

Aus dem Physiologischen Institut, Lehrstuhl: Physiologische Genomik im
Biomedizinischen Zentrum (BMC)
Institut der Ludwig-Maximilians-Universität München
Vorstand: Prof. Dr. Magdalena Götz



***Establishment of novel axial stem cell states recapitulating
neuromesodermal and dorsal neural tube progenitor identities***

Dissertation
zum Erwerb des Doktorgrades der Naturwissenschaften
an der Medizinischen Fakultät der
Ludwig-Maximilians-Universität München

vorgelegt von
Dolunay Kelle
aus
Izmir, Turkey

Jahr
2023

Mit Genehmigung der Medizinischen Fakultät
der Universität München

Betreuerin: Prof. Dr. Magdalena Götz

Zweitgutachter: Prof. Dr. Dominik Paquet

Dekan: Prof. Dr. med. Thomas Gudermann

Tag der mündlichen Prüfung: 23. April 2024

TABLE OF CONTENT

TABLE OF CONTENT	1
ZUSAMMENFASSUNG	5
ABSTRACT	5
LIST OF FIGURES	7
LIST OF TABLES	10
LIST OF ABBREVIATIONS	11
1. INTRODUCTION	14
1.1. Early embryonic development.....	14
1.2. Axial progenitors.....	15
1.3. Neuromesodermal progenitors.....	17
1.3.1. Signaling pathways involved in NMP maintenance	17
1.3.2. Molecular characteristics of steady-state NMPs.....	19
1.3.3. Lineage commitment of NMPs and development of axial structures.....	20
1.3.3.1. Neural fate	21
1.3.3.2. Mesodermal fate	24
1.3.3.3. Neural crest fate.....	26
1.4. Stem cells.....	28
1.4.1. Pluripotent stem cells.....	29
1.4.2. Multipotent stem cells.....	31
1.5. Neuromesodermal progenitors in vitro.....	32
1.5.1. Derivation of NMPs	32
1.5.1.1. Human NMPs.....	32
1.5.1.2. Mouse NMPs	33
1.5.2. Characteristics of NMP derivation cultures.....	34
1.5.3. NMP progeny.....	35
1.5.3.1. Neural differentiation.....	35
1.5.3.2. Mesodermal differentiation.....	36
1.5.3.3. Neural crest differentiation	37
1.6. Deciphering molecular dynamics during embryonic development	37
1.7. Aim and impact of this study	39

2. MATERIALS AND METHODS	41
2.1. Materials	41
2.1.1. Cell lines.....	41
2.1.2. Cell culture components.....	41
2.1.3. Chemicals and kits.....	43
2.1.4. Primers and probes	45
2.1.5. Antibodies.....	47
2.1.6. Instruments and software	48
2.2. Methods.....	49
2.2.1. Maintenance of PSCs	49
2.2.1.1. Human PSCs	49
2.2.1.2. Mouse EpiSCs.....	49
2.2.1.3. Orangutan iPSCs	50
2.2.2. Maintenance of primary human skeletal muscle cells	50
2.2.3. Derivation of axial stem cells.....	51
2.2.3.1. Derivation of axial stem cells from human ESCs	51
2.2.3.2. Derivation of axial stem cells from EpiSCs.....	52
2.2.3.3. Derivation of axial stem cells from orangutan iPSCs.....	52
2.2.4. Dissection of mouse embryos	52
2.2.5. Derivation of human neural progenitor cells	54
2.2.6. Neural differentiation from human axial stem cells.....	54
2.2.7. Neural differentiation from human neural progenitor cells.....	55
2.2.8. Skeletal muscle differentiation from human axial stem cells.....	55
2.2.9. Nephron differentiation from human axial stem cells	56
2.2.10. Cardiomyocyte differentiation from human axial stem cells	56
2.2.11. RNA extraction and cDNA synthesis.....	57
2.2.12. Quantitative RT-PCR.....	57
2.2.13. Immunostaining	57
2.2.14. Genomic DNA extraction and telomere length assay	58
2.2.15. Single-cell RNA sequencing.....	59
2.2.15.1. Preparation of single-cell suspensions	59
2.2.15.2. Library preparation and sequencing.....	59

