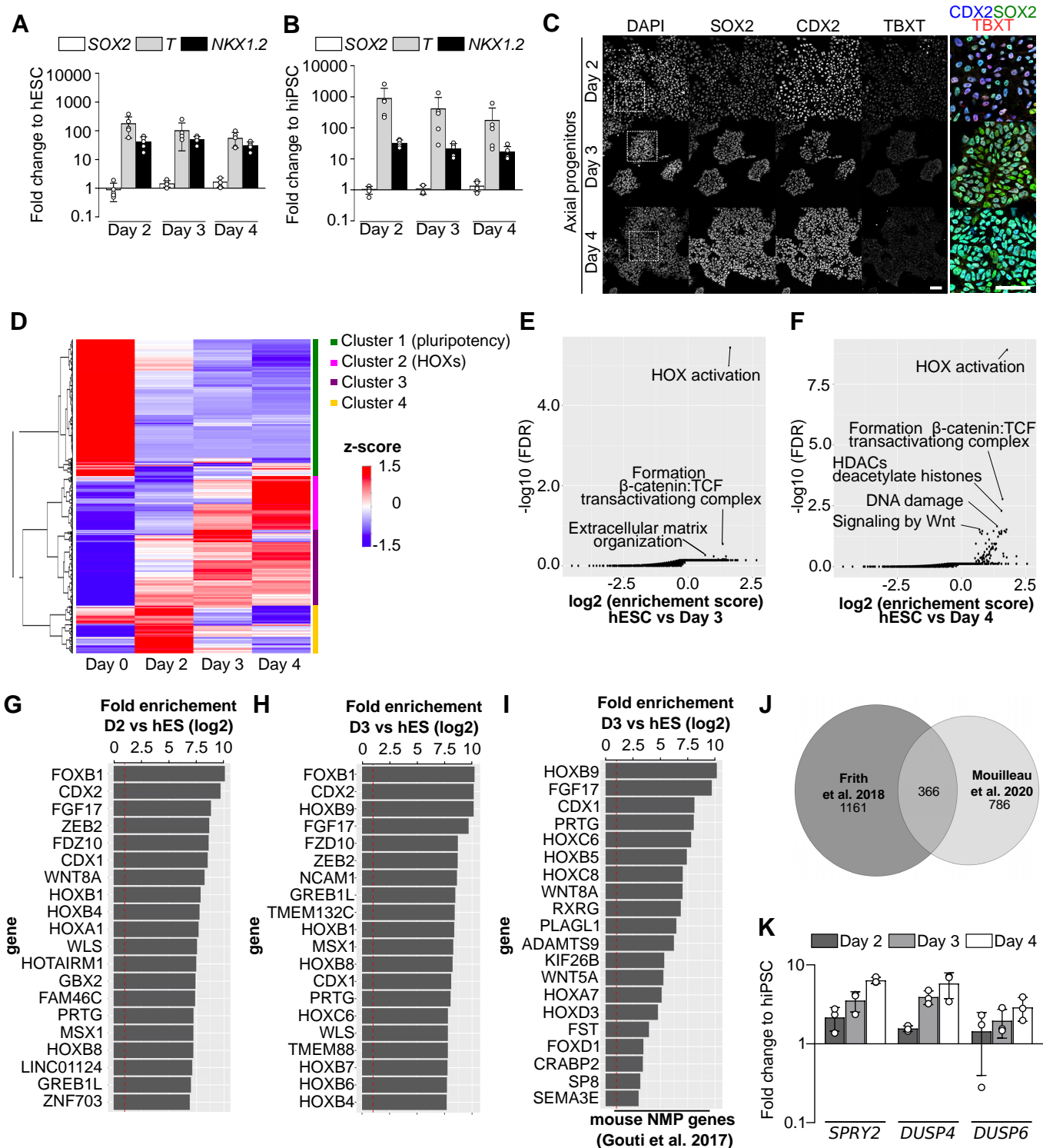


Figure S4 - The concentration of exposure to WNT agonist impacts the axial specification of stem cells and varies according cell lines. Related to Figure 1

(A) Quantification of HOXC6-9 immunostaining on Day 14 cryostat sections of EBs derived from different hESC (SA001 and H9) and hiPSC (AASV-GFP and WTSli002). Cells were exposed to CHIR99021 (CHIR) 3μM together with LDN193189 from Day 0 to 4, and SB431542 from Day 0 to 3, then RA 100nM was alternatively applied from Day 3 or Day 7, up to Day 9. EBs derived from WTSli002 line do not express caudal HOXs when treated with the usual concentration of CHIR 3μM. **(B)** Quantification of ISL1+ cells on immunostaining on Day 14 cryostat sections of Day 14 EBs derived from different hPSC lines as presented in (A). **(C)** Immunostaining on cryostat sections of D14 hiPSC (WTSli002)-derived EBs for HOXs showing the generation of HOXC8 and then HOXC9 MNs when addition of RA is delayed. Cells were exposed to CHIR 4μM together with LDN193189 from Day 0 to 4, and SB431542 from Day 0 to 3, then RA 100nM was alternatively applied from Day 3, 4 or 5, up to Day 9. Scale bar: 100 μm. **(D-E)** Quantification of indicated markers on immunostaining on Day 14 EBs cryostat sections. **(F)** Immunostaining on Day 3 EBs plated on matrigel coated coverslips, derived from hESC line SA001 or hiPSC line WTSli002. EBs were exposed either to CHIR 3 or 4 μM on Day 0 of differentiation together with LDN193189 and SB431542. Scale bar: 40 μm. Data are shown as mean ± SD. Each circle is an independent biological replicate: n=3 to 5 (D-E). * if $P \leq 0.05$ and ** if $P \leq 0.01$. ANOVA with KW post hoc.

Figure S5

**Figure S5 - Similarity between axial progenitors derived from hESC and hiPSC and mouse NMPs.**

(A-B) Real time qPCR analysis of axial progenitor markers in hESC- (SA001) (A) and hiPSC- (WTSli002) (B) -derived progenitors at Day 2, 3 and 4 of differentiation. Data are represented as mRNAs expression fold-changes to Day 0. **(C)** Immunostaining for axial progenitor markers on hiPSC (WTSli002)-derived Day 2, 3 and 4 progenitors. White square indicate the cells shown in higher magnification on the right of the panel Scale bars: 100 μ m (left), 40 μ m right, magnification). **(D)** Hierarchical clustering showing transcriptional dynamic (z-score) of a subset of genes from Day 0 to Day 4 and regrouped through 4 different clusters. **(E-F)** Reactome pathway analysis of genes upregulated (fold change ≥ 2.0 , p-value < 0.05) between **(E)** Day 0 and Day 3 or **(F)** Day 0 and Day 4. **(G-I)** 20 most enriched genes in D2 and D3 progenitors versus hESC. Progenitors acquired a transcriptomic signature resembling murine axial progenitors. **(J)** Comparison of genes up-regulated between Day 0 and Day 3 (fold change ≥ 2.0 , p-value < 0.05) and “NMP-like genes” reported in (Frith et al., 2018). **(K)** Real time-PCR analysis of FGF signaling markers mRNAs in hiPSC (WTSli002) derived progenitors at Day 2, 3 and 4 of differentiation. Data are represented as mRNAs expression fold-changes to Day 0, shown as mean \pm SD. Each circle is an independent biological replicate: n=5 (B, C), n=2 to 3 (K).

Figure S6

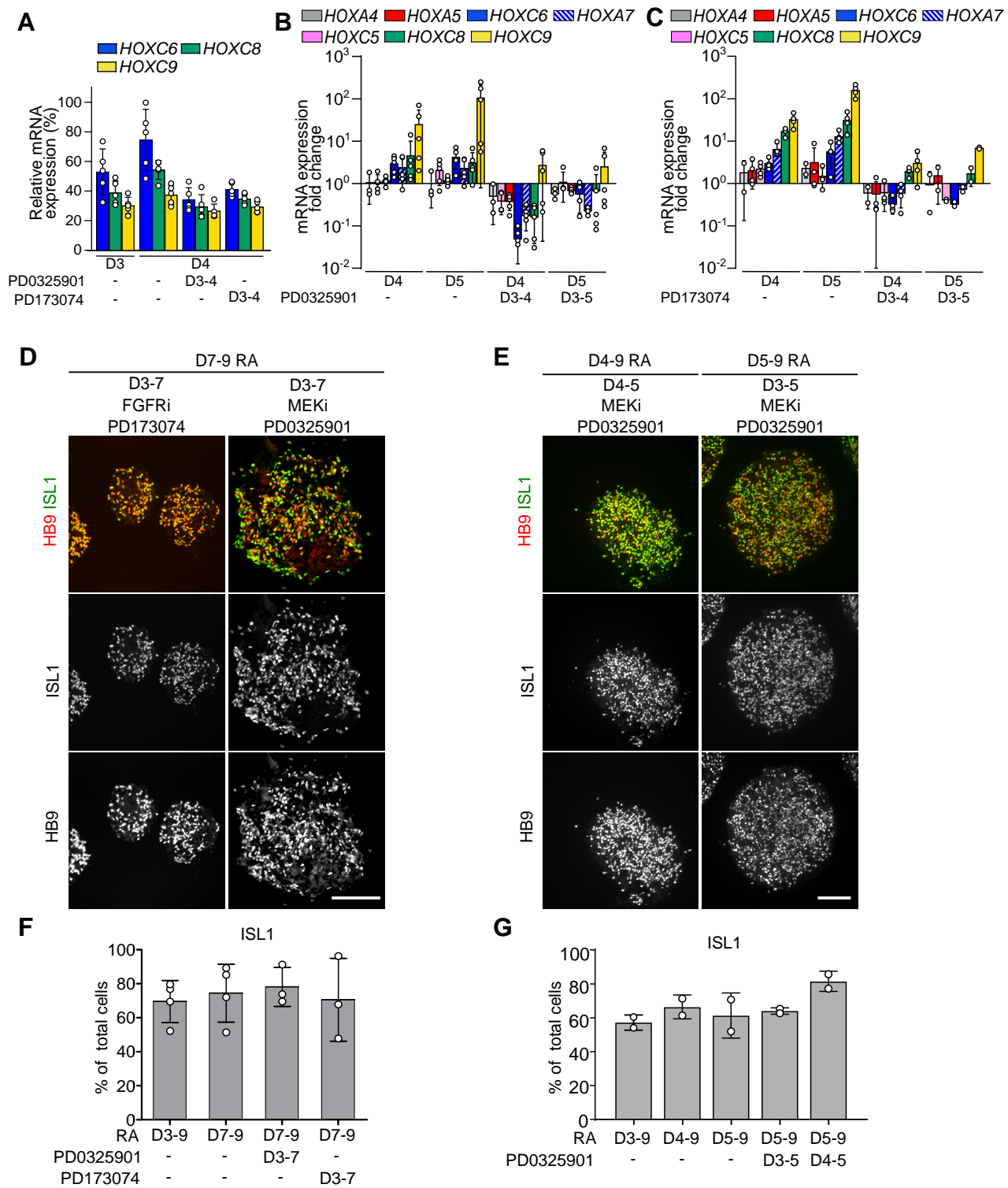


Figure S6 - FGFR and MEK1/2 inhibition in axial progenitors does not impact motor neuron specification

(A) Real time qPCR analysis of HOX mRNAs in hiPSC (WTSli002)-derived progenitors at Day 3, Day 4 and Day 5 of differentiation. Data are represented as relative expression to the highest expressed gene. MEK1/2 and FGFR inhibitors (PD173074), applied from Day 3 prevent the temporal increase in caudal HOX expression. **(B-C)** Real time PCR analysis of HOX mRNAs in progenitors at Day 4 and 5 of differentiation. Data are the same as in Fig.3B-C, but here represented as mRNA expression fold change to Day 3 controls to better visualize decreases in expression. MEK1/2 and FGFR inhibitors, applied from Day 3 prevent the temporal increase in caudal HOX expression and induce a repression of HOX genes compare to their initial expression level. **(D-E)** Immunostaining on Day 14 cryostat sections of embryoid bodies against ISL1 and HB9. Scale bars: 100 μ m. **(F-G)** Quantification of the proportion of ISL1 positive cells in Day 14 EBs from experiment presented respectively in A and B. Exposing axial progenitors to either FGFR or MEK1/2 inhibitors does not prevent the specification of motor neurons and have no effect on the proportion of ISL1+ and HB9+ cells. Data are shown as mean \pm SD; Each circle is an independent biological replicate: n=5 (A), n=3 to 5 (B), n=4 (C), n=3 to 4 (E), n=2 (F).

Figure S7

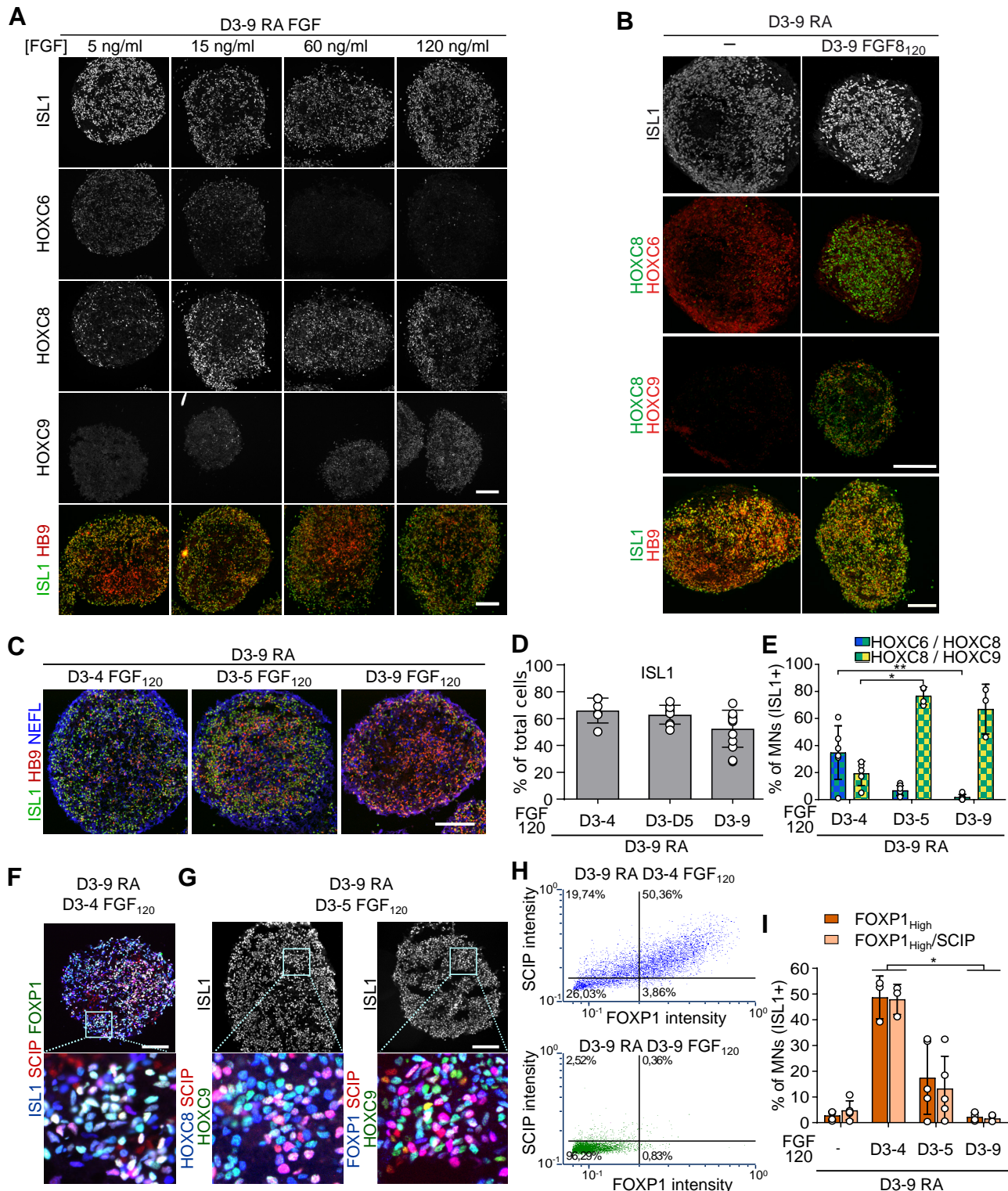


Figure S7 - Extrinsic FGF2 and FGF8 controls motor neuron subtype identity.

(A) Immunostaining on cryostat sections of hESC (SA001)-derived EBs at Day 14 of differentiation. Staining for ISL1, HOXC6, HOXC8 and HOXC9 to test the effect of varying FGF2 concentration on MN subtype identity. Scale bar: 100 μ m. (B) Immunostaining for ISL1, HB9, HOXC6, HOXC8, HOXC9 in control and upon addition of 120 ng/ml FGF8 between Day 3 and Day 9. Scale bar: 100 μ m. (C-I) Characterization on Day 14 of differentiation of the effect of varying FGF2 duration of exposure on MN generation and subtype identity. In all cases, FGF2 was added at Day 3 and the duration of exposure was modulated. (C) Immunostaining on cryostat sections of EBs for ISL1, HB9 and NEFL to test for MN specification. (D) Proportion of MNs on Day 14 of differentiation. (E) Quantification of HOXC6/HOXC8/ISL1 and HOXC8/HOXC9/ISL1 co-expression. (F-G) Immunostaining for ISL1, SCIP and FOXP1 (F) or HOXC8/HOXC9/SCIP/ISL1 (G) aiming at characterizing the specification of MNs acquiring the identity of limb and digit innervating motor neurons. (H) Quantification of FOXP1^{high} and FOXP1^{high}/SCIP motor neurons using FCS software. Intensities of staining were extracted with Cell Profiler software and plotted using FCS Express software. FOXP1 weak and negative cells were defined on the FGF D3-9 condition. The gate was then reported on the other conditions. (I) Proportion of FOXP1^{high} and FOXP1^{high}/SCIP upon modulation of FGF2 treatment duration. All scale bars: 100 μ m. Data are shown as mean \pm SD. Each circle is an independent biological replicate: n=3 to 9 (D), n=3 to 6 (E), n=3 to 5 (I). * if $P \leq 0.05$ and ** if $P \leq 0.01$. ANOVA with KW post hoc.

Figure S8

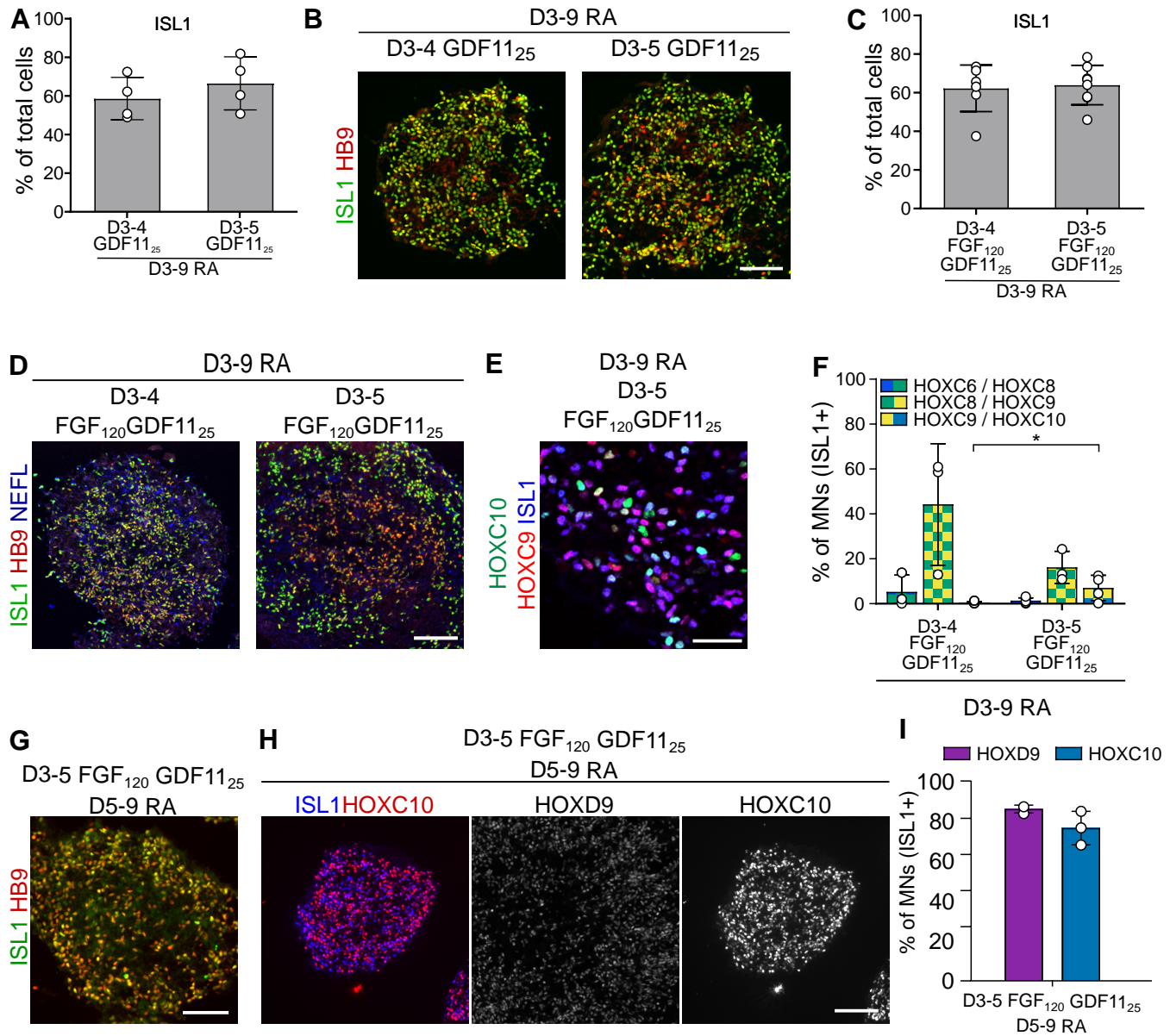


Figure S8 - FGF2 and GDF11 cooperate to induce caudal thoracic and lumbar motor neurons. RNA-seq analysis of in vitro derived axial progenitors. Related to Figure 2-4

(A) Quantification of the percentage of MNs on Day 14 of differentiation upon addition of GDF11. GDF11 (25 ng/ml) was added on Day 3 for different durations. **(C)** Quantification of the percentage of MNs (ISL1+ cells) on Day 14 of differentiation upon addition of FGF2 and GDF11. FGF2 (120 ng/ml) and GDF11 (25 ng/ml) were added on Day 3 of differentiation for different durations. **(B,D)** Representative images of immunostainings for the motor neuron markers ISL1, HB9 (B) and the pan neuronal marker NEFL on cryostat sections of EBs on Day 14 of differentiation. Scale bars: 100µm. **(E)** Representative image of immunostaining for ISL1, HOXC9 and HOXC10 on cryostat sections of EBs on Day 14 of differentiation. Scale bar: 30µm. **(F)** Quantification of the percentage of HOXs co-expression in the indicated conditions. **(G-H)** Representative images of immunostainings for ISL1, HB9, HOXC9 and HOXC10 on cryostat sections of EBs on Day 14 of differentiation. Scale bars: 100µm. **(I)** Quantification of the percentage of HOXD9 and HOXC10 in the indicated conditions. Data are shown as mean ± SD. Each circle is an independent biological replicate: n=3 to 4 (A), n=4 (C), n=3 to 5 (F), n=3 (I). * if P ≤ 0.05. ANOVA with KW post hoc.

La dynamique temporelle des voies de signalisation rythme l'activation séquentielle des gènes *HOX*

Dans cette étude, nous avons cherché à comprendre les mécanismes de signalisation cellulaire en aval de GDF11 qui sont capables de contrôler l'activation temporelle des gènes *HOX*.

a. Contexte

Pendant le développement de l'axe antéro-postérieur, les gènes *Hox* sont progressivement induits dans les progéniteurs axiaux, ce qui conduit à une expression spatiale différenciée et à l'acquisition d'identités positionnelles spécifiques pour leurs descendants.

Des études ont montré que chez les Vertébrés, l'induction des gènes *Hox10-13* semble dépendre de Gdf11, qui régule également la transition du tronc à la queue et la fin de l'élongation (Aires et al., 2019; Jurberg et al., 2013). De plus chez les mutants *Gdf11*, l'expression de *Hox10* est maintenue mais décalée vers la partie caudale (McPherron et al., 1999). A l'inverse, des études de gain de fonction dans le tube neural révèlent que Gdf11 peut favoriser un déplacement antérieur des domaines d'expression des gènes *HoxC8*, *C9* et *C10*, suggérant ainsi que cette voie régule également les gènes situés dans la partie centrale des complexes *Hox* (Liu, 2006). Cependant le rôle de la voie GDF dans les progéniteurs axiaux lors de l'induction de *HOX* demeure difficile à comprendre car les expériences démontrant l'influence de ce signal sur les gènes *HOX* reposaient sur l'électroporation dans le tube neural ou la culture d'explants du tube préneurale ou PSM, un stade de différenciation postérieur à celui des progéniteurs axiaux où l'initiation se produit.

C'est pourquoi nous avons différencié des CSPi humaines en progéniteurs axiaux dans notre récente étude (Mouilleau et al., 2021). Ces travaux ont permis de mettre en évidence que les progéniteurs axiaux peuvent interpréter des signaux extrinsèques pour produire différents types de motoneurones. En effet, nous avons observé que le GDF11 et en particulier la combinaison du FGF2 et du GDF11 étaient les signaux les plus efficaces pour induire les motoneurones HOXD9 et HOXC10 qui sont respectivement localisés dans le domaine caudal thoracique et lombaire de la moelle épinière. Suite à ces découvertes, nous avons cherché à étudier les mécanismes en aval par lesquels le GDF11 et le FGF2/GDF11 favorisent l'induction des gènes *HOX*.

b. Résultats

J'ai d'abord montré que dans les progéniteurs axiaux, le GDF11 interagit avec le récepteur ALK5 pour activer la voie SMAD2/3. Cette activation implique notamment la phosphorylation de SMAD2 et sa translocation nucléaire. Par la suite, nous avons effectué une caractérisation de la dynamique de la voie SMAD2/3 et mis en évidence une augmentation progressive du signal nucléaire de P-SMAD2 pendant les trois premières heures post-stimulation, suivie d'une forte diminution. Cependant, l'activation de la voie SMAD2/3 perdure dans les progéniteurs axiaux, car 12h après la stimulation au GDF11, presque toutes les cellules présentent encore un signal positif. De plus, le traitement des progéniteurs axiaux à différentes concentrations de GDF11 ne modifie pas la dynamique temporelle de la voie SMAD2/3. Pour confirmer cette dynamique au niveau cellulaire, nous avons utilisé une lignée de cellules souches embryonnaires humaines rapporteuses pour SMAD2. Nous avons observé une dynamique similaire en mesurant le rapport nucléaire/cytoplasmique avec une augmentation progressive du signal les trois premières heures puis une diminution. Pour déterminer si la diminution du signal de la voie SMAD2/3 était dû à une dégradation du ligand GDF11, nous avons re-stimulé les progéniteurs avec du GDF11 6h après la première stimulation. Cette seconde stimulation n'a pas eu d'impact significatif sur l'intensité nucléaire de P-SMAD2, démontrant que la diminution de l'activité de la voie SMAD2/3 est liée à des mécanismes adaptatifs.

Puis, j'ai caractérisé les événements transcriptionnels après la stimulation des progéniteurs axiaux à GDF11. Ces résultats ont révélé une sur-expression de gènes associés à la voie TGF- β , en particulier des régulateurs négatifs de cette voie. En parallèle de ce réseau transcriptionnel, plusieurs *HOX* caractérisant les régions thoraciques et lombaires sont induits dès 3h après le traitement GDF11. A partir de 6h, j'ai détecté l'activation transcriptionnelle d'un nouveau gène, *HOXC10*, tandis que les paralogues *HOX11* et *HOX12* restaient réprimés. Ces observations indiquent que la colinéarité temporelle est conservée dans les progéniteurs exposés au GDF11. Par ailleurs, mes données préliminaires suggèrent que cette induction transcriptionnelle est associée à une progression des marques histones permissives (H3K4me3) vers les gènes *HOX* en 5', concomitante avec une diminution des marques répressives (H3K27me3). L'activation colinéaire des gènes *HOX* en réponse aux facteurs GDF11 a également été confirmée par des hybridations *in situ*.

Pour évaluer l'importance de la dynamique de la voie de signalisation SMAD2/3 dans la régulation de l'expression des gènes *HOX*, les progéniteurs axiaux ont été soumis à une stimulation transitoire par le facteur GDF11, suivi de l'inhibition de la voie par un antagoniste se liant aux récepteurs ALK5. Ces résultats indiquent que plus la signalisation est longue et plus le niveau d'expression des gènes *HOX* est élevé. Par conséquent, les cellules intègrent la durée de la signalisation SMAD2/3 pour induire les gènes *HOX*.

Pour conclure, nous avons cherché à modifier la dynamique de la voie SMAD2/3 après son activation pour évaluer son impact sur l'activation séquentielle des gènes *HOX*. Nous avons donc réalisé des co-traitements GDF11 et CHX afin de bloquer la néo-traduction des régulateurs négatifs de la voie SMAD2/3. Dans ce contexte, nous avons détecté l'expression de *HOXA11*, *C11* et *D11* et des paralogues *I3* dans des conditions où les progéniteurs sont exposés à FGF2/GDF11/CHX. De plus, mes résultats encore préliminaires (n=1) indiquent que lorsque la voie GDF11 est activée et la néo-traduction bloquée, le recrutement de SMAD2/3 dans les complexes *HOXA* et *HOXC* est plus fort.

Dans l'ensemble, ces résultats suggèrent que le rythme d'activation des gènes *HOX* pourrait être une propriété émergente des dynamiques de voies de signalisation en aval de facteurs extrinsèques. Ces données sont encore préliminaires et sont présentées sous la forme d'un article en construction.

Signaling pathway temporal dynamics pace *HOX* sequential activation

Authors: Rémi Robert, Célia Vaslin, Helena Malek, Robin Rondon, Jacques van Helden, Vanessa Ribes, Pascale Gilardi, Benoit Sorre, Stéphane Nedelec

Results:

Temporal dynamics of SMAD2/3 signaling downstream GDF11

To investigate how GDF11 paces *HOX* activation, we sought to determine the pathway activated downstream GDF11 and its temporal dynamics. Previous studies have demonstrated that GDF11, an extrinsic factor belonging to the TGF- β family, signals through the ALK5 receptor (TGFB1R1) and then phosphorylates the transcription factors SMAD2/3 to induce the expression of target genes (Andersson et al., 2006; Huse et al., 2001; Liu, 2006). Moreover, GDF11 involvement has been demonstrated in initiating the trunk to tail transition (Jurberg et al., 2013), regulating the expression of *HOX8* to *13* depending on the time of development at which GDF11 is modulate (Aires et al., 2019; Andersson et al., 2006; Liu, 2006; McPherron et al., 1999)

To test whether GDF11 signals through this pathway in hPSC-derived axial progenitors, we exposed day 3 progenitors to GDF11 and stained for P-SMAD2 and SMAD2/3 (**Fig. 1A**). At 1h post exposure, we observed an accumulation of P-SMAD2 and SMAD2/3 in the nucleus (**Fig. 1B**). Inhibition of the ALK5 receptor inhibitor (SB431542) blocked both the nuclear accumulation of SMADs (**Fig. 1B**) and the specification of thoracic motoneurons (**Fig. S1A/B**). However, neither inhibitors of MAPK pathways (ERK1/2, JUNK, P38) or AKT blocked the GDF11-induced specification of caudal thoracic motoneurons (**Fig. S1C**). These results indicate that non-canonical signaling pathways downstream of GDF11 are not involved here, opposite to other systems (Clayton et al., 2020). Therefore, GDF11 signals through the ALK5 receptor to activate the SMAD2/3 signaling pathway and controls *HOX* expression in hPSC-derived axial progenitors.

Then, we characterized SMAD2/3 signaling dynamics induced by GDF11 as well as FGF2/GDF11 that potentiated the specification of caudal motoneurons in our previous work (Mouilleau et al., 2021). We exposed day 3 progenitors to different concentrations of GDF11 and measured the nuclear accumulation of P-SMAD2 (mean nuclear signal) over time (**Fig.**

1D). In all conditions, we observed the same temporal dynamics. An increased proportion of P-SMAD2 was observed in the majority of cells at 1h post exposure. At 3h post exposure, the proportion of P-SMAD2 positive cells and the nuclear intensity increased (**Fig. 1D/E**). At 6h post exposure, the nuclear intensity of P-SMAD2 had strongly decreased. However, cells never return to the initial baseline even at 12h post-exposure (**Fig. 1D/E**). Hence, the concentration of GDF11 do not modify the temporal dynamics but it modify the amplitude of the signal (at 3h mean signal intensity = 0.112 at 5 ng/ml, 0.198 at 10 ng/ml and 0.363 at 25 ng/ml) (**Fig. 1C/D**). An increase in P-SMAD2 was also observed with FGF2 and GDF11 exposure suggesting that the combination of the two factors also allows the activation of SMAD2 (**Fig. S2A**).

To verify the temporal dynamic of SMAD2/3 pathway at the single cell level, we used a hESC line in which SMAD2 is fused to mCitrine allowing live monitoring of the subcellular dynamics of SMAD2. We differentiated these cells into axial progenitors and recorded the cellular response to GDF11 over time. We observed an increase of the nucleo/cytoplasmic ratio from 1h to 3h post GDF11 treatment and a subsequent decreased in intensity (**Fig. S3**). The dynamics showed a similar profile in all recorded cells. Altogether, these results indicate that in axial progenitors exposed to GDF11, the nuclear localization of SMAD2/3 reaches its highest level around 3h post exposure, and then decrease over time without, however, returning to baseline level which indicate a perdurable low level of signaling.

To determine whether the rapid decrease in pathway activity after 3h was due to a degradation of GDF11 ligand, we added fresh GDF11 6h after the initial exposure, a time point at which nuclear P-SMAD2 has already severely decreased (**Fig. 1D**), and tested whether it re-induced P-SMAD2 nuclear localization. This second stimulation has no significant impact on nuclear P-SMAD2 intensity (**Fig. 1F**). Therefore, the decrease of nuclear SMAD2 is not due to a decrease in extracellular GDF11. Instead the SMAD2/3 pathway appear to be adaptive in hPSC axial progenitors, a property of many signaling pathways, such as ERK or SMAD pathways, characterized by a progressive decrease of pathway activity due to the induction of delayed negative feedbacks (Li and Elowitz, 2019). In other models, pathway adaptation plays a critical role in controlling pathway gene expression and cellular behavior (Ashall et al., 2009; Ryu et al., 2016; Sorre et al., 2014).

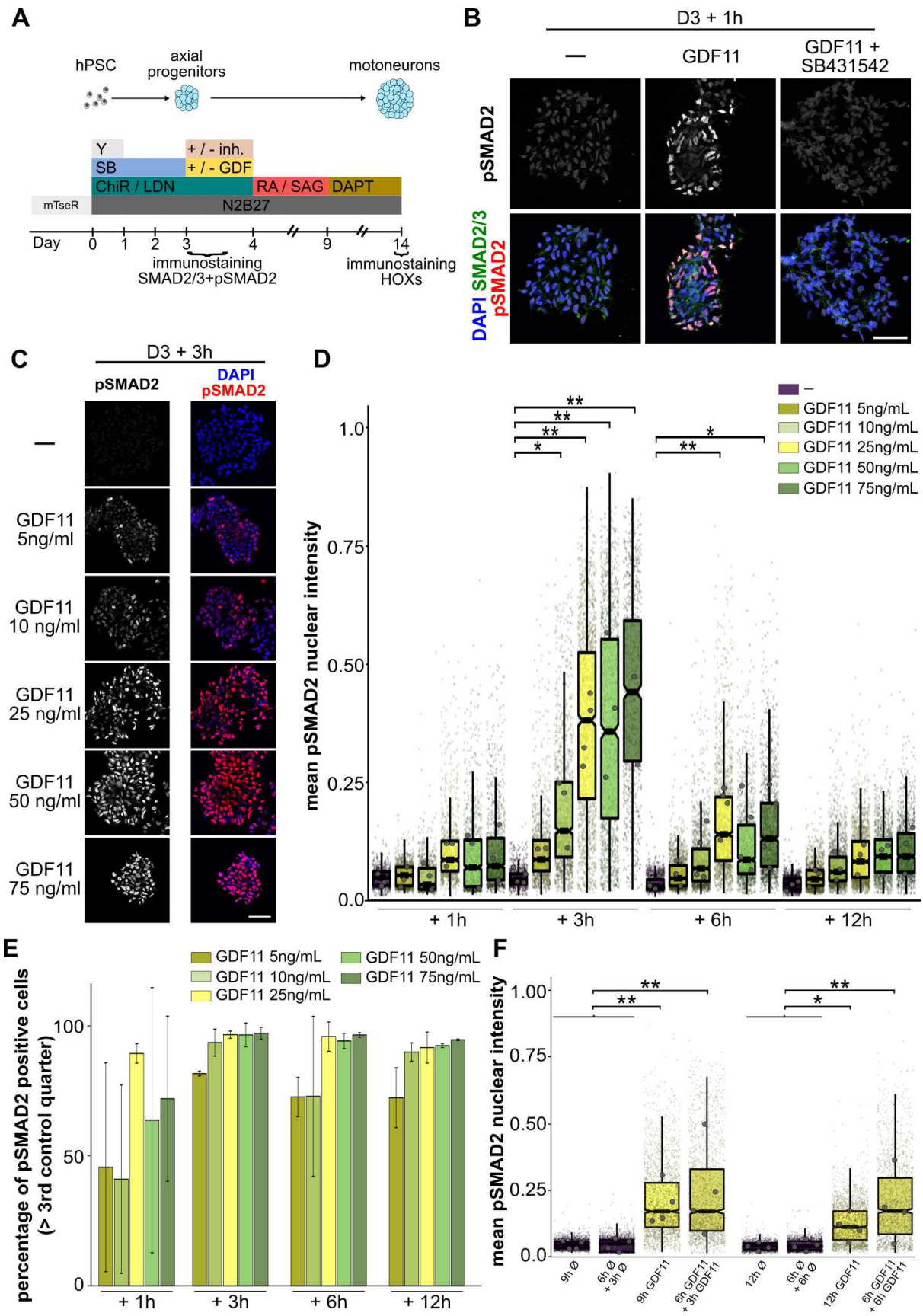


Figure 1 – GDF11 extrinsic factor activates SMAD2/3 pathway in axial progenitors. **A** - Differentiation conditions used in which GDF11 or pathways inhibitors are modulated. **B** - Immunostaining of SMAD2/3 and P-SMAD2 on plated axial progenitors. **C** - Immunostaining of P-SMAD2 on plated axial progenitors treated with different concentrations of GDF11. **D** - Box-plot of the mean nuclear signal intensity of pSMAD2 (n=2-4) for axial progenitors treated to different GDF11 concentrations. **E** - Percentage of axial progenitors staining for P-SMAD2 (n=2-4). For GDF11 conditions, cells with a mean

nuclear intensity of P-SMAD2 superior to the 3rd quarter of the untreated condition are considered as pSMAD2 positive cells. **F** - Box-plot of the mean nuclear signal intensity of pSMAD2 (n=4) for axial progenitors. Cells re-stimulated with fresh GDF11 are corresponding to “6h GDF11 + 3h GDF11” and “6h GDF11 + 6h GDF11”. In figure E, data are shown as mean±s.d. In figures D and F, each small dot corresponds to a single cell while the larger dots correspond to the mean of the biological replicates. The center line of box plot indicates the median; box limits the 1st and 3rd quartiles and whiskers the range. Kruskal-Wallis with non-parametric post hoc test: * if $P \leq 0.05$, ** if $P \leq 0.005$. Scale bar 50 μm .

Temporal dynamics of *HOX* induction downstream GDF11

To identify the temporal relationships between GDF11-SMAD2/3 dynamics and the downstream transcriptional programs including *HOX* genes, we exposed day 3 axial progenitors to GDF11 or FGF2/GDF11 and analyzed the changes in gene expression 3h and 6h post exposure (**Fig. 2A**).

We identified 28 GDF11-induced genes at 3h and 63 at 6h (fold change ≥ 3.0 and FDR ≤ 0.05) (**Fig. S4A**). As expected, we found an enrichment for genes associated with the TGF- β /GDF11 pathway (i.e. *PMEPA1*, *LEFTY1/2*, *PITX2*...) (**Fig. 2B, S4B**). Most of these genes upregulated at 3h and 6h are known targets of the TGF- β pathway (**Fig. 2B**) (Coda et al., 2017). Consistent with the SMAD2/3 adaptation occurring between 3 and 6h, several induced genes encode typical negative feedbacks of the TGF- β pathway including *LEFTY1/2*, *SMAD7*, *BAMBI* and *SKIL* (**Fig. 2B**). In addition, genes associated with the Wnt pathway (i.e. *WNT4/6/10A*, *SFRP5*) and FGF/MAPK pathway (i.e. *FGF18*, *PDGFB*, *DUSP9*) were also induced (**Fig. 2B, S4B**). All of these genes were also induced in the FGF2/GDF11 condition, along with other well-known targets of the FGF-MAPK pathways (i.e. *EGR1/2/3/4*, *ETS1*, *FOS*, *JUNB*...) in agreement with the addition of FGF2 (**Fig. S4B**).

In parallel to the induction of these canonical targets, we observed increased expression of *HOXA4*, *A9*, *A10*, *C9*, *D1*, *D8* and *D9* as early as 3h post GDF11 exposure (fold change ≥ 1.5 and FDR ≤ 0.1). Their expression was higher at 6h while *HOXC10* was induced only 6h post exposure but still with a low number of reads (**Fig. 2C, S4F**). Overall, the more 5' the genes are in the complexes, the lowest expressed they were post GDF11 or FGF2/GDF11 (**Fig. S4F**) while *HOX11/12* paralogs were not detected within this timeframe. These data indicated a conserved temporal collinearity downstream GDF11. We observed the same pattern of *HOX* induction upon FGF2/GDF11. However, FGF-dependent mid-cluster *HOX* genes were also increased (i.e. *HOXA5*, *C5*, *A6*, *A7*) likely due to an increase in FGF signaling by the exogenous FGF2 (**Fig. 2C, S2B/C, S4F/G**). This increase in *HOX* mRNAs upon GDF11 is transcription dependent and does not involve transcript stabilization since actinomycin D, a transcription inhibitor, fully blocked the induction (**Fig. S4H**).

To detect the chromatin changes associated with GDF11 treatment we performed preliminary results from Cut&Tag experiments aiming at detecting histone marks associated with permissive (H3K4me3) and repressive (H3K27me3) transcriptional states. Our preliminary results indicate that H3K4me3 was detected up to *HOXA5*, *C4* and *D3* at day 2, they were present up to *HOXA5*, *C8* and *D4* by day 3 consistent with the progressive induction of these genes over time. H3K4me3 was detected up to *HOXA10*, *C9* and *D9* 24h of after exposure to GDF11 (**Fig. S5A/C/D**). Permissive histone marks were not detected in weakly expressed genes such as *HOXC10* suggesting that this gene might be activated only in a subset of cells (**Fig. S5E**). Hence, in parallel to classic targets of the SMAD2/3 pathway, GDF11 induced a rapid increase in expression of *HOX9/10*.

Then, to verify the temporal dynamics of *HOX* activation at the single cell level we monitored *HOXC8*, *HOXC9* and *HOXC10* using fluorescent *in situ* hybridization (**Fig. 2D, S6A**). 6h after GDF11, 93.8% of the cells expressed *HOXC8* (60.8% in controls), 47.9% *HOXC9* (27.6% in controls) and 39.8% *HOXC10* (20.0% controls) (**Fig. 2E**). After 24h GDF11 exposure, almost all cells expressed *HOXC9* (85.0% versus 39.9% in controls) and 66.2% *HOXC10* (19.2% in controls) (**Fig. 2E**). At 48h, all cells appear to express *HOXC10* at a higher level than at 24h (**Fig. S6B**). FGF2/GDF11 induced the same dynamics of *HOX* induction (**Fig. 2E, S6B**). Overall, these data indicate that GDF11 induces *HOX9* and *10* but that the temporal collinearity is preserved with *HOXC10* peaking several hours later than *HOXC9*.