2.2.15.3. Processing of the datasets	60
2.2.15.4. Downstream analysis	61
2.2.16. Chromatome and proteome analysis	61
2.2.16.1. Nuclei isolation and sample preparation for chromatome analysis	62
2.2.16.2. Sample preparation for total proteome analysis.....	63
2.2.16.2. Nanoflow LC-MS/MS analysis for proteomes and chromatomes.....	64
2.2.16.3. MS data quantification	64
2.2.16.4. Statistical analyses.....	64
2.2.16.5. Web application development.....	65
3. RESULTS.....	66
3.1. Establishment of human axial stem cells.....	66
3.2. Transcriptomic characterization of axial stem cells	70
3.2. Transcriptional differences between neural lineage committed state of axial stem cells and conventional neural progenitor cells.....	79
3.3. Proteomic characterization of axial stem cells.....	81
3.4. Chromatome profiling in axial stem cells	86
3.5. Hierarchical regulation between axial stem cell states.....	88
3.6. Reproducible derivation of axial stem cells from different human pluripotent stem cell lines	90
3.7. Progeny profiling in axial stem cells.....	92
3.7.1. Neural differentiation from axial stem cells	92
3.7.2. Mesodermal differentiation from axial stem cells	106
3.7.2.1. Skeletal muscle differentiation targets the segmentation clock	106
3.7.2.2. Skeletal muscle differentiation targeting axial elongation-associated pathways	111
3.7.2.3. Differentiation into intermediate and lateral mesoderm derivatives	114
3.8. Reproducibility of axial stem cells in species	116
3.8.1. Generation of axial stem cells from mouse embryonic tissues	116
3.8.2. Generation of axial stem cells from pluripotent stem cells	118
3.8.3. Generation of CFS cells from mouse embryonic tissue.....	124
4. DISCUSSION.....	126
4.1. Establishment of axial stem cells relies on mimicking posterior embryonic development step-wise.....	126

4.2. Novel AxSC states recapitulate consecutive phases of axial development	127
4.2.1. Human CFS cells exhibit the unique features of NMPs	127
4.2.2. Human CS cells are committed to neural lineage	128
4.2.3. Human AxSCs share certain traits related to transcriptomic, metabolomic and epigenetic regulations	129
4.3. Human AxSC states can be distinguished by their lineage potential.....	130
4.3.1. Neural progeny of AxSCs	130
4.3.2. Mesodermal progeny of AxSCs	133
4.3.2.1. Skeletal muscle differentiation	133
4.3.2.1. Intermediate and lateral mesoderm differentiation	135
4.4. AxSCs are reproducible, depending on their in vivo counterparts and the different mechanisms of development in different species	137
4.5. The promise of AxSCs for clinical translation.....	139
LITERATURE	140
APPENDIX	154
ACKNOWLEDGEMENTS.....	161
AFFIDAVIT	162

ZUSAMMENFASSUNG

Die frühe Embryonalentwicklung beginnt mit aufeinander folgenden Teilungsereignissen, die schließlich zur Abgrenzung der Abstammungslinien und zur Bildung von Vorläufern der drei Keimblätter führen: Ektoderm, Endoderm und Mesoderm. Diese Vorläufer bringen spezifische Nachkommen hervor, z. B. neurale Zellen, die aus ektodermalen Vorläuferzellen hervorgehen, und Somiten, die aus mesodermalen Vorläuferzellen hervorgehen. In der Vergangenheit sind viele Wissenschaftler davon ausgegangen, dass die Aufteilung der pluripotenten Zellen in drei Abstammungslinien streng reguliert ist, ohne Vorhandensein von Zwischenphasen. Obwohl dieses Paradigma in der anterioren Embryonalentwicklung umgesetzt wird, findet im posterioren Embryo ein anderer Mechanismus statt, wie klonale Analysen und Linienverfolgungsexperimente zeigten.

Ein gemeinsamer Vorläufertyp für neurale und mesodermale Zellen wurde entdeckt, die neuromesodermalen Vorläuferzellen (NMP). Diese Vorläuferzellen treten in der Stammzellregion im hinteren Teil des Embryos auf und sind in Mensch, Huhn und Maus konserviert. Aus NMPs gehen axiale Strukturen hervor, d.h. das Neuralrohr und die Somiten. Moderne Methoden ermöglichen es, die molekularen Eigenschaften und die Umgebung der NMPs im stationären Zustand zu definieren und in vitro nachzuahmen. Die zu diesem Zweck entwickelten Verfahren erlauben jedoch nur die Erzeugung einer transienten Vorläuferpopulation und nicht die Etablierung des Stammzelläquivalents der NMPs.

Um die axiale Entwicklung zu rekapitulieren, habe ich zwei humane Stammzelllinien, CFS und CS, durch Posteriorisierung von humanen pluripotenten Stammzellen (hPSCs) etabliert. Ich bestätigte, dass CFS und CS NMPs bzw. dorsalen Neuralrohr-Vorläuferzellen entsprechen, und definierte diese Zellen daher zusammenfassend als axiale Stammzellen (AxSCs). Ich entdeckte eine hierarchische Regulation zwischen den AxSC-Zuständen, was darauf hindeutet, dass CS-Zellen ein weiteres Entwicklungsstadium darstellen, das aus CFS-Zellen, d.h. NMPs, hervorgeht. Mit Hilfe von Multiomics-Analysen identifizierte ich gemeinsame und einzigartige molekulare Eigenschaften von AxSCs auf Transkript- und Proteinebene. Durch Differenzierungsexperimente in neurale und mesodermale Nachkommen konnte ich zeigen, dass CFS-Zellen eine höhere Entwicklungsfähigkeit als CS-Zellen haben. Zudem fand ich heraus, dass CS-Zellen andere Eigenschaften haben als anteriore neurale Vorläuferzellen. Schließlich entdeckte ich, dass beide AxSC-Zustände zwischen verschiedenen hPSC-Linien reproduzierbar sind, aber dass ihre Reproduzierbarkeit zwischen verschiedenen Arten von den speziesspezifischen Entwicklungsmechanismen für die axiale Entwicklung abhängt.

Zusammenfassend habe ich zwei neue Zustände von Stammzellen entdeckt, die eine große Chance für die Entschlüsselung und Rekonstruktion der posterioren Entwicklung bieten. Die in meiner Arbeit gewonnenen Erkenntnisse und die vorgeschlagenen Modelle bilden die Grundlage für zukünftige Studien, um AxSCs nicht nur für das Verständnis von embryonalen Entwicklungsmechanismen, sondern darüber hinaus auch für therapeutische Zwecke in klinischen Anwendungen zu nutzen.

ABSTRACT

Early embryonic development occurs through successive division events that ultimately result in lineage segregation and the generation of progenitors for three lineages: ectoderm, endoderm, mesoderm. These progenitors give rise to specific descendants, for example, neural cells are derived from ectodermal progenitors, while somites are derived from mesodermal progenitors. Over the years, many scientists have published that the acquisition of cell fate from pluripotent cells into the three embryonic lineages does not pass through an intermediate multipotent progenitor phase. This paradigm is applicable to anterior embryonic development, but a different mechanism is in place for the posterior embryo, as shown by clonal analysis and lineage tracing experiments.