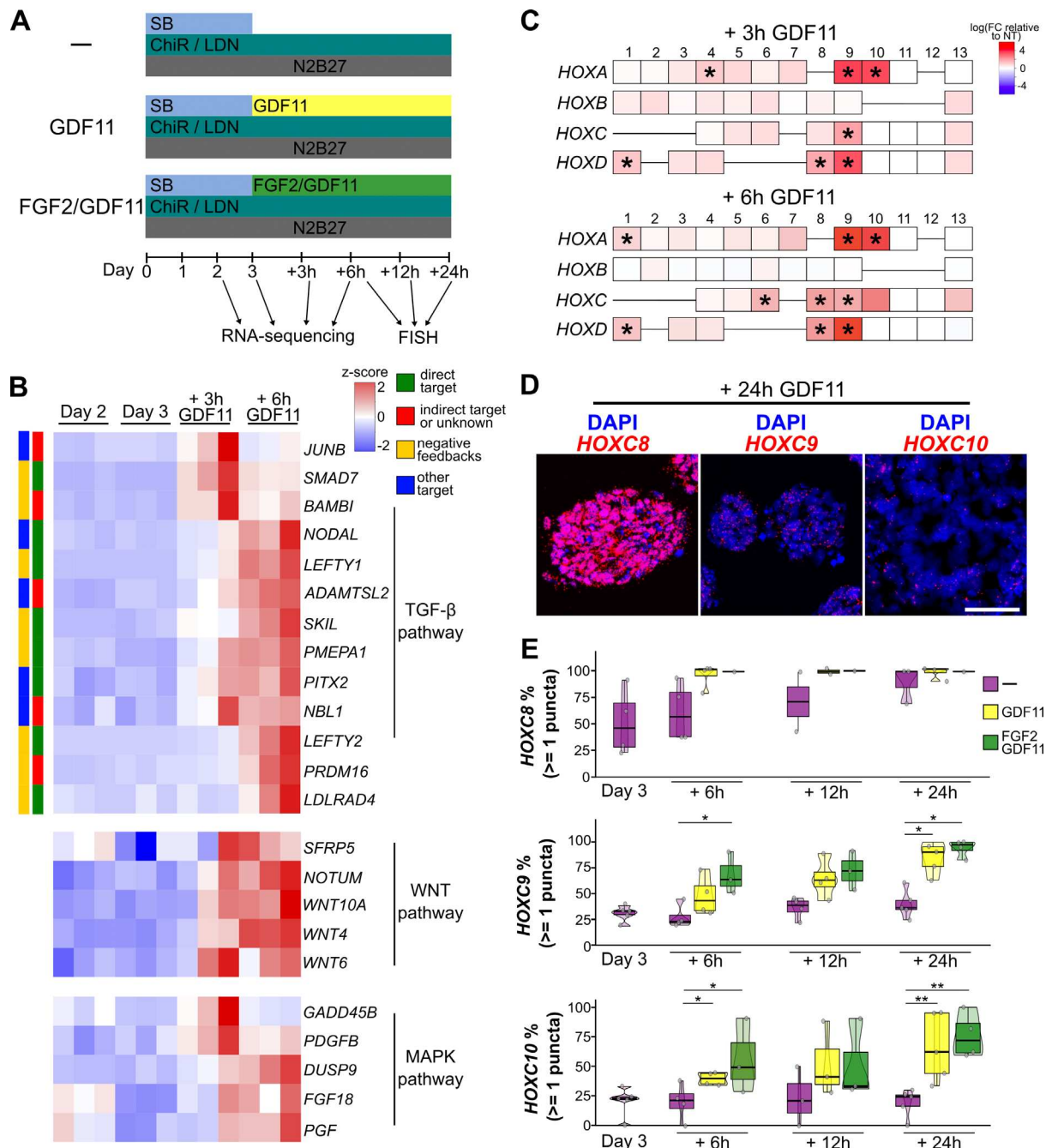


Figure 2 – GDF11 regulates *HOX* induction in axial progenitors. **A** - Differentiation conditions used in which *HOX* expressions were analyzed. **B** – Heatmap of the expression of genes (z-score) related to TGF- β , WNT and MAPK pathway target genes, which were identified in the functional enrichment analysis of Fig. S4B (n=3). The barcode for TGF- β target genes indicates whether they are direct/indirect and negative feedbacks/other. **C** - RNA-seq expression of *HOX* genes in which axial progenitors are treated to GDF11 during 3h or 6h (n=3). *HOX* expression is relative to untreated condition at the same time-point. The identification of genes differentially expressed “*” was performed using hyper-geometrical test (fold changes ≥ 1.5 , FDR ≤ 0.05). **D** - FISH staining for *HOXC8*, *C9* and *C10* of axial progenitors cryosections in day 3 differentiation treated 24h to GDF11. **E** - Percentage of *HOXC8*, *C9* and *C10* active cells. Cells with at least one puncta/transcript are accepted as transcriptionally active. Each circle is an independent biological replicate (n=1-5). Kruskal-Wallis with non-parametric post hoc test: * if $P \leq 0.05$, ** if $P \leq 0.005$. Scale bar 50 μ m.

HOX induction depends on the duration of SMAD2/3 signaling

We thus asked whether, the perdurable SMAD2/3 signaling that is maintained at low level post-adaptation (**Fig. 1D/E**) could be important to promote this increase in expression

of more 5' *HOX* genes in the 24h post GDF11. To test this hypothesis we first designed a strategy to be able to block SMAD2/3 signaling. We showed that blocking the ALK5 receptor using SB431542 after exposure to GDF11 was sufficient to revert nuclear P-SMAD2 to baseline level (**Fig. 3A-C**). We thus exposed cells to GDF11, added SB431542 at different time points (1h, 3h and 6h) post GDF11 addition and measured the expression of *HOXD9/A10/C10* at 24h (**Fig. 3D**). In all cases, the blockade of SMAD2/3 signaling resulted in lower level of *HOXD9*, *A10* and *C10* at 24h (**Fig. 3E**). Yet, the longer the signaling the highest the expression level. Therefore, cells integrate the duration of SMAD2/3 signaling to induce 5' *HOX* genes.

Hence, GDF11 induced a rapid increase in expression of *HOX9/10*, which continue to increase over time (level per cell and total number of cells) in a SMAD2/3-dependent manner. Yet of *HOX11* and *12* paralogs were not induced. Since we observed that SMAD2/3 pathway is rapidly down regulated, we reasoned that pathway adaptation, which limits pathway throughput, might play a role in slowing the induction of more and more 5' genes.

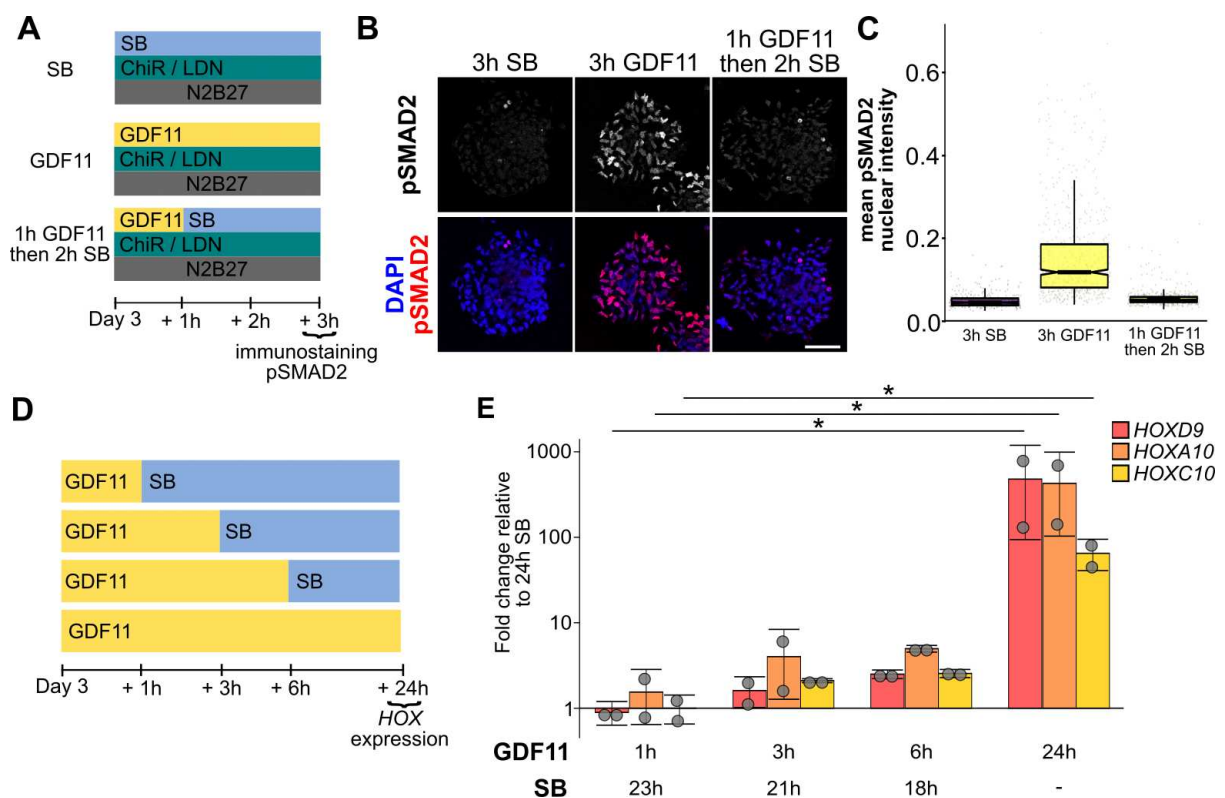


Figure 3 – Duration of GDF11 activity controls the expression *HOX* genes. **A** - Differentiation conditions used in **B** and **C**. **B** - Immunostaining of P-SMAD2 on plated axial progenitors. **C** - Box-plot of the mean nuclear signal intensity of P-SMAD2 (n=2) for axial progenitors treated to SB431542, GDF11 or GDF11 replaced by SB431542 after 1h. **D** - Experimental approach used in **E**. **E** - RT-qPCR of *HOXD9*, *C10* and *D9* level of expression in axial progenitors treated during different time by GDF11 (n=2). Fold change in expression is compared to axial progenitors treated 24h with SB431542. In figure **E**, data are shown as mean±s.d and each circle is an independent biological replicate. In figures **C**, each small dot corresponds to a single cell. The center line of box plot indicates the median; box limits the 1st and 3rd quartiles and whiskers the range. Kruskal-Wallis with non-parametric post hoc test: * if $P \leq 0.05$. Scale bar 50 μ m.

SMAD2/3 signaling pathway temporal dynamics might encode the tempo of *HOX* activation

Pathway adaptation is mediated by the neosynthesis of several pathway-induced negative feedbacks such as *LEFTY1/2*, *SMAD7*, *SKIL* (Li and Elowitz, 2019) that are activated upon GDF11 stimulation in axial progenitors (**Fig. 2B**). Previous data have demonstrated that pathway adaptation can be prevented by blocking negative feedbacks synthesis using protein synthesis inhibitors (Wilson et al., 2017).

To test this hypothesis, we first investigated whether *LEFTY1/2* protein was induced after GDF11 exposure and whether this neosynthesis was prevented by cycloheximide (CHX), a protein synthesis inhibitor. We indeed observed an induction of *LEFTY1/2* after exposure to GDF11 and its blocking by CHX (**Fig. 4A**). We then measured the level of nuclear P-SMAD2 at 6h, a time point at which adaptation occurred in control condition (**Fig. 4B**). We observed that in the progenitors exposed to GDF11/CHX, nuclear P-SMAD2 had a tendency to display a higher intensity as compared to the condition without CHX (6h GDF11 mean signal intensity = 0.180; 6h GDF11/CHX mean signal intensity = 0.328) (**Fig. 4B**).

We then exposed axial progenitors to CHX or anisomycin, two protein synthesis inhibitors together with FGF2/GDF11 (**Fig. 4C**). First, neither of the inhibitors blocked *HOXC6*, *C8*, *C9* and *C10* induction, indicating that *HOX* genes are direct targets of the signaling pathways activated by these factors. Second, not only translation inhibition did not block *HOX* induction but it lead, to the super activation of the *HOXC10* (**Fig. 4C**). *HOX* activation in the GDF11/CHX condition was prevented by ALK5 inhibition indicating that CHX effect on *HOX* genes is linked to the activation of the GDF pathway (**Fig. 4D**).

To further characterize the impact of CHX on *HOX* expression, we conducted a transcriptomic analysis comparing the FGF2/GDF11 and FGF2/GDF11/CHX conditions. Our analysis not only confirmed the robust activation of *HOXA10*, *C10* and *D10* paralogs but also revealed an additional induction of the *HOXA11*, *C11*, *D11* and *I3* paralogs, which were not expressed in the absence of CHX (**Fig. 4E**). We also observed a strong increase of canonical SMAD2/3 target genes in the CHX condition (*LEFTY1* 2.8 folds, *LEFTY2* 2.5 folds, *PMEPAI* 3.0 folds, *SMAD7* 2.7 folds) (**Fig. 4F**). Therefore, exposing axial progenitors to GDF11 while blocking translation induces a strong increase in the expression of the direct targets of the SMAD2/3 pathway including 5' *HOX* genes. As a major part of the GDF11-induced transcriptional program encodes negative feedbacks from the SMAD2/3 pathway, we reasoned

that CHX might prevent the translation of pathway inhibitors resulting in a sustained activation of the pathway and stronger induction of target genes including *HOX*.

Overall, these results suggest that CHX induces a strong induction of TGF- β canonical target genes as well as the induction of *HOX10-11* genes. To determine whether this increase in SMAD2/3 signaling induced SMAD2/3 genomic recruitment at target genes, we performed preliminary Cut&Tag against SMAD2/3 in GDF11 and GDF11/CHX condition at 3h post exposure. Our preliminary results indicate that SMAD2/3 appear to be detected near previously described binding sites at proximity of SMAD2/3 target genes such as *LEFTY1/2*, *SMAD7* or *PMEPA1* (**Fig. S7**). Importantly we also detected possible SMAD2/3 binding sites in the *HOX* clusters between *HOXA7-A10* and *HOXC8-C10* (**Fig. 4G**). It was previously demonstrated that the enhancer between *HOXC8* and *HOXC9* is bound by CDX2 and SMAD3 in a heterologous system (Gaunt, 2017). At all these sites, SMAD2/3 peaks were much higher in the GDF11/CHX condition than in the GDF11 condition alone, suggesting an increase in SMAD2/3 recruitment at population level (**Fig. 4G, S7A/B**). However, despite the strong induction of *HOXA11*, *C11* and *D11* at 6h in the transcriptomic analysis, at this stage we did not observe SMAD2/3 binding near these genes, including at the level of the *HOXD11* enhancer previously described to bind SMAD2/3 in a heterologous system (Gaunt et al., 2013). This might be due to the sensitivity of the assay or to the fact that we performed the Cut&Tag only at 3h so far while the transcriptomic analysis post CHX was performed at 6h.

These results indicate that blocking protein synthesis prevents SMAD2/3 feedbacks production which correlate with an increase in SMAD2/3 duration of signaling and an increase genomic recruitment near SMAD2/3 target genes including *HOX* that are super induced. Altogether, this indicates that the tempo of induction of more and more 5' *HOX* genes might be regulated by the amplitude and duration of activity of upstream signaling pathways. This mechanism may be an emerging property controlling the progressive induction of *HOX* genes during axial elongation.

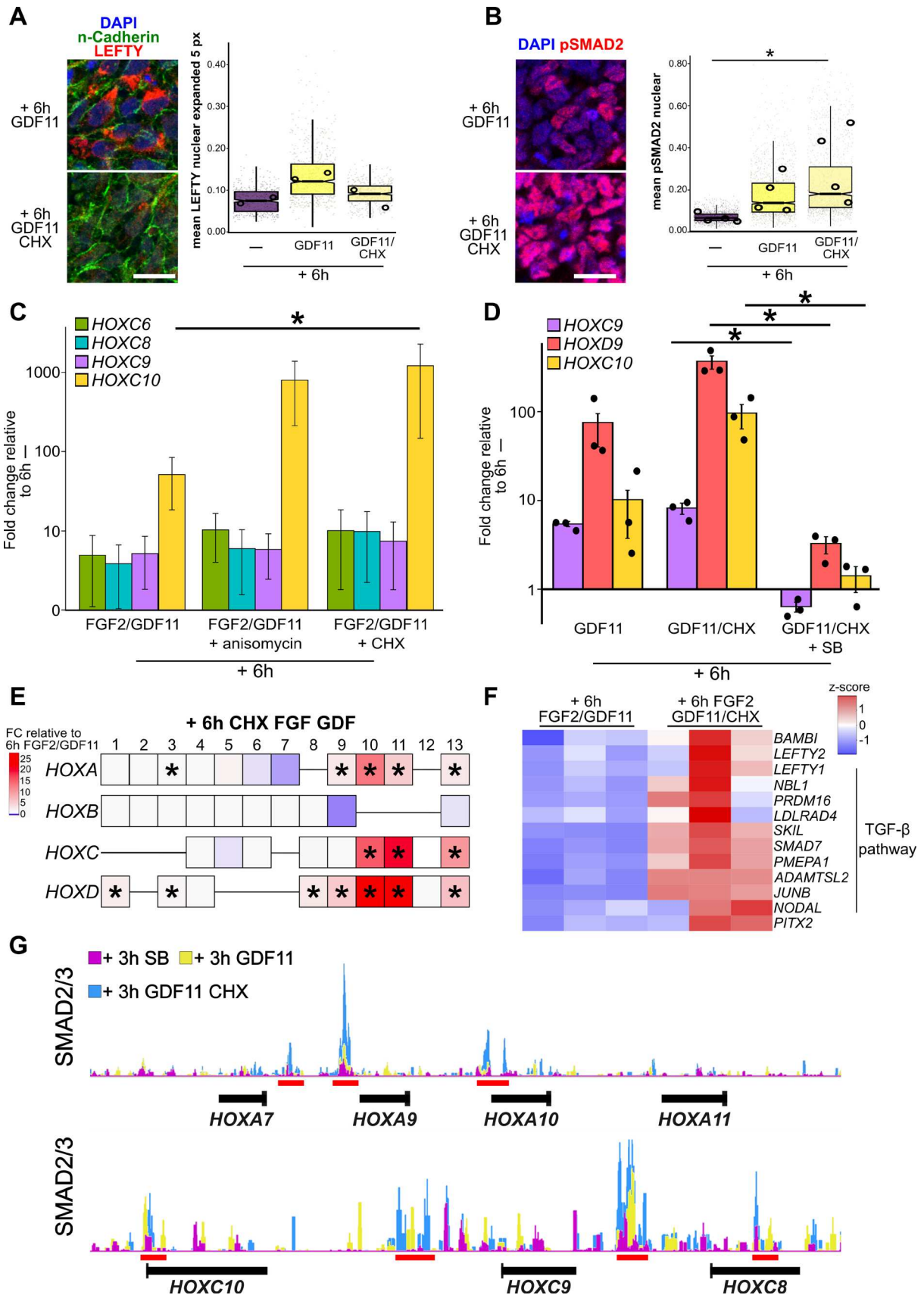


Figure 4 – Over activation of SMAD2/3 pathway accelerates the induction of more 5' *HOX* genes. **A** - Immunostaining of LEFTY on plated axial progenitors treated 6h to GDF11 or GDF11/CHX and box-plot of the mean intensity mean intensity in a 5px periphery to the nucleus (n=2). **B** - Immunostaining of P-SMAD2 on plated axial progenitors treated 6h to GDF11 or GDF11/CHX and box-plot of the mean nuclear signal intensity (n=4). **C** - RT-qPCR of *HOXC6/C8/C9/C10* level of expression in axial progenitors treated during 6h to FGF2/GDF11, FGF2/GDF11/anisomycin or FGF2/GDF11/CHX (n=3). Fold change

in expression is compared to non-treated progenitors at 6h. **D** - RT-qPCR of *HOXC9/D9/C10* level of expression in axial progenitors treated during 6h to GDF11, GDF11/CHX or GDF11/CHX/SB (n=3). Fold change in expression is compared to non-treated progenitors at 6h. **E** - Fold change expression of *HOX* genes after 6h FGF2/GDF11/CHX treatment. *HOX* expression is relative to the axial progenitors treated with FGF2/GDF11 (n=3). The identification of genes differentially expressed “**” was performed using hyper-geometrical test (fold changes ≥ 1.5 , FDR ≤ 0.05). **F** - Heatmap based on genes expression (z-score) for genes related to the TGF- β pathway in axial progenitors treated during 6h to FGF2/GDF11 or FGF2/GDF11/CHX (n=3). **G** - Cut&Tag showing tracks of SMAD2 and SMAD3 in *HOXA/C* clusters for the indicated treatments (n=1). Red bar indicates the position of potential binding sites for SMAD2/3. In figure A and B, data are shown as mean \pm s.d and each circle is an independent biological replicate. In figures E and F, each small dot corresponds to a single cell while the larger dots correspond to the mean of the biological replicates. The center line of box plot indicates the median; box limits the 1st and 3rd quartiles and whiskers the range. Kruskal-Wallis with non-parametric post hoc test: * if $P \leq 0.05$. Scale bar 20 μ m.

Supplementary data:

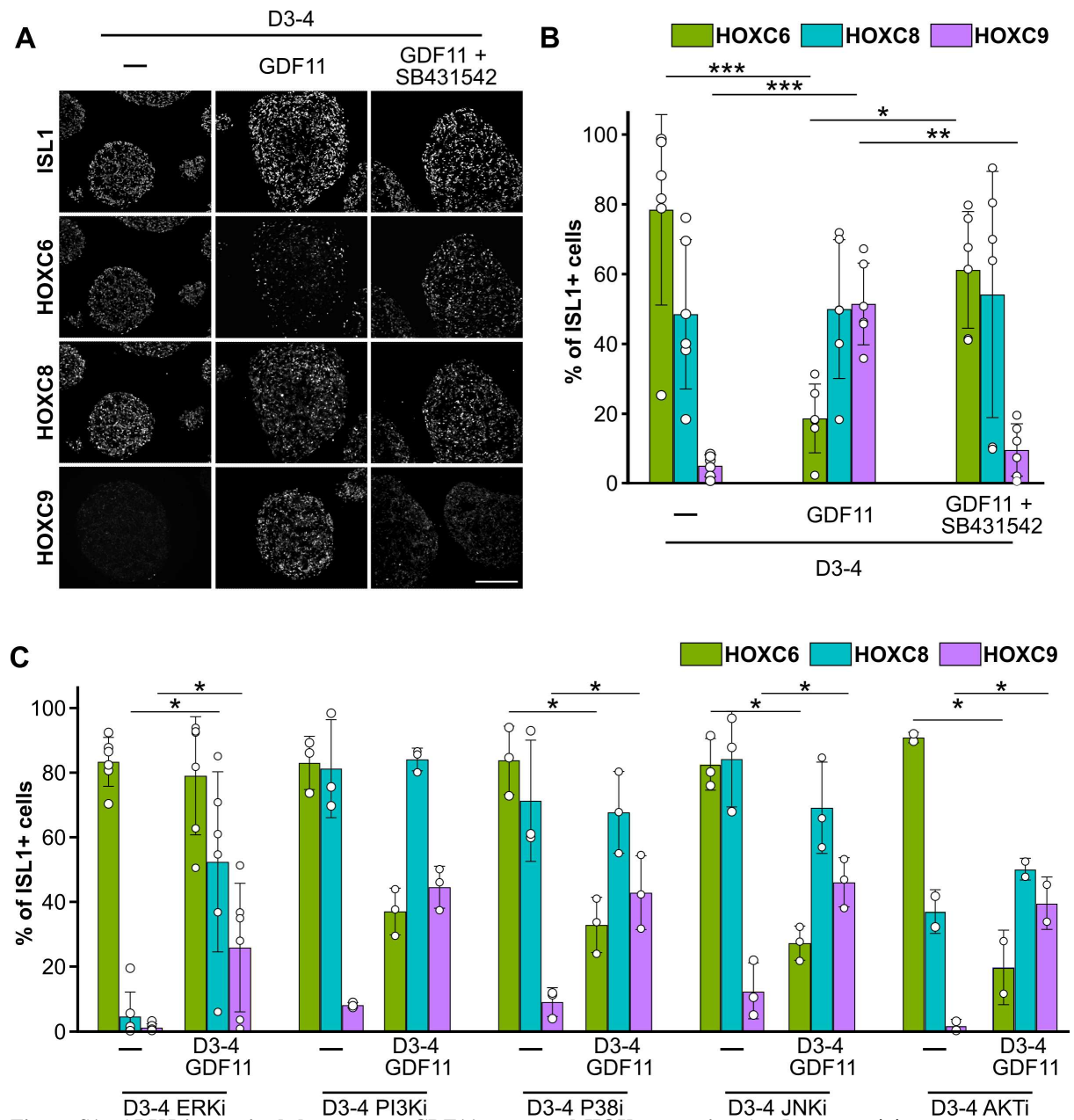


Figure S1 –ALK5 is required downstream GDF11 to control HOX patterning A - Immunostaining on motoneurons cryosections at day 14, against ISL1, HOXC6, HOXC8 and HOXC9. B,C - Proportion of motoneurons (ISL1⁺ cells) expressing the indicated markers (n=2-5). In figure B and C, data are shown as mean±s.d. and each circle is an independent biological replicate. Kruskal-Wallis with non-parametric post hoc test: * if $P \leq 0.05$, ** if $P \leq 0.005$, *** if $P \leq 0.001$. Scale bar 150 μ m.

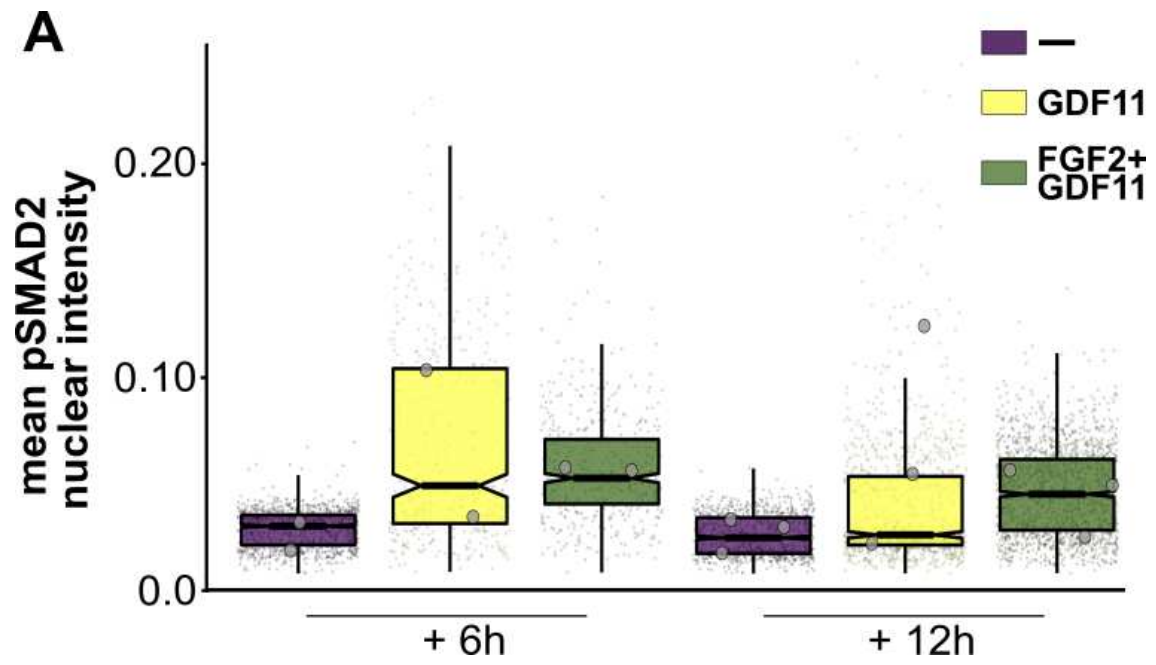


Figure S2 – FGF2/GDF11 co-stimulation activates SMAD2/3 pathway. A - Box-plot of the mean nuclear signal intensity of P-SMAD2 (n=2-4) for axial progenitors treated to GDF11 or FGF2/GDF11.

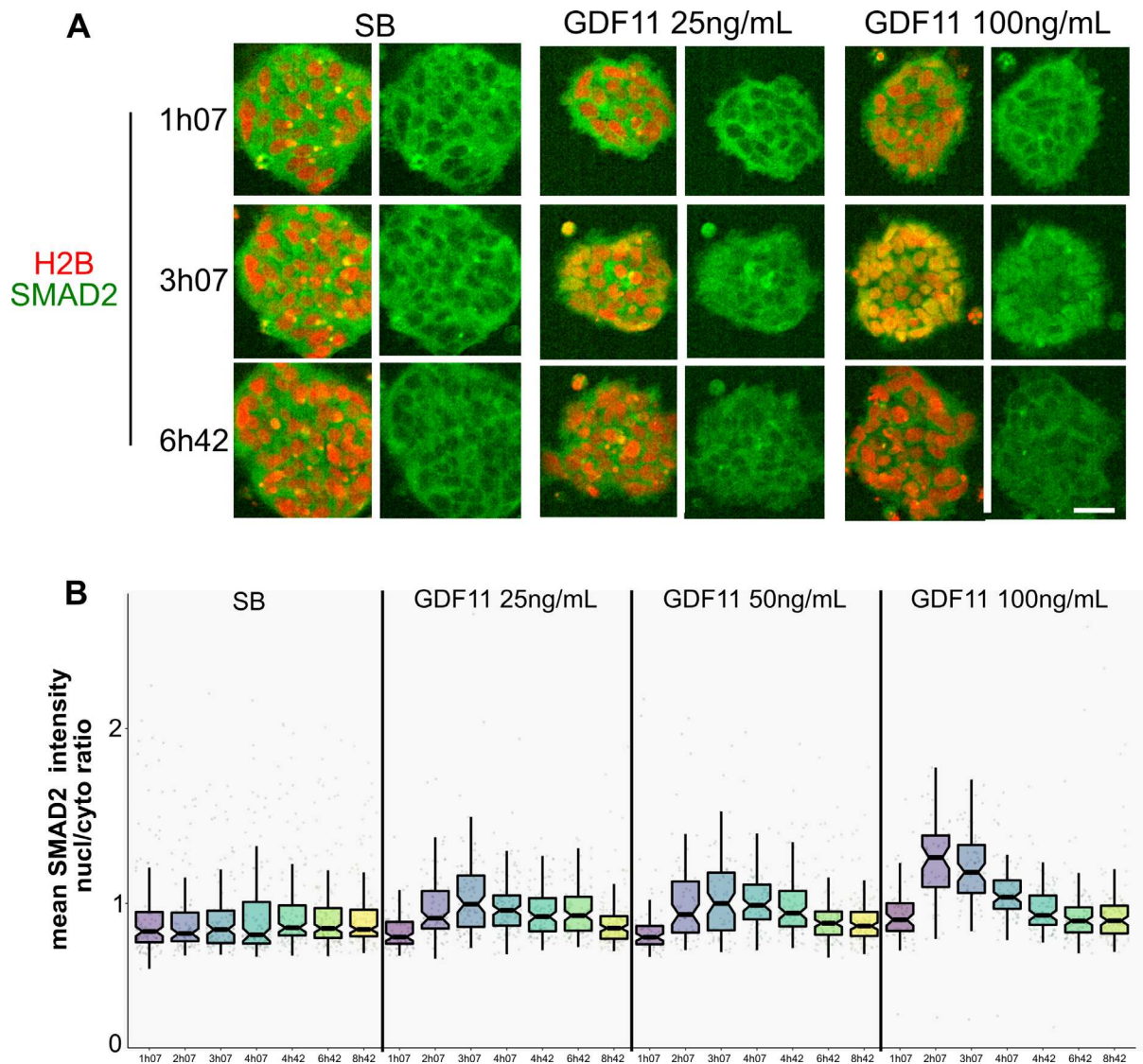


Figure S3 – Exposure of axial progenitors to GDF11 induces transient SMAD2 nuclear recruitment. **A** - Live imaging acquisition of hESC differentiated into axial progenitor expressing nuclear reporter H2B-mCherry (in red) and mCitrine-SMAD2 (in green) reporter constructs at different time points post medium change on day 3 of differentiation. **B** - Box-plot of mean intensity of mCitrine-Smad2 nucleocytoplasmic ratio within axial progenitors from SB431542 or GDF11 conditions at different time point (n=1). Each dot represents one measurement made on a single cell. Scale bar 20 μ m.

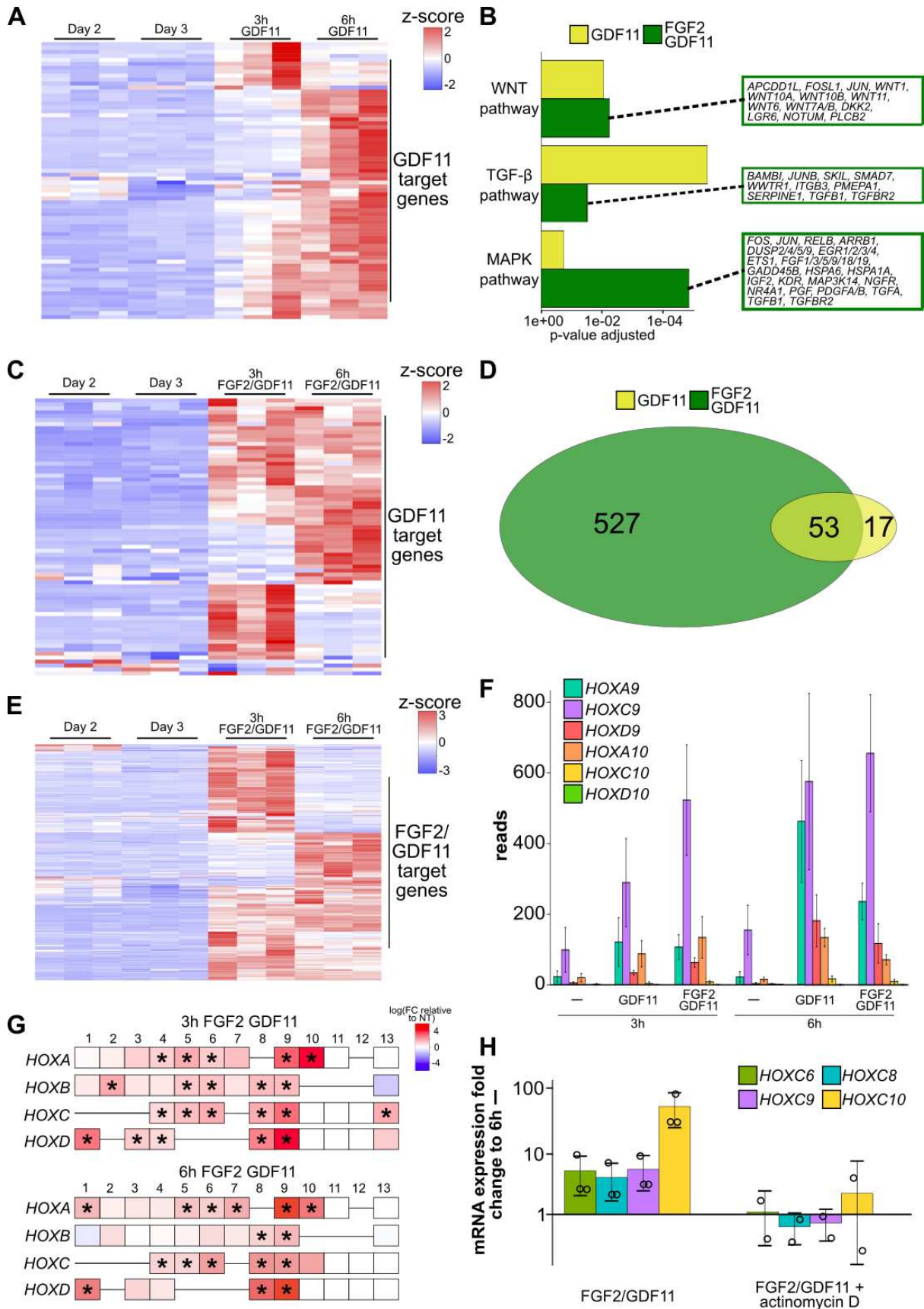


Figure S4 – GDF11 and FGF2/GDF11 upregulate transcriptional programs. **A** - Heatmap based on genes expression (z-score) for GDF11-upregulated genes (n=3). **B** - Functional enrichment analysis (DAVID) of the genes upregulated by GDF11 or FGF2/GDF11. **C** - Heatmap based on genes expression (z-score) for GDF11-upregulated genes in FGF2/GDF11 conditions. **D** - Venn diagram of genes upregulated by GDF11 and FGF2/GDF11 stimulations. **E** - Heatmap based on genes expression (z-score) for FGF2/GDF11-upregulated genes (n=3). **F** - Treatment and temporal evolution of *HOXA9/C9/D9/A10/C10/D10*

genes expression (reads) in axial progenitors (n=3). **G** - RNA-seq expression of *HOX* genes in which axial progenitors are treated to FGF2/GDF11 during 3h or 6h (n=3). *HOX* expression is relative to untreated condition at the same time-point. The identification of genes differentially expressed “*” was performed using hyper-geometrical test (fold changes ≥ 1.5 , FDR ≤ 0.05). **H** - RT-qPCR of *HOXC6/C8/C9/C10* level of expression in axial progenitors treated during 6h to FGF2/GDF11 or FGF2/GDF11/actinomycin D (n=2-3). Fold change in expression is compared to non-treated progenitors at 6h. In figure F and H, data are shown as mean \pm s.d and each circle is an independent biological replicate.

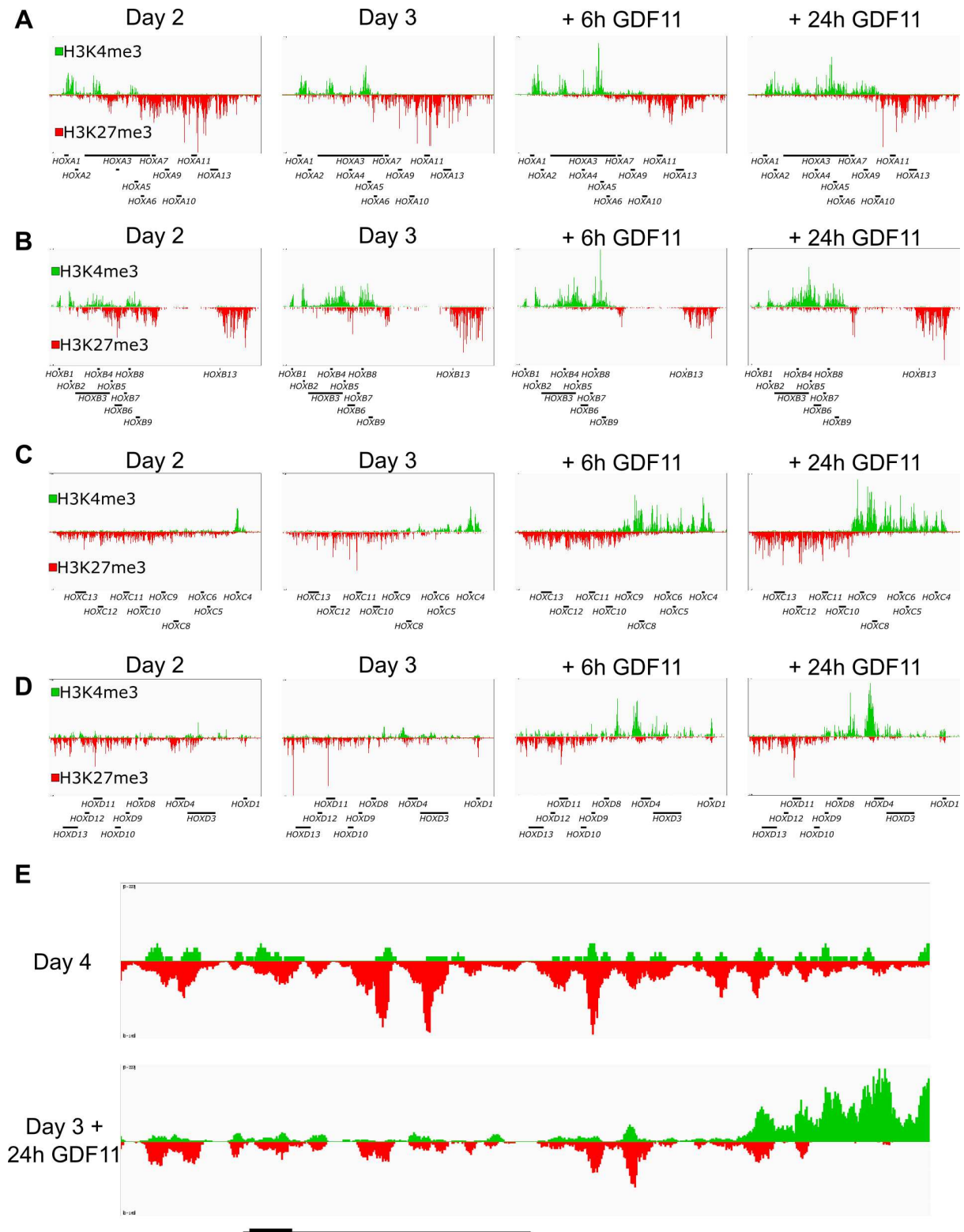


Figure S5 – GDF11 modify the histone marks along *HOX* clusters. A-D - Cut&Tag showing tracks of H3K4me3 (green) and H3K27me3 (red) at day 2, day 3, 6h GDF11 and 24h GDF11 in *HOXA* (A), *HOXB* (B), *HOXC* (C) and *HOXD* (D) (n=1). E – Zoom of *HOXC10* showing H3K4me3 and H3K27me3 histone marks (n=1).

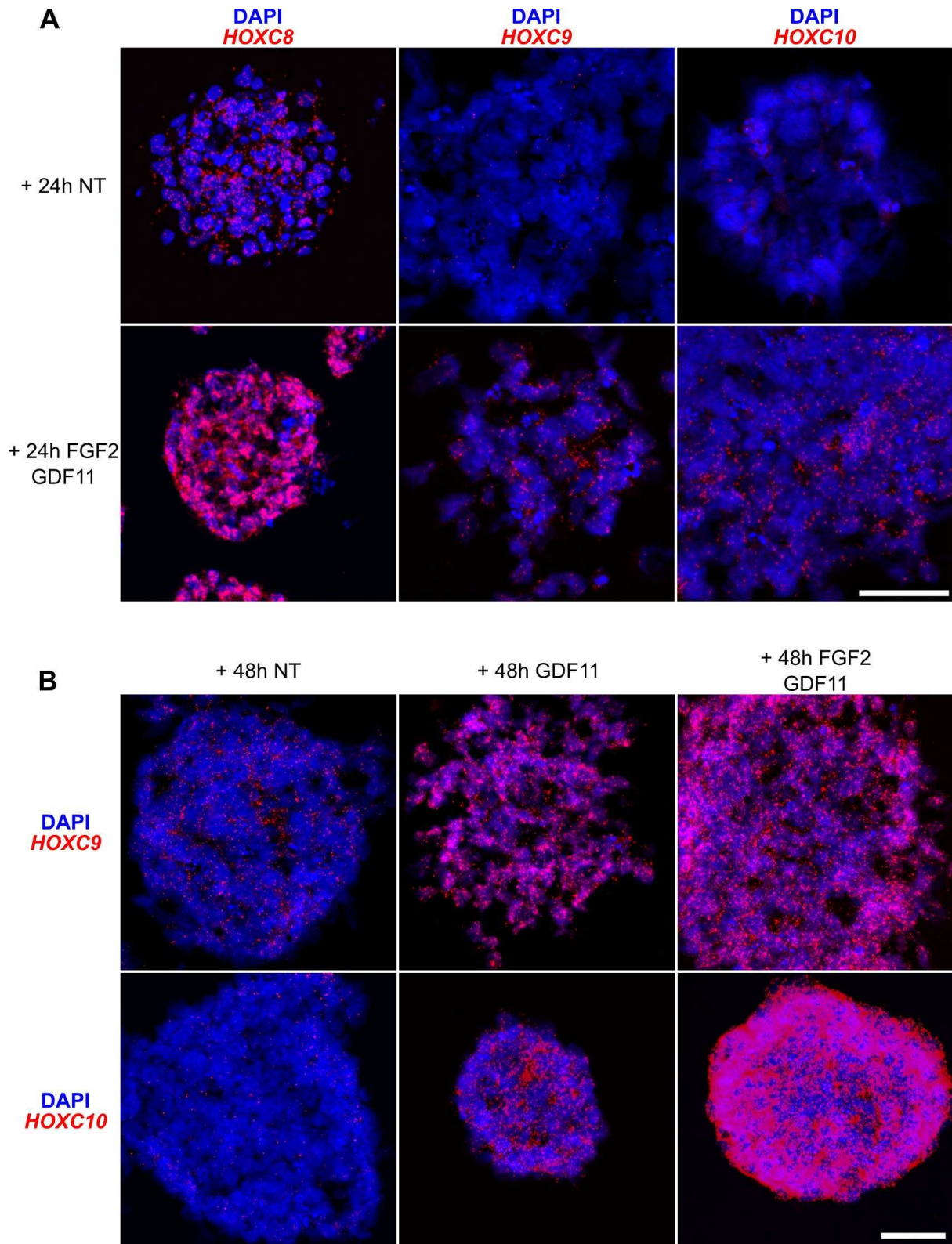


Figure S6 – Long durations of GDF11 exposure result in high and homogeneous expression of *HOX* genes. A - FISH staining for *HOXC8/C9/C10* of axial progenitors cryosections in day 3 differentiation treated 24h to FGF2/GDF11. B - FISH staining for *HOXC9/C10* of axial progenitors cryosections in day 3 differentiation treated 48h to GDF11 or FGF2/GDF11. Scale bar 50 μ m.

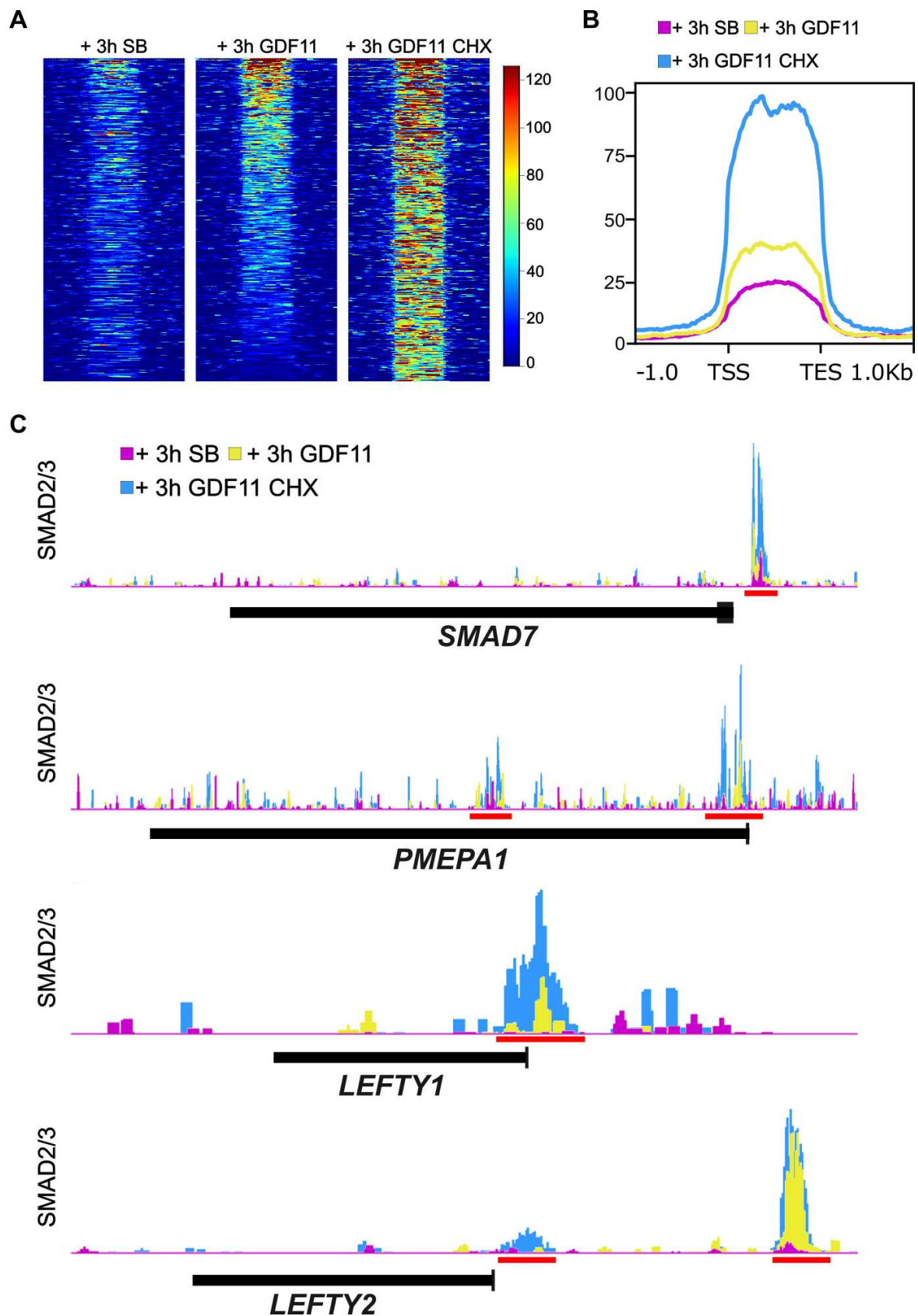


Figure S7 – SMAD2/3 bind near to different genes associated to TGF- β pathway. **A** – Heatmaps of normalized SMAD2/3 Cut&Tag signals obtained in 3h SB, GDF11 or GDF11/CHX. The ranking of the *cis*-regulatory modules is the same on all heatmaps and follows the occupancy levels of SMAD2/3 in descending order. **B** – Average profiles of SMAD2/3 signals at *cis*-regulatory modules belonging to the peaks defined in A in the indicated 3h SB, GDF11 or GDF11/CHX condition. **C** – Cut&Tag showing tracks of SMAD2 and SMAD3 near to *SMAD7*, *PMEPA1*, *LEFTY1* and *LEFTY2*. Red bar indicates the position of potential binding sites for SMAD2/3.

Materials and methods

Pluripotent stem cell

Experiments were performed with induced pluripotent stem cell (iPSC) line WTSli002-A obtained from the European Bank for Pluripotent Stem Cells (EBISC). iPSCs were cultured on matrigel diluted according to manufacturer instructions (Corning) in mTSER1 medium (Stem Cell Technologies) and amplified using EDTA (Life Technologies) clump-based passaging. They were tested for potential mycoplasma contamination every other week (MycoAlert™ Mycoplasma Detection Kit, Lonza, LT07-118). No contamination was detected during the study. iPSC frozen in Cryostor CS10 (Stem Cell Technologies) were thawed in presence of Y-27632 and the culture media was changed every day. Relevant ethic committees (declaration DC-2015-2559) approved experiments with hiPSC. For live imaging experiment with Citrine-SMAD2 hESC, cell were obtained from *Brivanlou Laboratory*.

hPSC differentiation

hPSC embryoid body-based differentiation was performed as previously described (Maury et al., 2015). hPSC were dissociated with cold accutase (Life Technologies) for 5 min at 37°C and resuspended in differentiation medium N2B27 (Advanced DMEM F12 and Neurobasal, Life Technologies), supplemented with N2 (Life Technologies), B27 without Vitamin A (Life Technologies), penicillin/streptomycin 1%, β-mercaptoethanol 0.1% (Life Technologies), Y-27632, CHIR-99021, SB-431542 and LDN-193189. Cells were seeded in ultra-low attachment 6 well plates (Corning) (2×10^5 cells per ml) to form embryoid bodies. All conditions of differentiation received the same medium, without Y-27632, at day 2 and at day 3, but SB431542 was removed at day 3. SAG, RA, GDF11 and DAPT were added at indicated time points. For motoneurons differentiation, the cell media were changed on day 4, day 7, day 9 and day 11, unless specified. Growth factors and small molecules are listed in Table 1.

Immunostaining

For P-SMAD2 and SMAD2/3 immunostaining on axial progenitors at day 3, embryoid bodies were plated on matrigel (Corning, diluted according to the manufacturer's recommendation) coated coverslips and allowed to adhere during 60 min at 37°C prior to fixation with 4% PFA containing phosphostop (Sigma) for 30 min, then washed 3 times for 10 min in 1X PBS. Cells were then treated with 1% SDS for 30 min at 37°C. The other staining for axial progenitors were fixed 15 min with 4% PFA at 4°C and rinsed with 1X PBS 3 times

for 5 min. For immunostaining on motoneurons at day 14, embryoid bodies were collected and fixed with 4% PFA 5 min at 4°C, rinsed with 1X PBS 3 times for 5 min. Embryoid bodies at day 14 were cryoprotected with 30% sucrose and embedded in OCT (Leica) prior sectioning using Cryostat Nx70 cryostat (13 µm thickness). Then, full embryoid bodies (axial progenitors) or slices (motoneurons) were incubated with saturation solution (1X PBS, FBS 10 %, 0,2% Triton) for 10 min, then the saturation solution was replaced by a solution containing primary antibodies diluted in 1X PBS, 2% FBS, 0,2% Triton and slices were incubated overnight at 4°C in a humidified chamber. Slices were washed at least 3 times 10 min with 1X PBS then exposed to secondary antibodies diluted in 1X PBS, 2% FBS, 0,2% Triton for 1h, washed 3 times 10 min with 1X PBS then incubated with DAPI for 10 min and washed with 1X PBS one last time. Slices were covered with fluoromount (Cliniscience) and a microscopy glass slide. Antibodies are listed in Table 2.

RNAscope

Embryoid bodies were fixed for 24h in 4% PFA at 4°C and rinsed with 1X PBS 3 times for 10 min. They were then cryoprotected with 30% sucrose and embedded in OCT (Leica) prior sectioning using Cryostat Nx70 cryostat (13 µm thickness). All reagents used for FISH are from the RNAscope® Multiplex Fluorescent Reagent v2 kit. Sections were first treated with hydrogen peroxide for 10 min at room temperature. A 99°C unmasking solution is then applied for 5 min. Protease III treatments lasting 30 min at 40°C are then carried out. *HOXC9* (ref: 584431-c3) and *HOXC10* (ref: 803141-C2) RNA probes were hybridized for 2h at 40°C. To amplify the hybridization signal, RNAscope® Multiplex FL v2 AMP1, AMP2 and AMP3 solutions were applied for 30 min, 30 min and 15 min respectively at 40°C. *HOXC9* and *HOXC10* were reveal using RNAscope® Multiplex FL v2 HRP-C3 and RNAscope® Multiplex FL v2 HRP-C2 which were used to bind 1:1000 Opal 520 (PerkinElmer, FP1487001KT). Finally, sections were mounted using Prolong Gold Antifade Mountant (Thermofisher).

Image acquisition and analysis

Slices embryoid bodies at day 14 were visualized and imaged either on an epifluorescent microscope DM6000 (Leica) with metamorph software (Metamorph inc), or with an automated inverted epifluorescent microscope Celldiscoverer 7 (Zeiss), using Zeiss Black software (Zeiss). Full embryoid bodies at day 3 were visualized and imaged using a confocal microscope TCS SP5 II (Leica) mounted on a DM6000 microscope stage on a single z. Live imaging experiments were acquired using a differential spinning disk DSD2 (Andor)

from the laboratory of B. Sorre. For RNAscope, the full thickness of the samples were imaged with a z-step of 1 μm using a confocal microscope TCS SP5 II (Leica) mounted on a DM6000 microscope stage. Acquisitions were done at the imaging facility of the Institut du Fer à Moulin. At least three different positions were imaged per biological replicate for the different experiments.

Then, the images were exported, processed (contrast/brightness adjustments) with Fiji software to display but were not modified for quantification. For P-SMAD2, LEFTY and HOXC6/C8/C9 nuclear intensity, quantitative analyses on images were performed using the CellProfiler software (Broad institute, open source). All images across conditions were automatically analyzed in batch ensuring unbiased analysis. DAPI stained nuclei were segmented into primary objects using CellProfiler segmentation pipeline and intensities of fluorescence were calculated in each primary object for HOXC6/C8/C8 and P-SMAD2 or in a periphery of 5px of each primary object for LEFTY. To determine the percentage of positive cells stained for P-SMAD2, the third quarter value of the untreated condition was calculated for each replicate, and in the GDF11 condition, cells with P-SMAD2 intensity greater than this value in their respective replicates were considered positive cells.

For live imaging, quantitative analyses were performed using the CellProfiler and the ratio between mean SMAD2 intensity within the nuclear mask area versus the cytoplasmic mask area was calculated in each cell. For the RNAscope, samples were analyzed manually in the full thickness of the section, and cells with at least one puncta were considered positive regardless of signal intensity.

Quantitative RT-qPCR analysis

Total RNA was extracted with a NucleoSpin RNA Mini kit (Macherey Nagel) following manufacturer instructions and cDNA synthesized using SuperScript III (Invitrogen). Real time quantitative PCR (RTqPCR) was performed using QuantStudio 5 Real-Time PCR System (Thermofisher Scientific) and a mix with qPCR Brilliant II SYBR MM with low ROX (Agilent). All expression data were normalized to Cyclophilin A mRNA. All analyses were performed with three technical replicates per plate. Relative expression levels were determined by calculating $2^{-\Delta\Delta C_t}$. Primers are listed in Table 3.

Sequencing and analysis

To generate samples for RNA-seq, total RNA was extracted from axial progenitors using NucleoSpin RNA Mini kit (Macherey Nagel) following manufacturer instructions. To

generate samples for Cut&Tag against H3K27me3 (Cell Signaling, 9733) and H3K4me3 (Siagenode, C154100003), each point was prepared from 100 000 axial progenitors following the procedure described in “Bench top CUT&Tag V.2” available on protocols.io. For Cut&Tag against SMAD2/3 (Cell Signaling, 8685), each point was prepared from 500 000 axial progenitors and cell nuclei have been isolated.

Libraries for RNA-seq and Cut&Tag samples were analysed using 4200 TapeStation system (Agilent) and sequenced 150pb (paired-end) with HiSeq X-ten (BGI). The quality of paired-end read sequences was assessed using FastQC (v0.11.9). Low quality nucleotides and Illumina adapter sequences were removed using Trimmomatic (v0.39) and PCR duplicate reads were removed using BBmap clumpify (v38.87). Filtered reads were aligned to the hg38 reference genome using Hisat2 (v2.4.1). The position of reads relative to genes and matrix generation were computed using StringTie.

For transcriptome dataset, the evaluation of the differential expression between two conditions was performed using the EdgeR package (v3.26) under R. For whole genome analysis, genes with fold changes ≥ 3.0 and $FDR \leq 0.05$ were considered as upregulated. For *HOX* cluster analysis, genes with fold changes ≥ 1.5 and $FDR \leq 0.1$ were considered as upregulated. Functional enrichment analysis was performed using DAVID.

For Cut&Tag against SMAD2/3, peaks were called using MACS2 (v2.2) with parameters `callpeak -t -f BAMPE --broad -g hs`. Reads coverage on defined CRMs was calculated using VisRseq (v0.9.42) and only peaks with a signal greater than 20 RPKM in 3h GDF11 or GDF11/CHX conditions, and a fold change greater than 2, compared with control conditions, were retained.

Statistical analysis

Independent experiments are differentiations performed on different weeks and started with newly thawed hPSCs. Number of n are indicated in figure legends. For qPCR and immunostaining data, each replicate corresponded to the mean value of the experiment, and Kruskal-Wallis and Dunn tests were performed following normality tests. For RNA-seq data, differential analysis were performed using hyper-geometrical test.

Table 1. Growth factors and small molecules

Molecule	Concentration	Supplier	Reference
Y-27632 (2HCl)	10 μ M	Stemgent	04-0012-H-10
Y-27632 (2HCl)	10 μ M	StemCell Technologies	72304
RA	100 nM	Sigma	R2625
LDN193189	100 nM	Stemgent	STE04-0074
SB431542	20 μ M	Stemgent	STE04-0010-10
CHIR99021 (CT99021)	4 μ M	Selleckchem	S1263
SAG	500 nM	Merk	566660
SCH772984	100 nM	Selleckchem	S7101
LY294002	10 μ M	Selleckchem	S1105
SB203580	25 μ M	Selleckchem	S1076
SP600125	5 μ M	Selleckchem	S1460
PD184352 (CI-1040)	10 μ M	Selleckchem	S1020
Recombinant Human FGF basic	120 ng/ml	Peprotech	100-18B
Recombinant Human/Murine/Rat GDF-11	5-75 ng/ml	Peprotech	120-11
DAPT	10 μ M	StemCell Technologies	72082

Table 2. Antibodies for immunostaining

Target	Host species	Dilution	References	Company
P-SMAD2	Rabbit	1:800	18338	Cell Signaling
SMAD2/3	Mouse	1:250	610842	BD Biosciences
LEFTY	Mouse	1:200	AF746	Rndsystems
HOXC6	Rabbit	1:8000	ARP38484_P050	Aviva Systems Biology
HOXC8	Mouse	1:10000	BLE920502	Biologend
HOXC9	Guinea Pig	1:300	-	J. Dasen
ISL1	Goat	1:1000	GT15051	Sigma-Aldrich

Table 3. Real time PCR Primers

Primer sequences	Primer name	Targeted gene	Forward/reverse primer
CCAGGACCAGAAAGCCAGTA	HOXC6-fw	HOXC6	forward

GTTAGGTAGCGATTGAAGTGA AA	HOXC6-rv	HOXC6	reverse
CCTCCGCCAACACTAACAGT	HOXC8-fw	HOXC8	forward
CAAGGTCTGATACCGGCTGT	HOXC8-rv	HOXC8	reverse
ACCTCCTAGCGTCCAGGTTT	HOXC9-fw	HOXC9	forward
CGAAGCTACAGGACGGAAA	HOXC9-rv	HOXC9	reverse
CCTGGGATAACGAAG	HOXC10-fw	HOXC10	forward
TTCCAGCGTCTGGTC	HOXC10-rw	HOXC10	reverse
TGAAGGAGGAGGAGA	HOXD9-fw	HOXD9	forward
CTTTCTCCAGCTCAA	HOXD9-rw	HOXD9	reverse
CCTTCCCAGAGCAGC	HOXA10-fw	HOXA10	forward
CTTCTCCAGCTCCAG	HOXA10-rw	HOXA10	reverse
CCCACCGTGTTCTTCGACAT	CYCOPHILIN_ A-fw	CYCOPHILI N A	forward
CCAGTGCTCAGAGCACGAAA	CYCOPHILIN_ A-rw	CYCOPHILI N A	reverse

Partie 2

Développement d'un modèle organoïde humain reproduisant la spécification de l'axe antéro- postérieur

Un modèle d'organoïde *in vitro* du tronc humain récapitule la co-morphogenèse de la moelle épinière et de la colonne vertébrale

Article publié dans le journal *Nature Biotechnology*

a. Contexte

Au cours de l'embryogenèse des Vertébrés, le tube neural et les somites émergent conjointement à partir d'un pool caudal de progéniteurs axiaux qui alimente l'extension axiale de l'embryon (Wymeersch et al., 2021). Parallèlement à la morphogenèse du tube neural et des somites, des mécanismes de structuration déterminent la localisation spatiale des différents types cellulaires le long de l'axe antéro-postérieur (Deschamps and Duboule, 2017; Nedelec and Martinez-Arias, 2021; Philippidou and Dasen, 2013; Wymeersch et al., 2021). Afin d'étudier la mise en place de ces structures au cours de l'embryogenèse, des modèles *in vitro* capables de récapituler certains aspects de la gastrulation d'embryons humains, de la somitogenèse ou du développement de la moelle épinière ont été développés. (Beccari et al., 2018; Libby et al., 2021; Miao et al., 2023; Moris et al., 2020; Veenvliet et al., 2020; Yaman and Ramanathan, 2023; Yamanaka et al., 2023). Cependant, ces différents modèles ne reproduisent pas correctement la co-morphogenèse et l'organisation des deux lignages. Cela indique que les paramètres définissant la formation de l'axe antéro-postérieur restent largement méconnus. La génération de modèles récapitulant la morphogenèse et l'organisation topographique des organes qui se développent progressivement et en synchronie lors de l'élongation de l'embryon reste donc un défi majeur pour la recherche fondamentale et translationnelle.

Dans de précédents travaux, nous avons décrit la génération de population homogènes de progéniteurs axiaux, à partir d'agrégats de CSPi humaines cultivés en 3D et exposés à un agoniste de la voie Wnt ainsi que des inhibiteurs à BMP et TGF- β /Activin/Nodal (Mouilleau et al., 2021). En présence de RA, ces progéniteurs se différencient de manière sélective en progéniteurs de motoneurones, qui, selon les conditions préalablement imposées, ont le potentiel de générer les différents types de motoneurones présents le long de l'axe antéro-postérieur *in vivo*.

Cependant, dans ce paradigme de différenciation, les structures ne s'allongent pas et génèrent uniquement des domaines discrets de l'axe antéro-postérieur. De plus, ces progéniteurs axiaux se différencient vers un destin exclusivement neuronal. Dans cette étude,

nous avons donc voulu identifier les paramètres qui permettent aux progéniteurs axiaux la cogénération du tube neural et des somites comme observé durant la formation du tronc humain.

b. Résultats

Dans le protocole précédemment utilisé pour la différenciation de CSPi en progéniteurs axiaux, les voies de signalisation TGF- β et BMP sont inhibées pour promouvoir la neurogenèse. En permettant l'activation endogène de ces voies Nodal, et en contrôlant la taille des agrégats cellulaires, nous avons observé que les structures initialement sphéroïdales perdent leur symétrie et s'allongent progressivement. Contrairement aux sphéroïdes exposés aux inhibiteurs des voies TGF- β et BMP, ces structures allongées sont composées de cellules exprimant des marqueurs des lignages neuraux ou mésodermiques. Par ailleurs, l'augmentation de la taille des agrégats cellulaires favorise fortement la différenciation mésodermique par rapport à la différenciation neuronale.

En maintenant ces structures, nous avons observé que les lignages neuraux et mésodermiques s'organisent dans l'espace. Les progéniteurs neuraux forment un neuroépithélium dans lequel les cellules ont acquis une polarité apico-basale inversée, caractérisée par la présence de cellules mitotiques du côté du milieu de culture.

La région plus interne des structures, sous le neuroépithélium, est composée de cellules somitiques. Cependant, ces structures ne présentent pas une organisation similaire à celle observée *in vivo*. Comme la matrice extracellulaire contrôle la polarité apico-basale ainsi que la morphogenèse du tube neural et des somites, ces structures ont été enrobées dans du matrigel. Dans ces nouvelles conditions, la morphologie et la polarité apico-basale des lignées neurales et somitiques sont radicalement modifiées et ressemblent à l'organisation du tronc au cours du développement embryonnaire.

Pour mieux déterminer les états cellulaires ainsi que l'organisation de ces organoïdes, nous avons réalisé des expériences de transcriptomique. Le séquençage unicellulaire a permis de révéler la présence d'un continuum de stades de différenciation allant des progéniteurs axiaux aux neurones spinaux ou aux cellules somitiques. Le séquençage spatial a ensuite démontré l'existence de domaines colinéaires d'expression des gènes *HOX* le long de l'axe antéro-postérieur. Ces différents résultats ont été confirmés par immunomarquage et hybridation *in situ*.

Lorsque ces organoïdes sont maintenus à des stades plus avancés de culture, les domaines d'expression *HOX* entraînent la génération d'un continuum de sous-types de

motoneurones le long de l'axe antéro-postérieur. De plus, nous avons observé que certains de ces motoneurones projetaient des axones vers des progéniteurs musculaires générés dans les tissus périphériques au tube neural.

Dans l'ensemble, cette étude propose un nouveau modèle *in vitro* permettant de récapituler certains aspects du développement du tronc humain. Par ailleurs, ce modèle offre de nouvelles opportunités afin d'étudier le lien entre la morphogenèse axiale et la mise en place des domaines d'expression des gènes *HOX* le long de l'axe antéro-postérieur.

c. Contributions

Pour cette étude, j'ai participé activement à la culture d'organoïdes, aux immunomarquages, à l'imagerie live et à l'analyse des données transcriptomiques spatiales. Ces résultats ont permis de mettre en évidence 1) l'élongation axiale à partir d'un pool de progéniteurs axiaux, 2) la spécification cellulaire le long de l'axe antéro-postérieur et 3) l'importance de la taille des agrégats cellulaires dans la différenciation neuronale et mésodermique. De plus, j'ai adapté une méthode d'hybridation *in situ* afin de marquer l'expression des gènes *HOX* dans les organoïdes ce qui a permis de révéler que dans la moitié des structures, *HOXC6* est exprimé tout le long du neuroépithélium alors que les domaines d'expression de *HOXC9* sont plus restreints aux parties caudales exprimant le facteur CDX2.

Self-organizing models of human trunk organogenesis recapitulate spinal cord and spine co-morphogenesis

Received: 6 August 2022

Accepted: 18 August 2023

Published online: 14 September 2023

 Check for updates

Simona Gribaudo^{1,2,3}, Rémi Robert^{1,2,3}, Björn van Sambeek^{4,5}, Camil Mirdass⁶, Anna Lyubimova^{4,5}, Kamal Bouhali⁷, Julien Ferent^{1,2,3}, Xavier Morin⁷, Alexander van Oudenaarden^{4,5} & Stéphane Nedelec^{1,2,3} ✉

Integrated *in vitro* models of human organogenesis are needed to elucidate the multi-systemic events underlying development and disease. Here we report the generation of human trunk-like structures that model the co-morphogenesis, patterning and differentiation of the human spine and spinal cord. We identified differentiation conditions for human pluripotent stem cells favoring the formation of an embryo-like extending antero-posterior (AP) axis. Single-cell and spatial transcriptomics show that somitic and spinal cord differentiation trajectories organize along this axis and can self-assemble into a neural tube surrounded by somites upon extracellular matrix addition. Morphogenesis is coupled with AP patterning mechanisms, which results, at later stages of organogenesis, in *in vivo*-like arrays of neural subtypes along a neural tube surrounded by spine and muscle progenitors contacted by neuronal projections. This integrated system of trunk development indicates that *in vivo*-like multi-tissue co-morphogenesis and topographic organization of terminal cell types can be achieved in human organoids, opening windows for the development of more complex models of organogenesis.

Organoids derived from human pluripotent stem cells (hPSCs) are transforming the study of human organ development and physiology. However, most organoids model a single organ, such as brain, intestine, pancreas, derived from a single lineage^{1,2}. Yet developmental and pathogenic processes, as well as drug actions, are multi-systemic and encompass several tissues and organs that often originate from different germ layers during embryogenesis. Therefore, the generation of scalable and tractable high-order models that recapitulate the multi-lineage decisions, morphogenesis and topographical organization of organs co-developing during human embryogenesis is a major challenge for basic and translational research^{1–4}.

Here we address this challenge for the multi-germ-layer developmental process that supports acquisition of human body shape and movements. After gastrulation, vertebrate embryos undergo a fundamental change in shape as they elongate posteriorly to form the antero-posterior (AP) axis of the body. This axial extension is fed by bipotent neuromesodermal axial progenitors that progressively generate both the caudal spinal neural tube, along which the distinct body sensory-motor circuits are formed, as well as mesoderm-derived somites that will generate the innervated adjacent muscles, the tendons or the spine^{5–9}. This coordinated development initiates the topographic organization of cell types interacting across germ layers along the

¹Institut du Fer à Moulin, Paris, France. ²Inserm, UMR-S 1270, Paris, France. ³Sorbonne Université, Science and Engineering Faculty, Paris, France. ⁴Oncode Institute, Utrecht, The Netherlands. ⁵Hubrecht Institute-KNAW (Royal Netherlands Academy of Arts and Sciences) and University Medical Center Utrecht, Utrecht, The Netherlands. ⁶Université Paris Cité, CNRS, Institut Jacques Monod, Paris, France. ⁷Institut de Biologie de l'ENS (IBENS), Ecole Normale Supérieure, CNRS, INSERM, Université PSL, Paris, France. ✉e-mail: stephane.nedelec@inserm.fr

body axis to support trunk shape, motor and sensory information processing^{5–10}.

In vitro models of somites and/or spinal cord development have been recently described^{11–32}, including complex multi-germ-layer models such as mouse and human gastruloids or trunk organoids^{14,18,23,26,29,33–36}, some characterized by the co-development of neural tubes and somites adopting an in vivo-like organization^{26,29} (reviewed, in part, in refs. 4,8,37). However, none of these integrated models recapitulates both the co-morphogenesis and AP patterning of the two lineages. Moreover, they do not develop beyond the early somite or neural progenitor stages to generate terminal cell types. In contrast, in models that progress through these early stages, such as neuromesodermal organoids or spinal organoids^{13,14,17,18,22}, the neural and/or somitic compartments do not undergo morphogenesis to form a neural tube or organized somites. In addition, cells in these organoids lack a typical topographical organization and do not form the organized array of cell types that ensures the formation of the distinct muscles, bones and neural circuits along the body axis. These different spinal-only, somite-only or bifurcated three-dimensional (3D) models are all based on the initial specification of axial progenitor-like cells by exposing hPSCs to CHIR-99021 (CHIR), a Wnt pathway activator in absence or in presence of exogenous FGF (for instance, refs. 13,17,20,25,26,29,33,34 or 14,18,23,24,30,35). However, the basis of the differences in cell type composition and morphogenesis of the resulting organoids remains unclear. Altogether, these observations indicate that key parameters controlling the co-differentiation and patterning of trunk neuro/mesodermal lineages, their morphogenesis and maturation are still incompletely understood, limiting the engineering of complex models of human trunk organogenesis.

We previously reported that 3D aggregates of hPSCs (spheroids) exposed to CHIR, together with inhibitors of the BMP and TGF- β /Activin/Nodal (from here, TGF- β) pathways, generated axial progenitor-like cells^{19,21}. As in embryos, these in vitro-generated progenitors differentiate selectively into spinal motor neurons in the presence of retinoic acid (RA), which favor neural differentiation, and a Hedgehog pathway agonist (SAG) that induces ventral cell fates^{19,21}. However, in this differentiation paradigm, the spheroids did not form a neural tube, and somitic cells were absent. Furthermore, spheroids did not show an axial organization of cell types as seen along the AP axis of embryos. Instead, as in other non-elongating spinal organoids or neuromuscular organoids, cells adopted an identity corresponding to a specific sub-domain of the trunk AP axis^{10,13,14,18,21}. These observations suggest that specific parameters restrict the morphogenesis and cell fate potential of axial-like progenitors in this differentiation paradigm. We, thus, sought to use this model to determine the parameters controlling human trunk axis formation, neural tube and somite morphogenesis.

Results

Axial extension coupled to neuro/mesodermal differentiation

Co-inhibition of TGF- β and BMP signaling pathways promotes neurogenesis in vivo and in vitro^{38–40}. We reasoned that one or both inhibitors might constrain the morphogenetic and cell fate potential of axial progenitor-like cells to promote the selective generation of neurons in our spinal motor neuron differentiation paradigm^{19,21}. To test this hypothesis, human induced pluripotent stem cells (hiPSCs) were dissociated to single cells and re-aggregated into aggregates to form size-controlled spheroids (Fig. 1a,b). Spheroids were exposed to the WNT agonist CHIR from day 0 to day 4, in presence or in absence of inhibitors of BMP (LDN193189) and TGF- β (SB431542) pathways. In all conditions, RA was added on day 3, followed by SAG on day 4 (Fig. 1a). As we previously described^{19,21}, inhibition of both BMP and TGF- β pathway, as well as the sole inhibition of the latter, resulted in spheroids that remained globular and in the specification of motor neuron progenitors at day 8 (Supplementary Fig. 1). In contrast, in the absence of both inhibitors, spheroids broke symmetry by day 4 and progressively

adopted an elongated shape over time (Fig. 1a–c, Supplementary Fig. 1 and Extended Data Fig. 1). By day 7, $63 \pm 16\%$ of the structures had length/width aspect ratio >1.3 . They reached a mean length of $462 \mu\text{m}$ with a maximum of $1,742 \mu\text{m}$ (Fig. 1c and Extended Data Fig. 1a–c). Similar numbers were obtained with a second cell line (WTC11; Extended Data Fig. 1d–f). Immunostainings of aggregates at day 4 revealed the presence of axial progenitor-like cells (CDX2/SOX2/TBXT) located selectively on one side of $51.3 \pm 17\%$ of the aggregates, together with the emergence of cells expressing markers of neural progenitors (SOX2/TBXT⁺) and mesoderm (CDX2/TBXT/SOX2⁺), including pre-somitic mesoderm (PSM) (TBX6/SOX2⁺) (Fig. 1d,e and Extended Data Fig. 2a–c), as seen in other models of trunk tissue development^{14,17,18,20,23,24,29,30,33–36}.

While the structures continued to develop (days 6, 7 and 8), axial progenitor-like cells remained located on one side, but TBXT level and the number of positive cells progressively decreased (Fig. 1f and Extended Data Fig. 2c). At the same time, the mesodermal and neural lineages further organized spatially. Neural progenitors (SOX2, SOX1/3 or PAX6) (Fig. 1g,h) formed an organized peripheral pseudo-stratified neuroepithelium in which cells acquired apical–basal polarity, revealed by the presence of mitotic cells (phospho-histone H3 (P-H3)), primary cilia (ARL13B) and accumulation of filamentous actin (F-actin) on the medium-facing side (Fig. 1i and Extended Data Fig. 2d). The neuroepithelium did not shape into a neural tube but, instead, kept an apical-out organization typical of the open neural plate. In agreement with the presence of SAG, the vast majority of neural cells expressed NKX6.1, which identifies the four most ventral progenitor domains of the caudal neural tube, and OLIG2, which is expressed by motor neuron progenitors within this region⁴¹ (Fig. 1j,k and Extended Data Fig. 2e). However, in the neuroepithelium of all the structures, we observed patches of cells that adopted a dorsal spinal progenitor identity (PAX3) (Extended Data Fig. 2f,h). The region underneath the neuroepithelium was composed almost exclusively of somitic cells, as shown by staining for somitic markers (MEOX1, FOXC2, PAX3 or SOX9) on 3D reconstructions or sections (Fig. 1l–o and Extended Data Fig. 2f–j). Except for rare patches of cells in some structures, all the cells adopted a somitic or neural fate. Also consistent with the presence of SAG⁴², mesodermal cells adopted mostly a ventral fate (SOX9/PAX3⁺) corresponding to the precursors of the axial skeleton or tendons⁴² (Fig. 1m,o and Extended Data Fig. 2f,h,i). However, at the region bordering the neuroepithelium, we systematically observed PAX3/FOXC2/SOX9⁺ cells that identify less committed paraxial mesoderm progenitors or dermomyotome progenitors, the precursors of the muscles (Fig. 1n,o and Extended Data Fig. 2f). Yet we did not observe an organization reminiscent of the embryonic segmented somites either on sections or on 3D reconstructions.

The morphogenesis, cell fate specification and organization described above were confirmed with additional hiPSC lines (SOX2-GFP, WTC11 and KUTE4; Extended Data Fig. 3a–d). However, the initial number of cells for optimal elongation and dual neuro/mesodermal specification varied between cell lines (between 80 and 160 cells). With all the tested cell lines, increasing the initial number of cells strongly promoted mesodermal over neural differentiation (Extended Data Fig. 3e–h). As endogenous TGF- β and BMP signaling appears to favor neural differentiation (Supplementary Fig. 1), we reasoned that inhibiting these pathways within the larger mesoderm-biased aggregates might restore the specification of neural progenitors. Notably, the addition of TGF- β and BMP inhibitors from day 0 in normally mesoderm-biased aggregates (>300 cells) increased neural progenitors that organized in a neuroepithelium in elongating structures (Extended Data Fig. 3i). This observation further reinforces the importance of these pathways, endogenous to the aggregates, in controlling neuro/mesodermal differentiation and elongation. Also, adding exogenous FGF2 between day 0 and day 2, as done in several protocols aiming at generating axial progenitors^{5,7,15}, strongly promoted mesodermal over neural specification (Extended Data Fig. 3j),

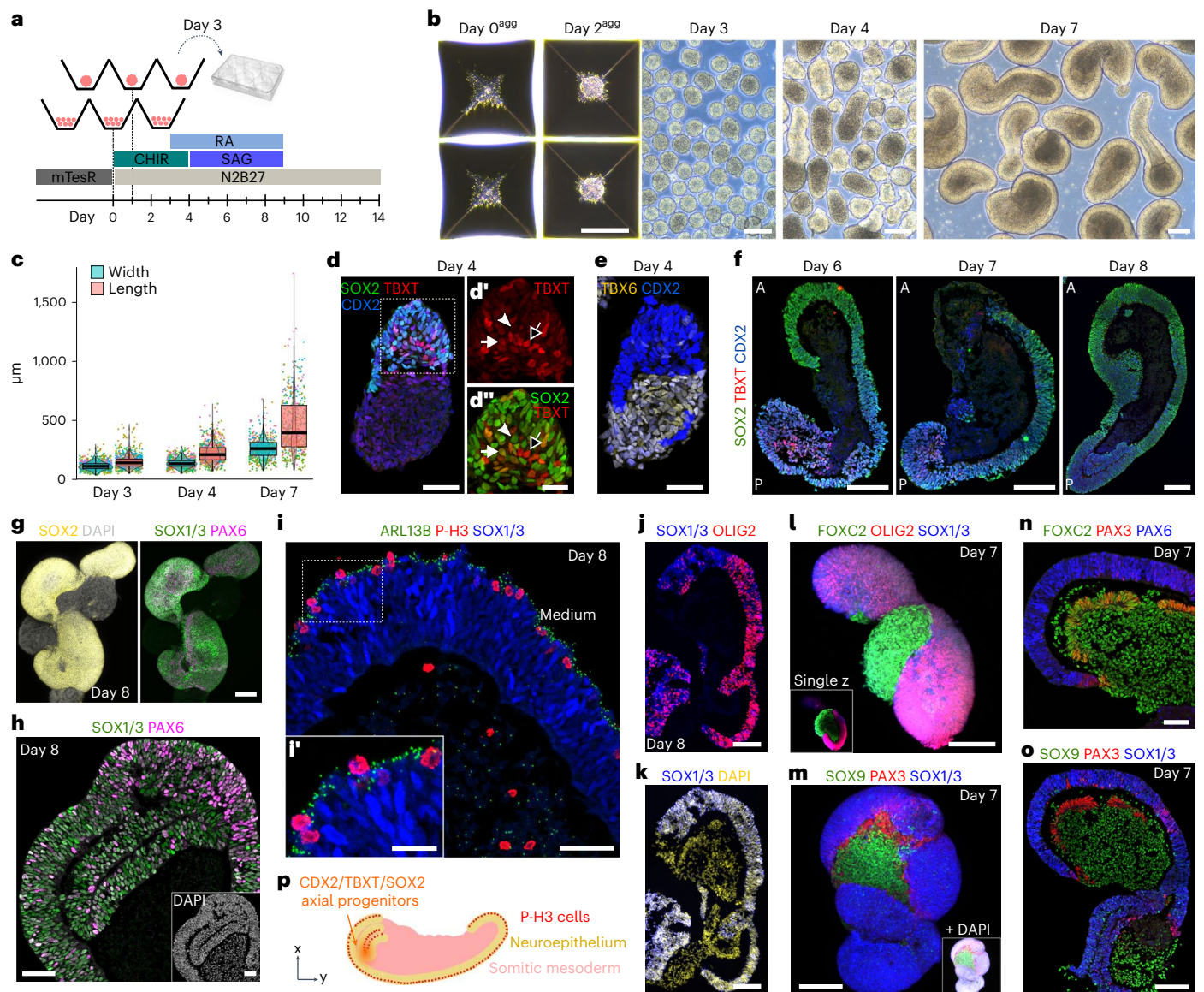


Fig. 1 | Axial elongation coupled to spinal neuroepithelium and somitic mesoderm specification. **a**, Experimental design. From day 3, spheroids were transferred in six-well plates. **b**, Representative bright-field images of aggregates from day 0 to day 7, $n = 11$. **c**, Box plot of width and length over time (WTS2, day 3, $n = 11$, 770 aggregates; day 4, $n = 11$, 705 aggregates; day 7, $n = 13$, 447 aggregates). In the box plots, the center line indicates the median; box limits indicate the 1st and 3rd quartiles; whiskers indicate the range. Each dot corresponds to a single structure, and different colors indicate independent experiments. **d-f**, Immunostaining of sections for markers of axial progenitor-like cells (SOX2/TBXT/CDX2), neural/axial (SOX2) and mesodermal (TBXT and TBX6) progenitors, in **d** $n = 7$, in **e** $n = 3$ and in **f** $n = 4$. In **d'** and **d''**, arrow: SOX2/TBXT⁺ cells; arrowhead: SOX2/TBXT⁻ cells; black arrow: TBXT/SOX2⁻ cells. **g**, 3D maximum intensity projection of whole-mount immunostaining for neural progenitor

markers (SOX2, SOX1/3 and PAX6). **h**, Markers as in **g** on cryostat section, $n = 3$. **i**, Mitotic cells (P-H3) and primary cilia (ARL13B) reside on the external surface of the neuroepithelium. Higher magnification in **i'**, $n = 3$. **j,k**, Immunostaining for markers of motor neuron (OLIG2) and neural (SOX1/3) progenitor, $n = 6$ in **j** and **k**. **l,m**, 3D maximum intensity projection of whole-mount immunostaining for markers of neural (SOX1/3), motor neuron (OLIG2) or dorsal (PAX3/SOX1/3) progenitors and somitic mesoderm (FOXC2, SOX9 or PAX3, negative for SOX1/3). **n,o**, Adjacent sections stained for markers of PSM/dermomyotome (PAX3), sclerotome (SOX9 and FOXC2) and neural progenitors (SOX1/3 and PAX6), $n = 4$ in **n** and **o**. **p**, Schematic of an elongated structure at day 7/8. Scale bars, 200 μm (**b**), 100 μm (**f,g,j-m,o**), 50 μm (**d,e,h,i,n**), 25 μm (**d',d''**) and 20 μm (**i'**). agg, aggregates.

in line with the role of FGF in embryos⁴³. Finally, some commercial B27 batches severely reduced mesoderm specification (TBX6) at day 3, correlating with a decrease in the number of elongated structures at day 7 (Extended Data Fig. 3k,l). Expectedly, with these batches, increasing the initial cell number partially rescued neuro/mesodermal co-specification at day 3 and elongation at day 7 (Extended Data Fig. 3k,l).

Overall, our mechanistic analysis shows that, in size-controlled aggregates, the WNT agonist CHIR cooperates with endogenous TGF- β signaling to control axis elongation and neuro/mesodermal

differentiation. As during vertebrate trunk development, axial progenitor-like cells locate at one pole of the growing structure, and a spinal neuroepithelium formed progressively together with somitic cells. However, the two lineages do not adopt an *in vivo*-like morphology. The neuroepithelium does not form a neural tube with cells' apical side facing a central lumen. Instead, it maintains an apical-out/basal-in organization typical of the neural plate. Similarly, instead of the segmented somites found around the neural tube in embryos, somitic cells form a large cluster underneath the neuroepithelium.

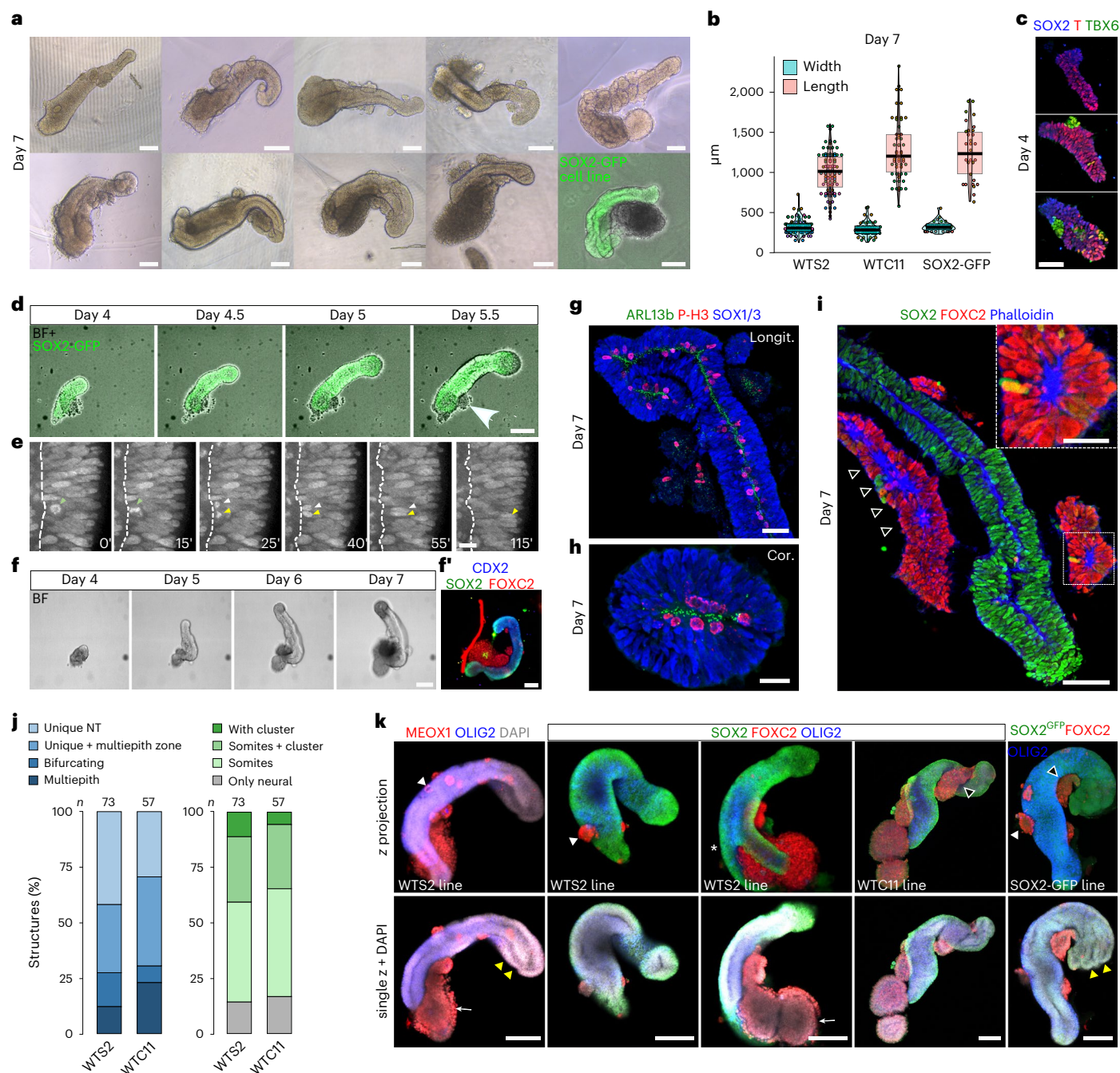


Fig. 2 | ECM favors an in vivo-like organization of the neural tube and somites.

a, Representative bright-field images of Matrigel-embedded structures (WTS2 or SOX2-GFP lines). **b**, Box plot of width and length for three cell lines (WTS2, $n = 8,100$ total structures; WTC11, $n = 4,57$ structures; SOX2-GFP, $n = 3,34$ structures). Box plots: center line is the median; box limits indicate the 1st and 3rd quartiles; whiskers indicate the range. Each dot is a structure. Each color is an independent experiment. **c**, 3D maximum intensity projections of whole-mount immunostaining at day 4 of axial progenitors (SOX2 and TBX6) and PSM (SOX2-GFP) markers ($n = 3,38$ total structures). **d**, Time-lapse imaging of a Matrigel-embedded SOX2-GFP structure (Supplementary Video 1). Arrow indicates non-neural cells. **e**, Time-lapse recording of cell behavior in day 7 neural tube (SOX2-GFP line). Green arrowhead indicates a mitotic cell; yellow and white arrowheads indicate daughter cells (Supplementary Video 3). **f, f'**, Time-lapse imaging of a WTS2 structure (**f**) and post-imaging immunostaining (**f'**) at day 7, identifying axial (SOX2/CDX2) and PSM (TBX6) progenitors (Supplementary Video 4).

g, h, Immunostaining on longitudinal (**g**) and coronal (**h**) sections for markers of mitotic cells (P-H3), primary cilia (ARL13B) and neural progenitors (SOX1/3), $n = 3$ in **g** and $n = 2$ in **h**. **i**, Staining on sections for F-actin (Phalloidin), markers of somites (FOXC2) and neural/axial progenitors (SOX2), $n = 3$. **j**, Quantification of somite and neural tube organization (WTS2 line, $n = 3,73$ total structures; WTC11 hiPSC line, $n = 3,57$ total structures). **k**, 3D maximum intensity projection (upper panel) or single optical plane (lower panel) of day 7 Matrigel-embedded organoids immunostained for markers of somitic (MEOX1 or FOXC2) and neural (OLIG2 and/or SOX2) progenitors. White arrowheads indicate examples of single somites; black arrowheads indicate examples of somite strings; white arrows indicate examples of clusters of somitic cells; yellow arrowheads indicate examples of multiple epithelia contiguous with a unique tube; asterisk indicates example of a bifurcating structure. Scale bars, 200 μm (**a**), 50 μm (**c, g, i**), 100 μm (**d, f, f', k**) and 20 μm (**e, h**). BF, bright-field; Multiepipith, multiple epithelia; NT, neural tube.