A common progenitor type for neural and mesodermal cells, called neuromesodermal progenitors (NMPs), has been discovered. These progenitors occur temporally in the stem zone region at the posterior of the embryo and are conserved in human, chick, and mouse embryos. NMPs are the source of the axial structures, the neural tube and the somites. State-of-the-art modalities allow the definition of the molecular features and the environment of steady-state NMPs, paving the way to mimic this developmental state *in vitro*. However, the methods developed for this purpose only allow the generation of the transient progenitor population and have not succeeded in modeling the stem cell counterpart of NMPs.

To recapitulate axial development, I established two human stem cell lines, named CFS and CS, through posteriorization of human pluripotent stem cells (hPSCs). I confirmed that CFS and CS correspond to NMPs and dorsal neural tube progenitors, respectively, and therefore collectively named these cells axial stem cells (AxSCs). I detected a hierarchical regulation between AxSC states, indicating that CS cells represent a further developmental stage derived from CFS cells, hence NMPs. Using multiomics analysis, I identified common and unique molecular features of AxSCs at the transcript and protein level. I validated that CFS cells have a higher developmental capacity than CS cells by performing differentiation experiments into neural and mesodermal progeny. Furthermore, I demonstrated that CS cells have distinct characteristics differentiating from anterior neural progenitors. Finally, I discovered that both AxSC states are reproducible between different hPSC lines, but their reproducibility across species depends on species-specific developmental mechanisms for axial elongation.

In summary, I have discovered two novel stem cell states that offer great opportunities to unravel and reconstruct posterior development. The findings and proposed models within my study will provide a basis for future studies to utilize AxSCs not only to understand developmental mechanisms, but also to better utilize them in future therapeutic approaches.

LIST OF FIGURES

Figure 1: Comparison human and mouse early developmental timelines.....	15
Figure 2: The map of tailbud progenitors.....	17
Figure 3: Axial structures and molecular signatures of the bipotent tailbud region.....	20
Figure 4: Secondary neural tube formation in chick and mouse embryos.....	22
Figure 5: Identification of spinal cord domains	23
Figure 6: Molecular basis of somite subgroup formation and skeletal muscle development.....	26
Figure 7: Neural crest subtypes classified by Hox gene expression.....	27
Figure 8: Summary of published human and mouse NMP derivation protocols.....	34
Figure 9: Establishment of human AxSC lines and their long-term morphology	67
Figure 10: Transcriptional characterization of human AxSCs by the markers of different developmental stages	68
Figure 11: Expression of key transcription factors in human AxSCs on the protein level	69
Figure 12: Relative quantification of telomere length in H9 and human AxSC lines	70
Figure 13: Expression of pluripotency and AxSC marker genes in single-cell resolution	71
Figure 14: Differentially expressed genes in human CFS and CS cells	72
Figure 15: Expression of HES5 and ZIC2 in human AxSCs	73
Figure 16: Embryo illustration showing the key developmental markers for different stages of axial development.....	74
Figure 17: Expression of NMP-associated genes and HOX profile in human AxSC	75
Figure 18: Expression of mesoderm development-associated genes in human AxSCs and H9 cells.....	76
Figure 19: Expression of neural development-associated genes in human AxSCs and H9 cells	78
Figure 20: Expression of spinal cord markers in human AxSCs and H9 cells.....	79
Figure 21: Derivation and characterization of human NPCs from H9 cells	80
Figure 22: Transcriptional differences between human NPCs and human CS cells.....	81
Figure 23: Characterization of AxSCs by global proteome analysis	83
Figure 24: Mitochondrial staining in human AxSCs and H9 cells.....	85
Figure 25: Expression of CD markers at transcript and protein level in human AxSCs and H9 cells included datasets	86
Figure 26: Analysis of chromatin-bound proteins in human AxSCs and H9 cells	87
Figure 27: Expression of pluripotency-related genes in human AxSCs and H9 cells	88

Figure 28: Hierarchical regulation between human AxSC states	89
Figure 29: Reproducibility of AxSC states in hESC and hiPSC lines.....	91
Figure 30: Characterization of neural differentiation cultures from AxSCs	94
Figure 31: Characterization of neural differentiation from NPCs.....	95
Figure 32: Cell cycle analysis time course during the neural differentiation of AxSCs and NPCs	97
Figure 33: Expression of ventral spinal cord markers in human AxSCs and human NPC differentiation datasets.....	99
Figure 34: Expression of neural tube and neural crest-development associated genes in human AxSC and human NPC differentiation datasets.....	100
Figure 35: Expression of undifferentiated human CS cell markers in neural differentiation datasets.....	102
Figure 36: Experimental validation of the markers used to annotate the neural differentiation cultures obtained from human AxSCs and human NPCs	103
Figure 37: Summary of the progeny analysis in human AxSCs and human NPCs differentiation cultures.....	105
Figure 38: HOX expression profile in neural differentiation trajectory from AxSCs and H9 cells as well as in vivo spinal cord cells	106
Figure 39: Characterization of skeletal muscle differentiation cultures from human CFS cells by mimicking segmentation clock.....	108
Figure 40: The effect of horse serum on maturation of skeletal muscle cultures from human CFS cells	109
Figure 41: Immunostaining of skeletal muscle differentiation cultures mimicking segmentation clock in human CFS cells	110
Figure 42: The effect of Notch inhibition with or without SHH activation in human CS cells	111
Figure 43: Characterization of skeletal muscle differentiation cultures from human CFS by mimicking axial elongation steps	112
Figure 44: Immunostaining of skeletal muscle differentiation cultures from human CFS cells by mimicking axial elongation.....	113
Figure 45: Characterization of skeletal muscle differentiation culture from human CS by mimicking axial elongation steps	114
Figure 46: Characterization of nephron differentiation cultures from human CFS and CS cells.....	115
Figure 47: Characterization of cardiomyocyte differentiation from CFS and CS AxSCs.....	116
Figure 48: Characterization of the AxSC cultures derived from mouse embryonic tissues.	117

Figure 49: Establishment and characterization of mouse AxSCs from E14 mESC-based EpiSCs	119
Figure 50: The effect of BMP inhibition and modulation of cytokine concentration on mouse AxSC establishment from E14 mESC-based EpiSCs.....	121
Figure 51: Establishment and characterization of CS lines from orangutan iPSCs	123
Figure 52: Establishment and characterization of CFS cells from mouse embryos	125
Figure 53: Summary of AxSC and NPC differentiation trajectory towards neural lineage ..	133
Figure 54: Proposed model for developmental capacity of human AxSCs	136

LIST OF TABLES

Table 1: Basal media, supplements, ligands, coating and passaging reagents used in cell culture experiments in this study	41
Table 2: Chemicals and kits used in this study	43
Table 3: Probes and primers used in this study.....	45
Table 4: Antibodies used in this study	47
Table 5: Instrument and software used in this study	48
Table 6: Number of embryos per somite stage and number of embryos pooled	53
Table 7: Top 40 differentially expressed proteins	84
Table 8: Marker genes used for annotation of the cellular populations in scRNA sequencing time course datasets	96
Table 9: Conditions tested for skeletal muscle differentiation from human CFS lines	107

LIST OF ABBREVIATIONS

2D	Two Dimensional
3D	Three Dimensional
AA	Ascorbic Acid
aMN	Accessory Motor Neuron
AxSC	Axial Stem Cell
BDNF	Brain-derived Neurotrophic Factor
BMP	Bone Morphogenic Protein
BSA	Bovine Serum Albumin
CAA	Chloroacetamide
cAMP	Cyclic Adenosine Monophosphate
CD	Cluster of Differentiation
CDM	Central Dermomyotome
CHIR	CHIR99021
CLE	Caudal Lateral Epiblast
CM	Cardiomyocyte Differentiation
CNH	Chordo Neural Hinge
CNS	Central Nervous System
Ct	Cycle Threshold
DDA	Data-Dependent Acquisition
DIA	Data-Independent Acquisition
DAPI	4',6-Diamidino-2-Phenylindole
DEG	Differentially Expressed Genes
DEP	Differentially Expressed Proteins
DMEM	Dulbecco's Modified Eagle's Medium
DMH1	Dorsomorphin Homolog 1
DML	Dorsomedial Lip
DNA	Deoxyribonucleic Acid
DNP	Dorsal Neural Progenitor
DRG	Dorsal Root Ganglia
DTT	Dithiothreitol
EB	Embryoid Body
ECM	Extracellular Matrix
EDTA	Ethylenediaminetetraacetic Acid
EMT	Epithelial-to-Mesenchymal Transition
Epi	Epiblast
EpiSC	Mouse Epiblast Stem Cell