This striking inverted organization provided an opportunity to investigate mechanisms controlling neuroepithelium and somite morphogenesis.

Extracellular matrix promotes neural tube and segmented somite formation

During embryogenesis, the extracellular matrix (ECM) controls epithelial apical–basal polarity as well as neural tube and somite morphogenesis^{44–47}. Embedding of *in vitro* models of somites and/or spinal cord development in Matrigel, a basement membrane matrix, promotes an *in vivo*-like organization with the formation of epithelial somites and neural tubes^{20,25,26,29,30}. To test the impact of ECM on morphogenesis, individual spheroids were embedded in a drop of 50% Matrigel at day 3, before elongation, and cultured free floating. More than 60% of aggregates elongated between day 4 and day 7 (Fig. 2a and Extended Data Fig. 4a), reaching a maximal extension from 1,600 μm to 2,300 μm depending on the cell line (Fig. 2b and Extended Data Fig. 4b–e). At day 4, 1 day after embedding, axial progenitor-like cells (SOX2/TBXT) were detected in all structures, whereas PSM (TBX6) was observed in $51 \pm 15\%$ of them, and very few contained somitic cells (MEOX1) at this early stage (Fig. 2c).

To monitor the dynamics of morphogenesis, we performed live imaging from day 4. We first differentiated an hiPSC line in which SOX2, an axial and neural progenitor marker, is fused to GFP (Fig. 2d and Extended Data Fig. 5a). We observed a highly dynamic elongation reminiscent of the process of vertebrate embryo caudal extension. In contrast to the non-Matrigel condition, the neuroepithelium adopted a tubular shape with a central lumen (Fig. 2d, Extended Data Fig. 5a and Supplementary Videos 1 and 2). The neuroepithelium was strikingly dynamic, with multiple mitoses occurring on the lumen side, followed by interkinetic nuclear migration toward the external region as seen in embryos (Fig. 2e and Supplementary Video 3). We confirmed the dynamics of elongation with another cell line (Fig. 2f, Extended Data Fig. 5b and Supplementary Video 4). Post-imaging immunostaining showed the presence of a neural tube in all elongated organoids, 80% of which had surrounding somitic cells (Fig. 2f' and Extended Data Fig. 5b). As in mouse and chick embryos^{48–50} and in recent *in vitro* models of AP axis extension^{25,30,34,51}, axis elongation was dependent on active WNT and FGF signaling as well as on balanced level of RA. Indeed, the inhibition of WNT secretion (IWP2), FGF receptors (PD173074) or their downstream effector ERK1/2 kinase (SCH772984) from day 3 prevented elongation (Extended Data Fig. 6). Similarly, increasing RA concentration severely affected elongation, consistent with its teratogenic effect on neural tube formation (Extended Data Fig. 6)⁵².

We confirmed the tubular organization of the neuroepithelium on longitudinal and coronal sections at day 7 (Fig. 2g,h). In contrast to the non-Matrigel condition, neural progenitor cells now always acquired an *in vivo*-like apical-in/basal-out organization indicated by the presence of F-actin, primary cilia (ARL13B) and mitotic cells (P-H3) on the internal lumen side (Fig. 2g,h and Extended Data Fig. 7a). Similar to the neural compartment, somitic cells underwent a radical change in their organization, with individual segmented somites or a string of somites distributed around the neural tube (MEOX1 and FOXC2; Fig. 2i–k, Extended Data Figs. 5 and 7b–d and Supplementary Fig. 2). The somitic cell apical side faced the center of the rosette (F-actin enrichment; Fig. 2i). This reorganization of the neural and somitic cells was confirmed with two other cell lines (WTC11 and SOX2-GFP) (Extended Data Fig. 7e–g). To characterize the neural tube and somite organization more systematically, we performed whole-mount staining on randomly selected structures (73 from three independent experiments) and monitored their shape and location (Fig. 2j,k and Supplementary Fig. 2). We observed a unique neural tube in $88 \pm 5\%$ of the cases. The neural tube can sometimes be in continuity with multiple epithelia at one extremity ($31 \pm 6\%$) or split into two well-identified tubes ($15 \pm 3\%$) (Fig. 2k). The remaining displayed several lumens all along the structure ($12 \pm 3\%$) (Fig. 2j and Supplementary Fig. 2). We observed the presence of somites

in various numbers and shapes around the tubes (Fig. 2i–k and Supplementary Fig. 2) and a cluster of somitic cells in 40% of the organoids (white arrow in Fig. 2k) while $14 \pm 5\%$ were devoid of somitic cells. We obtained similar results with the WTC11 line (Fig. 2j,k and Supplementary Fig. 2). Hence, whereas axial elongation and cell fate specification are similar in the presence or absence of Matrigel (Figs. 1 and 2), in the latter condition, the morphology and apical–basal polarity of the neural and somitic lineages are substantially modified and mimic several aspects of the *in vivo* organization as previously observed with the generation of trunk-like structure from mouse embryonic stem cells²⁶.

AP-like organization of the extending axis

To further determine cell states and their organization, we performed single-cell and spatial transcriptomic sequencing on Matrigel-embedded organoids at day 7 of differentiation. We first analyzed the transcriptome of 9,619 cells isolated from 26 organoids. Clustering identified eight main cell states. The identity of the clusters was determined based on a combination of previously identified markers and using random forest classifiers trained on single-cell transcriptomic data⁵³ from human embryos at gastrulation stage (CS7) (ref. 54) and human embryonic spinal cord⁵⁵ (Extended Data Fig. 8a,b). This analysis revealed the presence of axial progenitors, neural progenitors, neurons, neural crest cells as well as mesodermal derivatives, including somitic, endothelial, cardiac and nephron progenitors (Fig. 3a,b, Extended Data Fig. 8c and Supplementary Table 1). Anterior neural cell types, notochord cells or endoderm derivatives were absent. Further analysis of specific markers (Extended Data Fig. 8c) revealed the presence of a continuum of differentiation stages from axial neuromesodermal progenitors to spinal neurons or somitic cells. The axial progenitor cluster was composed at its tip by progenitors displaying a molecular signature associated with bipotent neuromesodermal progenitors (*TBXT*, *WNT5A*, *WNT3*, *FGF8/17*, *CYP26A1* and *RAR γ* ; some shown in Extended Data Fig. 8c). The other cells of this cluster resembled progenitors engaged toward a neural fate, as indicated by decreasing levels of neuromesodermal progenitor markers and the progressive appearance of neural markers (*NKX6-1*, *HES5*...), an identity observed in the caudal lateral epiblast or the pre-neural tube⁵ (Extended Data Figs. 2a and 8c). Axial progenitor/pre-neural tube markers (*TBXT*, *NKX1-2*...) and neural markers, respectively, decreased and increased as cells were closer from the neural progenitor cluster. This cluster was composed mostly of motor neuron and interneuron progenitors located ventrally *in vivo* (*NKX6-1*, *OLIG2* and *DBX2*) (Extended Data Fig. 8c). Similarly, the somite cluster was composed of differentiation stages spanning from posterior PSM (*TBX6* and *MESP2*) to a preponderant sclerotome fate (*SOX9*, *PAX1*, *PAX9* and *NKX3-2*) with some syndetome progenitors, the precursors of the tendons (*SCX* and *FMOD*) and chondrocyte progenitors (*SOX5/6/9* and *COL2A1*). Finally, we detected a small cluster of endothelial cells (*CDH5*, *KDR*, *SOX17* and *ETV2*) (Extended Data Figs. 2a and 8c) that derive, in embryos, either from the lateral sclerotome^{56,57} or the lateral plate mesoderm⁵⁸.

To determine whether these differentiation trajectories spatially organized as in vertebrate embryos or adopt an *in vitro*-specific organization, we performed spatial transcriptomic using RNA tomography (tomo-seq)⁵⁹ at day 7 (Fig. 3c). Two organoids were serially cross-sectioned all along their longest axis, and individual slices were processed for bulk RNA sequencing (Fig. 3c). We detected roughly 1,000 genes displaying reproducible expression patterns along the main axis between replicates (Fig. 3c). These genes segregated into four clusters. We projected the z-scores of the genes composing each tomo-cluster on the single-cell uniform manifold approximation and projection (UMAP) (Fig. 3d). The four tomo-clusters also segregated in the UMAP, indicating that the distinct cell states were differentially present along the AP axis (Fig. 3d). Indeed, axial progenitor (*TBXT*, *CYP26A1*, *FGF8*...) and PSM genes (*TBX6*...) were detected at the posterior end (tomo-cluster 1) (Fig. 3c,d and Supplementary Table 2).

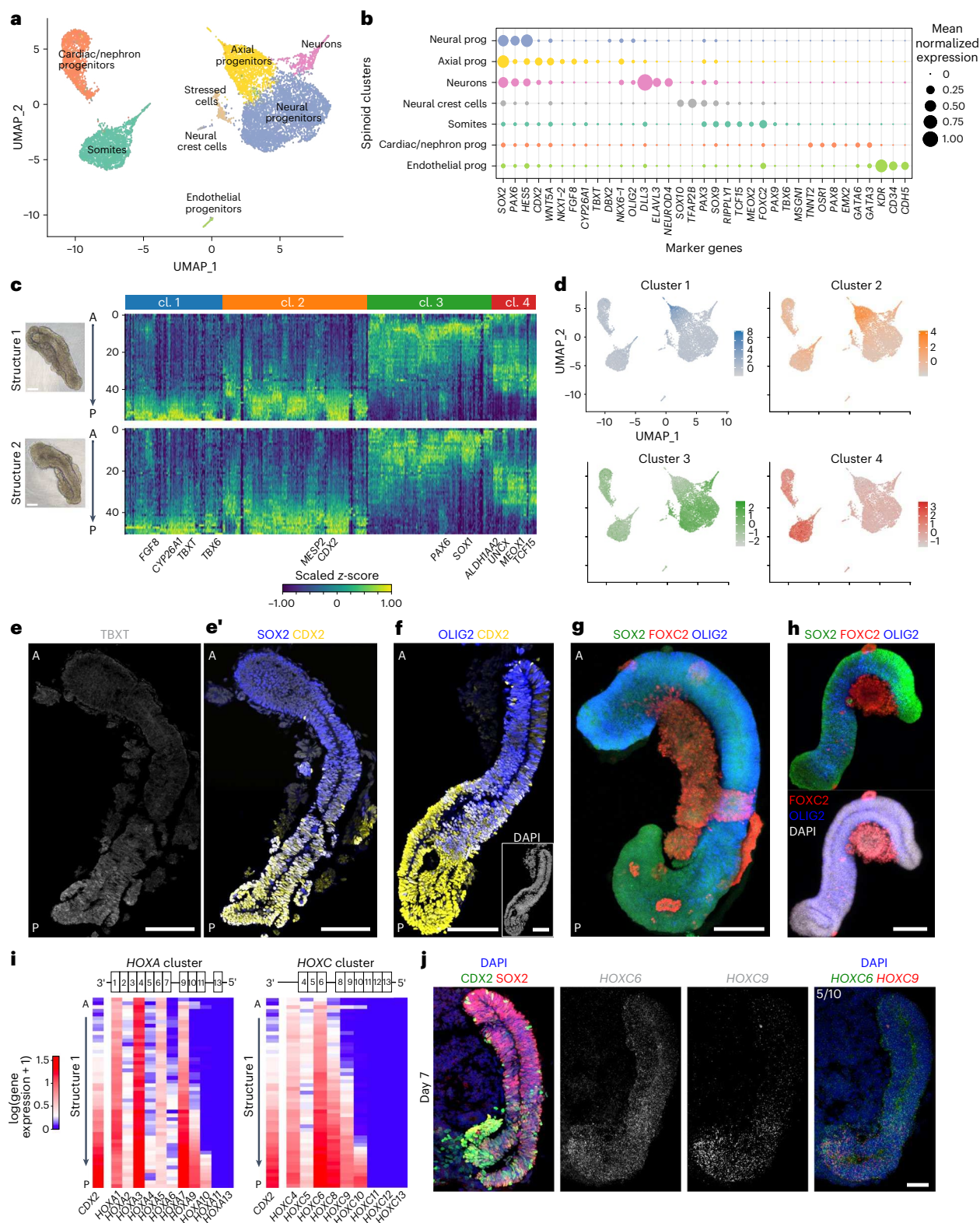


Fig. 3 | Neuro/mesodermal trajectories are organized and patterned along an in vivo-like AP axis. **a**, UMAP plot of 9,619 cells from 26 day 7 organoids. **b**, Dot plot of marker genes in the indicated clusters. **c**, Bright-field images of the structures tomo-sequenced and the corresponding heat map showing the expression pattern of 1,000 genes showing reproducible pattern between the two samples ($n = 2$). **d**, z-scores of the tomo-seq cluster genes in **c** plotted on the UMAP plot. **e, f**, Representative images of sections of day 7 structures. **e, e'**, Immunostaining for axial progenitor markers (TBXT, CDX2 and SOX2), $n = 5$. **f**, Immunostaining for markers of caudal progenitor

progenitors (OLIG2), $n = 4$. $CDX2^{high}$ region identifies the posterior (P) versus the anterior (A) pole. **g, h**, 3D maximum intensity projection of whole-mount immunostaining for markers of neural/axial (SOX2), motor neuron (OLIG2) and somitic (FOXC2) progenitors. **i**, Tomo-seq data analysis of *CDX2*, *HOXA* and *HOXC* clusters. $CDX2^{high}$ identifies the posterior region. **j**, Immunostaining of sections for markers of neural/axial (SOX2) and caudal (CDX2) progenitors; FISH for *HOXC6* and *HOXC9* mRNAs was performed on the adjacent sections, $n = 3$. Scale bars, 200 μ m (**c**), 100 μ m (**e–h**) and 50 μ m (**j**). For TBXT in **e**, a median filter of 0.1 was applied to reduce background noise. cl., cluster; prog, progenitor.

Anterior to this region, cells expressed genes corresponding to the early pre-neural tube and PSM (tomo-cluster 2), followed by neural and somitic progenitor genes (tomo-clusters 3 and 4). Neural progenitors extended up to the most anterior region (tomo-cluster 4) (Fig. 3c,d and Supplementary Table 2). Thus, spatial transcriptomic coupled to single-cell RNA sequencing (scRNA-seq) indicates that neuro/mesodermal differentiation trajectories are spatially organized along an AP-like axis as observed in vertebrate embryos. To confirm this organization, we performed immunostaining for key markers. At day 7, a cluster of axial progenitor-like cells (CDX2/TBXT/SOX2) was located at one extremity in $73 \pm 6\%$ of the organoids (Fig. 3e,e'). In the remaining, TBXT was no longer detectable at that stage (Fig. 1f). In $85 \pm 6\%$ of the cases, CDX2 progressively decreased anteriorly together with the appearance of more committed neural progenitor markers (OLIG2 and PAX3) or somitic markers (PAX3/FOXC2 or SOX9/FOXC2) (Fig. 3f–h and Extended Data Fig. 7b–g). For instance, OLIG2 showed a strong posterior to anterior gradient of expression in $48 \pm 5\%$ of the neural tube (Fig. 3g) and was already expressed along most of its length in the rest (Fig. 3h). As in gastruloids^{33,60}, cardiac (GATA6) and intermediate mesoderm (LHX1) progenitors (which include nephron progenitors) were observed in a variable subset of organoids in various quantities ($38 \pm 30\%$ of the structures). When present, they formed a cluster abutting somitic cells (Extended Data Fig. 9a).

In embryos, the extension of the AP axis is coupled to its patterning by HOX transcription factors, which allocate cell diversity along the body^{9,10}. HOX genes are clustered in the genome and are differentially expressed along the extending AP axis according to their chromosomal location, with 3' genes expressed more anteriorly than their 5' neighbors⁹ (Extended Data Fig. 9b). To determine whether in vitro trunk elongation and morphogenesis were coupled to AP patterning mechanisms, we monitored the expression profiles of HOX clusters using the tomo-seq data. We observed AP patterns of expression reminiscent of those in vertebrate embryos (Fig. 3i and Extended Data Fig. 9b–e)^{9,10}. For instance, *HOXC6*, whose anterior boundary of expression is seen in the rostral brachial spinal cord in embryos, was expressed along most of the axis (Fig. 3i and Extended Data Fig. 9c). By contrast, *HOXC8* and then *HOXA9/C9*, whose expression domains start, respectively, in the mid-brachial and anterior thoracic regions of the mouse spinal cord, were progressively enriched in more posterior regions (Fig. 3i and Extended Data Fig. 9c)¹⁰. *HOXC6* and *HOXC9* segregated expression patterns were also observed by fluorescence in situ hybridization (FISH) on sections spanning the anterior (CDX2) and the most caudal (CDX2) regions of the neural tube (five on 10 organoids; Fig. 3j and Extended Data Fig. 9f,g). Finally, *HOXA10/C10*, whose anterior boundaries are in the lumbar spinal cord and that were restricted to axial progenitors in the scRNA-seq dataset (Extended Data Fig. 8c), were observed by tomo-seq only in the most caudal region of one of the two organoids (Fig. 3i and Extended Data Fig. 9c). These results indicate that different organoids are at distinct degrees of axial extension up to a caudal thoracic or lumbar level.

Overall, a highly organized and patterned AP axis is forming during differentiation. Both neural and somitic differentiation trajectories emerge from a posterior pole containing axial progenitors. They morph into somites surrounding a patterned neural tube. The observed developmental sequence and spatial organization is thus similar to the ones described in vertebrate embryos (Extended Data Figs. 2a and 9b).

Self-organization propagates at later stages of organogenesis. Current mouse and human models mimicking aspects of the morphogenesis of the trunk do not appear to develop beyond the early neural and somite progenitor stages^{20,24,26,29,30,34}, limiting their possible use to study trunk organogenesis. We sought to test whether our differentiation paradigm allowed differentiation and morphogenesis to proceed.

Inspired by the timing of human embryo development^{61,62}, we cultured Matrigel-embedded and non-embedded trunk-like structures for an additional week, up to day 14. In the non-Matrigel condition, the external neuroepithelium was maintained. Neural progenitors were detected on the medium-facing side, whereas post-mitotic motor neurons (HB9/ISL1) and extending neurites (NEFL) were detected on the inner basal side, in agreement with the apical–basal polarity of the neuroepithelium at day 7 (Extended Data Fig. 10a). However, these non-embedded organoids were prone to fold on themselves complicating the analysis of their spatial organization. In contrast, in the Matrigel-embedded condition, the neural tube organization remained detectable in all elongated organoids (Fig. 4). In comparison to the day 7 stage, neural progenitors (OLIG2) were now present on the internal side of the neural tubes, whereas post-mitotic motor neurons (HB9/ISL1) were located on the external basal side, an organization reminiscent of the one observed in vivo around the fourth week of embryogenesis^{61,62} (Fig. 4a,b). The tempo of neurogenesis and the integrity of the neural tube were dependent on Notch signaling as shown by the massive neurogenesis (HB9/ISL1/NEFL) across the epithelium and the loss of the apical organization upon treatment with a Notch pathway inhibitor (DAPT) at day 8 (Extended Data Fig. 10b,c).

Consistent with the presence of post-mitotic neurons at the periphery of the neural tube, analysis on cryostat sections as well as 3D reconstruction of cleared organoids revealed the presence of neurites (NEFL and β III-tubulin) projecting outside the neural tube and invading the surrounding tissues (Fig. 4c–e and Supplementary Videos 5 and 6). This peripheral tissue was composed, in its vast majority, of chondrocytes or their progenitors (SOX9), in agreement with the preponderant sclerotome fate observed at day 7 (Fig. 4c). We also observed clusters of myoblast cells, the precursors of muscle fibers, in $38 \pm 13\%$ of the organoids (MYOD; Fig. 4e,g). When present, these clusters were always surrounded by neurites (Fig. 4e,g).

The propagation of neural tube morphological and topographical features at these later stages pushed us to test whether large-scale embryo-like patterns of terminal cell types were generated, a feature poorly observed in most current self-organizing organoid models. During development, spatial patterns of HOX mRNAs are translated into stereotypic expression domains of HOX transcription factors, which instruct cell identities along the body axis. For instance, the differential expression of HOXA, HOXC or HOXD paralogs in post-mitotic motor neurons orchestrates their diversification to form distinct locomotor circuits along the spinal cord¹⁰. We previously showed that, as in the mouse, HOXC6/C8/C9/C10 identify distinct motor neuron subpopulations along the human embryonic spinal cord²¹. We, thus, monitored the spatial distribution of HOXC6, HOXC8 and HOXC9 at day 14. Eighty percent of the organoids showed spatial patterns of expression within motor neurons (Fig. 4h,i). In particular, $54 \pm 16\%$ showed collinear expression of HOXC6, HOXC8 and HOXC9 (Fig. 4h and Extended data Fig. 10d; WTS2 cell line), which encompassed the brachial to thoracic region in embryos^{10,21}. The rest displayed more restricted HOXC6/HOXC8 or HOXC8/HOXC9 patterns (Fig. 4i). These patterns were always collinear with HOX genomic location. Hence, AP patterns are formed identical in their sequence to those observed in vivo.

Overall, our results show that neural tube and somites can develop beyond the early progenitor stages and that the self-organizing capacity of the system continues at later stages of trunk development. After 2 weeks, progenitors and neurons adopted an organization reminiscent of the one observed along the AP axis of human embryos around the fourth week of embryogenesis (Carnegie stage 13) (refs. 61,62).

Discussion

We describe here the generation of a self-organizing, multi-germ-layer system derived from hiPSCs that serves as a tractable and scalable model of human spine and spinal cord co-development.

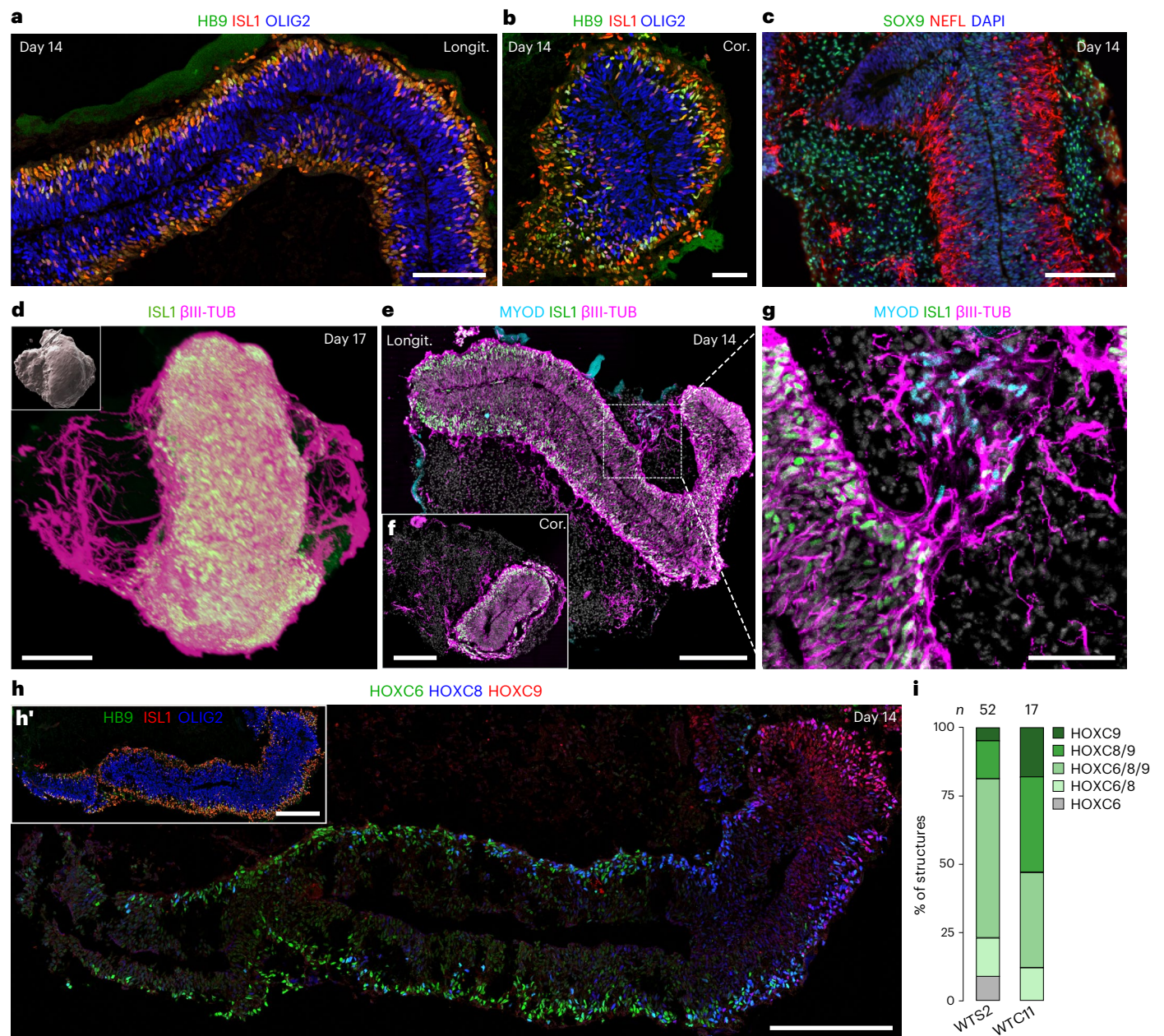


Fig. 4 | Neural tube maturation and generation of terminal cell types.

a,b, Immunostaining of longitudinal (**a**, $n = 7$) and coronal (**b**, $n = 3$) sections for markers of motor neuron (HB9/ISL1) and their progenitors (OLIG2). **c**, Immunostaining of sections for chondrocytes/chondrocyte progenitors (SOX9) and projecting neurons (NEFL), $n = 5$. **d**, 3D reconstruction of whole-mount immunostaining for markers of neurons (β III-tubulin) and motor neurons (ISL1) at day 17; in the inset, rendered surface to show the full volume of the organoid, $n = 2$, seven structures. **e–g**, Immunostaining of longitudinal (**e**) and coronal (**f**)

sections for markers of motor neurons (ISL1), projecting neurons (β III-tubulin) and myoblast (MYOD) ($n = 3$, 94 structures). In **g**, example of a myoblast cluster (MYOD) with neurites (β III-tubulin) within and around. **h**, Immunostaining of longitudinal sections for HOXC6, HOXC8 and HOXC9; **h'**, adjacent section stained for markers of motor neurons (ISL1) and their progenitors (OLIG2), $n = 6$ in **h** and **h'**. **i**, Percentage of organoids displaying the indicated HOXC patterns (WTS2, $n = 4$; WTC11, $n = 2$). Scale bars, 100 μ m (**a,c**), 50 μ m (**b,g**) and 200 μ m (**d–f,h**). Cor., coronal; Longit., longitudinal; TUB, tubulin.

As occurs during embryo caudal development, these structures elongate to form an AP axis along which somitic and spinal neural lineages organize anteriorly and self-assemble into a patterned neural tube surrounded by somites in an ECM-dependent manner. After 2 weeks, the organization mimics many aspects of the human spinal cord at the end of the fourth gestation week, 2 weeks after implantation, suggesting that the embryonic developmental tempo is recapitulated in this model⁵⁵.

This model differs from previous 3D PSC-derived models of spinal cord and/or somite development^{11–14,17–26,28–30,32,35,63} by recapitulating

the co-morphogenesis and AP patterning of the two lineages, which adopt an in vivo-like tissue and cell type organization up to terminal stages of differentiation. Our study identified several parameters that control these processes, which help rationalize the differences between previous models of trunk tissue generation (Supplementary Table 3).

The generation of AP-like cells from mouse or human PSCs depends on WNT pathway activation (exposure to CHIR) with or without addition of exogenous FGF^{5,7,15,19,21,31}. Increasing WNT^{15,24} or FGF (Extended Data Fig. 3j) favors mesodermal over neural specification, in line with the role of the WNT and FGF pathways in embryos (our data and

refs. 5,43). We further show that the size of the aggregates plays a critical role in this process. Balanced neuro/mesodermal co-specification occurs in a narrow range of initially aggregated PSCs (80–160), with a strong bias toward mesodermal differentiation in larger aggregates (>300 cells) (Extended Data Fig. 3e–h). In agreement with these observations, previous protocols using CHIR or CHIR+FGF or high concentration of CHIR (>5 μM) with aggregates formed with more than 200 cells generated somite-only elongating models^{24,30} (Supplementary Table 3). The effect of cell number appears to be mediated, at least in part, by a change in the level of endogenous TGF- β and BMP signaling (Extended Data Fig. 3i), indicating an underappreciated role of these pathways in the trunk neuro/mesodermal balance and in the effect of aggregate size on cell fate specification. This result may explain why recent protocols using inhibitors of these two pathways required high concentrations of CHIR (>5 μM) (refs. 24,30) to generate somite-only structures while neural cells were cogenerated at lower concentrations^{24,29} (Supplementary Table 3). Together, these observations indicate that the co-emergence of neural and somitic cells along the extending axis is, in *in vitro* models, as in embryos, under the control of a tightly balanced combinatorial action of WNT, FGF, TGF- β , BMP and RA pathways, which, *in vitro*, is influenced by the number of cells. In a tissue engineering perspective, as the level of endogenous signaling pathways, such as WNT, vary between hPSC cultures (for instance, ref. 64), the reproducible generation of patterned trunk-like structures from multiple lines will require careful adaptation. Our detailed analysis of the different parameters regulating the emergence of these structures provides a way to rationalize this process. The use of microwells (here, aggrewwells), which allow the production of thousands of organoids in parallel, instead of the 96-well plates used in most protocols, should simplify the control of these parameters and the scaling of trunk-like structure production.

In addition, by comparing Matrigel-embedded versus non-embedded conditions, we observed a dominant role of the ECM in determining large-scale tissue organization in line with other studies on mouse and human models^{20,24–26,30,33,65}. In the absence of Matrigel, the neuroepithelium resembled an open neural plate with an apical-out organization, and somitic cells did not shape into defined apical-in somitic clusters. The starting cell number might have an influence on these processes because differentiation based on large aggregates (>3,000 cells) results in structures composed of multiple internal neuroepithelia instead of the more peripheral observed here¹⁸ (Supplementary Table 3). Matrigel addition induces a re-organization of cells' apical–basal polarity, a key parameter for epithelia morphogenesis⁴⁵, to promote the formation of a neural tube surrounded by segmented somites. The two alternative models of human trunk development controlled by Matrigel addition should help to decipher the mechanisms underlying neural tube formation in these *in vitro* systems and their relation to the *in vivo* situation as well as the links among ECM, cell apical–basal polarity and high-order tissue organization across germ layers, a question that is difficult to address in animals, let alone in humans.

The surprisingly high level of self-organization in the absence of extra-embryonic tissues, endoderm, notochord or anterior neural cell types indicates that these processes are, in part, autonomous within the neuro/mesodermal lineage. However, we observed that neural tubes, to some extent, and, more particularly, the somites, are not as robustly organized as *in vivo*. Their number and organization vary between organoids (Figs. 2 and 3). The shape of the embryo and the mechanical constraints provided by the missing embryonic and extra-embryonic tissues likely participate in the robustness of this stereotypical organization^{66–68}. This model will, thus, provide a system to interrogate the mechanisms controlling the morphogenesis of the somites and the neural tube as well as the interdependence of the two processes^{67,69} during normal but also pathological development, such as neural tube defects⁷⁰.

Notably, these trunk organoids continue to develop past the progenitor stage while maintaining aspects of the morphological organization observed in embryos around the fourth week of embryogenesis, a step in organogenesis poorly reached in previous models recapitulating aspects of trunk morphogenesis^{18,20,24–26,29,30}. In particular, we observed the birth of post-mitotic neurons at the periphery of the neuroepithelium and the generation of patterns of neuronal subtypes along the AP axis expressing distinct HOX transcription factors, which, in embryos, instruct the formation of distinct locomotor circuits¹⁰. Furthermore, we show that motor neuron axons exit the neural tube in abundance and, in some instances, contact muscle progenitors. The use of a high concentration of Matrigel in comparison to all previous studies (Supplementary Table 3) might be key in providing a scaffold for an organized development at these later stages. Future work will determine whether functional neuro-muscular circuits form, as previously observed in other models^{14,71,72}. However, signaling centers essential for dorso-ventral patterning, such as the notochord, are absent in these trunk-like structures, as in most other models of trunk tissue development^{14,17,18,20,25,26,30,33,34}. The consequence is a poor recapitulation of the topographical organization of cell types observed along the embryo dorso-ventral axis, a process essential for the proper formation of spinal sensory-motor circuits and associated musculo-skeletal system. The recent generation of notochord-like cells in an *in vitro* model of human trunk development may represent a step forward in overcoming this limitation³⁵.

Overall, the self-organizing model of human trunk development that we describe here should help capture the multi-systemic aspects of spine and sensory-motor neural circuit co-development and disease. Previous organoid models based on self-organization, including brain organoids, poorly recapitulate high-order tissue morphogenesis together with large-scale topographic organization of terminal cell types^{13,14,18,73}, an important limitation in the field. Our work indicates that 3D self-organizing models can achieve many aspects of these complex events upon definition of the pathways controlling cell fate and morphogenesis opening the avenue for the development of more advanced models of human organogenesis.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41587-023-01956-9>.

References

1. Sahu, S. & Sharan, S. K. Translating embryogenesis to generate organoids: novel approaches to personalized medicine. *iScience* **23**, 101485 (2020).
2. Shao, Y. & Fu, J. Engineering multiscale structural orders for high-fidelity embryoids and organoids. *Cell Stem Cell* **29**, 722–743 (2022).
3. Moris, N., Alev, C., Pera, M. & Martinez Arias, A. Biomedical and societal impacts of *in vitro* embryo models of mammalian development. *Stem Cell Rep.* **16**, 1021–1030 (2021).
4. Sozen, B., Conkar, D. & Veenvliet, J. V. Carnegie in 4D? Stem-cell-based models of human embryo development. *Semin. Cell Dev. Biol.* **131**, 44–57 (2022).
5. Wymeersch, F. J., Wilson, V. & Tsakiridis, A. Understanding axial progenitor biology *in vivo* and *in vitro*. *Development* **148**, dev180612 (2021).
6. Pourquié, O., Al Tanoury, Z. & Chal, J. The long road to making muscle *in vitro*. In *Current Topics in Developmental Biology* Vol. 129 (ed Brivanlou, A. H.) 123–142 (Academic Press, 2018).
7. Henrique, D., Abranches, E., Verrier, L. & Storey, K. G. Neuromesodermal progenitors and the making of the spinal cord. *Development* **142**, 2864–2875 (2015).

8. Nedelec, S. & Martinez-Arias, A. In vitro models of spinal motor circuit's development in mammals: achievements and challenges. *Curr. Opin. Neurobiol.* **66**, 240–249 (2021).
9. Deschamps, J. & Duboule, D. Embryonic timing, axial stem cells, chromatin dynamics, and the Hox clock. *Genes Dev.* **31**, 1406–1416 (2017).
10. Philippidou, P. & Dasen, J. S. *Hox* genes: choreographers in neural development, architects of circuit organization. *Neuron* **80**, 12–34 (2013).
11. Abdel Fattah, A. R. et al. Actuation enhances patterning in human neural tube organoids. *Nat. Commun.* **12**, 3192 (2021).
12. Budjan, C. et al. Paraxial mesoderm organoids model development of human somites. *eLife* **11**, e68925 (2022).
13. Duval, N. et al. BMP4 patterns Smad activity and generates stereotyped cell fate organization in spinal organoids. *Development* **146**, dev175430 (2019).
14. Faustino Martins, J.-M. et al. Self-organizing 3D human trunk neuromuscular organoids. *Cell Stem Cell* **26**, 172–186 (2020).
15. Gouti, M. et al. In vitro generation of neuromesodermal progenitors reveals distinct roles for Wnt signalling in the specification of spinal cord and paraxial mesoderm identity. *PLoS Biol.* **12**, e1001937 (2014).
16. Zheng, Y. et al. Dorsal-ventral patterned neural cyst from human pluripotent stem cells in a neurogenic niche. *Sci. Adv.* **5**, eaax5933 (2019).
17. Lee, J.-H. et al. Production of human spinal-cord organoids recapitulating neural-tube morphogenesis. *Nat. Biomed. Eng.* **6**, 435–448 (2022).
18. Libby, A. R. G. et al. Axial elongation of caudalized human organoids mimics aspects of neural tube development. *Development* **148**, dev198275 (2021).
19. Maury, Y. et al. Combinatorial analysis of developmental cues efficiently converts human pluripotent stem cells into multiple neuronal subtypes. *Nat. Biotechnol.* **33**, 89–96 (2015).
20. Miao, Y. et al. Reconstruction and deconstruction of human somitogenesis in vitro. *Nature* **614**, 500–508 (2023).
21. Mouilleau, V. et al. Dynamic extrinsic pacing of the HOX clock in human axial progenitors controls motor neuron subtype specification. *Development* **148**, dev194514 (2021).
22. Ogura, T., Sakaguchi, H., Miyamoto, S. & Takahashi, J. Three-dimensional induction of dorsal, intermediate and ventral spinal cord tissues from human pluripotent stem cells. *Development* **145**, dev162214 (2018).
23. Olmsted, Z. T. & Paluh, J. L. Co-development of central and peripheral neurons with trunk mesendoderm in human elongating multi-lineage organized gastruloids. *Nat. Commun.* **12**, 3020 (2021).
24. Sanaki-Matsumiya, M. et al. Periodic formation of epithelial somites from human pluripotent stem cells. *Nat. Commun.* **13**, 2325 (2022).
25. van den Brink, S. C. et al. Single-cell and spatial transcriptomics reveal somitogenesis in gastruloids. *Nature* **582**, 405–409 (2020).
26. Veenvliet, J. V. et al. Mouse embryonic stem cells self-organize into trunk-like structures with neural tube and somites. *Science* **370**, eaba4937 (2020).
27. Xu, P.-F. et al. Construction of a mammalian embryo model from stem cells organized by a morphogen signalling centre. *Nat. Commun.* **12**, 3277 (2021).
28. Xue, W. et al. Generation of dorsoventral human spinal cord organoids via functionalizing composite scaffold for drug testing. *iScience* **26**, 105898 (2022).
29. Yaman, Y. I. & Ramanathan, S. Controlling human organoid symmetry breaking reveals signaling gradients drive segmentation clock waves. *Cell* **186**, 513–527 (2023).
30. Yamanaka, Y. et al. Reconstituting human somitogenesis in vitro. *Nature* **614**, 509–520 (2023).
31. Diaz-Cuadros, M. et al. In vitro characterization of the human segmentation clock. *Nature* **580**, 113–118 (2020).
32. Park, J., Hsiung, H.-A., Khven, I., La Manno, G. & Lutolf, M. P. Self-organizing in vitro mouse neural tube organoids mimic embryonic development. *Development* **149**, dev201052 (2022).
33. Beccari, L. et al. Multi-axial self-organization properties of mouse embryonic stem cells into gastruloids. *Nature* **562**, 272–276 (2018).
34. Moris, N. et al. An in vitro model of early anteroposterior organization during human development. *Nature* **582**, 410–415 (2020).
35. Rito, T., Libby, A. R. G., Demuth, M. & Briscoe, J. Notochord and axial progenitor generation by timely BMP and NODAL inhibition during vertebrate trunk formation. Preprint at *bioRxiv* <https://doi.org/10.1101/2023.02.27.530267> (2023).
36. Van Den Brink, S. C. et al. Symmetry breaking, germ layer specification and axial organisation in aggregates of mouse embryonic stem cells. *Development* **141**, 4231–4242 (2014).
37. Olmsted, Z. T. & Paluh, J. L. Stem cell neurodevelopmental solutions for restorative treatments of the human trunk and spine. *Front. Cell. Neurosci.* **15**, 667590 (2021).
38. Chambers, S. M. et al. Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. *Nat. Biotechnol.* **27**, 275–280 (2009).
39. Hemmati-Brivanlou, A., Kelly, O. G. & Melton, D. A. Follistatin, an antagonist of activin, is expressed in the Spemann organizer and displays direct neuralizing activity. *Cell* **77**, 283–295 (1994).
40. Camus, A., Perea-Gomez, A., Moreau, A. & Collignon, J. Absence of Nodal signaling promotes precocious neural differentiation in the mouse embryo. *Dev. Biol.* **295**, 743–755 (2006).
41. Sagner, A. & Briscoe, J. Establishing neuronal diversity in the spinal cord: a time and a place. *Development* **146**, dev182154 (2019).
42. Tani, S., Chung, U.-I., Ohba, S. & Hojo, H. Understanding paraxial mesoderm development and sclerotome specification for skeletal repair. *Exp. Mol. Med.* **52**, 1166–1177 (2020).
43. Goto, H., Kimmey, S. C., Row, R. H., Matus, D. Q. & Martin, B. L. FGF and canonical Wnt signaling cooperate to induce paraxial mesoderm from tailbud neuromesodermal progenitors through regulation of a two-step epithelial to mesenchymal transition. *Development* **144**, 1412–1424 (2017).
44. Araya, C., Carmona-Fontaine, C. & Clarke, J. D. W. Extracellular matrix couples the convergence movements of mesoderm and neural plate during the early stages of neurulation. *Dev. Dyn.* **245**, 580–589 (2016).
45. Buckley, C. E. & St Johnston, D. Apical–basal polarity and the control of epithelial form and function. *Nat. Rev. Mol. Cell Biol.* **23**, 559–577 (2022).
46. Long, K. R. & Huttner, W. B. How the extracellular matrix shapes neural development. *Open Biol.* **9**, 180216 (2019).
47. Mongera, A., Michaut, A., Guillot, C., Xiong, F. & Pourquié, O. Mechanics of anteroposterior axis formation in vertebrates. *Annu. Rev. Cell Dev. Biol.* **35**, 259–283 (2019).
48. Boulet, A. M. & Capecchi, M. R. Signaling by FGF4 and FGF8 is required for axial elongation of the mouse embryo. *Dev. Biol.* **371**, 235–245 (2012).
49. Cunningham, T. J., Kumar, S., Yamaguchi, T. P. & Duester, G. *Wnt8a* and *Wnt3a* cooperate in the axial stem cell niche to promote mammalian body axis extension. *Dev. Dyn.* **244**, 797–807 (2015).
50. Ribes, V., Le Roux, I., Rhinn, M., Schuhbauer, B. & Dollé, P. Early mouse caudal development relies on crosstalk between retinoic acid, Shh and Fgf signalling pathways. *Development* **136**, 665–676 (2009).