ESC	Embryonic Stem Cell
ExESC	Extra Embryonic Stem Cell
FBS	Fetal Bovine Serum
FGF2	Fibroblast Growth Factor
FP	Floor Plate
GDNF	Glial Cell Line-derived Neurotrophic Factor
GO	Gene Ontology
hESC	Human Embryonic Stem Cell
HGF	Hepatocyte Growth Factor
hiPSC	Human Induced Pluripotent Stem Cell
HS	Horse Serum
ICM	Inner Cell Mass
IGF	Insulin-like Growth Factor
IgG	Immunoglobulin G
KSR	Knockout Serum Replacement
LIF	Leukemia Inhibitory Factor
LPMP	Lateral/paraxial Mesoderm Progenitor
mESC	Mouse Embryonic Stem Cell
MN	Motor Neuron
MSC	Multipotent Stem Cell
N	Neural differentiation
NC	Neural Crest
NC_d	Neural Crest derivatives
NEAA	Non-Essential Amino Acid
NMP	Neuromesodermal Progenitor
NotoP	Notochord Progenitor
Np	Nephron differentiation
NPC	Neural Progenitor Cell
NSB	Node Streak Border
oriPSC	Orangutan Induced Pluripotent Stem Cell
P	Parental
p	Passage
PBS	Phosphate-buffered Saline
PG	Paralogous Group
PMA	Purmorphamine
PNC	Pre-EMT Neural Crest Cell
PNP	Posterior Neural Progenitors
PNS	Peripheral Nervous System

PS	Primitive Streak
PSC	Pluripotent Stem Cell
PXM	Paraxial Mesoderm
RA	Retinoic Acid
RN	Rostral Node
RNA	Ribonucleid Acid
RT-qPCR	Quantitative Reverse Transcription Polymerase Chain Reaction
SB	SB431542
SC	Spinal Cord
Sch	Schwann Cell
scRNA	Single-cell RNA
SEM	Standard Error of Mean
SHH	Sonic Hedgehog
SKM	Skeletal Muscle Differentiation
SN	Sensory Neuron
SS	Somite Stage
TCEP	Tris(2-carboxyethyl)phosphine
TE	Trophectoderm
TF	Transcription Factor
TFA	Trifluoroacetic Acid
TGF β	Transforming Growth Factor β
UMAP	Uniform Manifold Approximation and Projection
VLL	Ventrolateral Lip
VNP	Ventral Neural Progenitor
WT	Wild Type