51. Anand, G. M. et al. Controlling organoid symmetry breaking uncovers an excitable system underlying human axial elongation. *Cell* **186**, 497–512 (2023).
52. Padmanabhan, R. Retinoic acid-induced caudal regression syndrome in the mouse fetus. *Reprod. Toxicol.* **12**, 139–151 (1998).
53. Tan, Y. & Cahan, P. SingleCellNet: a computational tool to classify single cell RNA-seq data across platforms and across species. *Cell Syst.* **9**, 207–213 (2019).
54. Tyser, R. C. V. et al. Single-cell transcriptomic characterization of a gastrulating human embryo. *Nature* **600**, 285–289 (2021).
55. Rayon, T., Maizels, R. J., Barrington, C. & Briscoe, J. Single-cell transcriptome profiling of the human developing spinal cord reveals a conserved genetic programme with human-specific features. *Development* **148**, dev199711 (2021).
56. Mayeuf-Louchart, A. et al. Endothelial cell specification in the somite is compromised in Pax3-positive progenitors of *Foxc1/2* conditional mutants, with loss of forelimb myogenesis. *Development* **143**, 872–879 (2016).
57. Mayeuf-Louchart, A. et al. Notch regulation of myogenic versus endothelial fates of cells that migrate from the somite to the limb. *Proc. Natl Acad. Sci. USA* **111**, 8844–8849 (2014).
58. Prummel, K. D., Nieuwenhuize, S. & Mosimann, C. The lateral plate mesoderm. *Development* **147**, dev175059 (2020).
59. Junker, J. P. et al. Genome-wide RNA tomography in the zebrafish embryo. *Cell* **159**, 662–675 (2014).
60. Rossi, G. et al. Capturing cardiogenesis in gastruloids. *Cell Stem Cell* **28**, 230–240 (2021).
61. Rayon, T. et al. Species-specific pace of development is associated with differences in protein stability. *Science* **369**, eaba7667 (2020).
62. Marklund, U. et al. Detailed expression analysis of regulatory genes in the early developing human neural tube. *Stem Cells Dev.* **23**, 5–15 (2014).
63. Matsuda, M. et al. Recapitulating the human segmentation clock with pluripotent stem cells. *Nature* **580**, 124–129 (2020).
64. Blauwkamp, T. A., Nigam, S., Ardehali, R., Weissman, I. L. & Nusse, R. Endogenous Wnt signalling in human embryonic stem cells generates an equilibrium of distinct lineage-specified progenitors. *Nat. Commun.* **3**, 1070 (2012).
65. Kyprianou, C. et al. Basement membrane remodelling regulates mouse embryogenesis. *Nature* **582**, 253–258 (2020).
66. Rifes, P. et al. Redefining the role of ectoderm in somitogenesis: a player in the formation of the fibronectin matrix of presomitic mesoderm. *Development* **134**, 3155–3165 (2007).
67. Radice, G. L. et al. Developmental defects in mouse embryos lacking N-cadherin. *Dev. Biol.* **181**, 64–78 (1997).
68. Nikolopoulou, E., Galea, G. L., Rolo, A., Greene, N. D. E. & Copp, A. J. Neural tube closure: cellular, molecular and biomechanical mechanisms. *Development* **144**, 552–566 (2017).
69. Ensini, M., Tsuchida, T. N., Belting, H. G. & Jessell, T. M. The control of rostrocaudal pattern in the developing spinal cord: specification of motor neuron subtype identity is initiated by signals from paraxial mesoderm. *Development* **125**, 969–982 (1998).
70. Wolujewicz, P., Steele, J. W., Kaltschmidt, J. A., Finnell, R. H. & Ross, M. E. Unraveling the complex genetics of neural tube defects: From biological models to human genomics and back. *Genesis* **59**, e23459 (2021).
71. Andersen, J. et al. Generation of functional human 3D cortico-motor assembloids. *Cell* **183**, 1913–1929 (2020).
72. Barbeau, S., Tahraoui-Bories, J., Legay, C. & Martinat, C. Building neuromuscular junctions in vitro. *Development* **147**, dev193920 (2020).
73. Cederquist, G. Y. et al. Specification of positional identity in forebrain organoids. *Nat. Biotechnol.* **37**, 436–444 (2019).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2023

Methods

Culture of hiPSCs

The iPSC lines WTSli002 (male, alternative name HPSI0913i-eika_2, RRID: CVCL_AH30, abbreviated WTS2) and HPSI0714i-KUTE_4 (female, RRID: CVCL_EF20, abbreviated KUTE4) were obtained from the European Bank for Pluripotent Stem Cells. SOX2-GFP (male, WTC-mEGFP-SOX2-cl26 mono allelic tag, RRID: CVCL_WM15) and WTC11 (male, GM25256, RRID: CVCL_Y803) were obtained from the National Institute of General Medical Sciences Human Genetic Cell Repository at the Coriell Institute for Medical Research. Experiments with hiPSCs were approved by relevant ethics committees (declaration DC-2015-2559). For the references of reagents, see Supplementary Table 4. All hiPSC lines were cultured at 37 °C on Matrigel in mTSE1 medium and amplified using EDTA (Life Technologies) clump-based passaging. Cells were tested for potential mycoplasma contamination every other week (MycAlert Mycoplasma Detection Kit). No contamination was detected during the study. hiPSCs were thawed in the presence of Y-27632 (10 μM), and the culture media were changed every day.

Differentiation of hiPSCs

For the references of products and chemicals, see Supplementary Table 4. hiPSC spheroid-based differentiation was performed using a modified version of refs. 19,21. hiPSCs were dissociated to single cells and counted. For dissociation, cold accutase (4 °C) was applied on hiPSCs. Dishes were incubated at 37 °C for 3–5 min (until clusters of cells start to detach from the bottom of the plate). Colonies were dissociated by gentle pipetting. Then, cells were centrifuged (300g, 5 min), resuspended in N2B27 differentiation medium (Advanced DMEMF12/Neurobasal v/v, supplemented with N2, B27 without vitamin A, penicillin–streptomycin 1% and β-mercaptoethanol 0.1%) with Y-27632 (10 μM). For differentiation, single cells (50–600 cells per well as indicated) were seeded in 800 or 400 aggrewell plates (preferably 400) in N2B27, CHIR-99021 (4 μM) and, when indicated, SB431542 (20 μM) and LDN193189 (100 nM) or FGF2 (20 ng ml⁻¹). Plates were centrifuged (100g, 3 min) and transferred at 37 °C/5% CO₂. Medium change at day 2 was performed in the aggrewells with the same medium except Y-27632. At day 3, aggregates were transferred in ultra-low adhesion six-well plates (Corning). RA (10 nM) was added on day 3 and SAG (Smoothed Agonist, 500 nM) on day 4. DAPT (10 μM) was added at day 8 when indicated. Media were changed every other day unless specified.

For Matrigel embedding, spheroids at day 3 were individually embedded in 50% growth factor-reduced Matrigel (Corning) and left for 45 min at 37 °C for the Matrigel to solidify. Drops were transferred in ultra-low adhesion six-well plates, and differentiation proceeded as described above. For live imaging, test of inhibitors or analysis of Matrigel-embedded structures on day 4, the day 3 spheroids (15 to 20 structures) were embedded in 50% Matrigel and deposited on glass either on Petri dishes (MatTek) or 96-well plates (Cellvis). Dishes were flipped upside down, transferred at 37 °C for 15 min and turned back before adding medium in wells. Spheroids were then cultured as described above until imaging (see below). For the modulation of signaling pathways (Extended Data Fig. 6), the following compounds were added between day 3 and day 7: PD173074 (FGFR inhibitor): 100 nM; SCH772984 (ERK inhibitor): 100 nM; FGF2: 30 ng ml⁻¹; RA: 500 nM; and IWP2 (WNT pathway inhibitor): 5 μM.

Tissue preparation for immunofluorescence analysis

The day 4 or day 7 organoids embedded in Matrigel placed on glass coverslip were fixed with 4% paraformaldehyde (PFA) at room temperature, rinsed with PBS and kept at room temperature before processing. Free-floating organoids, embedded or not, were fixed at 4 °C with 4% cold PFA for 15 min (days 4–8), 5 min (days 10–14) or 20 min (whole mount), rinsed with PBS and processed for whole mount or cryoprotected with 30% sucrose. For cryostat sectioning, structures were embedded in OCT (TissueTek, Leica), organized for longitudinal

or sagittal sectioning at the bottom of cryomolds and frozen on dry ice. Then, 14-μm or 16-μm sections (MICROM MICROTOME) were placed on SuperFrost slides (EpreDia) and kept at –80 °C. For the day 3 analysis, spheroids were plated on Matrigel-coated glass coverslips and fixed 60 min afterwards with 4% cold PFA for 15 min.

Array tomography

Tomo-seq was performed using a robotized (SORT-seq2-based) version of a previously published tomo-seq protocol⁵⁹. In brief, fresh organoids ($n = 2$, WTS2 cell line) at day 7 were washed in cold PBS, transferred in plastic mold pre-filled with ice-cooled OCT cryosolution (TissueTek, Leica), snap frozen on dry ice and stored at –80 °C until sectioning using a cryotome. Sections of 20 μm were collected in the wells of a hard-shell polymerase chain reaction (PCR) low-profile, semi-skirted 96-well plate (Bio-Rad) that was pre-filled with mineral oil (Sigma-Aldrich) and CEL-seq2 primers. For each well, a unique, barcoded CEL-seq2 primer was used, which allowed us to pool the content of the wells after second-strand synthesis. To sequence the mRNA content of the wells, SORT-seq (robotized CEL-seq2-based scRNA-seq) was performed using a NanoDrop II liquid handling platform (GC biotech).

10x Genomics scRNA-seq

For single-cell preparation, 26 organoids at day 7 coming from one differentiation of the WTS2 cell line were collected in a 1.5-ml Eppendorf tube and cooled on ice for 10 min. Organoids were washed with PBS w/o Ca²⁺/Mg²⁺ and incubated with trypsin-EDTA 4 °C (0.05%) for 10 min. After trypsin removal, mechanical dissociation was performed using a p1000 pipette in N2B27 supplemented with DNase (25 μg ml⁻¹) and 0.1% BSA. The supernatant containing the cells was filtered through a 40-μm filter and centrifuged for 6 min at 300g. Cell number, together with the absence of dead cells, debris or doublets, was determined after resuspension in N2B27 using a hemocytometer and trypan blue staining. Cells were then pelleted (6 min, 300g), resuspended in CryoStor medium, frozen and stored at –80 °C. Frozen cells were thawed in a 37 °C water bath, resuspended in culture medium (N2B27) and spun at 500g for 7 min at room temperature. Cells were resuspended in N2B27 medium with 10 μM Y-27632 at a density of 700–1,200 cells per microliter and processed for scRNA-seq (10x Genomics) by Single Cell Discoveries, following the manufacturer's protocol and as previously described²⁵.

Sequencing data analysis

Sequencing, processing array tomography data, comparison between tomo-seq and scRNA-seq, clustering genes on the basis of AP expression patterns and gene reproducibility analysis between replicates were performed as described in ref. 25.

Specifically, for tomo-seq data analysis, sections with at least 9,000 genes were included (code is available at <https://github.com/anna-alemany>). Specifically, for the single-cell data, FASTQ files were generated from BCL files with 10x Genomics Cell Ranger1 mkfastq 4.0. FASTQ files were mapped with Cell Ranger1 count 4.0. Read 1 was used for identification of cell barcode and unique molecular identifier (UMI). Read 2 was used to map to the reference genome, CRCh38-2020-A. Filtering of empty barcodes was done following the standard Cell Ranger procedure. Gene annotations of GENCODE v32/Ensembl 98 were used. Processing and analyzing single-cell data was done in R Studio (version 4.1.0) using the Seurat package (version 4.1.0). In brief, cells with at least 3,000 transcripts, and with between 2,000 and 9,500 genes, were selected. Additionally, cells with a log₁₀ (genes-per-transcript) value higher than 0.8, and a ratio of reads originating from mitochondrial genes between 0.01 and 0.1, were selected. Finally, only genes present in at least 10 cells were used for analysis. Data were normalized using SCTransform in Seurat and clustered using the Leiden algorithm with the first 50 dimensions of reduction. Differentially expressed genes for each cluster were determined using a Wilcoxon rank-sum test on genes showing at least 0.25 log₂(fold change).

Comparison of organoid cell clusters to primary cell types was based on a previously published analysis²⁵. In brief, common marker genes ($P < 0.01$ and $\log_2(\text{fold change}) > 1.01$) between organoid cell clusters and cell clusters of primary human embryonic tissues as defined in previously published scRNA-seq datasets^{54,55} were found. Statistical significance for this amount of common marker genes was calculated using a binomial test, where probabilities were determined by randomizing the primary tissue cluster marker gene list, based on the full list of primary tissue marker genes ($n = 10,000$).

Immunostaining

For immunostainings, cryostat sections were incubated with a blocking solution (0.2% Triton X-100/PBS (PBST); 10% FBS) for 10 min at room temperature. Primary antibodies (Supplementary Table 5) were diluted in staining solution (0.2% PBST; 2% FBS) and incubated overnight at 4 °C in a humidified chamber. For Matrigel-embedded organoids in drops deposited on glass, Triton was 0.3%, and primary and secondary antibodies were incubated overnight at room temperature. Secondary antibodies and/or Phalloidin (Supplementary Table 5) were incubated for 1 h at room temperature in the staining solution with DAPI when needed. Slices were mounted in Fluoromount (Sigma-Aldrich or CliniSciences). For whole-mount staining, organoids were incubated in 0.5% PBST/10% FBS solution at room temperature for 4–6 h. Primary antibodies were diluted in 0.5% PBST/2% FBS, incubated for 24–48 h at 4 °C and washed 12–16 h with PBST. Secondary antibodies were diluted in 0.5% PBST/2% FBS and incubated for 4–6 h at room temperature. All steps were performed on a rocking platform. All structures were mounted using Fluoromount before imaging. We observed that the widely used goat anti-SOX1 antibody (R&D Systems, AF3369) cross-reacts with SOX3 protein and gives staining consistent with SOX3 expression profile in scRNA-seq data (expression in the axial progenitor/preneural tube transition zone where SOX1 is absent). An anti-SOX1-specific antibody (Cell Signaling Technology, 4194) gives a weaker staining selectively in the neuroepithelium consistent with the low expression of SOX1 in motor neuron progenitors in mouse and human embryos⁵⁵. Therefore, we re-labelled the staining obtained with the R&D systems antibody as SOX1/3 throughout the article.

FISH assay

FISH for *HOXC6/9* genes was performed using RNAscope Multiplex Fluorescent version 2 (Advanced Cell Diagnostics) according to the manufacturer's instructions, on 14- μm -thick cryostat longitudinal sections of day 7 organoids.

In brief, cryostat sections were post-fixed for 30 min in 4% PFA and then rinsed with PBS. Slides were dehydrated in sequential baths of 50%, 70% and 100% EtOH and then rehydrated with 70% and 50% EtOH before being incubated in PBS. Slides were treated with hydrogen peroxide for 10 min at room temperature, washed with fresh distilled water and incubated with protease III at room temperature for 10 min. Hybridization, amplification and signal development were performed following kit instructions using *HOXC6* probe/Opal 520 (1:1,000) and *HOXC9* probe/Opal 620 (1:1,000) combinations (Supplementary Table 4). CDX2 staining was performed on the adjacent slice to identify the caudal region of the organoid. Sections were selected based on the presence of an extended neural tube and a caudal pole of CDX2^{high} cells ($n = 3$, for a total of 10 organoids).

Clearing protocol

Immunolabeled organoids were embedded into 1.5% low-melting agarose (in ultra-pure Tris-acetate-EDTA solution at 1%). The embedded samples were cleared following a modified version of the original iDISCO+ protocol⁷⁴. After being placed in 15-ml tubes (TPP, 91515), samples were dehydrated in 50% methanol (in ddH₂O) for 1 h, 80% methanol/water for 1 h and 100% methanol for 1 h twice. Samples were incubated for 90 min in one volume of methanol/two volumes of dichloromethane (DCM) until they sank at the bottom of the tube.

Samples were then washed for 30 min twice in 100% DCM. Finally, samples were incubated (without shaking) in dibenzyl ether (DBE) overnight and then stored in DBE at room temperature.

Image acquisition

Samples were visualized and imaged using a DMIL LED fluorescence microscope (Leica), coupled to an MC120 HD camera (Leica) and LAS version 4.8 software (bright-field images); a confocal microscope, TCS SP5 II (Leica), coupled to LASAF; a DM6000 microscope (Leica) with a CoolSNAP EZ CDD camera controlled by MetaMorph software; or a CellDiscoverer 7 (Zeiss), equipped with an AxioCam 506m camera controlled by ZEN 2.6 Blue edition software (Zeiss) or a stereo-microscope (Linx, Vision Engineering).

For time-lapse experiments, images were acquired either with a $\times 10$ dry objective (CFI Plan Apo Lambda $\times 10$, Nikon MRD00105) or a $\times 63$ oil immersion objective (HCX PL APO $\times 63/1.32-0.6$ Oil CS, Leica 506180) on an inverted microscope (Nikon Ti Eclipse) equipped with a heating enclosure (Digital Pixel, a CO₂ controller (5%)), a spinning disk confocal head (Yokogawa, CSU-W1), MicroManager software 1.4 and an sCMOS Camera (Orca Flash4LT, Hamamatsu). $\times 10$ images were recorded over 78- μm -thick z-stacks (14 steps of 6 μm) for 48 h (1-h time lapse). At $\times 63$, images were recorded over 45- μm -thick z-stacks (nine steps of 5 μm) for 9.5 h (5-min time lapse) (13% laser power, 100-ms exposure/z-step). For image processing and data analysis, we used ImageJ/Fiji software. When needed, images were subjected to brightness and contrast adjustments, by equal application across the entire image, to equilibrate channel intensities and background using ImageJ/Fiji or Adobe Photoshop CS4 software.

Cleared samples were imaged ($\times 6.3$ zoom magnification) on a light-sheet microscope (LaVision Biotec) equipped with an sCMOS camera (Andor Neo) and Inspector Microscope controller software. Imaris (Bitplane) was used for 3D reconstructions, snapshots and movies.

Quantification and data analysis

Across the article, n indicates independent differentiations started from newly thawed hiPSCs, and each individual experiment contained technical replicates. All quantifications are expressed as mean \pm s.d. unless specified in figure legends.

To analyze aggregate morphology, width and length were manually measured on $\times 5$ bright-field images using ImageJ/Fiji software. Twenty to 100 aggregates were analyzed per n . For the WTS2 cell line, $n = 11$ at day 3, $n = 12$ at day 4 and $n = 13$ at day 7; for the WTC11 cell line, $n = 2$ at day 3 and $n = 3$ at day 4 and day 7. For the Matrigel-embedded structures, 8–20 organoids per experiment were analyzed (WTS2, $n = 8$; WTC11, $n = 4$; SOX2-GFP, $n = 3$). The aspect ratio was calculated by dividing the maximum length by the maximum width of individual aggregates, and the extension length was calculated by subtracting the width from the length. The aggregates with an aspect ratio > 1.3 were considered elongated for further analysis. To quantify the percentage of spheroids that elongated after Matrigel embedding, cell aggregates (100–200 per experiment) were individually embedded at day 3 and cultured in an ultra-low adhesion six-well plate with about 50 structures per well. The percentage of elongated structures in each well was evaluated at day 7 under a stereo-microscope (Linx, Vision Engineering) (WTS2, $n = 9$; WTC11, $n = 3$; SOX2-GFP, $n = 2$).

The percentage of structures expressing TBX6 at day 3 was determined on aggregates plated on Matrigel-coated coverslips (see tissue preparation subsections, $n = 2$, 82 structures total, ~ 40 structures/ n). The percentage of non-Matrigel-embedded aggregates with SOX2/TBXT/CDX2 triple-positive axial progenitors was quantified on series of cryostat sections spanning the entire organoid. All aggregates were included at day 4 ($n = 7$, 159 structures). Only elongated structures were included at day 6 ($n = 3$, 74 structures), day 7 ($n = 4$, 48 structures) and day 8 ($n = 3$, 35 structures). The percentage of Matrigel-embedded aggregates expressing SOX2/TBXT/TBX6 at day 4 was determined on

structures placed in Matrigel drops on coverslips after whole-mount immunostaining. All aggregates were included in the analysis ($n = 3$, 38 total structures). To quantify the neural tube and somite organization, day 7 batches were processed for whole-mount immunostaining, and randomly selected elongated structures were imaged using confocal microscopy (WTS2 cell line: $n = 3$, 73 total structures; WTC11 cell line: $n = 3$, 57 total structures). For each structure, the following criteria were scored: (1) neural tube shapes (unique; unique in continuity with a multi-epithelia region on one extremity; unique then bifurcating; and multiple epithelia all along the neural region); (2) OLIG2 expression (strong gradient or restricted negative regions at the extremities); and (3) somitic organization (single rosette-like somites or a string of somites; presence of somitic tissue organized at the anterior level). To quantify the impact of signaling pathway modulation in 96-well plates, day 7 organoids were fixed and stained for SOX2 and DAPI. The aspect ratio of all structures was calculated based on SOX2 staining using ImageJ/Fiji software ($n = 3$ independent experiments, with 2 to 6 technical replicates per n , 5–20 organoids per technical replicate with the SOX2-GFP cell line; $n = 1$, with two technical replicates per n , 5–20 organoids per technical replicate for the WTS2 cell line). The quantifications of organoids displaying a specific pattern of HOX transcription factors were performed on day 14. Serial longitudinal sections across the entire organoid width were performed. Then, 14- μm sections were collected. Four sections spaced by about 170 μm were collected on each slide. HOX patterns were analyzed for each organoid over the four slices ($n = 4$, 52 structures in the WTS2 cell line; $n = 2$, 17 total structures in the WTC11 cell line). The percentage of organoids with MYOD clusters present in the tissue surrounding the neural regions was performed on day 14. Two sections per organoid, spaced by about 170 μm , were scored ($n = 3$, 94 total structures, WTS2 cell line). Aggregates that failed to elongate (remained globular) at day 7 were excluded from the following analysis: tomo-seq, scRNA-seq, FISH and HOX expression at day 14.

Statistical analyses were performed using GraphPad Prism 5.0 software. Statistically significant differences between conditions were assessed using the Kruskal–Wallis non-parametric test, followed by Dunn's multiple comparison test because data did not fit normal distribution (Shapiro–Wilk normality test). For bar plots and box plots, data were illustrated using ggplot2 and gplot packages in R Studio (version 4.1.0). No statistical methods were used to predetermine sample sizes. The experiments were not randomized, and the investigators were not blinded to sample allocation during experiments.

Ethics

Experiments with hiPSCs were approved by relevant ethics committees and appropriately declared. Informed consent was signed by donors of cells, complying with the ethical guidelines of the institutions and with the legislation requirements of the country of origin. The trunk organoids described do not contain cells forming extra-embryonic tissues required for implantation, neither anterior neural progenitors generating the brain nor endoderm cells forming the gut or primordial germ cells. Therefore, these organoids cannot give rise to human embryos.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All scRNA-seq and tomo-seq data are available at the Gene Expression Omnibus: [GSE215983](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE215983) (SuperSeries composed of two SubSeries (single-cell and tomo-seq files))—the tomo-seq data ([GSE215982](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE215982)) and the scRNA-seq data ([GSE215981](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE215981)). The human embryo scRNA-seq data used in Extended Data Fig. 10 can be found at <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE171892> and <https://www.ebi.ac.uk/biostudies/arrayexpress/studies/E-MTAB-9388>. For single-cell analysis, we used GRCh38-2020-A (from 10x Genomics) based on hg38.

The human genome GRCh38 is available at https://www.ncbi.nlm.nih.gov/datasets/genome/GCF_000001405.26/. Additional data supporting the findings of this study are available in the article and the Supplementary Information. Further information for resources or technical information are available upon reasonable request from the corresponding author. Requests for transcriptomic analysis should be sent to a.vanoudenaarden@hubrecht.eu.

Code availability

The tomo-seq analysis script is available at <https://github.com/anna-alemany>.

References

74. Renier, N. et al. Mapping of brain activity by automated volume analysis of immediate early genes. *Cell* **165**, 1789–1802 (2016).

Acknowledgements

We thank M. Russeau for help and assistance on pluripotent stem cell culture; P. Lacour, F. Luciani, T. Rashed and H. Malek for their help during the project; J. Dasen for providing guinea pig anti-HOXC9 antibody; C. Metin for sharing ARL13B antibody; and L. Goutebroze for providing the Phalloidin probes. We thank H. Wichterle and the Ribes laboratory, in particular V. Ribes and P. Gilardi, as well as F. Causseret for helpful discussions along the project and critical reading of the manuscript. We thank J.-A. Girault and I. Caillé for their comments on the manuscripts. Imaging experiments and iPSC cultures were carried out, respectively, at the Imaging Platform and the Cell and Tissue Engineering Facility of the Institut du Fer à Moulin. This project was supported by two grants from the ATIP/Avenir program (S.N. and J.F.), Laboratoire d'Excellence (LabEx) Biopsy (ANR-10-LABX-73, S.N.), ANR SYMASYM (ANR-18-CE16-0021-03, S.N. and X.M.) as well as the NWO Gravitation Project: BRAINSCAPES: A Roadmap from Neurogenetics to Neurobiology (NWO: 024.004.012, B.v.S. and A.v.O.) and a ZonMW PSIDER grant (GREAT, 40-46800-98-015, A.L. and A.v.O.).

Author contributions

S.N. and S.G. conceived the project. S.N., S.G., X.M., A.v.O. and R.R. designed experiments. S.G., S.N., R.R. and C.M. performed most wet lab experiments. K.B., R.R., S.N. and X.M. performed time-lapse imaging. J.F. performed clearing and 3D imaging. A.v.O. and A.L. performed scRNA-seq and tomo-seq experiments. A.v.O. and B.v.S. performed scRNA-seq and tomo-seq analysis. Project administration: S.N., X.M. and A.v.O. Supervision: S.N., A.v.O. and X.M. Writing—original draft: S.G. and S.N. Writing—review and editing: all authors.

Competing interests

The authors declare no competing interests.

Additional information

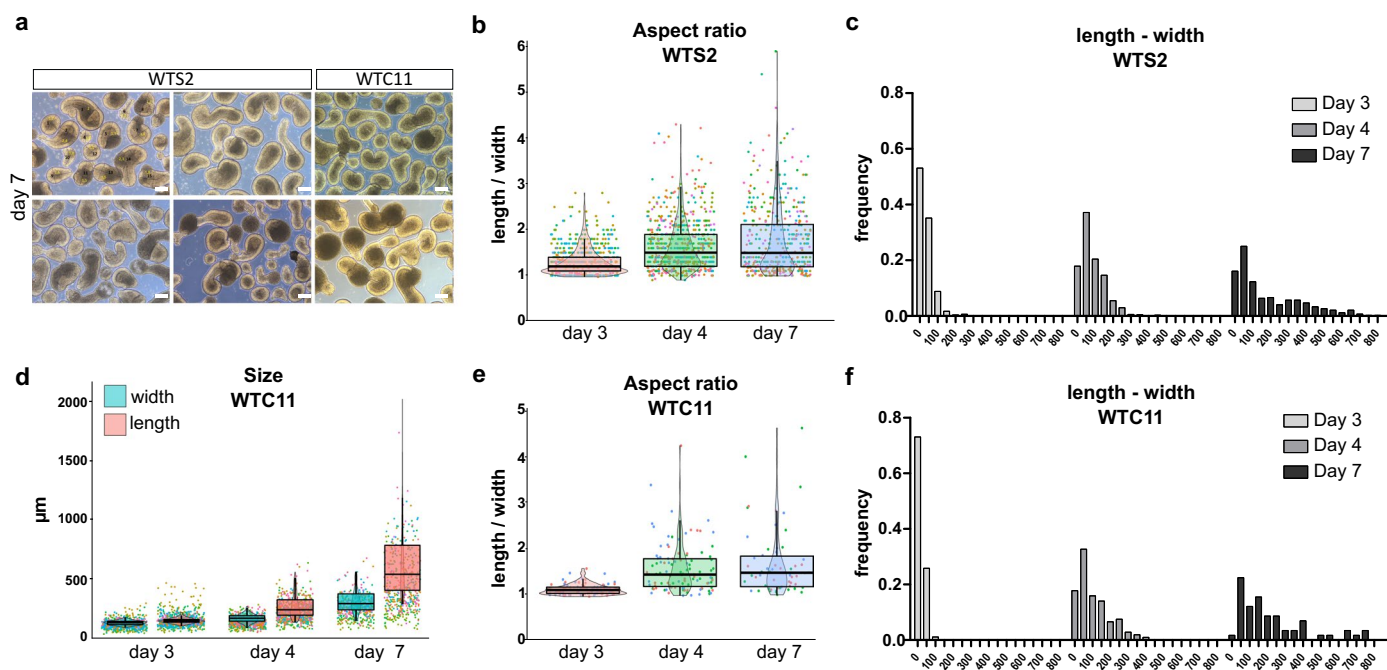
Extended data is available for this paper at <https://doi.org/10.1038/s41587-023-01956-9>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41587-023-01956-9>.

Correspondence and requests for materials should be addressed to Stéphane Nedelec.

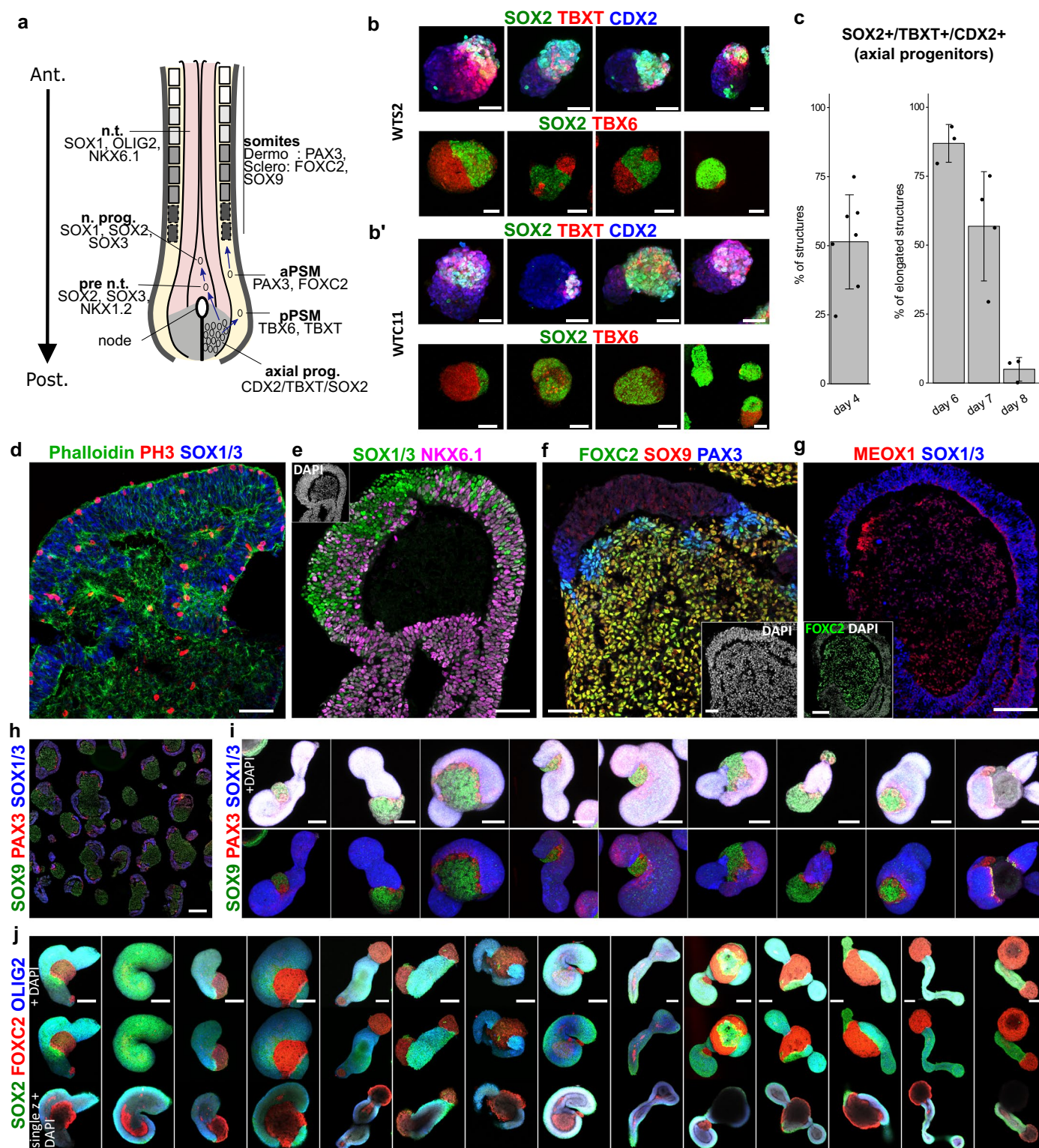
Peer review information *Nature Biotechnology* thanks Alexandre Mayran and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.



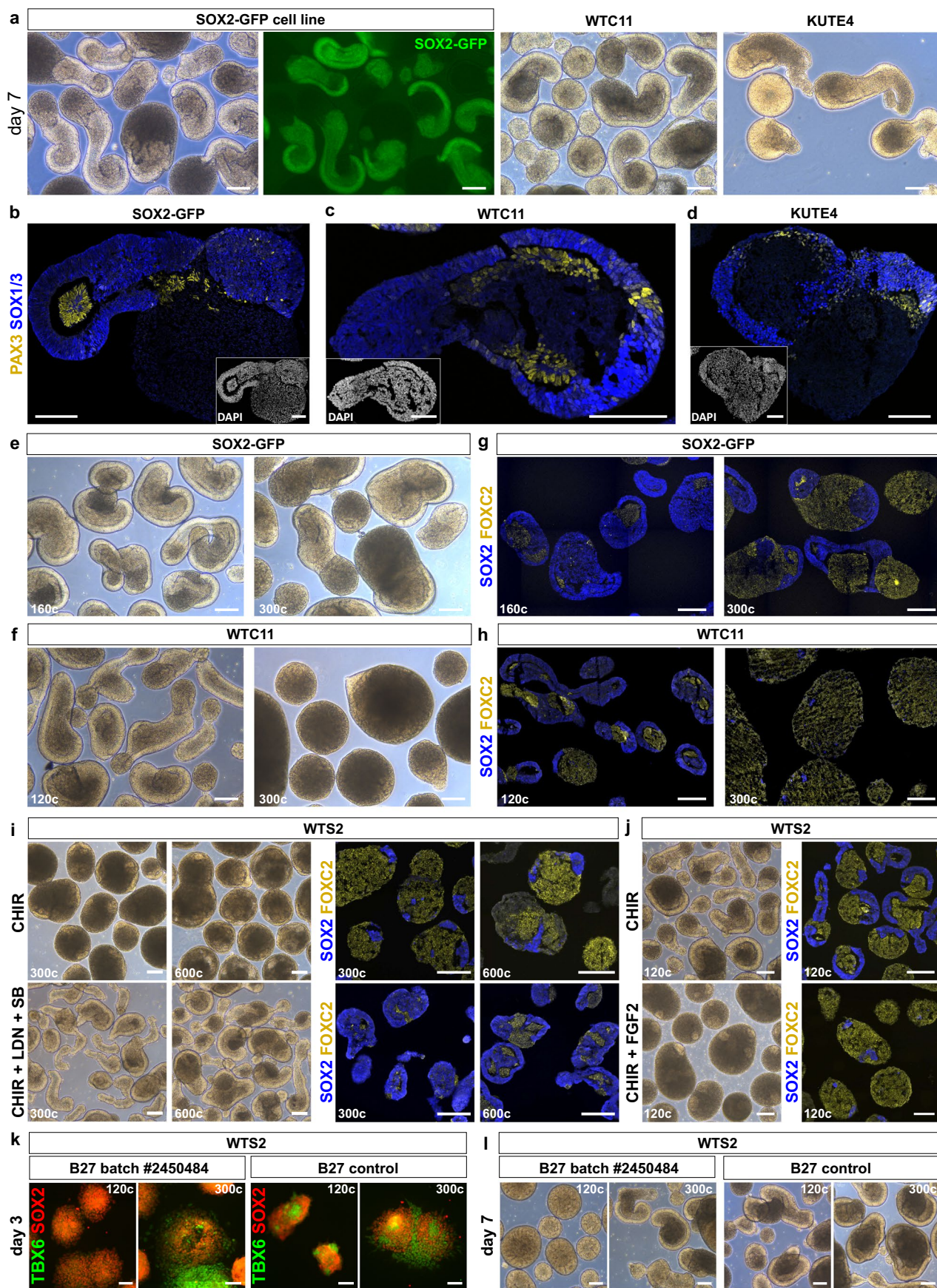
Extended Data Fig. 1 | Morphological evolution of hiPSC-derived aggregates over time. **a**, Representative bright field images of day 7 structures derived from WTS2 or WTC11 cell lines and from independent experiments, $n = 13$ for WTS2 and $n = 7$ for WTC11. **b**, Aspect ratio (Length (L)/Width (W)) of WTS2-derived structures over time. At day 7, the mean aspect ratio is 1.7 ± 0.76 . **c**, Distribution of (Lmax-Wmax) of WTS2-derived structures over time: bin center = $50 \mu\text{m}$. **d**, Box-plot of WTC11-derived structures width and length over time. At day 7, the mean length is $516 \pm 285 \mu\text{m}$. **e**, WTC11-derived structures aspect ratio over

time. At day 7, the mean aspect ratio is 1.7 ± 0.75 . **f**, Distribution of (Lmax-Wmax) of WTC11-derived structures over time: bin center = $50 \mu\text{m}$. For box plots, the central line indicates the median; box limits the 1st and 3rd quartiles and whiskers indicate the range. Each dot corresponds to a single structure and different colors indicate independent experiments. For WTS2, $n = 11$ for day 3; $n = 12$ for day 4; $n = 13$ for day 7. For WTC11, $n = 2$ for day 3; $n = 3$ for day 4; $n = 3$ for day 7. Scale bar, $200 \mu\text{m}$ (a).

**Extended Data Fig. 2 | Cell fate specification and topographical organization.**

a, Schematic of the caudal region of the mammalian embryo. Axial progenitors generate both spinal neural and somitic lineages. n.t.: neural tube. pPSM, aPSM: posterior and anterior PSM. **b-b'**, Maximum-intensity projection of whole-mount immunostaining on day 4 for markers of axial (SOX2/TBXT/CDX2), neural (SOX2) and mesodermal (TBX6) progenitors. WTS2 (**b**) and WTC11 (**b'**) derived structures. **c**, Percentage of structures with regionalized SOX2/TBXT/CDX2 progenitors in the entire population (day 4, $n = 7$, 159 structures) or among the elongated structures (at day 6, $n = 3$, 74 structures; day 7, $n = 4$, 48 structures and day 8, $n = 3$, 35 structures). Error bars represents the standard deviation.

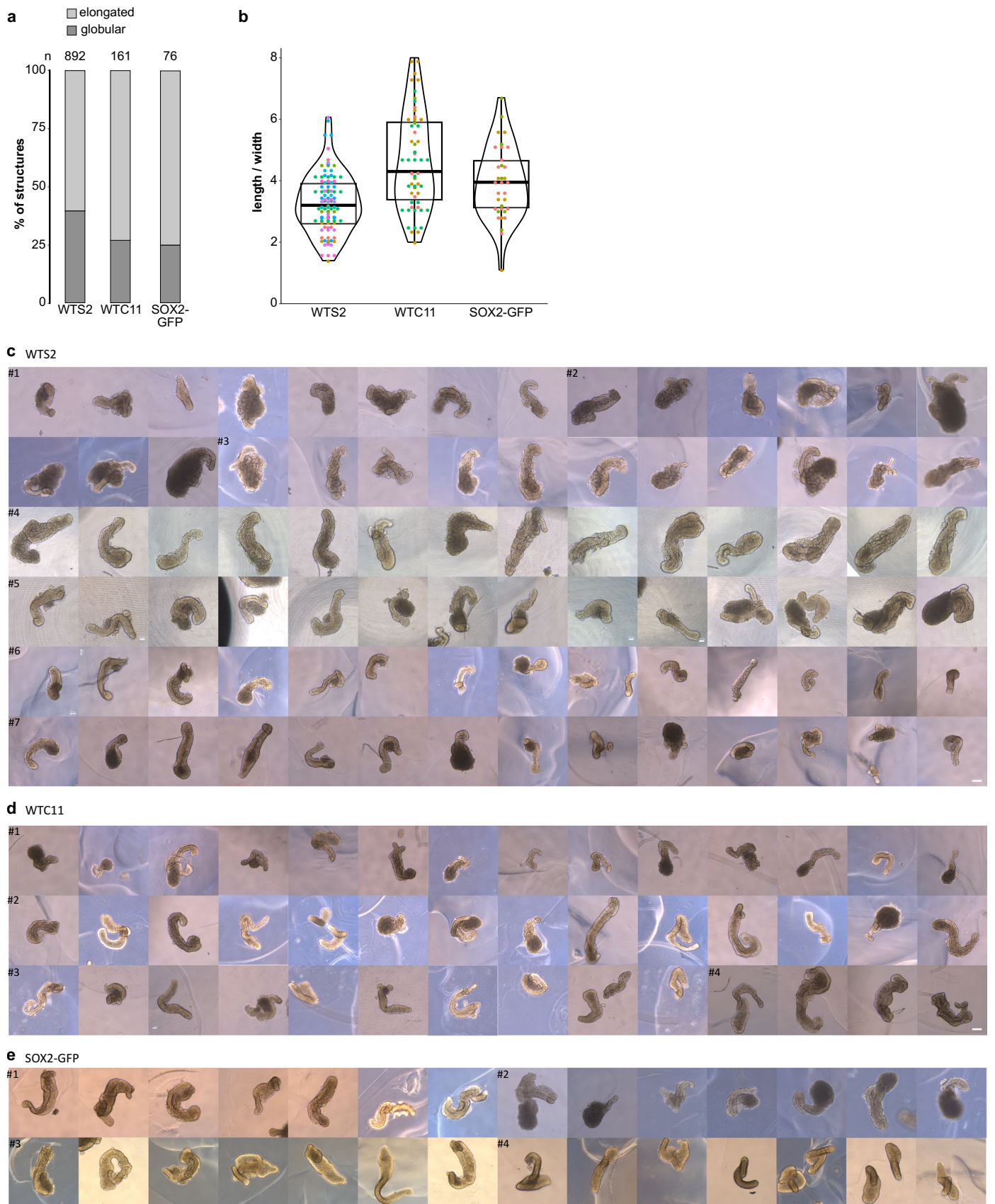
d-g, Immunostaining of sections at day 7 for mitotic cells (P-H3), F-Actin (Phalloidin) and neural progenitors (SOX1/3) markers (**d**, $n = 2$), or for the ventral neural progenitor marker (NKX6.1) (**e**, $n = 4$), or for the somitic markers FOXC2, SOX9, PAX3 (**f**, $n = 2$) and MEOX1 (**g**, $n = 5$). **h**, Low magnification view of sections of structures at day 7, immunostained for SOX9, PAX3 and SOX1/3, $n = 4$. **i, j**, 3D maximum-intensity projections of whole mount immunostaining at day 7 for the neuronal (SOX1/3, PAX3 in **i**) or SOX1/3, OLIG2 in (**j**) and somitic (FOXC2, SOX9, PAX3) markers. Optical plane in **j**. Scale bars, 50 μm (**b**, **d-f**), 100 μm (**g**, **i, j**), 500 μm (**h**).



Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | Analysis of structures derived from other hiPSC lines and factors impacting neuro/mesodermal differentiation. **a**, Representative bright field images of day 7 structures generated from SOX2-GFP, WTC11 and KUTE4 cell lines (n = 3, n = 3, n = 2). **b-d**, Immunostaining of sections of day 7 structures for neural progenitors (SOX1/3) and somitic (PAX3) markers, n = 3 for SOX2-GFP and WTC11, n = 2 for KUTE4. **e-h**, Representative bright field images (e, f; n = 3 and n = 4) and sections immunostained for neural progenitors (SOX2) and somitic (FOXC2) markers (g, h; n = 3 and n = 4) of day 7 structures generated by aggregating different cell numbers at day 0 (120, 160 and 300 cells). **i**, Representative bright field images (left panel) and sections immunostained for SOX2 and FOXC2 markers (right panel) of day 7 structures generated by aggregating different cell numbers at day 0 (300 or 600 cells) in presence of CHIR or with the addition of LDN and SB (n = 3). **j**, Representative bright field

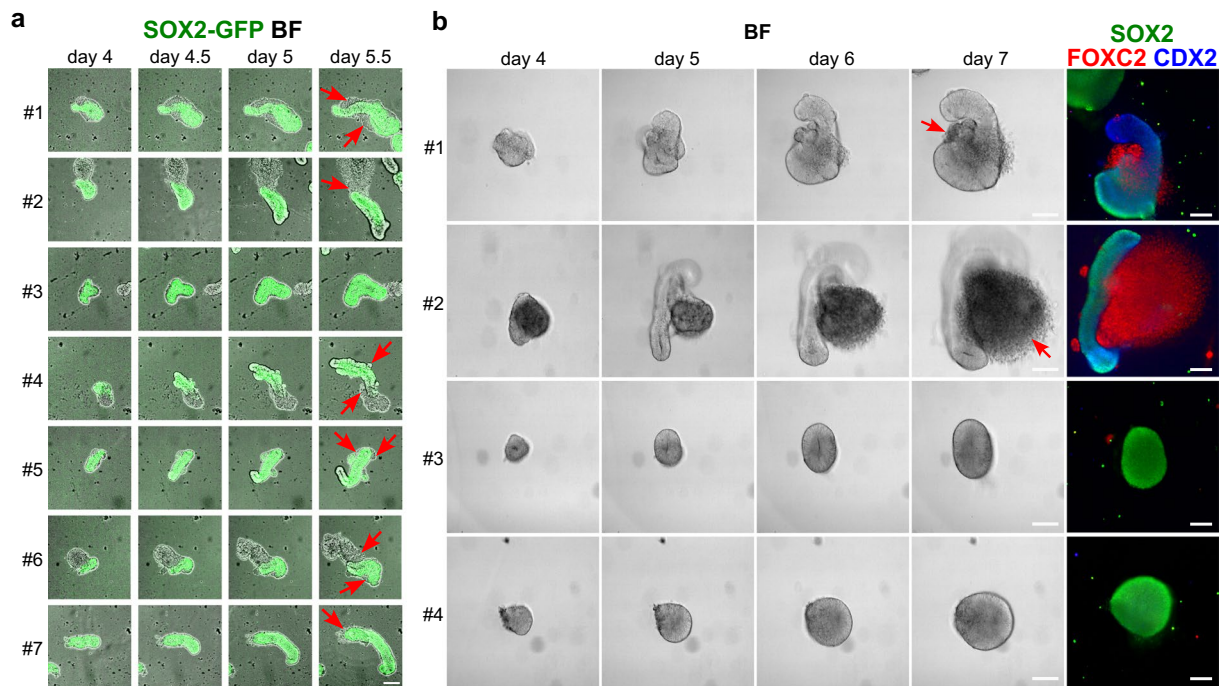
images (left panel) and sections immunostained for SOX2 and FOXC2 markers (right panel) of day 7 structures in presence of CHIR or with the addition of bFGF (20 ng/ml) for the first 2 days of differentiation (n = 3). **k**, Immunostaining of plated day 3 aggregates for neural progenitor (SOX2) and PSM (TBX6) markers. Aggregates obtained with different initial cell numbers were cultured using different B27 supplement (percentage of aggregates with TBX6 staining at day 3: B27 control, 50 ± 13% (120 cells) versus 83 ± 8% (300 cells), n = 2; B27 batch #2450484, 14 ± 9% (120 cells) versus 32 ± 21% (300 cells), n = 2; B27 batch #2450486, 2% (120 cells) versus 8% (300 cells), n = 1). **l**, Representative bright field images of day 7 structures obtained with the different initial cell numbers and different B27 supplements (n as in k). Scale bars, 200 µm (a, e, f, g, h, i, j, l), 100 µm (b, c, d), 50 µm (k).



Extended Data Fig. 4 | See next page for caption.

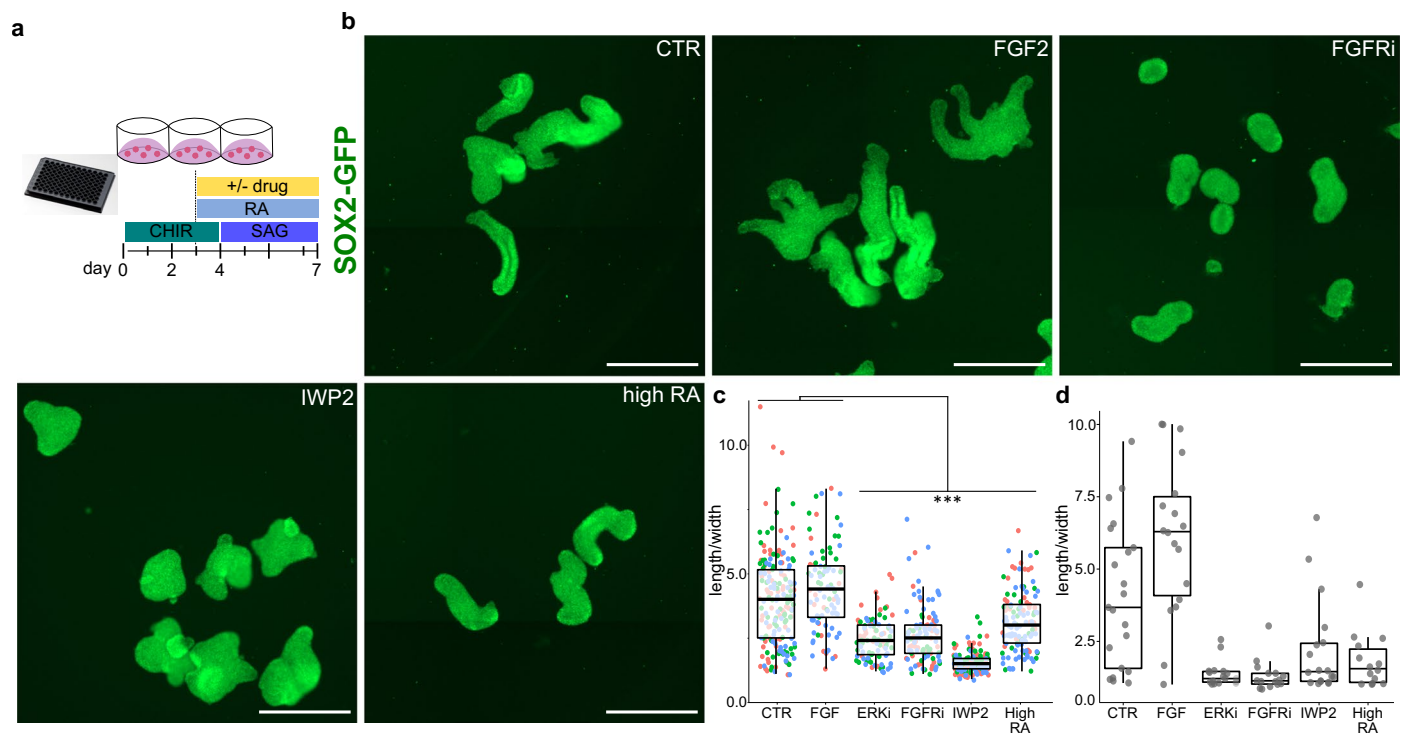
Extended Data Fig. 4 | Morphometric analysis of day 7 matrigel embedded structures. **a.** Percentage of elongation for structures derived from 3 different cells lines (WTS2, n = 9, 892 structures total; WTC11, n = 3, 161 structures total; and SOX2-GFP, n = 2, 76 structures total). **b.** Box-plot of aspect ratio (length/width) of elongated structures at day 7 (WTS2, n = 8, 100 structures total; WTC11, n = 4, 57 structures total; SOX2-GFP, n = 3, 34 structures total). In the box plots, the

central line indicates the median; box limits the 1st and 3rd quartiles and whiskers the range. Each dot corresponds to a single aggregate and different colors indicate independent differentiations. **c-e.** Gallery of bright field images of day 7 structures derived from 3 different hiPSC lines. Numbers indicate series from independent experiments. Scale bars, 200 μm (c, d, e).



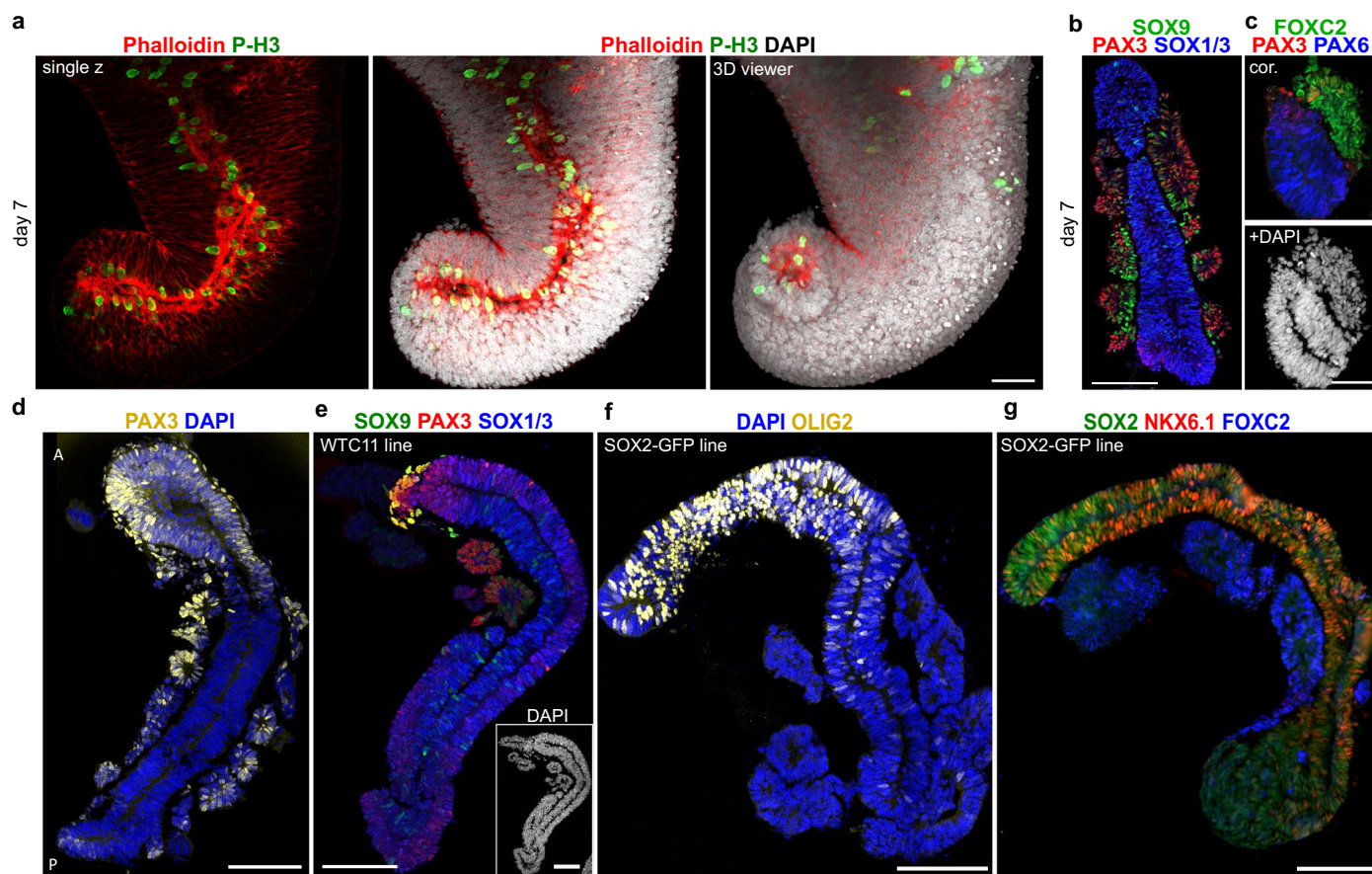
Extended Data Fig. 5 | Live imaging of matrigel embedded structures. a, Time-lapse imaging showing dynamic elongation of matrigel-embedded SOX2-GFP derived structures between day 4 and 5.5. Red arrows pinpoint SOX2 negative cells. **b,** Time-lapse of matrigel-embedded WTS2 derived structures between day

4 and 7. Post-imaging whole mount immunostaining at day 7 for axial progenitors (SOX2, CDX2) and somitic (FOXC2) markers for the corresponding structure. #3 and #4, examples of non-elongating aggregates. Scale bars, 100 μ m (a, b).



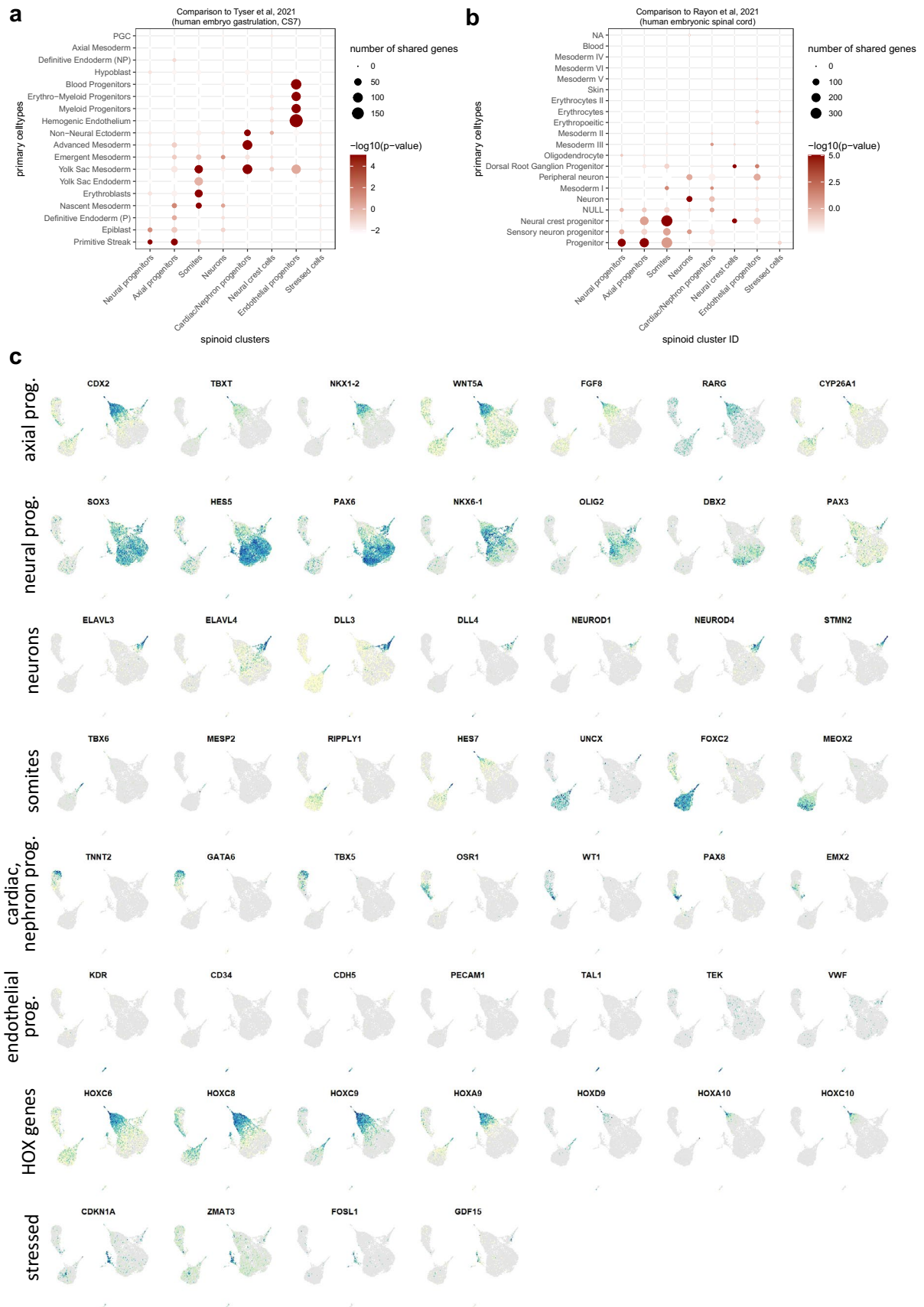
Extended Data Fig. 6 | WNT, FGF and RA pathways control neural tube formation. **a**, Experimental design. **b**, Representative images of SOX2-GFP whole mount structures in the different culture conditions at day 7, $n = 3$. **c-d**, Quantification of the aspect ratio (length/width) in the different conditions for day 7 organoids derived from SOX2-GFP (**c**) ($n = 3$, 2 to 6 wells/ n /condition)

and WTS2 (**d**) ($n = 1$, 2 wells/ n /condition) hiPSC lines. ERKi and FGFRi = FGF pathways inhibitors. IWP2 = Wnt pathway inhibitor. In the box plots, the central line indicates the median; box limits the 1st and 3rd quartiles and whiskers the range. Each dot represents a single organoid. Statistical test in **c**: Kruskal-Wallis followed by Dunn's multiple comparison test. *** $P < 0.001$. Scale bar, 500 μm .



Extended Data Fig. 7 | Organization of the neural and somitic compartments in matrigel-embedded structures. **a**, Confocal virtual slice and 3D visualization of whole mount immunostaining for F-Actin (Phalloidin) and mitotic cells (P-H3) indicating the apical side of the neuroepithelium. **b,c**, Representative images of longitudinal (**b**, $n = 3$) and coronal (**c**, $n = 2$) sections of day 7 structures immunostained for neural progenitors (SOX1/3, PAX6) and mesodermal (PAX3,

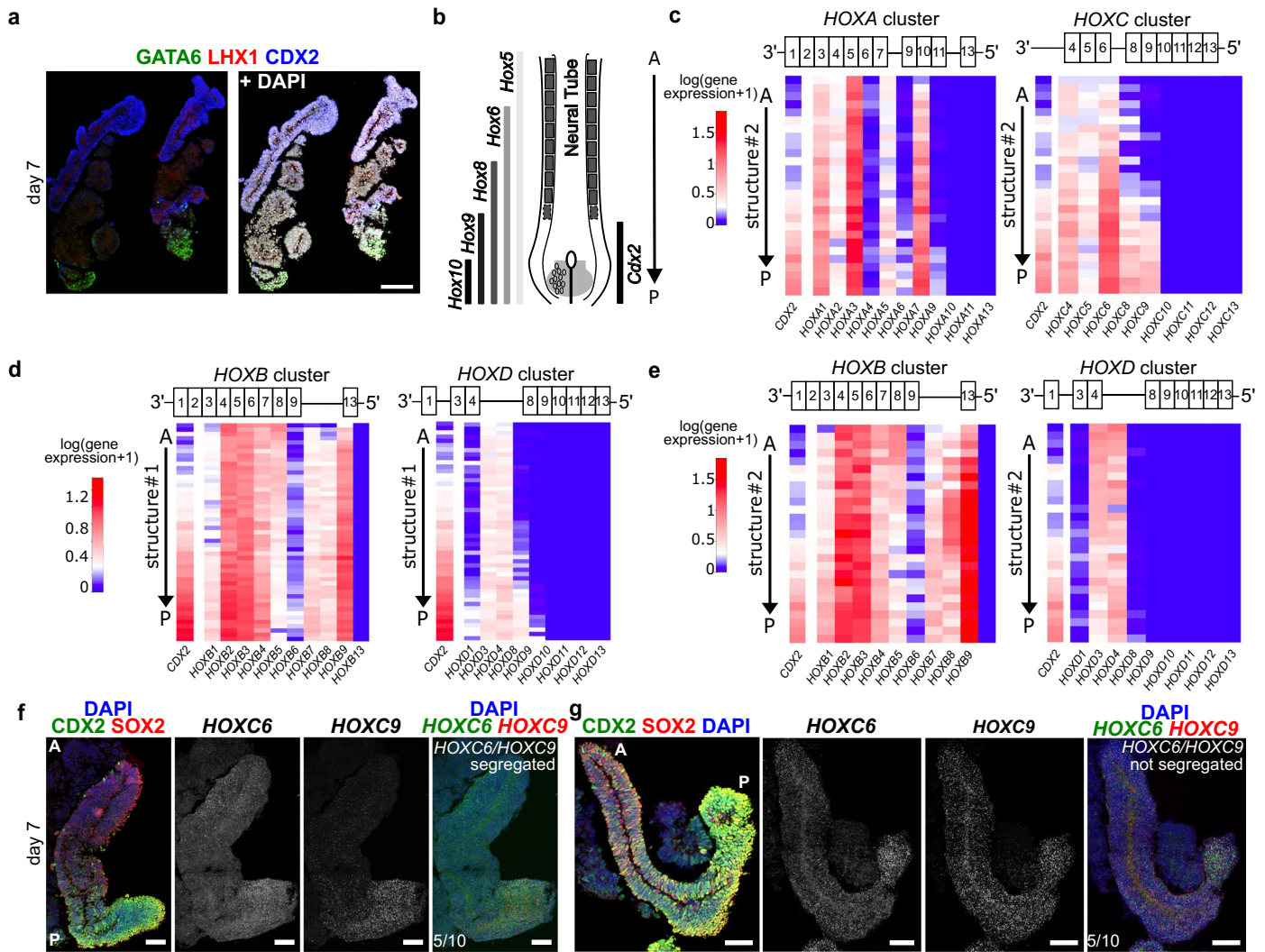
SOX9, FOXC2) markers. **d**, Immunostaining of a section consecutive to the one in Fig. 3e for the neural progenitor and PSM/somitic marker PAX3. **e-g**, Sections of day 7 structures derived from WTC11 (**e**, $n = 2$) or SOX2-GFP (**f**, $n = 4$ and $n = 2$) cell lines immunostained for markers of neural progenitors (SOX1/3, PAX3, OLIG2, NKX6.1) and somitic (SOX9, PAX3, FOXC2). Scale bars, 50 μm (**a**, **c**), 100 μm (**b**, **d-g**).



Extended Data Fig. 8 | See next page for caption.

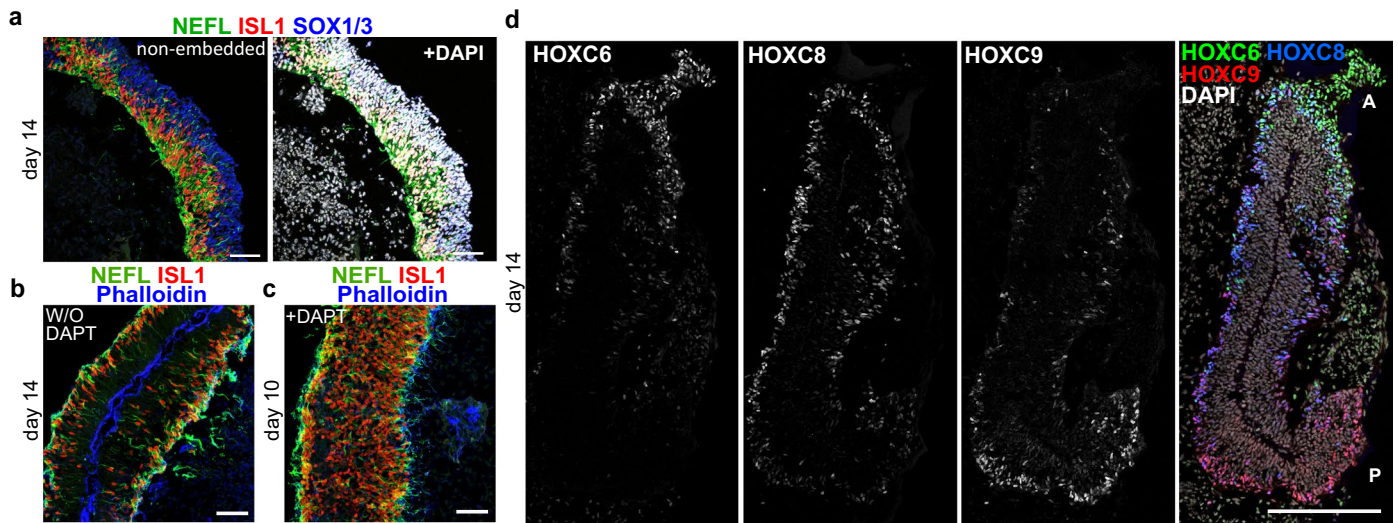
Extended Data Fig. 8 | ScRNAseq analysis of day 7 matrigel-embedded structures. a-b, ScRNA-seq clusters generated from day 7 structures compared with annotated scRNA-seq datasets from primary human gastrulating embryo⁵³, and human embryonic spinal cord tissue⁵⁴. Statistical significance was calculated using a binomial test, where probabilities were determined by randomizing the

primary tissue cluster marker gene list, based on the full list of primary tissue marker genes (n = 10,000). **c**, UMAPs colored by the expression of specific marker genes of each of the detected cell types and the expression of *HOX* genes differentially expressed along the spinal cord and somites.



Extended Data Fig. 9 | Cardiac progenitor and HOX mRNA patterns in trunk organoids. **a**, Representative images of day 7 sections immunostained for markers of cardiac (GATA6) and intermediate (LHX1) mesoderm progenitors and the caudal progenitor marker (CDX2), $n = 2$. GATA6/LHX1 clusters are in 38,1% of the structures ($n = 3, 222$ non-Matrigel-embedded structures). **b**, Expression pattern of *Hoxa/c* and *Cdx2* genes in mouse embryo. **c-e**, Tomo-seq data analysis showing the expression patterns of *HOXA* and *HOXC* clusters along structure #2 (c) and for *HOXB* and *HOXD* clusters along the structure #1 (d) and #2 (e).

f-g, Related to Fig. 3j, FISH for *HOXC6* and *HOXC9* mRNAs on additional neural tubes at day 7. **f**, Example of neural tube spanning from CDX2-low/negative region (anterior) to CDX2-high region (posterior) with segregated *HOXC6* and *HOXC9* mRNA expression domains. **g**, example of CDX2 highly expressed throughout the neural tube, with co-expressed *HOXC6* and *HOXC9* mRNA domains (5/10 organoids with co-expression, $n = 3$) (g). Scale bars, 100 μm (a), 50 μm (f-g).



Extended Data Fig. 10 | Characterization of trunk organoids at later stages.
a, Sections of day 14 non-embedded organoids immunostained for neuronal (NEFL), motor neurons (ISL1) and neuronal progenitors (SOX1/3) markers, $n = 3$.
b,c, Immunostaining for NEFL, ISL1 and the apical marker F-Actin (Phalloidin) at

day 14 in the control condition (**b**, $n = 3$), and at day 10 after DAPT addition on day 8 (**c**, $n = 2$). **d**, Related to Fig. 4h, other example of collinear expression patterns of HOXC6, HOXC8 and HOXC9 along the neural tube (longitudinal cryostat sections at day 14). Scale bars, 50 μm (a-c), 200 μm (d).

Discussion

Bien que des facteurs extrinsèques contrôlent les schémas d'expression des *Hox*, des changements intrinsèques à la cellule, passant d'un état chromatinien répressif à un état chromatinien permissif, se produisent au sein des complexes *Hox* et sont en corrélation avec la séquence temporelle d'induction des *Hox* (Bel-Vialar et al., 2002; Deschamps and Duboule, 2017; Liu et al., 2001; Mazzone et al., 2013; Narendra et al., 2015; Noordermeer et al., 2014). Il reste donc à savoir si l'ouverture progressive de la chromatine le long des complexes sert de minuterie interne actionnée par des signaux extrinsèques ou si des séquences de facteurs scrétés qui activent progressivement des gènes *Hox* plus caudaux définissent le rythme de l'induction. Afin d'aborder cette question, j'ai utilisé des progéniteurs axiaux dérivés de CSPi humains et montré que l'activation progressive des gènes *HOX* dépend des facteurs extrinsèques FGFs qui sont produits par les progéniteurs eux-mêmes et que l'ajout extrinsèque de FGF2 ou de GDF11 accélère l'expression des *HOX* thoraciques et lombaires (Mouilleau et al., 2021). Suite à ces résultats, j'ai étudié les mécanismes moléculaires par lesquels la voie GDF fonctionne pour initier l'expression des gènes *HOX*. Mes données indiquent que 1) GDF11 induit une accumulation nucléaire transitoire du facteur de transcription SMAD2 qui corrèle avec le début de l'induction des gènes *HOX* thoraciques et lombaires, 2) l'induction des *HOX* dépend de la durée de signalisation de la voie SMAD et 3) la dynamique temporelle de la voie SMAD pourrait coder le rythme d'activation colinéaire de ces gènes *HOX*. En parallèle de ces résultats, j'ai également participé au développement d'un nouveau modèle organoïde qui reproduit la formation du tronc humain. Ce modèle se caractérise par la mise en place de domaines d'expression des gènes *HOX* le long de l'axe antéro-postérieur et représente une avancée prometteuse pour mieux comprendre les mécanismes de régulation des gènes *HOX* pendant le développement humain.

Dans l'ensemble, ce travail fournit de nouvelles informations sur le rôle des facteurs extrinsèques dans les progéniteurs axiaux et ouvre de nouvelles perspectives quant aux mécanismes sous-jacents par lesquels ces signaux agissent afin de réguler l'expression des gènes *HOX*. Ci-dessous, je discute de la signification de mes résultats et je suggère des expériences alternatives qui pourraient compléter ces données et répondre à des questions encore non résolues.

- I. Est-ce que les facteurs SMAD2/3 régulent directement l'expression des gènes *HOX* ?

En aval de GDF11, mes données indiquent que l'activation des *HOX* thoraciques et lombaires dépend du récepteur ALK5. Par ailleurs, mes résultats démontrent que l'induction des gènes *HOX* persiste même après l'inhibition de la traduction, mais cette activation ne se produit plus en présence de l'inhibiteur de ALK5, suggérant ainsi que les voies de signalisation en aval de ce récepteur sont directement impliquées dans la régulation des gènes *HOX*. Mes données montrent également que la stimulation des progéniteurs axiaux à GDF11 conduit à l'activation des effecteurs SMAD2/3 qui sont connus pour médier la signalisation ALK5. Pour confirmer ces résultats, il est essentiel de réaliser des expériences complémentaires et d'évaluer s'il y a bien un lien direct entre le recrutement nucléaire de SMAD2/3 et l'induction des gènes *HOX*, ou bien si d'autres voies non-canoniques (e.g. Pi3K, P38, JNK) signalent à travers ALK5 pour réguler l'expression des gènes *HOX* (Clayton et al., 2020).

Pour répondre à cette question, une approche pourrait consister à empêcher l'activité de SMAD2 ou SMAD3 par le biais d'une perte d'expression de ces gènes médiée par dCas9-KRAB et de tester si, dans ces conditions, GDF11 est toujours capable d'induire les gènes *HOX* thoraciques et lombaires (Genga et al., 2019; Gilbert et al., 2013). En effet, si SMAD2 ou SMAD3 sont effectivement nécessaire pour l'induction des gènes *HOX* par GDF11, alors je m'attendrais à l'absence d'induction des *HOX* dans ces conditions (e.g. *HOXA10/C10/D9*). Une approche alternative consisterait à transfecter une forme constitutivement active de SMAD2 ou SMAD3, puis à observer si l'expression de ces protéines, même en l'absence de GDF11, permet d'induire l'expression des gènes *HOX*.

Par ailleurs, ces pertes ou gains de fonction de SMAD2 et/ou SMAD3 permettront de discriminer l'action de ces effecteurs sur l'expression des gènes *HOX*. En effet, si la perte d'expression des gènes *HOX* est plus forte lorsque SMAD2 et SMAD3 sont tous deux non fonctionnels que lorsqu'un seul de ces facteurs est réprimé, cela suggérera que ces derniers coopèrent pour réguler l'expression des gènes *HOX*.

II. Comment les voies de signalisation activées par les facteurs extrinsèques régulent l'expression des gènes *HOX* ?

Les modifications de la chromatine associées avec l'ouverture des complexes *Hox* et leur activation séquentielle a été en partie liée à l'action des facteurs extrinsèques (Deschamps and Duboule, 2017; Mazzoni et al., 2013). Mes données préliminaires suggèrent une possible

liaison de SMAD2/3 dans le complexe *HOXA* et *HOXC* lorsque les progéniteurs axiaux sont exposés 3h au GDF11. L'un de ces enhancers est situé en amont du gène *HOXC8*, et des précédents travaux sur des cellules HepG2 ont montré que la synergie de la voie GDF11 et du facteur CDX2 permet l'activation de ce site (Gaunt, 2017). En parallèle, nos données transcriptomiques indiquent que les progéniteurs axiaux exposés au GDF11 activent l'expression de *HOXA10/C10*. Ces résultats suggèrent donc que la liaison de SMAD2/3 sur les complexes pourrait être nécessaire pour activer l'expression des *HOX* en 5'. Cependant, le lien entre les facteurs extrinsèques, les effecteurs qu'ils activent et les changements transcriptionnels induits reste inconnu.

Pour examiner cet aspect, il serait intéressant d'évaluer si SMAD2/3 joue un rôle actif dans l'ouverture de la chromatine au niveau des complexes *HOX*. En effet, de précédentes études indiquent que les SMADs peuvent agir en tant que facteurs recrutant des acétyltransférases (e.g. p300, CBP) sur certains sites de la chromatine et ainsi favoriser l'activation transcriptionnelle des gènes voisins (Feng et al., 1998; Janknecht et al., 1998; Pouponnot et al., 1998; Ross et al., 2006). Pour déterminer si la liaison de SMAD2/3 aux complexes *HOX* est nécessaire pour permettre l'ouverture de la chromatine, deux approches complémentaires pourraient être envisagées. La première consisterait à induire la perte de fonction des SMADs puis à mesurer l'accessibilité de la chromatine après la stimulation des progéniteurs axiaux au GDF11. Pour ce faire, nous envisageons de faire des expériences ATAC-seq dans les cellules mutantes *SMAD2/3*. Si dans ces cellules il n'y a pas d'ouverture de la chromatine vers l'extrémité 5' des complexes *HOX*, cela suggèrera que les SMAD2/3 agissent probablement en tant que facteurs pionniers. La deuxième approche pour caractériser les mécanismes par lesquels SMAD2/3 pourrait ouvrir la chromatine est de déterminer si la stimulation des progéniteurs avec du GDF11 permet le recrutement de p300 ou CBP sur les complexes *HOX*. Si cela s'avère exact, il sera intéressant d'identifier si les sites de liaison de p300 et CBP coïncident avec les sites de liaison de SMAD2/3, et de déterminer si la perte de fonction de SMAD2/3 empêche le recrutement de p300 et CBP. Cependant si ces deux approches n'indiquent pas une modification de l'ouverture de la chromatine ou le recrutement d'acétyltransférases, cela indiquera que des mécanismes indépendants participent à l'ouverture de la chromatine afin de permettre dans un second temps la liaison des SMADs sur les complexes et l'induction des gènes *HOX*.

Au cours de ma thèse, je me suis également intéressé au facteur extrinsèque FGF2 et montré que celui-ci participe à la régulation des gènes *HOX* brachiaux et thoraciques. En aval

du facteur FGF2, nous avons montré que la régulation de l'expression des gènes *HOX* dépend des récepteurs FGFR1-3 et de la voie MEK1/2-ERK1/2. Ces résultats soulèvent donc plusieurs questions concernant les mécanismes par lesquels ERK1/2 pourrait agir pour réguler l'expression des gènes *HOX*.

Une possibilité serait que les facteurs ERK1/2 et CDX interagissent pour activer les gènes *HOX*. En effet, de précédentes études indiquent que la voie Fgf et le facteur Cdx2 agissent en synergie pour réguler l'expression des gènes *Hox* (Bel-Vialar et al., 2002; Mazzone et al., 2013). De plus, de nombreux sites de liaison Cdx2 ont été retrouvés au sein des complexes allant jusqu'aux paralogues *Hox9* (Amin et al., 2016; Neijts et al., 2017). Enfin, il a été démontré que la phosphorylation du résidu sérine en position 60 de la protéine Cdx2 dépend de l'activité de la voie de signalisation MAPK-ERK1/2 lors de la différenciation des cellules intestinales (Houde et al., 2001; Rings et al., 2001). Dans les progéniteurs axiaux, il est donc possible que les changements d'expression des *HOX* dépendent de la phosphorylation de Cdx2 par la voie MAPK-ERK1/2. Pour tester cette idée, il serait intéressant dans un premier temps de faire un western blot contre phospho-Cdx2 lorsque la voie ERK1/2 est activée ou inhibée dans les progéniteurs axiaux. De plus, la mutagenèse du site de phosphorylation Cdx2-Ser60 en alanine pourrait permettre de maintenir la conformation dé-phosphorylée de Cdx2 et confirmer le rôle fonctionnel de cette protéine pour induire l'expression des gènes *HOX*. Par ailleurs, il est également possible que ERK1/2 participe à la phosphorylation d'autres facteurs de transcription tels que ceux appartenant à la famille ETS (Gille et al., 1995; Waas and Dalby, 2001). Pour déterminer si l'activité de la voie ERK1/2 passe par ces protéines, il serait intéressant de faire un gain ou perte de fonction de ces protéines et voir si l'induction des *HOX* est modifié.

D'autre part, nous avons également envisagé que la voie FGF pourrait activer la signalisation SMAD2/3, comme cela a été démontré dans d'autres systèmes (Aragón et al., 2011; Gao et al., 2009; Millet et al., 2009). Or nous n'avons pas observé de translocation nucléaire de SMAD2/3 après avoir exposé les progéniteurs axiaux à FGF2. Par conséquent, il y aurait des effecteurs terminaux permettant la régulation des *HOX* et qui seraient spécifiques au signal FGF ou GDF. Cependant, nous n'avons pas étudié en détail si la dynamique de la voie SMAD activée par le GDF11 pouvait être modifiée par une co-stimulation de la voie FGF. En effet, il est possible que la voie FGF induise une sur-expression d'activateurs ou de répresseurs de la voie SMAD, ce qui pourrait alors modifier la dynamique du signal initié par le GDF11.

Une telle interaction entre ces voies de signalisation pourrait potentiellement influencer l'induction des gènes *HOX* et mérite d'être explorée plus en profondeur.

III. Est-ce que le rythme d'activation des gènes *HOX* dépend de la dynamique temporelle des voies de signalisation ?

La dynamique des voies de signalisation est définie en fonction de l'amplitude et la durée des signaux (Purvis and Lahav, 2013). Mes résultats montrent que dans les progéniteurs axiaux exposés à GDF11, le profil temporel de la voie SMAD est caractérisé par 1) une phase d'activation au cours de laquelle les effecteurs SMAD2/3 sont transloqués dans le noyau, 2) une phase d'adaptation caractérisée par une diminution du signal nucléaire et 3) une phase de maintenance au cours de laquelle le signal, bien que réduit, persiste dans le temps. En parallèle de l'activation de la voie SMAD2/3, nous avons observé l'induction des gènes *HOX9-10*. Cependant les gènes *HOX11-12* restent réprimés. Suite à ces observations, nous nous sommes demandé si la dynamique temporelle des voies de signalisation et le rythme d'activation des gènes *HOX* étaient associés.

En effet, plusieurs études ont mis en évidence que la dynamique des voies de signalisation pouvait influencer l'expression de gènes cibles, entraînant ainsi diverses réponses cellulaires (Ebisuya and Briscoe, 2018; Purvis and Lahav, 2013). À titre d'exemple, lorsque les cellules PC12 sont stimulées par l'EGF, l'activation de la voie ERK est transitoire et conduit à la prolifération cellulaire (Marshall, 1995). En revanche, lorsque la signalisation ERK est maintenue grâce au facteur NGF, les cellules rentrent en différenciation. Il serait donc plausible que dans les progéniteurs axiaux, la dynamique temporelle de la voie SMAD joue un rôle crucial dans l'activation progressive des gènes *HOX*. L'absence d'adaptation de cette voie entraînerait alors la perte de l'activation séquentielle et colinéaire des gènes *HOX*.

Des études complémentaires ont montré que dans les cellules PC12, l'activation transitoire de la voie ERK par EGF est médiée par des rétrocontrôles négatifs (Sasagawa et al., 2005). Au contraire, le NGF permet l'induction de rétrocontrôles positifs permettant une dynamique forte et maintenue de ERK. Les rétrocontrôles participent donc à l'émergence des durées, des fréquences et des amplitudes des voies afin de permettre à un signal d'initier différents processus biologiques. Dans les progéniteurs axiaux, la régulation des gènes *HOX* pourrait donc être influencée par des rétrocontrôles négatifs qui modulent la dynamique temporelle des voies de signalisation.

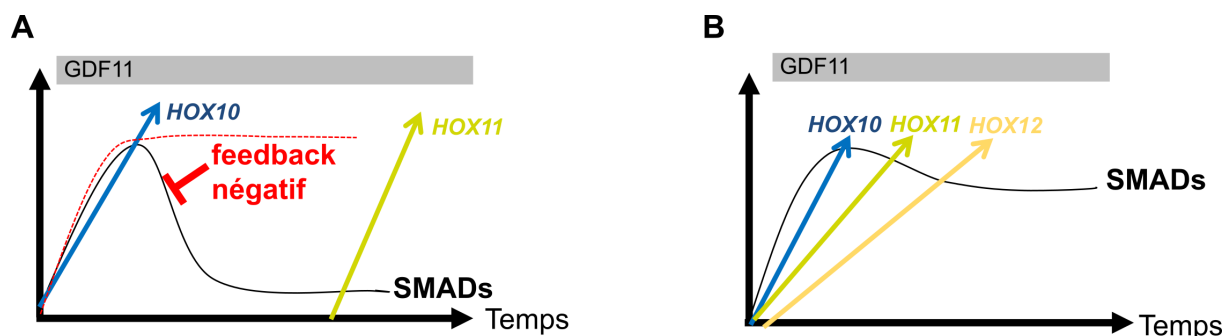


Figure 22 – Illustration d'un modèle théorique mettant en évidence l'importance de la dynamique temporelle des voies de signalisation dans le contrôle du rythme d'activation des gènes *HOX*. (a) L'activation de la voie SMAD induit l'expression des gènes *HOX10* ainsi que de rétrocontrôles négatifs. Ces rétrocontrôles entraînent une adaptation de la voie SMAD, retardant ainsi l'activation des gènes *HOX11*. (b) En absence de rétrocontrôles négatifs, l'activité de la voie SMAD est maintenue dans le temps ce qui entraîne une induction précoce des gènes *HOX11/12*.

Au cours de mes travaux, plusieurs résultats obtenus suggèrent un rôle de la dynamique des voies de signalisation dans la régulation des gènes *HOX*. Tout d'abord mes données transcriptomiques et des immunomarquages indiquent que la voie GDF induit l'expression de nombreux régulateurs négatifs. Ces rétrocontrôles négatifs agissent à différents niveaux de la voie puisque des facteurs réprimant la voie au niveau des ligands de GDF11 (*LEFTY1/2*), des récepteurs ALKs (*BAMBI*, *SMAD7*) et des effecteurs *SMAD2/3* dans le cytoplasme (*PMEPA1*) ou dans le noyau (*SKIL*) ont été identifiés (Chen and Meng, 2004).

De plus, nous avons montré que le blocage de la néo-traduction des cibles de la voie GDF, dont font partie les rétrocontrôles négatifs, entraînent l'induction de nouveau *HOX* situés à l'extrémité 5' des complexes (e.g. *HOX11* et paralogues *13*). Cependant, ces résultats n'ont pas encore permis de démontrer de manière concluante que la dynamique de la voie de signalisation GDF, médiée par les *SMADs*, est effectivement modifiée après le blocage de la néo-traduction. En effet, nous avons caractérisé l'intensité moyenne nucléaire de P-*SMAD2* sur un seul temps et la variabilité des résultats obtenus n'a pas permis de montrer une différence significative. Afin de mieux caractériser le potentiel impact du blocage de la néo-traduction sur la dynamique de la voie, il est nécessaire de caractériser l'amplitude de la voie sur plusieurs temps après le co-traitement des progéniteurs axiaux au GDF11 et à la cycloheximide. Par ailleurs, l'inhibition de la traduction pourrait également avoir un impact indirect sur l'activation des gènes *HOX*. En effet, des études antérieures sur des cellules murines exposées à la cycloheximide ont montré une altération de la structure de la chromatine en phase S, associée à une activation transcriptionnelle générale dépendant de l'ARN polymérase II (Cesari et al., 1998). Or, selon nos résultats transcriptomiques, 472 gènes sont deux fois plus exprimés après le traitement FGF2/GDF11, et 2142 après le traitement FGF2/GDF11/CHX. Si, dans le cas où

la traduction est bloquée, certains de ces gènes correspondent bien à des cibles des voies FGF et GDF, d'autres ne semblent pas être associés à l'une ou l'autre de ces voies. Afin de s'assurer que la cycloheximide n'a pas une activité indirecte sur les complexes *HOX*, il serait donc important d'analyser l'accessibilité chromatinienne des complexes *HOX* et le niveau d'expression de ces gènes dans des CSPi exposés à la cycloheximide. Si, dans ces conditions, la chromatine reste fermée et les gènes *HOX* réprimés, cela suggérerait que la cycloheximide n'a pas d'effet direct sur la régulation des gènes *HOX*. En revanche, si l'ajout de la cycloheximide conduit à l'ouverture de la chromatine dans les CSPi, cela indiquerait un effet indirect dans la régulation des gènes *HOX*. Dans ce cas, cela pourrait signifier l'existence d'un mécanisme indépendant des facteurs extrinsèques qui participe à l'ouverture de la chromatine afin de permettre dans un second temps le recrutement de facteurs de signalisation.

De plus, pour valider correctement et directement l'impact des dynamiques des voies de signalisation sur le rythme d'activation des gènes *HOX*, il serait important d'effectuer des pertes fonctionnelles ciblant spécifiquement les régulateurs négatifs que nous avons identifiés précédemment dans notre transcriptome (e.g. *SKIL*, *LEFTY1/2*, *SMAD7*...). De cette manière, nous serons en mesure de déterminer si les régulateurs négatifs de ces voies contribuent au rythme d'activation des gènes *HOX* en modulant la dynamique des voies de signalisation.