1. INTRODUCTION

1.1. Early embryonic development

Embryonic development starts with the fertilization process; fusion of haploid oocyte and sperm to generate diploid zygote. The zygote can give rise to all types of embryonic and extraembryonic cells required for the gestation of the embryo. Developmental timelines vary between species, but the division events for furcation are conserved stepwise in mammals (Fig. 1), and they are denoted with different terminology in different species, e.g., Carnegie stages (CS) in human, Theiler stages (TS) in mouse and Hamburger–Hamilton stages (HH) in chick embryos, corresponding to different embryonic days (E)¹⁻⁶. The first cell-fate decision occurs at blastocyst stage (E5 in human, E3.5 in mouse), when the blastula divides itself into trophectoderm cells (TE) and inner cell mass (ICM), both cell types showcase morphological differences, and possess vastly different developmental capacity⁷. TE generates the placenta and enables blastocyst implantation in the uterus. ICM gives rise to epiblast (EPI) and primitive endoderm cells (PE or hypoblast)⁸. PE forms yolk sac, which is known as the second extra-embryonic tissue, and EPI differentiates into the three embryonic lineages post gastrulation. Gastrulation starts with formation of a transient structure, the primitive streak (PS), and ends with the segregation of three germ layers: ectoderm, endoderm and mesoderm. PS originates from posterior EPI and differentiates into mesoderm and endoderm layers while anterior EPI gives rise to ectoderm. The three germ layers differ in their developmental potential to generate distinct cell types⁹. Ectoderm is the source of neural and skin cells. Endoderm cells are divided into three groups: foregut, midgut and hindgut. Foregut cells differentiate into lung, liver, pancreas, stomach and thyroid. Midgut and hindgut cells generate small and large intestines respectively. Mesoderm gives rise to the heart, blood, vasculature and musculature.

Over the years, gastrulation theory caused disagreement among scientists¹⁰. One group claimed that gastrulation governs the development from head to tail, whereas the other group suggested that a unique process, separate from gastrulation, is responsible for the posterior development¹⁰. This position was based on evidence indicating that mutations disrupting posterior structures did not affect anterior development severely¹¹⁻¹³. Fate mapping studies in *Xenopus*, chick and mouse embryos showed that, the developing tail contains a progenitor pool capable to differentiate further into the relevant posterior structures: notochord, neural tube and somites, which are required for axial elongation¹⁴⁻¹⁶. The newly identified progenitors were denominated as axial progenitors. The region, where axial progenitors are located, has become known as posterior embryonic growth zone¹⁷.