Une autre approche possible pour influencer la dynamique des signaux pourrait consister à activer la voie GDF de manière pulsatile en modulant sa fréquence. De précédentes études ont en effet montré que la modulation de ce paramètre permet d'induire une expression pulsatile des gènes et d'empêcher l'adaptation des voies de signalisation (Sorre et al., 2014). Il serait donc intéressant d'explorer si l'activation séquentielle des gènes *HOX* est altérée lorsqu'on compare un signal maximal (avec une stimulation pulsative sans adaptation) à un signal minimal (avec une stimulation unique où la voie est rapidement réprimée). Cependant, nos données préliminaires semblent indiquer le contraire (données non montrées). En effet, l'expression des gènes *HOX* thoraciques et lombaires semble être plus forte lorsque l'activité des voies de signalisation est prolongée (une stimulation GDF11 maintenue) en comparaison à une activation pulsatile des voies (4 séquences de stimulation de 1h avec du GDF11, suivies de 2h de répression par un inhibiteur de ALK5). Toutefois, il est possible que ce résultat soit lié à des limitations techniques car les changements multiples de milieu de culture dans un intervalle de temps restreint peuvent stresser les cellules. Afin de tester cette idée de manière plus précise, il serait donc nécessaire d'adopter une autre approche technique. Par exemple, il serait

intéressant de développer des lignées de CSPi humaines exprimant des constructions optogénétiques permettant l'activation des effecteurs SMAD2/3 en réponse à la lumière (Toettcher et al., 2013). De plus, il est également possible que le paradigme que nous avons utilisé pour générer ces résultats préliminaires ne soit pas approprié. Pour identifier préalablement le meilleur protocole de stimulation par fréquence permettant d'activer la voie de signalisation tout en évitant son adaptation, il serait judicieux de développer un modèle biophysique basé sur les résultats obtenus précédemment. Grâce à ce modèle, nous pourrions alors tester une multitude de paramètres tels que la concentration, la durée de stimulation ou de répression, et le nombre de stimulations afin d'identifier le signal maximal avec une faible rétroaction et le signal minimal avec une rétroaction élevée.

En plus de l'influence de la dynamique des voies de signalisation, nos résultats suggèrent que la durée de signalisation de la voie SMAD est un paramètre essentiel pour contrôler la régulation des gènes *HOX*. En effet, lorsque les progéniteurs sont exposés à GDF11 sur une courte durée, alors le niveau d'expression des gènes *HOX* est partiel. A l'inverse lorsque le facteur GDF11 est maintenu sur une période plus longue, alors l'expression des *HOX* est plus forte. Pour compléter ces résultats il serait intéressant de déterminer si le moment où les progéniteurs sont exposés aux facteurs extrinsèques est également un paramètre important dans la régulation de l'expression des gènes *HOX*. Une expérience à envisager serait donc d'exposer des progéniteurs axiaux à différents stades de différenciation à GDF11. Si l'ajout de facteurs extrinsèques sur des progéniteurs « vieux » permet l'activation de nouveaux *HOX* situés en 5' des complexes qui ne sont pas activés dans des progéniteurs « jeunes » exposés sur la même durée à GDF11, cela suggérerait que le moment où les progéniteurs sont exposés aux facteurs extrinsèques est un paramètre crucial. Ce résultat suggérerait l'existence d'autres mécanismes complémentaires des facteurs extrinsèques participant à la régulation des *HOX*.

IV. Comment la dynamique temporelle des voies de signalisation est décodée par les gènes *HOX* ?

D'après mes résultats transcriptomique, il y a une augmentation de l'expression de *HOXC9* 3h après traitement au GDF11. L'expression de *HOXC10* est ensuite induite à 6h post-traitement à GDF11 tandis que les gènes *HOX11/12* restent réprimés. De plus, mes données d'hybridation *in situ* semblent corroborer ces résultats, car l'activation de *HOXC10* dans tous les progéniteurs axiaux est observée plus tardivement que celle de *HOXC9*. Ces données

indiquent donc que le GDF11 induit l'expression des gènes *HOXC9* et *HOXC10*, mais que la colinéarité temporelle reste préservée. Il serait donc intéressant de comprendre comment les signaux extrinsèques sont interprétés par les complexes afin de maintenir l'activation séquentielle et colinéaire des gènes *HOX*.

Une possibilité à envisager serait que plusieurs enhancers se trouvent le long des complexes *HOX*, et que ces derniers ont des affinités variables pour les facteurs de transcription en fonction de leur position. Ainsi, les enhancers qui se trouvent en 3' des complexes *HOX* auraient une affinité élevée pour lier les facteurs activés par les voies de signalisation. Une courte activation des voies de signalisation serait alors suffisante pour permettre le recrutement de ces facteurs sur les enhancers en 3', puis d'activer les gènes *HOX* à proximité. En revanche, les enhancers situés en 5' des complexes *HOX* auraient une affinité plus faible pour ces mêmes facteurs de signalisation. Par conséquent, le recrutement de ces facteurs nécessiterait une activation plus longue ou plus intense des voies de signalisation afin que les gènes *HOX* se trouvant à proximité puissent ensuite être activés. Cette idée est en partie suggérée par nos données préliminaires où nous avons mesuré le recrutement de SMAD2/3 dans des progéniteurs axiaux exposés à du GDF11 et à la cycloheximide. Lorsque nous comparons cette condition à celle où les progéniteurs ont été uniquement stimulés par du GDF11, nous observons dans ces deux cas le recrutement de SMAD2/3 entre les gènes *HOXA7* et *HOXA9*. Cependant un nouveau site de recrutement pour SMAD2/3 est observé entre les gènes *HOXA9* et *HOXA10* lorsque les progéniteurs sont co-stimulés par GDF11 et la cycloheximide. Les immunomarquages de P-SMAD2 suggérant une activité plus élevée de la voie dans les progéniteurs stimulés par GDF11/CHX, les résultats de Cut&Tag indiqueraient donc que l'augmentation de l'amplitude de la signalisation pourrait permettre le recrutement de SMAD2/3 sur des enhancers situés plus en 5' des complexes, facilitant ainsi l'induction de nouveaux gènes *HOX*. Cependant le recrutement de SMAD2/3 sur des nouveaux sites en 5' pourrait également être associé à un effet indirect de la cycloheximide qui induirait l'ouverture de la chromatine le long des complexes *HOX* comme nous l'avons mentionné précédemment (Cesari et al., 1998). Dans ce cas, le maintien de l'expression séquentiel et colinéaire des gènes *HOX* après l'ajout de facteur extrinsèque dépendrait de l'ouverture progressive de la chromatine, permettant ainsi le recrutement graduel de facteurs de signalisation de plus en plus en 5' des complexes *HOX*. Pour explorer d'avantage ces différentes hypothèses, il faudrait traiter les progéniteurs avec des concentrations de GDF11 variables afin de moduler l'amplitude

de la voie SMAD2/3, ou bien d'étudier le recrutement de SMAD2/3 et l'accessibilité de la chromatine sur les complexes *HOX* à différents moments post-stimulation.

V. Quelles sont les possibilités offertes par les organoïdes pour étudier les mécanismes de régulation des *HOX* ?

Mes travaux ont également permis de générer un nouveau modèle d'organoïde récapitulant la formation du tronc humain. Ce modèle *in vitro* est caractérisé par la ségrégation spatiale de l'expression des gènes *HOX* le long de l'axe antéro-postérieur. Effectivement, des expériences d'hybridation *in situ* et de séquençage spatial ont démontré que les gènes *HOX* situés en 3' des complexes sont exprimés tout le long de l'axe antéro-postérieur, tandis que les gènes en 5' des complexes présentent une expression plus restreinte aux régions caudales. Ces observations sont similaires à celles faites chez les Vertébrés au cours de l'embryogenèse (Dasen et al., 2005; Gaunt, 1988; Gaunt et al., 1989; Liu et al., 2001; Nishimoto et al., 2014). Lorsque ces organoïdes sont maintenus à des stades tardifs de différenciation, les profils d'expression des gènes *HOX* sont traduits et entraînent la formation d'un continuum de sous-types de motoneurones exprimant des facteurs de transcription HOX distincts le long de l'axe antéro-postérieur. Ce modèle offre donc de nouvelles opportunités afin d'étudier comment la perturbation de signaux régulant l'expression des gènes *HOX* est retranscrite le long de l'axe antéro-postérieur.

À l'avenir, il serait intéressant d'utiliser ces organoïdes pour étudier comment les voies de signalisation peuvent moduler les domaines d'expression des gènes *HOX* ainsi que la spécification des motoneurones le long de l'axe antéro-postérieur. Pour y parvenir, une approche possible consisterait à utiliser une lignée cellulaire dans laquelle l'expression de *SMAD2/3* est réprimée, puis à mesurer les domaines d'expression des gènes *HOX* et les sous-types de motoneurones. Selon nos résultats précédents, la partie la plus caudale des organoïdes où sont localisés les progéniteurs axiaux exprime *HOXA10/C10/D9*. De plus, les motoneurones qui sont organisés le long de l'axe antéro-postérieur au stade tardif de différenciation expriment les facteurs HOXC6 dans la région antérieure et HOXC9 dans la partie postérieure. Par conséquent, il est possible que des organoïdes dérivés de cellules souches pluripotentes présentant une perte de fonction de SMAD2/3 n'expriment plus *HOXA10/C10/D9* dans les progéniteurs axiaux et perdent HOXC9 dans les motoneurones. Une autre expérience possible serait de muter les enhanceurs de SMAD2/3 que nous avons trouvés dans les complexes *HOXA/C*

afin de valider leur rôle fonctionnel dans la mise en place des domaines d'expression des *HOX* le long de l'axe.

De plus, ce nouveau modèle *in vitro* offre une approche unique afin de déterminer si les domaines d'expression des *HOX* le long de l'axe émergent de la dynamique des voies de signalisation. Effectivement, il serait intéressant d'étudier si la suppression de la rétroaction négative des voies de signalisation induit des modèles d'expression des *HOX* plus caudaux ainsi que de nouveaux sous-types de motoneurones spécifiques aux régions lombaires et sacrées. Si tel est le cas, cela pourrait indiquer que la dynamique des voies de signalisation joue un rôle crucial dans la spécification de l'axe antéro-postérieur en permettant la génération des types cellulaires appropriés au bon moment et au bon endroit.

En résumé, ce nouveau modèle d'organoïde permettra à l'avenir de caractériser plus rapidement les mécanismes de régulation des gènes *HOX* impliqués dans la spécification axiale et la diversité cellulaire au cours du développement embryonnaire.

VI. Quel impact pour l'ingénierie cellulaire ?

Grâce à mes travaux de thèse, j'ai pu développer de nouveaux modèles qui contribuent à une meilleure compréhension du rôle des facteurs extrinsèques dans le développement et la diversification de l'axe antéro-postérieur. En plus des apports fondamentaux sur les mécanismes participant au développement embryonnaire, ces travaux ont permis d'améliorer le contrôle de la production des motoneurones humains *in vitro*. Étant donné que les facteurs *HOX* sont également impliqués dans l'acquisition de l'identité positionnelle des cellules dérivées du mésoderme et de l'ectoderme (Carapuço et al., 2005; Deschamps and Van Nes, 2005; Grapin-Botton and Melton, 2000), un meilleur contrôle de leur expression aura des conséquences importantes pour une conception plus rationnelle des organoïdes et donc pour l'ingénierie des tissus humains.

De plus, certaines pathologies telles que les maladies amyotrophiques et Parkinson affectent spécifiquement certains sous-types de neurones. Par exemple, la sclérose latérale amyotrophique est caractérisée par la mort des motoneurones se trouvant dans la colonne motrice latérale, qui sont responsables de l'innervation des membres. (Kanning et al., 2010). La différenciation ciblée à partir de cellules souches pluripotentes provenant de patients permettra de mieux comprendre les mécanismes sous-jacents associées à ces pathologies. Cela permettra

également d'identifier de nouvelles cibles thérapeutiques et de réaliser des criblages à haut débit pour découvrir des molécules à potentiel thérapeutique (Sun et al., 2018).

Au cours de ma thèse, j'ai également contribué au développement d'un nouveau type d'organoïdes qui permet de reproduire les premières phases du développement du tronc humain. En combinant les modèles organoïdes avec les avancées récentes en imagerie, génomique et modélisation physique, nous avons maintenant la possibilité d'acquérir une compréhension inégalée des processus de développement. En effet, ces modèles organoïdes offrent une plateforme expérimentale puissante pour étudier les mécanismes moléculaires, cellulaires et tissulaires impliqués dans le développement humain. De plus, ces modèles organoïdes ont un potentiel thérapeutique prometteur puisqu'en utilisant des CSPi dérivées de patients porteurs de mutations génétiques spécifiques liées à des maladies, nous pouvons recréer *in vitro* les caractéristiques de ces maladies et développer des stratégies thérapeutiques ciblées pour les traiter. Cela ouvre de nouvelles perspectives pour la médecine personnalisée et la recherche de thérapies innovantes. Par ailleurs, le développement de nouveaux organoïdes multicellulaires et de plus en plus complexe laisse entrevoir de nouvelles perspectives dans notre façon d'exercer la science, notamment en ne s'appuyant plus uniquement sur l'utilisation de modèles animaux.

Pour conclure, mon projet a permis de fournir de nouvelles informations sur la manière dont les voies de signalisation encodent des informations afin de contrôler le destin des cellules. Le rôle de la dynamique de signalisation dans l'expansion du répertoire d'encodage des morphogènes est un domaine en émergence, offrant un grand potentiel pour la recherche fondamentale et appliquée. Mon projet prépare donc le terrain techniquement et conceptuellement pour déchiffrer le rôle de la dynamique de signalisation dans le développement en utilisant des modèles *in vitro* humains 3D parfaitement adaptés à l'approche des événements temporels cellulaires rapides.

Bibliographie

Aires, R., De Lemos, L., Nóvoa, A., Jurberg, A.D., Mascrez, B., Duboule, D., Mallo, M., 2019. Tail Bud Progenitor Activity Relies on a Network Comprising Gdf11, Lin28, and Hox13 Genes. *Developmental Cell* 48, 383-395.e8. <https://doi.org/10.1016/j.devcel.2018.12.004>

Aires, R., Jurberg, A.D., Leal, F., Nóvoa, A., Cohn, M.J., Mallo, M., 2016. Oct4 Is a Key Regulator of Vertebrate Trunk Length Diversity. *Developmental Cell* 38, 262–274. <https://doi.org/10.1016/j.devcel.2016.06.021>

Akai, J., Halley, P.A., Storey, K.G., 2005. FGF-dependent Notch signaling maintains the spinal cord stem zone. *Genes Dev.* 19, 2877–2887. <https://doi.org/10.1101/gad.357705>

Akasaka, T., Kanno, M., Balling, R., Mieza, M.A., Taniguchi, M., Koseki, H., 1996. A role for *mel-18*, a Polycomb group-related vertebrate gene, during the anteroposterior specification of the axial skeleton. *Development* 122, 1513–1522. <https://doi.org/10.1242/dev.122.5.1513>

Akasaka, T., Lohuizen, M.V., Lugt, N.V.D., 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. <https://doi.org/10.1242/dev.128.9.1587>

Akker, E.V.D., Fromental-Ramain, C., Graaff, W.D., Le Mouellic, H., Brûlet, P., Chambon, P., Deschamps, J., 2001. Axial skeletal patterning in mice lacking all paralogous group 8 Hox genes. *Development* 128, 1911–1921. <https://doi.org/10.1242/dev.128.10.1911>

Amândio, A.R., Beccari, L., Lopez-Delisle, L., Mascrez, B., Zakany, J., Gitto, S., Duboule, D., 2021. Sequential *in cis* mutagenesis *in vivo* reveals various functions for CTCF sites at the mouse *HoxD* cluster. *Genes Dev.* 35, 1490–1509. <https://doi.org/10.1101/gad.348934.121>

Amin, S., Neijts, R., Simmini, S., Van Rooijen, C., Tan, S.C., Kester, L., Van Oudenaarden, A., Creighton, M.P., Deschamps, J., 2016. Cdx and T Brachyury Co-activate Growth Signaling in the Embryonic Axial Progenitor Niche. *Cell Reports* 17, 3165–3177. <https://doi.org/10.1016/j.celrep.2016.11.069>

Andersson, O., Reissmann, E., Ibáñez, C.F., 2006. Growth differentiation factor 11 signals through the transforming growth factor- β receptor ALK5 to regionalize the anterior–posterior axis. *EMBO Reports* 7, 831–837. <https://doi.org/10.1038/sj.embor.7400752>

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, 1234167. <https://doi.org/10.1126/science.1234167>

Aragón, E., Goerner, N., Zaromytidou, A.-I., Xi, Q., Escobedo, A., Massagué, J., Macias, M.J., 2011. A Smad action turnover switch operated by WW domain readers of a phosphoserine code. *Genes Dev.* 25, 1275–1288. <https://doi.org/10.1101/gad.2060811>

- Aranda, S., Mas, G., Di Croce, L., 2015. Regulation of gene transcription by Polycomb proteins. *Sci. Adv.* 1, e1500737. <https://doi.org/10.1126/sciadv.1500737>
- Attisano, L., Silvestri, C., Izzi, L., Labbé, E., 2001. The transcriptional role of Smads and FAST (FoxH1) in TGF β and activin signalling. *Molecular and Cellular Endocrinology* 180, 3–11. [https://doi.org/10.1016/S0303-7207\(01\)00524-X](https://doi.org/10.1016/S0303-7207(01)00524-X)
- Beccari, L., Moris, N., Girgin, M., Turner, D.A., Baillie-Johnson, P., Cossy, A.-C., Lutolf, M.P., Duboule, D., Arias, A.M., 2018. Multi-axial self-organization properties of mouse embryonic stem cells into gastruloids. *Nature* 562, 272–276. <https://doi.org/10.1038/s41586-018-0578-0>
- Beck, F., 2002. Homeobox genes in gut development. *Gut* 51, 450–454. <https://doi.org/10.1136/gut.51.3.450>
- 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, 5103–5115. <https://doi.org/10.1242/dev.129.22.5103>
- Bertrand, N., Médevielle, F., Pituello, F., 2000. FGF signalling controls the timing of Pax6 activation in the neural tube. *Development* 127, 4837–4843. <https://doi.org/10.1242/dev.127.22.4837>
- Blassberg, R., Patel, H., Watson, T., Gouti, M., Metzis, V., Delás, M.J., Briscoe, J., 2022. Sox2 levels regulate the chromatin occupancy of WNT mediators in epiblast progenitors responsible for vertebrate body formation. *Nat Cell Biol* 24, 633–644. <https://doi.org/10.1038/s41556-022-00910-2>
- Boulet, A.M., Capecchi, M.R., 2012. Signaling by FGF4 and FGF8 is required for axial elongation of the mouse embryo. *Developmental Biology* 371, 235–245. <https://doi.org/10.1016/j.ydbio.2012.08.017>
- Boyer, L.A., Plath, K., Zeitlinger, J., Brambrink, T., Medeiros, L.A., Lee, T.I., Levine, S.S., Wernig, M., Tajonar, A., Ray, M.K., Bell, G.W., Otte, A.P., Vidal, M., Gifford, D.K., Young, R.A., Jaenisch, R., 2006. Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* 441, 349–353. <https://doi.org/10.1038/nature04733>
- Brafman, D., Willert, K., 2017. Wnt/ β -catenin signaling during early vertebrate neural development: Wnt Signaling in Neural Development. *Devel Neurobio* 77, 1239–1259. <https://doi.org/10.1002/dneu.22517>
- Breen, T.R., Harte, P.J., 1993. *trithorax* regulates multiple homeotic genes in the bithorax and Antennapedia complexes and exerts different tissue-specific, parasegment-specific and promoter-specific effects on each. *Development* 117, 119–134. <https://doi.org/10.1242/dev.117.1.119>

- Brown, J.M., Storey, K.G., 2000. A region of the vertebrate neural plate in which neighbouring cells can adopt neural or epidermal fates. *Current Biology* 10, 869–872. [https://doi.org/10.1016/S0960-9822\(00\)00601-1](https://doi.org/10.1016/S0960-9822(00)00601-1)
- Busby, L., Nájera, G.S., Steventon, B., 2023. A population intrinsic timer controls *Hox* gene expression and cell dispersion during progenitor addition to the body axis (preprint). *Developmental Biology*. <https://doi.org/10.1101/2023.05.10.540133>
- Byrd, K.N., Shearn, A., 2003. ASH1, a *Drosophila* trithorax group protein, is required for methylation of lysine 4 residues on histone H3. *Proc. Natl. Acad. Sci. U.S.A.* 100, 11535–11540. <https://doi.org/10.1073/pnas.1933593100>
- Cadigan, K.M., Waterman, M.L., 2012. TCF/LEFs and Wnt Signaling in the Nucleus. *Cold Spring Harbor Perspectives in Biology* 4, a007906–a007906. <https://doi.org/10.1101/cshperspect.a007906>
- Cambray, N., Wilson, V., 2007. Two distinct sources for a population of maturing axial progenitors. *Development* 134, 2829–2840. <https://doi.org/10.1242/dev.02877>
- Cambray, N., Wilson, V., 2002. Axial progenitors with extensive potency are localised to the mouse chordoneural hinge. *Development* 129, 4855–4866. <https://doi.org/10.1242/dev.129.20.4855>
- Cao, K., Collings, C.K., Marshall, S.A., Morgan, M.A., Rendleman, E.J., Wang, L., Sze, C.C., Sun, T., Bartom, E.T., Shilatifard, A., 2017. SET1A/COMPASS and shadow enhancers in the regulation of homeotic gene expression. *Genes Dev.* 31, 787–801. <https://doi.org/10.1101/gad.294744.116>
- Carapuço, M., Nóvoa, A., Bobola, N., Mallo, M., 2005. *Hox* genes specify vertebral types in the presomitic mesoderm. *Genes Dev.* 19, 2116–2121. <https://doi.org/10.1101/gad.338705>
- Carpenter, E.M., 2002. *Hox* Genes and Spinal Cord Development. *Dev Neurosci* 24, 24–34. <https://doi.org/10.1159/000064943>
- Cesari, M., Héliot, L., Meplan, C., Pabion, M., Khochbin, S., 1998. S-phase-dependent action of cycloheximide in relieving chromatin-mediated general transcriptional repression. *Biochemical Journal* 336, 619–624. <https://doi.org/10.1042/bj3360619>
- Chambeyron, S., Bickmore, W.A., 2004. Chromatin decondensation and nuclear reorganization of the *HoxB* locus upon induction of transcription. *Genes Dev.* 18, 1119–1130. <https://doi.org/10.1101/gad.292104>
- Chawengsaksophak, K., James, R., Hammond, V.E., Köntgen, F., Beck, F., 1997. Homeosis and intestinal tumours in *Cdx2* mutant mice. *Nature* 386, 84–87. <https://doi.org/10.1038/386084a0>
- Chen, Y.G., Meng, A.M., 2004. Negative regulation of TGF- β signaling in development. *Cell Res* 14, 441–449. <https://doi.org/10.1038/sj.cr.7290246>

- Clayton, S.W., Ban, G.I., Liu, C., Serra, R., 2020. Canonical and noncanonical TGF- β signaling regulate fibrous tissue differentiation in the axial skeleton. *Sci Rep* 10, 21364. <https://doi.org/10.1038/s41598-020-78206-4>
- Core, N., Bel, S., Gaunt, S.J., Aurrand-Lions, M., Pearce, J., Fisher, A., Djabali, M., 1997. Altered cellular proliferation and mesoderm patterning in Polycomb-M33-deficient mice. *Development* 124, 721–729. <https://doi.org/10.1242/dev.124.3.721>
- Cunningham, T.J., Kumar, S., Yamaguchi, T.P., Duester, G., 2015. *Wnt8a* and *Wnt3a* cooperate in the axial stem cell niche to promote mammalian body axis extension: Role of *Wnt8a* During Body Axis Extension. *Dev. Dyn.* 244, 797–807. <https://doi.org/10.1002/dvdy.24275>
- Dasen, J.S., Liu, J.-P., Jessell, T.M., 2003. Motor neuron columnar fate imposed by sequential phases of Hox-c activity. *Nature* 425, 926–933. <https://doi.org/10.1038/nature02051>
- Dasen, J.S., Tice, B.C., Brenner-Morton, S., Jessell, T.M., 2005. A Hox Regulatory Network Establishes Motor Neuron Pool Identity and Target-Muscle Connectivity. *Cell* 123, 477–491. <https://doi.org/10.1016/j.cell.2005.09.009>
- Del Álamo, D., Rouault, H., Schweisguth, F., 2011. Mechanism and Significance of cis-Inhibition in Notch Signalling. *Current Biology* 21, R40–R47. <https://doi.org/10.1016/j.cub.2010.10.034>
- Del Corral, R.D., Olivera-Martinez, I., Goriely, A., Gale, E., Maden, M., Storey, K., 2003. Opposing FGF and Retinoid Pathways Control Ventral Neural Pattern, Neuronal Differentiation, and Segmentation during Body Axis Extension. *Neuron* 40, 65–79. [https://doi.org/10.1016/S0896-6273\(03\)00565-8](https://doi.org/10.1016/S0896-6273(03)00565-8)
- Delfino-Machín, M., Lunn, J.S., Breitkreuz, D.N., Akai, J., Storey, K.G., 2005. Specification and maintenance of the spinal cord stem zone. *Development* 132, 4273–4283. <https://doi.org/10.1242/dev.02009>
- Denans, N., Imura, T., Pourquié, O., 2015. Hox genes control vertebrate body elongation by collinear Wnt repression. *eLife* 4, e04379. <https://doi.org/10.7554/eLife.04379>
- Deschamps, J., Duboule, D., 2017. Embryonic timing, axial stem cells, chromatin dynamics, and the Hox clock. *Genes Dev.* 31, 1406–1416. <https://doi.org/10.1101/gad.303123.117>
- Deschamps, J., Van Nes, J., 2005. Developmental regulation of the Hox genes during axial morphogenesis in the mouse. *Development* 132, 2931–2942. <https://doi.org/10.1242/dev.01897>
- Deschamps, J., Wijgerde, M., 1993. Two Phases in the Establishment of HOX Expression Domains. *Developmental Biology* 156, 473–480. <https://doi.org/10.1006/dbio.1993.1093>
- Diez Del Corral, R., Morales, A.V., 2017. The Multiple Roles of FGF Signaling in the Developing Spinal Cord. *Front. Cell Dev. Biol.* 5, 58. <https://doi.org/10.3389/fcell.2017.00058>

- Doetschman, T.C., Eistetter, H., Katz, M., Schmidt, W., Kemler, R., 1985. The *in vitro* development of blastocyst-derived embryonic stem cell lines: formation of visceral yolk sac, blood islands and myocardium. *J Embryol Exp Morphol* 87, 27–45.
- Duboule, D., Dollé, P., 1989. The structural and functional organization of the murine HOX gene family resembles that of *Drosophila* homeotic genes. *The EMBO Journal* 8, 1497–1505. <https://doi.org/10.1002/j.1460-2075.1989.tb03534.x>
- Duboule, D., Morata, G., 1994. Colinearity and functional hierarchy among genes of the homeotic complexes. *Trends in Genetics* 10, 358–364. [https://doi.org/10.1016/0168-9525\(94\)90132-5](https://doi.org/10.1016/0168-9525(94)90132-5)
- Dunwoodie, S.L., Clements, M., Sparrow, D.B., Sa, X., Conlon, R.A., Beddington, R.S.P., 2002. Axial skeletal defects caused by mutation in the spondylocostal dysplasia/pudgy gene *Dll3* are associated with disruption of the segmentation clock within the presomitic mesoderm. *Development* 129, 1795–1806. <https://doi.org/10.1242/dev.129.7.1795>
- Dupé, V., Davenne, M., Brocard, J., Dollé, P., Mark, M., Dierich, A., Chambon, P., Rijli, F.M., 1997. *In vivo* functional analysis of the *Hoxa -1* 3' retinoic acid response element (3' RARE). *Development* 124, 399–410. <https://doi.org/10.1242/dev.124.2.399>
- Ebisuya, M., Briscoe, J., 2018. What does time mean in development? *Development* 145, dev164368. <https://doi.org/10.1242/dev.164368>
- Echevarria, D., Martinez, S., Marques, S., Lucas-Teixeira, V., Belo, J.A., 2005. *Mkp3* is a negative feedback modulator of *Fgf8* signaling in the mammalian isthmic organizer. *Developmental Biology* 277, 114–128. <https://doi.org/10.1016/j.ydbio.2004.09.011>
- Edri, S., Hayward, P., Jawaid, W., Martinez Arias, A., 2019. NeuroMesodermal Progenitors (NMPs): a comparative study between Pluripotent Stem Cells and Embryo derived populations. *Development* dev.180190. <https://doi.org/10.1242/dev.180190>
- Eskeland, R., Leeb, M., Grimes, G.R., Kress, C., Boyle, S., Sproul, D., Gilbert, N., Fan, Y., Skoultschi, A.I., Wutz, A., Bickmore, W.A., 2010. *Ring1B* Compacts Chromatin Structure and Represses Gene Expression Independent of Histone Ubiquitination. *Molecular Cell* 38, 452–464. <https://doi.org/10.1016/j.molcel.2010.02.032>
- Essalmani, R., Zaid, A., Marcinkiewicz, J., Chamberland, A., Pasquato, A., Seidah, N.G., Prat, A., 2008. *In vivo* functions of the proprotein convertase PC5/6 during mouse development: *Gdf11* is a likely substrate. *Proc. Natl. Acad. Sci. U.S.A.* 105, 5750–5755. <https://doi.org/10.1073/pnas.0709428105>
- Evans, A.L., Faial, T., Gilchrist, M.J., Down, T., Vallier, L., Pedersen, R.A., Wardle, F.C., Smith, J.C., 2012. Genomic Targets of Brachyury (T) in Differentiating Mouse Embryonic Stem Cells. *PLoS ONE* 7, e33346. <https://doi.org/10.1371/journal.pone.0033346>
- Evans, M.J., Kaufman, M.H., 1981. Establishment in culture of pluripotential cells from mouse embryos. *Nature* 292, 154–156. <https://doi.org/10.1038/292154a0>

- Feng, X.-H., Zhang, Y., Wu, R.-Y., Derynck, R., 1998. The tumor suppressor Smad4/DPC4 and transcriptional adaptor CBP/p300 are coactivators for Smad3 in TGF- β -induced transcriptional activation. *Genes Dev.* 12, 2153–2163. <https://doi.org/10.1101/gad.12.14.2153>
- Ferraiuolo, M.A., Rousseau, M., Miyamoto, C., Shenker, S., Wang, X.Q.D., Nadler, M., Blanchette, M., Dostie, J., 2010. The three-dimensional architecture of Hox cluster silencing. *Nucleic Acids Research* 38, 7472–7484. <https://doi.org/10.1093/nar/gkq644>
- Forlani, S., Lawson, K.A., Deschamps, J., 2003. Acquisition of Hox codes during gastrulation and axial elongation in the mouse embryo. *Development* 130, 3807–3819. <https://doi.org/10.1242/dev.00573>
- Frith, T.J., Granata, I., Wind, M., Stout, E., Thompson, O., Neumann, K., Stavish, D., Heath, P.R., Ortmann, D., Hackland, J.O., Anastassiadis, K., Gouti, M., Briscoe, J., Wilson, V., Johnson, S.L., Placzek, M., Guarracino, M.R., Andrews, P.W., Tsakiridis, A., 2018. Human axial progenitors generate trunk neural crest cells in vitro. *eLife* 7, e35786. <https://doi.org/10.7554/eLife.35786>
- Gao, S., Alarcón, C., Sapkota, G., Rahman, S., Chen, P.-Y., Goerner, N., Macias, M.J., Erdjument-Bromage, H., Tempst, P., Massagué, J., 2009. Ubiquitin Ligase Nedd4L Targets Activated Smad2/3 to Limit TGF- β Signaling. *Molecular Cell* 36, 457–468. <https://doi.org/10.1016/j.molcel.2009.09.043>
- Garriock, R.J., Chalamalasetty, R.B., Kennedy, M.W., Canizales, L.C., Lewandoski, M., Yamaguchi, T.P., 2015. Lineage tracing of neuromesodermal progenitors reveals novel Wnt-dependent roles in trunk progenitor cell maintenance and differentiation. *Development* 142, 1628–1638. <https://doi.org/10.1242/dev.111922>
- Gaunt, S.J., 2017. Gdf11/Smad signalling and Cdx proteins cooperate to activate the Hoxc8 early enhancer in HepG2 cells. *Int. J. Dev. Biol.* 61, 427–432. <https://doi.org/10.1387/ijdb.170066sg>
- Gaunt, S.J., 1988. Mouse homeobox gene transcripts occupy different but overlapping domains in embryonic germ layers and organs: a comparison of *Hox-3.1* and *Hox-1.5*. *Development* 103, 135–144. <https://doi.org/10.1242/dev.103.1.135>
- Gaunt, S.J., George, M., Paul, Y.-L., 2013. Direct activation of a mouse Hoxd11 axial expression enhancer by Gdf11/Smad signalling. *Developmental Biology* 383, 52–60. <https://doi.org/10.1016/j.ydbio.2013.08.025>
- Gaunt, S.J., 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, 131–141. <https://doi.org/10.1242/dev.107.1.131>
- Gaunt, S.J., Strachan, L., 1996. Temporal colinearity in expression of anterior hox genes in developing chick embryos. *Dev. Dyn.* 207, 270–280. [https://doi.org/10.1002/\(SICI\)1097-0177\(199611\)207:3<270::AID-AJA4>3.0.CO;2-E](https://doi.org/10.1002/(SICI)1097-0177(199611)207:3<270::AID-AJA4>3.0.CO;2-E)

Gaunt, S.J., Strachan, L., 1994. Forward spreading in the establishment of a vertebrate Hox expression boundary: The expression domain separates into anterior and posterior zones, and the spread occurs across implanted glass barriers. *Dev. Dyn.* 199, 229–240. <https://doi.org/10.1002/aja.1001990307>

Gavalas, A., Studer, M., Lumsden, A., Rijli, F.M., Krumlauf, R., Chambon, P., 1998. *Hoxa1* and *Hoxb1* synergize in patterning the hindbrain, cranial nerves and second pharyngeal arch. *Development* 125, 1123–1136. <https://doi.org/10.1242/dev.125.6.1123>

Genga, R.M.J., Kernfeld, E.M., Parsi, K.M., Parsons, T.J., Ziller, M.J., Maehr, R., 2019. Single-Cell RNA-Sequencing-Based CRISPRi Screening Resolves Molecular Drivers of Early Human Endoderm Development. *Cell Reports* 27, 708–718.e10. <https://doi.org/10.1016/j.celrep.2019.03.076>

Gérard, M., Duboule, D., Zákány, J., 1993. Structure and activity of regulatory elements involved in the activation of the *Hoxd-11* gene during late gastrulation. *The EMBO Journal* 12, 3539–3550. <https://doi.org/10.1002/j.1460-2075.1993.tb06028.x>

Gilbert, L.A., Larson, M.H., Morsut, L., Liu, Z., Brar, G.A., Torres, S.E., Stern-Ginossar, N., Brandman, O., Whitehead, E.H., Doudna, J.A., Lim, W.A., Weissman, J.S., Qi, L.S., 2013. CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes. *Cell* 154, 442–451. <https://doi.org/10.1016/j.cell.2013.06.044>

Gille, H., Kortenjann, M., Thomae, O., Moomaw, C., Slaughter, C., Cobb, M.H., Shaw, P.E., 1995. ERK phosphorylation potentiates Elk-1-mediated ternary complex formation and transactivation. *The EMBO Journal* 14, 951–962. <https://doi.org/10.1002/j.1460-2075.1995.tb07076.x>

Gillespie, R.F., Gudas, L.J., 2007. Retinoid Regulated Association of Transcriptional Co-regulators and the Polycomb Group Protein SUZ12 with the Retinoic Acid Response Elements of *Hoxa1*, *RARβ2*, and *Cyp26A1* in F9 Embryonal Carcinoma Cells. *Journal of Molecular Biology* 372, 298–316. <https://doi.org/10.1016/j.jmb.2007.06.079>

Goto, H., Kimmey, S.C., Row, R.H., Matus, D.Q., Martin, B.L., 2017. FGF and canonical Wnt signaling cooperate to induce paraxial mesoderm from tailbud neuromesodermal progenitors through regulation of a two-step EMT. *Development* dev.143578. <https://doi.org/10.1242/dev.143578>

Gouti, M., Delile, J., Stamatakis, D., Wymeersch, F.J., Huang, Y., Kleinjung, J., Wilson, V., Briscoe, J., 2017. A Gene Regulatory Network Balances Neural and Mesoderm Specification during Vertebrate Trunk Development. *Developmental Cell* 41, 243–261.e7. <https://doi.org/10.1016/j.devcel.2017.04.002>

Gouti, M., Metzis, V., Briscoe, J., 2015. The route to spinal cord cell types: a tale of signals and switches. *Trends in Genetics* 31, 282–289. <https://doi.org/10.1016/j.tig.2015.03.001>

Gouti, M., Tsakiridis, A., Wymeersch, F.J., Huang, Y., Kleinjung, J., Wilson, V., Briscoe, J., 2014. In Vitro Generation of Neuromesodermal Progenitors Reveals Distinct Roles for Wnt

Signalling in the Specification of Spinal Cord and Paraxial Mesoderm Identity. *PLoS Biol* 12, e1001937. <https://doi.org/10.1371/journal.pbio.1001937>

Graham, A., Papalopulu, N., Krumlauf, R., 1989. The murine and *Drosophila* homeobox gene complexes have common features of organization and expression. *Cell* 57, 367–378. [https://doi.org/10.1016/0092-8674\(89\)90912-4](https://doi.org/10.1016/0092-8674(89)90912-4)

Grapin-Botton, A., Melton, D.A., 2000. Endoderm development: from patterning to organogenesis. *Trends in Genetics* 16, 124–130. [https://doi.org/10.1016/S0168-9525\(99\)01957-5](https://doi.org/10.1016/S0168-9525(99)01957-5)

Guillot, C., Djeflal, Y., Michaut, A., Rabe, B., Pourquié, O., 2021. Dynamics of primitive streak regression controls the fate of neuromesodermal progenitors in the chicken embryo. *eLife* 10, e64819. <https://doi.org/10.7554/eLife.64819>

Guo, Y., Xu, Q., Canzio, D., Shou, J., Li, J., Gorkin, D.U., Jung, I., Wu, H., Zhai, Y., Tang, Y., Lu, Y., Wu, Y., Jia, Z., Li, W., Zhang, M.Q., Ren, B., Krainer, A.R., Maniatis, T., Wu, Q., 2015. CRISPR Inversion of CTCF Sites Alters Genome Topology and Enhancer/Promoter Function. *Cell* 162, 900–910. <https://doi.org/10.1016/j.cell.2015.07.038>

Henrique, D., Abranches, E., Verrier, L., Storey, K.G., 2015. Neuromesodermal progenitors and the making of the spinal cord. *Development* 142, 2864–2875. <https://doi.org/10.1242/dev.119768>

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

Houde, M., Laprise, P., Jean, D., Blais, M., Asselin, C., Rivard, N., 2001. Intestinal Epithelial Cell Differentiation Involves Activation of p38 Mitogen-activated Protein Kinase That Regulates the Homeobox Transcription Factor CDX2. *Journal of Biological Chemistry* 276, 21885–21894. <https://doi.org/10.1074/jbc.M100236200>

Iimura, T., Denans, N., Pourquié, O., 2009. Chapter 7 Establishment of Hox Vertebral Identities in the Embryonic Spine Precursors, in: *Current Topics in Developmental Biology*. Elsevier, pp. 201–234. [https://doi.org/10.1016/S0070-2153\(09\)88007-1](https://doi.org/10.1016/S0070-2153(09)88007-1)

Iimura, T., Pourquié, O., 2006. Collinear activation of *Hoxb* genes during gastrulation is linked to mesoderm cell ingression. *Nature* 442, 568–571. <https://doi.org/10.1038/nature04838>

Iimura, T., Yang, X., Weijer, C.J., Pourquié, O., 2007. Dual mode of paraxial mesoderm formation during chick gastrulation. *Proc. Natl. Acad. Sci. U.S.A.* 104, 2744–2749. <https://doi.org/10.1073/pnas.0610997104>

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. *Mechanisms of Development* 103, 27–33. [https://doi.org/10.1016/S0925-4773\(01\)00338-0](https://doi.org/10.1016/S0925-4773(01)00338-0)

- Izpisúa-Belmonte, J.C., 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. *The EMBO Journal* 10, 2279–2289. <https://doi.org/10.1002/j.1460-2075.1991.tb07764.x>
- Janknecht, R., Wells, N.J., Hunter, T., 1998. TGF- β -stimulated cooperation of Smad proteins with the coactivators CBP/p300. *Genes Dev.* 12, 2114–2119. <https://doi.org/10.1101/gad.12.14.2114>
- Jegalian, B.G., De Robertis, E.M., 1992. Homeotic transformations in the mouse induced by overexpression of a human Hox3.3 transgene. *Cell* 71, 901–910. [https://doi.org/10.1016/0092-8674\(92\)90387-R](https://doi.org/10.1016/0092-8674(92)90387-R)
- Juan, A.H., Ruddle, F.H., 2003. Enhancer timing of Hox gene expression: deletion of the endogenous *Hoxc8* early enhancer. *Development* 130, 4823–4834. <https://doi.org/10.1242/dev.00672>
- Jung, H., Lacombe, J., Mazzoni, E.O., Liem, K.F., Grinstein, J., Mahony, S., Mukhopadhyay, D., Gifford, D.K., Young, R.A., Anderson, K.V., Wichterle, H., Dasen, J.S., 2010. Global Control of Motor Neuron Topography Mediated by the Repressive Actions of a Single Hox Gene. *Neuron* 67, 781–796. <https://doi.org/10.1016/j.neuron.2010.08.008>
- Jurberg, A.D., Aires, R., Varela-Lasheras, I., Nóvoa, A., Mallo, M., 2013. Switching Axial Progenitors from Producing Trunk to Tail Tissues in Vertebrate Embryos. *Developmental Cell* 25, 451–462. <https://doi.org/10.1016/j.devcel.2013.05.009>
- Jürgens, G., 1985. A group of genes controlling the spatial expression of the bithorax complex in *Drosophila*. *Nature* 316, 153–155. <https://doi.org/10.1038/316153a0>
- Kanning, K.C., Kaplan, A., Henderson, C.E., 2010. Motor Neuron Diversity in Development and Disease. *Annu. Rev. Neurosci.* 33, 409–440. <https://doi.org/10.1146/annurev.neuro.051508.135722>
- Karabagli, H., Karabagli, P., Ladher, R., Schoenwolf, G., 2002. Comparison of the expression patterns of several fibroblast growth factors during chick gastrulation and neurulation. *Anatomy and Embryology* 205, 365–370. <https://doi.org/10.1007/s00429-002-0264-7>
- Keynes, R., Krumlauf, R., 1994. Hox Genes and Regionalization of the Nervous System. *Annu. Rev. Neurosci.* 17, 109–132. <https://doi.org/10.1146/annurev.ne.17.030194.000545>
- Kmita, M., Duboule, D., 2003. Organizing Axes in Time and Space; 25 Years of Colinear Tinkering. *Science* 301, 331–333. <https://doi.org/10.1126/science.1085753>
- Kmita, M., Fraudeau, N., Héroult, Y., Duboule, D., 2002. Serial deletions and duplications suggest a mechanism for the collinearity of Hoxd genes in limbs. *Nature* 420, 145–150. <https://doi.org/10.1038/nature01189>

- Koch, F., Scholze, M., Wittler, L., Schifferl, D., Sudheer, S., Grote, P., Timmermann, B., Macura, K., Herrmann, B.G., 2017. Antagonistic Activities of Sox2 and Brachyury Control the Fate Choice of Neuro-Mesodermal Progenitors. *Developmental Cell* 42, 514-526.e7. <https://doi.org/10.1016/j.devcel.2017.07.021>
- Kondoh, H., Takemoto, T., 2012. Axial stem cells deriving both posterior neural and mesodermal tissues during gastrulation. *Current Opinion in Genetics & Development* 22, 374–380. <https://doi.org/10.1016/j.gde.2012.03.006>
- Krumlauf, R., 1994. Hox genes in vertebrate development. *Cell* 78, 191–201. [https://doi.org/10.1016/0092-8674\(94\)90290-9](https://doi.org/10.1016/0092-8674(94)90290-9)
- Kumar, S., Duester, G., 2014. Retinoic acid controls body axis extension by directly repressing *Fgf8* transcription. *Development* 141, 2972–2977. <https://doi.org/10.1242/dev.112367>
- Lacombe, J., Hanley, O., Jung, H., Philippidou, P., Surmeli, G., Grinstein, J., Dasen, J.S., 2013. Genetic and Functional Modularity of Hox Activities in the Specification of Limb-Innervating Motor Neurons. *PLoS Genet* 9, e1003184. <https://doi.org/10.1371/journal.pgen.1003184>
- Langston, A.W., Gudas, L.J., 1992. Identification of a retinoic acid responsive enhancer 3' of the murine homeobox gene Hox-1.6. *Mechanisms of Development* 38, 217–227. [https://doi.org/10.1016/0925-4773\(92\)90055-O](https://doi.org/10.1016/0925-4773(92)90055-O)
- Langston, A.W., Thompson, J.R., Gudas, L.J., 1997. Retinoic Acid-responsive Enhancers Located 3' of the Hox A and Hox B Homeobox Gene Clusters. *Journal of Biological Chemistry* 272, 2167–2175. <https://doi.org/10.1074/jbc.272.4.2167>
- Lawson, K.A., Meneses, J.J., Pedersen, R.A., 1991. Clonal analysis of epiblast fate during germ layer formation in the mouse embryo. *Development* 113, 891–911. <https://doi.org/10.1242/dev.113.3.891>
- Lee, Y.J., McPherron, A., Choe, S., Sakai, Y., Chandraratna, R.A., Lee, S.-J., Oh, S.P., 2010. Growth differentiation factor 11 signaling controls retinoic acid activity for axial vertebral development. *Developmental Biology* 347, 195–203. <https://doi.org/10.1016/j.ydbio.2010.08.022>
- Lewis, E.B., 1998. The bithorax complex: the first fifty years. *Int J Dev Biol* 42, 403–415.
- Lewis, E.B., 1978. A gene complex controlling segmentation in *Drosophila*. *Nature* 276, 565–570. <https://doi.org/10.1038/276565a0>
- Li, L., Liu, B., Wapinski, O.L., Tsai, M.-C., Qu, K., Zhang, J., Carlson, J.C., Lin, M., Fang, F., Gupta, R.A., Helms, J.A., Chang, H.Y., 2013. Targeted Disruption of Hotair Leads to Homeotic Transformation and Gene Derepression. *Cell Reports* 5, 3–12. <https://doi.org/10.1016/j.celrep.2013.09.003>
- Libby, A.R.G., Joy, D.A., Elder, N.H., Bulger, E.A., Krakora, M.Z., Gaylord, E.A., Mendoza-Camacho, F., Butts, J.C., McDevitt, T.C., 2021. Axial elongation of caudalized human

organoids mimics aspects of neural tube development. *Development* 148, dev198275. <https://doi.org/10.1242/dev.198275>

Lippmann, E.S., Williams, C.E., Ruhl, D.A., Estevez-Silva, M.C., Chapman, E.R., Coon, J.J., Ashton, R.S., 2015. Deterministic HOX Patterning in Human Pluripotent Stem Cell-Derived Neuroectoderm. *Stem Cell Reports* 4, 632–644. <https://doi.org/10.1016/j.stemcr.2015.02.018>

Liu, J.-P., 2006. The function of growth/differentiation factor 11 (Gdf11) in rostrocaudal patterning of the developing spinal cord. *Development* 133, 2865–2874. <https://doi.org/10.1242/dev.02478>

Liu, J.-P., Laufer, E., Jessell, T.M., 2001. Assigning the Positional Identity of Spinal Motor Neurons. *Neuron* 32, 997–1012. [https://doi.org/10.1016/S0896-6273\(01\)00544-X](https://doi.org/10.1016/S0896-6273(01)00544-X)

Liu, P., Wakamiya, M., Shea, M.J., Albrecht, U., Behringer, R.R., Bradley, A., 1999. Requirement for Wnt3 in vertebrate axis formation. *Nat Genet* 22, 361–365. <https://doi.org/10.1038/11932>

Lorente, M.D.M., Marcos-Gutiérrez, C., Pérez, C., Schoorlemmer, J., Ramírezp, A., Magin, T., Vidal, M., 2000. Loss- and gain-of-function mutations show a Polycomb group function for Ring1A in mice. *Development* 127, 5093–5100. <https://doi.org/10.1242/dev.127.23.5093>

Luis De La Pompa, J., Wakeham, A., Correia, K.M., Samper, E., Brown, S., Aguilera, R.J., Nakano, T., Honjo, T., Mak, T.W., Rossant, J., Conlon, R.A., 1997. Conservation of the Notch signalling pathway in mammalian neurogenesis. *Development* 124, 1139–1148. <https://doi.org/10.1242/dev.124.6.1139>

Lunn, J.S., Fishwick, K.J., Halley, P.A., Storey, K.G., 2007. A spatial and temporal map of FGF/Erk1/2 activity and response repertoires in the early chick embryo. *Developmental Biology* 302, 536–552. <https://doi.org/10.1016/j.ydbio.2006.10.014>

Mallo, M., Vinagre, T., Carapuo, M., 2009. The road to the vertebral formula. *Int. J. Dev. Biol.* 53, 1469–1481. <https://doi.org/10.1387/ijdb.072276mm>

Margueron, R., Li, G., Sarma, K., Blais, A., Zavadil, J., Woodcock, C.L., Dynlacht, B.D., Reinberg, D., 2008. Ezh1 and Ezh2 Maintain Repressive Chromatin through Different Mechanisms. *Molecular Cell* 32, 503–518. <https://doi.org/10.1016/j.molcel.2008.11.004>

Mariani, L., Guo, X., Menezes, N.A., Drozd, A.M., Çakal, S.D., Wang, Q., Ferretti, E., 2021. A TALE/HOX code unlocks WNT signalling response towards paraxial mesoderm. *Nat Commun* 12, 5136. <https://doi.org/10.1038/s41467-021-25370-4>

Mark, M., Rijli, F.M., Chambon, P., 1997. Homeobox Genes in Embryogenesis and Pathogenesis. *Pediatr Res* 42, 421–429. <https://doi.org/10.1203/00006450-199710000-00001>

Marshall, C.J., 1995. Specificity of receptor tyrosine kinase signaling: Transient versus sustained extracellular signal-regulated kinase activation. *Cell* 80, 179–185. [https://doi.org/10.1016/0092-8674\(95\)90401-8](https://doi.org/10.1016/0092-8674(95)90401-8)

- Martin, B.L., Kimelman, D., 2012. Canonical Wnt Signaling Dynamically Controls Multiple Stem Cell Fate Decisions during Vertebrate Body Formation. *Developmental Cell* 22, 223–232. <https://doi.org/10.1016/j.devcel.2011.11.001>
- Massagué, J., Seoane, J., Wotton, D., 2005. Smad transcription factors. *Genes Dev.* 19, 2783–2810. <https://doi.org/10.1101/gad.1350705>
- Mathis, L., François Nicolas, J., 2000. Different clonal dispersion in the rostral and caudal mouse central nervous system. *Development* 127, 1277–1290. <https://doi.org/10.1242/dev.127.6.1277>
- Maury, Y., Côme, J., Piskorowski, R.A., Salah-Mohellibi, N., Chevaleyre, V., Peschanski, M., Martinat, C., Nedelec, S., 2015. Combinatorial analysis of developmental cues efficiently converts human pluripotent stem cells into multiple neuronal subtypes. *Nat Biotechnol* 33, 89–96. <https://doi.org/10.1038/nbt.3049>
- Mazzoni, E.O., Mahony, S., Peljto, M., Patel, T., Thornton, S.R., McCuine, S., Reeder, C., Boyer, L.A., Young, R.A., Gifford, D.K., Wichterle, H., 2013. Saltatory remodeling of Hox chromatin in response to rostrocaudal patterning signals. *Nat Neurosci* 16, 1191–1198. <https://doi.org/10.1038/nn.3490>
- McGinnis, W., Levine, M.S., Hafen, E., Kuroiwa, A., Gehring, W.J., 1984. A conserved DNA sequence in homoeotic genes of the *Drosophila* Antennapedia and bithorax complexes. *Nature* 308, 428–433. <https://doi.org/10.1038/308428a0>
- McGrew, M.J., Sherman, A., Lillico, S.G., Ellard, F.M., Radcliffe, P.A., Gilhooley, H.J., Mitrophanous, K.A., Cambray, N., Wilson, V., Sang, H., 2008. Localised axial progenitor cell populations in the avian tail bud are not committed to a posterior Hox identity. *Development* 135, 2289–2299. <https://doi.org/10.1242/dev.022020>
- McPherron, A.C., Lawler, A.M., Lee, S.-J., 1999. Regulation of anterior/posterior patterning of the axial skeleton by growth/differentiation factor 11. *Nat Genet* 22, 260–264. <https://doi.org/10.1038/10320>
- Metzis, V., Steinhäuser, S., Pakanavicius, E., Gouti, M., Stamatakis, D., Ivanovitch, K., Watson, T., Rayon, T., Mousavy Gharavy, S.N., Lovell-Badge, R., Luscombe, N.M., Briscoe, J., 2018. Nervous System Regionalization Entails Axial Allocation before Neural Differentiation. *Cell* 175, 1105–1118.e17. <https://doi.org/10.1016/j.cell.2018.09.040>
- Miao, Y., Djeflal, Y., De Simone, A., Zhu, K., Lee, J.G., Lu, Z., Silberfeld, A., Rao, J., Tarazona, O.A., Mongera, A., Rigoni, P., Diaz-Cuadros, M., Song, L.M.S., Di Talia, S., Pourquié, O., 2023. Reconstruction and deconstruction of human somitogenesis in vitro. *Nature* 614, 500–508. <https://doi.org/10.1038/s41586-022-05655-4>
- Millet, C., Yamashita, M., Heller, M., Yu, L.-R., Veenstra, T.D., Zhang, Y.E., 2009. A Negative Feedback Control of Transforming Growth Factor- β Signaling by Glycogen Synthase Kinase 3-mediated Smad3 Linker Phosphorylation at Ser-204. *Journal of Biological Chemistry* 284, 19808–19816. <https://doi.org/10.1074/jbc.M109.016667>

- Montavon, T., Le Garrec, J.-F., Kerszberg, M., Duboule, D., 2008. Modeling *Hox* gene regulation in digits: reverse collinearity and the molecular origin of thumbness. *Genes Dev.* 22, 346–359. <https://doi.org/10.1101/gad.1631708>
- 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, 1132–1145. <https://doi.org/10.1016/j.cell.2011.10.023>
- Moris, N., Anlas, K., Van Den Brink, S.C., Alemany, A., Schröder, J., Ghimire, S., Balayo, T., Van Oudenaarden, A., Martinez Arias, A., 2020. An in vitro model of early anteroposterior organization during human development. *Nature* 582, 410–415. <https://doi.org/10.1038/s41586-020-2383-9>
- Mouilleau, V., Vaslin, C., Robert, R., Gribaudo, S., Nicolas, N., Jarrige, M., Terray, A., Lesueur, L., Mathis, M.W., Croft, G., Daynac, M., Rouiller-Fabre, V., Wichterle, H., Ribes, V., Martinat, C., Nedelec, S., 2021. Dynamic extrinsic pacing of the HOX clock in human axial progenitors controls motor neuron subtype specification. *Development* 148, dev194514. <https://doi.org/10.1242/dev.194514>
- Murry, C.E., Keller, G., 2008. Differentiation of Embryonic Stem Cells to Clinically Relevant Populations: Lessons from Embryonic Development. *Cell* 132, 661–680. <https://doi.org/10.1016/j.cell.2008.02.008>
- Naiche, L.A., Holder, N., Lewandoski, M., 2011. FGF4 and FGF8 comprise the wavefront activity that controls somitogenesis. *Proc. Natl. Acad. Sci. U.S.A.* 108, 4018–4023. <https://doi.org/10.1073/pnas.1007417108>
- Nakashima, M., Toyono, T., Akamine, A., Joyner, A., 1999. Expression of growth/differentiation factor 11, a new member of the BMP/TGF β superfamily during mouse embryogenesis. *Mechanisms of Development* 80, 185–189. [https://doi.org/10.1016/S0925-4773\(98\)00205-6](https://doi.org/10.1016/S0925-4773(98)00205-6)
- Narendra, V., Bulajić, M., Dekker, J., Mazzoni, E.O., Reinberg, D., 2016. CTCF-mediated topological boundaries during development foster appropriate gene regulation. *Genes Dev.* 30, 2657–2662. <https://doi.org/10.1101/gad.288324.116>
- Narendra, V., Rocha, P.P., An, D., Raviram, R., Skok, J.A., Mazzoni, E.O., Reinberg, D., 2015. CTCF establishes discrete functional chromatin domains at the *Hox* clusters during differentiation. *Science* 347, 1017–1021. <https://doi.org/10.1126/science.1262088>
- Nedelec, S., Martinez-Arias, A., 2021. In vitro models of spinal motor circuit’s development in mammals: achievements and challenges. *Current Opinion in Neurobiology* 66, 240–249. <https://doi.org/10.1016/j.conb.2020.12.002>
- Neijts, R., Amin, S., Van Rooijen, C., Deschamps, J., 2017. *Cdx* is crucial for the timing mechanism driving colinear *Hox* activation and defines a trunk segment in the *Hox* cluster topology. *Developmental Biology* 422, 146–154. <https://doi.org/10.1016/j.ydbio.2016.12.024>

- Neijts, R., Amin, S., Van Rooijen, C., Tan, S., Creyghton, M.P., De Laat, W., Deschamps, J., 2016. Polarized regulatory landscape and Wnt responsiveness underlie Hox activation in embryos. *Genes Dev.* 30, 1937–1942. <https://doi.org/10.1101/gad.285767.116>
- Niederreither, K., Dollé, P., 2008. Retinoic acid in development: towards an integrated view. *Nat Rev Genet* 9, 541–553. <https://doi.org/10.1038/nrg2340>
- Nishimoto, S., Minguillon, C., Wood, S., Logan, M.P.O., 2014. A Combination of Activation and Repression by a Colinear Hox Code Controls Forelimb-Restricted Expression of Tbx5 and Reveals Hox Protein Specificity. *PLoS Genet* 10, e1004245. <https://doi.org/10.1371/journal.pgen.1004245>
- Nishita, J., Ohta, S., Bleyl, S.B., Schoenwolf, G.C., 2011. Detection of isoform-specific fibroblast growth factor receptors by whole-mount in situ hybridization in early chick embryos. *Dev. Dyn.* 240, 1537–1547. <https://doi.org/10.1002/dvdy.22616>
- Noordermeer, D., Leleu, M., Schorderet, P., Joye, E., Chabaud, F., Duboule, D., 2014. Temporal dynamics and developmental memory of 3D chromatin architecture at Hox gene loci. *eLife* 3, e02557. <https://doi.org/10.7554/eLife.02557>
- Noordermeer, D., Leleu, M., Splinter, E., Rougemont, J., De Laat, W., Duboule, D., 2011. The Dynamic Architecture of *Hox* Gene Clusters. *Science* 334, 222–225. <https://doi.org/10.1126/science.1207194>
- Oginuma, M., Moncuquet, P., Xiong, F., Karoly, E., Chal, J., Guevorkian, K., Pourquié, O., 2017. A Gradient of Glycolytic Activity Coordinates FGF and Wnt Signaling during Elongation of the Body Axis in Amniote Embryos. *Developmental Cell* 40, 342–353. <https://doi.org/10.1016/j.devcel.2017.02.001>
- Oh, S.P., Yeo, C.-Y., Lee, Y., Schrewe, H., Whitman, M., Li, E., 2002. Activin type IIA and IIB receptors mediate Gdf11 signaling in axial vertebral patterning. *Genes Dev.* 16, 2749–2754. <https://doi.org/10.1101/gad.1021802>
- Olivera-Martinez, I., Harada, H., Halley, P.A., Storey, K.G., 2012. Loss of FGF-Dependent Mesoderm Identity and Rise of Endogenous Retinoid Signalling Determine Cessation of Body Axis Elongation. *PLoS Biol* 10, e1001415. <https://doi.org/10.1371/journal.pbio.1001415>
- Olivera-Martinez, I., Schurch, N., Li, R.A., Song, J., Halley, P.A., Das, R.M., Burt, D.W., Barton, G.J., Storey, K.G., 2014. Major transcriptome re-organisation and abrupt changes in signalling, cell cycle and chromatin regulation at neural differentiation *in vivo*. *Development* 141, 3266–3276. <https://doi.org/10.1242/dev.112623>
- Olivera-Martinez, I., Storey, K.G., 2007. Wnt signals provide a timing mechanism for the FGF-retinoid differentiation switch during vertebrate body axis extension. *Development* 134, 2125–2135. <https://doi.org/10.1242/dev.000216>
- Osorno, R., Tsakiridis, A., Wong, F., Cambray, N., Economou, C., Wilkie, R., Blin, G., Scotting, P.J., Chambers, I., Wilson, V., 2012. The developmental dismantling of pluripotency

is reversed by ectopic Oct4 expression. *Development* 139, 2288–2298. <https://doi.org/10.1242/dev.078071>

Pang, D., Thompson, D.N.P., 2011. Embryology and bony malformations of the craniovertebral junction. *Childs Nerv Syst* 27, 523–564. <https://doi.org/10.1007/s00381-010-1358-9>

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. <https://doi.org/10.1101/gad.12.15.2332>

Philippidou, P., Dasen, J.S., 2013. Hox Genes: Choreographers in Neural Development, Architects of Circuit Organization. *Neuron* 80, 12–34. <https://doi.org/10.1016/j.neuron.2013.09.020>

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

Pilon, N., Oh, K., Sylvestre, J.-R., Bouchard, N., Savory, J., Lohnes, D., 2006. Cdx4 is a direct target of the canonical Wnt pathway. *Developmental Biology* 289, 55–63. <https://doi.org/10.1016/j.ydbio.2005.10.005>

Pilon, N., Oh, K., Sylvestre, J.-R., Savory, J.G.A., Lohnes, D., 2007. Wnt signaling is a key mediator of *Cdx1* expression in vivo. *Development* 134, 2315–2323. <https://doi.org/10.1242/dev.001206>

Pinglay, S., Bulajić, M., Rahe, D.P., Huang, E., Brosh, R., Mamrak, N.E., King, B.R., German, S., Cadley, J.A., Rieber, L., Easo, N., Lionnet, T., Mahony, S., Maurano, M.T., Holt, L.J., Mazzoni, E.O., Boeke, J.D., 2022. Synthetic regulatory reconstitution reveals principles of mammalian *Hox* cluster regulation. *Science* 377, eabk2820. <https://doi.org/10.1126/science.abk2820>

Pollock, R.A., Jay, G., Bieberich, C.J., 1992. Altering the boundaries of Hox3.1 expression: Evidence for antipodal gene regulation. *Cell* 71, 911–923. [https://doi.org/10.1016/0092-8674\(92\)90388-S](https://doi.org/10.1016/0092-8674(92)90388-S)

Pöpperl, H., Featherstone, M.S., 1993. Identification of a Retinoic Acid Response Element Upstream of the Murine *Hox-4.2* Gene. *Molecular and Cellular Biology* 13, 257–265. <https://doi.org/10.1128/mcb.13.1.257-265.1993>

Postlethwait, J.H., Schneiderman, H.A., 1971. Pattern formation and determination in the antenna of the homoeotic mutant *Antennapedia* of *Drosophila melanogaster*. *Developmental Biology* 25, 606–640. [https://doi.org/10.1016/0012-1606\(71\)90008-X](https://doi.org/10.1016/0012-1606(71)90008-X)

Pouponnot, C., Jayaraman, L., Massagué, J., 1998. Physical and Functional Interaction of SMADs and p300/CBP. *Journal of Biological Chemistry* 273, 22865–22868. <https://doi.org/10.1074/jbc.273.36.22865>

- Psychoyos, D., Stern, C.D., 1996. Fates and migratory routes of primitive streak cells in the chick embryo. *Development* 122, 1523–1534. <https://doi.org/10.1242/dev.122.5.1523>
- Purvis, J.E., Lahav, G., 2013. Encoding and Decoding Cellular Information through Signaling Dynamics. *Cell* 152, 945–956. <https://doi.org/10.1016/j.cell.2013.02.005>
- Rekaik, H., Lopez-Delisle, L., Hintermann, A., Mascrez, B., Bochaton, C., Mayran, A., Duboule, D., 2023. Sequential and directional insulation by conserved CTCF sites underlies the Hox timer in stem embryos. *Nat Genet.* <https://doi.org/10.1038/s41588-023-01426-7>
- Rhinn, M., Dollé, P., 2012. Retinoic acid signalling during development. *Development* 139, 843–858. <https://doi.org/10.1242/dev.065938>
- Ribes, V., Le Roux, I., Rhinn, M., Schuhbauer, B., Dollé, P., 2009. Early mouse caudal development relies on crosstalk between retinoic acid, Shh and Fgf signalling pathways. *Development* 136, 665–676. <https://doi.org/10.1242/dev.016204>
- Rings, E.H.H.M., Boudreau, F., Taylor, J.K., Moffett, J., Suh, E.R., Traber, P.G., 2001. Phosphorylation of the serine 60 residue within the Cdx2 activation domain mediates its transactivation capacity. *Gastroenterology* 121, 1437–1450. <https://doi.org/10.1053/gast.2001.29618>
- Rinn, J.L., Kertesz, M., Wang, J.K., Squazzo, S.L., Xu, X., Bruggmann, S.A., Goodnough, L.H., Helms, J.A., Farnham, P.J., Segal, E., Chang, H.Y., 2007. Functional Demarcation of Active and Silent Chromatin Domains in Human HOX Loci by Noncoding RNAs. *Cell* 129, 1311–1323. <https://doi.org/10.1016/j.cell.2007.05.022>
- Rivera-Pérez, J.A., Magnuson, T., 2005. Primitive streak formation in mice is preceded by localized activation of Brachyury and Wnt3. *Developmental Biology* 288, 363–371. <https://doi.org/10.1016/j.ydbio.2005.09.012>
- Robinton, D.A., Chal, J., Lummertz Da Rocha, E., Han, A., Yermalovich, A.V., Oginuma, M., Schlaeger, T.M., Sousa, P., Rodriguez, A., Urbach, A., Pourquié, O., Daley, G.Q., 2019. The Lin28/let-7 Pathway Regulates the Mammalian Caudal Body Axis Elongation Program. *Developmental Cell* 48, 396–405.e3. <https://doi.org/10.1016/j.devcel.2018.12.016>
- Roelen, B.A.J., 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. *Mechanisms of Development* 119, 81–90. [https://doi.org/10.1016/S0925-4773\(02\)00329-5](https://doi.org/10.1016/S0925-4773(02)00329-5)
- Romanos, M., Allio, G., Roussigné, M., Combres, L., Escalas, N., Soula, C., Médevielle, F., Steventon, B., Trescases, A., Bénazéraf, B., 2021. Cell-to-cell heterogeneity in Sox2 and Bra expression guides progenitor motility and destiny. *eLife* 10, e66588. <https://doi.org/10.7554/eLife.66588>
- Ross, S., Cheung, E., Petrakis, T.G., Howell, M., Kraus, W.L., Hill, C.S., 2006. Smads orchestrate specific histone modifications and chromatin remodeling to activate transcription. *EMBO J* 25, 4490–4502. <https://doi.org/10.1038/sj.emboj.7601332>

- Sanaki-Matsumiya, M., Matsuda, M., Gritti, N., Nakaki, F., Sharpe, J., Trivedi, V., Ebisuya, M., 2022. Periodic formation of epithelial somites from human pluripotent stem cells. *Nat Commun* 13, 2325. <https://doi.org/10.1038/s41467-022-29967-1>
- Sasagawa, S., Ozaki, Y., Fujita, K., Kuroda, S., 2005. Prediction and validation of the distinct dynamics of transient and sustained ERK activation. *Nat Cell Biol* 7, 365–373. <https://doi.org/10.1038/ncb1233>
- Savory, J.G.A., Bouchard, N., Pierre, V., Rijli, F.M., De Repentigny, Y., Kothary, R., Lohnes, D., 2009. *Cdx2* regulation of posterior development through non-Hox targets. *Development* 136, 4099–4110. <https://doi.org/10.1242/dev.041582>
- Schöck, F., Reischl, J., Wimmer, E., Taubert, H., Purnell, B.A., Jäckle, H., 2000. Phenotypic suppression of empty spiracles is prevented by *buttonhead*. *Nature* 405, 351–354. <https://doi.org/10.1038/35012620>
- Schuettengruber, B., Cavalli, G., 2009. Recruitment of Polycomb group complexes and their role in the dynamic regulation of cell fate choice. *Development* 136, 3531–3542. <https://doi.org/10.1242/dev.033902>
- Selleck, M.A.J., Stern, C.D., 1991. Fate mapping and cell lineage analysis of Hensen's node in the chick embryo. *Development* 112, 615–626. <https://doi.org/10.1242/dev.112.2.615>
- Shashikant, C.S., Bieberich, C.J., Belting, H.-G., Wang, J.C.H., Borbély, M.A., Ruddle, F.H., 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. <https://doi.org/10.1242/dev.121.12.4339>
- Shashikant, C.S., Kim, C.B., Borbély, M.A., Wang, W.C.H., Ruddle, F.H., 1998. Comparative studies on mammalian *Hoxc8* early enhancer sequence reveal a baleen whale-specific deletion of a cis-acting element. *Proc. Natl. Acad. Sci. U.S.A.* 95, 15446–15451. <https://doi.org/10.1073/pnas.95.26.15446>
- Shi, Y., Massagué, J., 2003. Mechanisms of TGF- β Signaling from Cell Membrane to the Nucleus. *Cell* 113, 685–700. [https://doi.org/10.1016/S0092-8674\(03\)00432-X](https://doi.org/10.1016/S0092-8674(03)00432-X)
- Shimizu, T., Bae, Y.-K., Muraoka, O., Hibi, M., 2005. Interaction of Wnt and caudal-related genes in zebrafish posterior body formation. *Developmental Biology* 279, 125–141. <https://doi.org/10.1016/j.ydbio.2004.12.007>
- Shin, S., Dalton, S., Stice, S.L., 2005. Human Motor Neuron Differentiation from Human Embryonic Stem Cells. *Stem Cells and Development* 14, 266–269. <https://doi.org/10.1089/scd.2005.14.266>
- Simeone, A., Acampora, D., Nigro, V., Faiella, A., D'Esposito, M., Stornaiuolo, A., Mavilio, F., Boncinelli, E., 1991. Differential regulation by retinoic acid of the homeobox genes of the four HOX loci in human embryonal carcinoma cells. *Mechanisms of Development* 33, 215–227. [https://doi.org/10.1016/0925-4773\(91\)90029-6](https://doi.org/10.1016/0925-4773(91)90029-6)

- Simon, J., Chiang, A., Bender, W., 1992. Ten different *Polycomb* group genes are required for spatial control of the *abdA* and *AbdB* homeotic products. *Development* 114, 493–505. <https://doi.org/10.1242/dev.114.2.493>
- Sive, H.L., Draper, B.W., Harland, R.M., Weintraub, H., 1990. Identification of a retinoic acid-sensitive period during primary axis formation in *Xenopus laevis*. *Genes Dev.* 4, 932–942. <https://doi.org/10.1101/gad.4.6.932>
- Small, K.M., Potter, S.S., 1993. Homeotic transformations and limb defects in Hox A11 mutant mice. *Genes & Development* 7, 2318–2328. <https://doi.org/10.1101/gad.7.12a.2318>
- Sorre, B., Warmflash, A., Brivanlou, A.H., Siggia, E.D., 2014. Encoding of Temporal Signals by the TGF- β Pathway and Implications for Embryonic Patterning. *Developmental Cell* 30, 334–342. <https://doi.org/10.1016/j.devcel.2014.05.022>
- Soshnikova, N., Duboule, D., 2009. Epigenetic Temporal Control of Mouse *Hox* Genes in Vivo. *Science* 324, 1320–1323. <https://doi.org/10.1126/science.1171468>
- Spitz, F., Gonzalez, F., Duboule, D., 2003. A Global Control Region Defines a Chromosomal Regulatory Landscape Containing the HoxD Cluster. *Cell* 113, 405–417. [https://doi.org/10.1016/S0092-8674\(03\)00310-6](https://doi.org/10.1016/S0092-8674(03)00310-6)
- Steventon, B., Martinez Arias, A., 2017. Evo-engineering and the cellular and molecular origins of the vertebrate spinal cord. *Developmental Biology* 432, 3–13. <https://doi.org/10.1016/j.ydbio.2017.01.021>
- Stifani, N., 2014. Motor neurons and the generation of spinal motor neuron diversity. *Front. Cell. Neurosci.* 8. <https://doi.org/10.3389/fncel.2014.00293>
- 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, 1728–1732. <https://doi.org/10.1126/science.7916164>
- Suh, J., Eom, J., Kim, N., Woo, K.M., Baek, J., Ryoo, H., Lee, S., Lee, Y., 2019. Growth differentiation factor 11 locally controls anterior–posterior patterning of the axial skeleton. *Journal Cellular Physiology* 234, 23360–23368. <https://doi.org/10.1002/jcp.28904>
- Sun, X., Song, J., Huang, H., Chen, H., Qian, K., 2018. Modeling hallmark pathology using motor neurons derived from the family and sporadic amyotrophic lateral sclerosis patient-specific iPS cells. *Stem Cell Res Ther* 9, 315. <https://doi.org/10.1186/s13287-018-1048-1>
- Szumaska, D., Pielas, G., Essalmani, R., Bilski, M., Mesnard, D., Kaur, K., Franklyn, A., El Omari, K., Jefferis, J., Bentham, J., Taylor, J.M., Schneider, J.E., Arnold, S.J., Johnson, P., Tymowska-Lalanne, Z., Stammers, D., Clarke, K., Neubauer, S., Morris, A., Brown, S.D., Shaw-Smith, C., Cama, A., Capra, V., Ragoussis, J., Constam, D., Seidah, N.G., Prat, A., Bhattacharya, S., 2008. VACTERL/caudal regression/Currarino syndrome-like malformations in mice with mutation in the proprotein convertase *Pcsk5*. *Genes Dev.* 22, 1465–1477. <https://doi.org/10.1101/gad.479408>

- Takada, S., Stark, K.L., Shea, M.J., Vassileva, G., McMahon, J.A., McMahon, A.P., 1994. Wnt-3a regulates somite and tailbud formation in the mouse embryo. *Genes Dev.* 8, 174–189. <https://doi.org/10.1101/gad.8.2.174>
- Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., Yamanaka, S., 2007. Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. *Cell* 131, 861–872. <https://doi.org/10.1016/j.cell.2007.11.019>
- Takahashi, K., Yamanaka, S., 2006. Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell* 126, 663–676. <https://doi.org/10.1016/j.cell.2006.07.024>
- Takemoto, T., Uchikawa, M., Kamachi, Y., Kondoh, H., 2006. Convergence of Wnt and FGF signals in the genesis of posterior neural plate through activation of the *Sox2* enhancer N-1. *Development* 133, 297–306. <https://doi.org/10.1242/dev.02196>
- Takemoto, T., Uchikawa, M., Yoshida, M., Bell, D.M., Lovell-Badge, R., Papaioannou, V.E., Kondoh, H., 2011. Tbx6-dependent Sox2 regulation determines neural or mesodermal fate in axial stem cells. *Nature* 470, 394–398. <https://doi.org/10.1038/nature09729>
- Tam, P.P.L., Tan, S.-S., 1992. The somitogenetic potential of cells in the primitive streak and the tail bud of the organogenesis-stage mouse embryo. *Development* 115, 703–715. <https://doi.org/10.1242/dev.115.3.703>
- Ten Berge, D., Koole, W., Fuerer, C., Fish, M., Eroglu, E., Nusse, R., 2008. Wnt Signaling Mediates Self-Organization and Axis Formation in Embryoid Bodies. *Cell Stem Cell* 3, 508–518. <https://doi.org/10.1016/j.stem.2008.09.013>
- Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S., Jones, J.M., 1998. Embryonic Stem Cell Lines Derived from Human Blastocysts. *Science* 282, 1145–1147. <https://doi.org/10.1126/science.282.5391.1145>
- Tie, F., Banerjee, R., Saiakhova, A.R., Howard, B., Monteith, K.E., Scacheri, P.C., Cosgrove, M.S., Harte, P.J., 2014. Trithorax monomethylates histone H3K4 and interacts directly with CBP to promote H3K27 acetylation and antagonize Polycomb silencing. *Development* 141, 1129–1139. <https://doi.org/10.1242/dev.102392>
- Tiret, L., Le Mouellic, H., Maury, M., Brûlet, P., 1998. Increased apoptosis of motoneurons and altered somatotopic maps in the brachial spinal cord of *Hoxc-8* -deficient mice. *Development* 125, 279–291. <https://doi.org/10.1242/dev.125.2.279>
- Toettcher, J.E., Weiner, O.D., Lim, W.A., 2013. Using Optogenetics to Interrogate the Dynamic Control of Signal Transmission by the Ras/Erk Module. *Cell* 155, 1422–1434. <https://doi.org/10.1016/j.cell.2013.11.004>
- Tsakiridis, A., Huang, Y., Blin, G., Skylaki, S., Wymeersch, F., Osorno, R., Economou, C., Karagianni, E., Zhao, S., Lowell, S., Wilson, V., 2014. Distinct Wnt-driven primitive streak-

like populations reflect *in vivo* lineage precursors. *Development* 141, 1209–1221. <https://doi.org/10.1242/dev.101014>

Tschopp, P., Christen, A.J., Duboule, D., 2012. Bimodal control of *Hoxd* gene transcription in the spinal cord defines two regulatory subclusters. *Development* 139, 929–939. <https://doi.org/10.1242/dev.076794>

Turner, D.A., Girgin, M., Alonso-Crisostomo, L., Trivedi, V., Baillie-Johnson, P., Glodowski, C.R., Hayward, P.C., Collignon, J., Gustavsen, C., Serup, P., Steventon, B., Lutolf, M., Martinez, A.A., 2017. Anteroposterior polarity and elongation in the absence of extraembryonic tissues and spatially localised signalling in *Gastruloids*, mammalian embryonic organoids. *Development* dev.150391. <https://doi.org/10.1242/dev.150391>

Turner, D.A., Hayward, P.C., Baillie-Johnson, P., Rué, P., Broome, R., Faunes, F., Martinez Arias, A., 2014. Wnt/ β -catenin and FGF signalling direct the specification and maintenance of a neuromesodermal axial progenitor in ensembles of mouse embryonic stem cells. *Development* 141, 4243–4253. <https://doi.org/10.1242/dev.112979>

Tzouanacou, E., Wegener, A., Wymeersch, F.J., Wilson, V., Nicolas, J.-F., 2009. Redefining the Progression of Lineage Segregations during Mammalian Embryogenesis by Clonal Analysis. *Developmental Cell* 17, 365–376. <https://doi.org/10.1016/j.devcel.2009.08.002>

Van Den Akker, E., Forlani, S., Chawengsaksophak, K., De Graaff, W., Beck, F., Meyer, B.I., Deschamps, J., 2002. *Cdx1* and *Cdx2* have overlapping functions in anteroposterior patterning and posterior axis elongation. *Development* 129, 2181–2193. <https://doi.org/10.1242/dev.129.9.2181>

Van Den Brink, S.C., Baillie-Johnson, P., Balayo, T., Hadjantonakis, A.-K., Nowotschin, S., Turner, D.A., Martinez Arias, A., 2014. Symmetry breaking, germ layer specification and axial organisation in aggregates of mouse embryonic stem cells. *Development* 141, 4231–4242. <https://doi.org/10.1242/dev.113001>

Van Der Lugt, N.M., Domen, J., Linders, K., Van Roon, M., Robanus-Maandag, E., Te Riele, H., Van Der Valk, M., Deschamps, J., Sofroniew, M., Van Lohuizen, M., 1994. Posterior transformation, neurological abnormalities, and severe hematopoietic defects in mice with a targeted deletion of the *bmi-1* proto-oncogene. *Genes Dev.* 8, 757–769. <https://doi.org/10.1101/gad.8.7.757>

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. <https://doi.org/10.1242/dev.02216>

Van Rooijen, C., Simmini, S., Bialecka, M., Neijts, R., Van De Ven, C., Beck, F., Deschamps, J., 2012. Evolutionarily conserved requirement of *Cdx* for post-occipital tissue emergence. *Development* 139, 2576–2583. <https://doi.org/10.1242/dev.079848>

- Veenvliet, J.V., Bolondi, A., Kretzmer, H., Haut, L., Scholze-Wittler, M., Schifferl, D., Koch, F., Guignard, L., Kumar, A.S., Pustet, M., Heimann, S., Buschow, R., Wittler, L., Timmermann, B., Meissner, A., Herrmann, B.G., 2020. Mouse embryonic stem cells self-organize into trunk-like structures with neural tube and somites. *Science* 370, eaba4937. <https://doi.org/10.1126/science.aba4937>
- Vermot, J., Pourquié, O., 2005. Retinoic acid coordinates somitogenesis and left–right patterning in vertebrate embryos. *Nature* 435, 215–220. <https://doi.org/10.1038/nature03488>
- Vermot, J., Schuhbaur, B., Mouellic, H.L., McCaffery, P., Garnier, J.-M., Hentsch, D., Brûlet, P., Niederreither, K., Chambon, P., Dollé, P., Roux, I.L., 2005. Retinaldehyde dehydrogenase 2 and *Hoxc8* are required in the murine brachial spinal cord for the specification of *Lim1*⁺ motoneurons and the correct distribution of *Islet1*⁺ motoneurons. *Development* 132, 1611–1621. <https://doi.org/10.1242/dev.01718>
- Verrier, L., Davidson, L., Gierliński, M., Dady, A., Storey, K.G., 2018. Neural differentiation, selection and transcriptomic profiling of human neuromesodermal progenitors-like cells in vitro. *Development* dev.166215. <https://doi.org/10.1242/dev.166215>
- Viswanathan, S.R., Daley, G.Q., 2010. *Lin28*: A MicroRNA Regulator with a Macro Role. *Cell* 140, 445–449. <https://doi.org/10.1016/j.cell.2010.02.007>
- Waas, W.F., Dalby, K.N., 2001. Purification of a Model Substrate for Transcription Factor Phosphorylation by ERK2. *Protein Expression and Purification* 23, 191–197. <https://doi.org/10.1006/prep.2001.1491>
- Wahl, M.B., Deng, C., Lewandoski, M., Pourquié, O., 2007. FGF signaling acts upstream of the NOTCH and WNT signaling pathways to control segmentation clock oscillations in mouse somitogenesis. *Development* 134, 4033–4041. <https://doi.org/10.1242/dev.009167>
- Wellik, D.M., 2009. Chapter 9 *Hox* Genes and Vertebrate Axial Pattern, in: *Current Topics in Developmental Biology*. Elsevier, pp. 257–278. [https://doi.org/10.1016/S0070-2153\(09\)88009-5](https://doi.org/10.1016/S0070-2153(09)88009-5)
- Wellik, D.M., Capecchi, M.R., 2003. *Hox10* and *Hox11* Genes Are Required to Globally Pattern the Mammalian Skeleton. *Science* 301, 363–367. <https://doi.org/10.1126/science.1085672>
- Wichterle, H., Lieberam, I., Porter, J.A., Jessell, T.M., 2002. Directed Differentiation of Embryonic Stem Cells into Motor Neurons. *Cell* 110, 385–397. [https://doi.org/10.1016/S0092-8674\(02\)00835-8](https://doi.org/10.1016/S0092-8674(02)00835-8)
- Wilcock, A.C., Swedlow, J.R., Storey, K.G., 2007. Mitotic spindle orientation distinguishes stem cell and terminal modes of neuron production in the early spinal cord. *Development* 134, 1943–1954. <https://doi.org/10.1242/dev.002519>
- Wilson, V., Olivera-Martinez, I., Storey, K.G., 2009. Stem cells, signals and vertebrate body axis extension. *Development* 136, 1591–1604. <https://doi.org/10.1242/dev.021246>

- Wittler, L., Shin, E., Grote, P., Kispert, A., Beckers, A., Gossler, A., Werber, M., Herrmann, B.G., 2007. Expression of *Msn1* in the presomitic mesoderm is controlled by synergism of WNT signalling and *Tbx6*. *EMBO Reports* 8, 784–789. <https://doi.org/10.1038/sj.embor.7401030>
- Woo, C.J., Kharchenko, P.V., Daheron, L., Park, P.J., Kingston, R.E., 2010. A Region of the Human HOXD Cluster that Confers Polycomb-Group Responsiveness. *Cell* 140, 99–110. <https://doi.org/10.1016/j.cell.2009.12.022>
- Wu, Y., Wang, G., Scott, S.A., Capecchi, M.R., 2008. Hoxc10 and Hoxd10 regulate mouse columnar, divisional and motor pool identity of lumbar motoneurons. *Development* 135, 171–182. <https://doi.org/10.1242/dev.009225>
- Wymeersch, F.J., Huang, Y., Blin, G., Cambray, N., Wilkie, R., Wong, F.C., Wilson, V., 2016. Position-dependent plasticity of distinct progenitor types in the primitive streak. *eLife* 5, e10042. <https://doi.org/10.7554/eLife.10042>
- Wymeersch, F.J., Skylaki, S., Huang, Y., Watson, J.A., Economou, C., Marek-Johnston, C., Tomlinson, S.R., Wilson, V., 2018. Transcriptionally dynamic progenitor populations organised around a stable niche drive axial patterning. *Development* dev.168161. <https://doi.org/10.1242/dev.168161>
- Wymeersch, F.J., Wilson, V., Tsakiridis, A., 2021. Understanding axial progenitor biology *in vivo* and *in vitro*. *Development* 148, dev180612. <https://doi.org/10.1242/dev.180612>
- Yamaguchi, T.P., Bradley, A., McMahon, A.P., Jones, S., 1999. A *Wnt5a* pathway underlies outgrowth of multiple structures in the vertebrate embryo. *Development* 126, 1211–1223. <https://doi.org/10.1242/dev.126.6.1211>
- Yaman, Y.I., Ramanathan, S., 2023. Controlling human organoid symmetry breaking reveals signaling gradients drive segmentation clock waves. *Cell* 186, 513–527.e19. <https://doi.org/10.1016/j.cell.2022.12.042>
- Yamanaka, Y., Hamidi, S., Yoshioka-Kobayashi, K., Munira, S., Sunadome, K., Zhang, Y., Kurokawa, Y., Ericsson, R., Mieda, A., Thompson, J.L., Kerwin, J., Lisgo, S., Yamamoto, T., Moris, N., Martinez-Arias, A., Tsujimura, T., Alev, C., 2023. Reconstituting human somitogenesis in vitro. *Nature* 614, 509–520. <https://doi.org/10.1038/s41586-022-05649-2>
- Yang, M., Yang, S.-L., Herrlinger, S., Liang, C., Dzieciatkowska, M., Hansen, K.C., Desai, R., Nagy, A., Niswander, L., Moss, E.G., Chen, J.-F., 2015. Lin28 promotes the proliferative capacity of neural progenitor cells in brain development. *Development* 142, 1616–1627. <https://doi.org/10.1242/dev.120543>
- Yoshida, K., Nakahata, A., Treen, N., Sakuma, T., Yamamoto, T., Sasakura, Y., 2017. *Hox* - mediated endodermal identity patterns the pharyngeal muscle formation in the chordate pharynx. *Development* dev.144436. <https://doi.org/10.1242/dev.144436>

- Yoshikawa, Y., Fujimori, T., McMahon, A.P., Takada, S., 1997. Evidence That Absence of Wnt-3a Signaling Promotes Neuralization Instead of Paraxial Mesoderm Development in the Mouse. *Developmental Biology* 183, 234–242. <https://doi.org/10.1006/dbio.1997.8502>
- Young, T., Rowland, J.E., Van De Ven, C., Bialecka, M., Novoa, A., Carapuco, M., Van Nes, J., De Graaff, W., Duluc, I., Freund, J.-N., Beck, F., Mallo, M., Deschamps, J., 2009. Cdx and Hox Genes Differentially Regulate Posterior Axial Growth in Mammalian Embryos. *Developmental Cell* 17, 516–526. <https://doi.org/10.1016/j.devcel.2009.08.010>
- Zakany, J., 1997. Deletion of a HoxD enhancer induces transcriptional heterochrony leading to transposition of the sacrum. *The EMBO Journal* 16, 4393–4402. <https://doi.org/10.1093/emboj/16.14.4393>
- Zawel, L., Le Dai, J., Buckhaults, P., Zhou, S., Kinzler, K.W., Vogelstein, B., Kern, S.E., 1998. Human Smad3 and Smad4 Are Sequence-Specific Transcription Activators. *Molecular Cell* 1, 611–617. [https://doi.org/10.1016/S1097-2765\(00\)80061-1](https://doi.org/10.1016/S1097-2765(00)80061-1)
- Zeitlinger, J., 2020. Seven myths of how transcription factors read the cis-regulatory code. *Current Opinion in Systems Biology* 23, 22–31. <https://doi.org/10.1016/j.coisb.2020.08.002>
- Zhang, X., Peterson, K.A., Liu, X.S., McMahon, A.P., Ohba, S., 2013. Gene Regulatory Networks Mediating Canonical Wnt Signal-Directed Control of Pluripotency and Differentiation in Embryo Stem Cells. *Stem Cells* 31, 2667–2679. <https://doi.org/10.1002/stem.1371>
- Zhang, Y.E., 2009. Non-Smad pathways in TGF- β signaling. *Cell Res* 19, 128–139. <https://doi.org/10.1038/cr.2008.328>