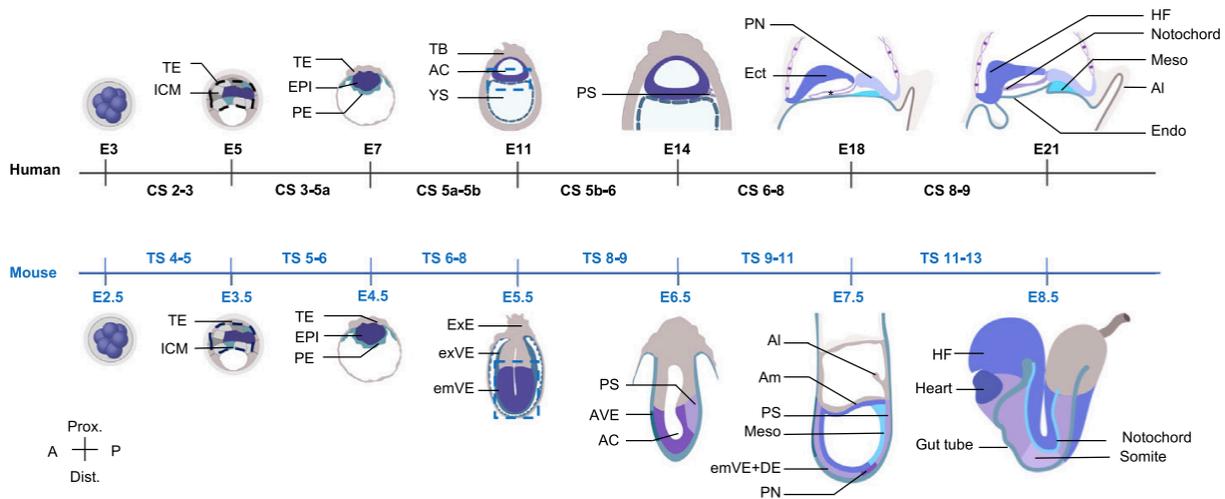


Figure 1: Comparison human and mouse early developmental timelines

Embryonic development is classified by Carnegie stages (CS) in human and Theiler stages (TS) in mouse. Each CS or TS is defined by morphological hallmarks. Embryonic day (E) describes only the period of days post coitum (A: anterior, AC: amnion cavity, Al: allantois, Am: amnion, AVE: anterior visceral endoderm, DE: definitive endoderm, Ect: ectoderm, emVE: embryonic ventral endoderm, Endo: endoderm, EPI: epiblast, ExE: extraembryonic ectoderm, exVE: extraembryonic ventral endoderm, HF: headfold, ICM: inner cell mass, Meso: mesoderm, P: posterior, PE: primitive endoderm, PN: primitive node, PS: primitive streak, TB: trophoblast, TE: trophectoderm, YS: yolk sac) (adapted from¹⁸).

1.2. Axial progenitors

Developing tail, known as tailbud region, in mouse embryos has four main regions: primitive streak (PS), rostral node (RN), node-streak border (NSB) and caudal lateral epiblast (CLE) (Fig. 2a). Homotrophic grafting experiments in mouse showed that each tailbud region has a distinct developmental capacity with regards to their capacity to give rise to axial structures^{19,20}. PS differentiates into somites, intermediate and ventrolateral mesoderm. RN cells are able to give rise to notochord. NSB descendants are located in the neural tube and somites, both at the trunk and tail level. The cells in NSB region migrate to chordo-neural hinge (CNH) region after E10.5²⁰ (Fig. 2b). It has been suggested that the CNH region harbors self-renewing cells. The existence of such population of cells was shown via serial transplantation experiments. Following 48 hours of grafting CNH cells from E10.5-12.5 embryos to anterior PS of E8.5 embryos, second and third generations revealed the presence of cells in the CNH region. This finding suggests that CNH cells retain the developmental potential¹⁵. CLE is the most heterogenous compartment of the tailbud due to its differentiation propensity towards various lineages^{19,20}. It has been divided into five parts named L1 to L5. L1-L3 cells further colonize the somites and neural tube. L1-L3 progeny contribute substantially to the dorsal neural tube with minor integrations in the ventral

neural tube. The observed developmental potential is the opposite of NSB cells. L4-L5 cells can differentiate into lateral mesoderm. L4 progeny is present in the CNH, neural tube and somitic mesoderm while L5 cells are not able to contribute to the neural tube or CNH. Cambrey and Wilson showed that the tailbud cells exhibit a high degree of plasticity^{15,19}. This could be done by performing transplantation of mouse NSB cells into the primitive streak, resulting in the rerouting of its progeny towards lateral mesoderm.

The neuroectodermal and mesodermal cells originating from the same tailbud regions (NSB and CLE) raised the question: Can these cells emanate from a common progenitor? This hypothesis would not fall in line with the traditional gastrulation theory. A breakthrough experiment conducted by Tzouanacou and colleagues investigated lineage segregation in the mouse embryos by performing single cell labeling experiments in utero²¹. The clonal analysis proved that part of the posterior neuroectodermal and mesodermal cells share a common origin in tailbud. The common source was identified as a self-renewing, bipotent progenitor cells named neuromesodermal progenitors (NMPs).

In addition to NMPs, lineage tracing experiments conducted by Wymeersch et al.^{20,22} revealed two other progenitor types in the tailbud region in conjunction with former studies¹⁹. These progenitors were named notochord progenitors (NotoP) and lateral/paraxial mesoderm progenitors (LPMP) based on their developmental axial capacity. NMP, NotoP and LPMP are collectively named axial progenitors, due to their developmental contribution to axial structures and their elongation. Axial progenitors emerge at E7.5 in mouse embryos. LPMPs are exhausted at E8.5 while NotoP and NMPs can be identified until E13.5 coinciding with the completion of somitogenesis²². Initially, NMPs are located in both NSB and L1-L3 CLE regions, NotoPs are present only in NSB, and LPMPs are found in L5 CLE region. Both NMPs and NotoPs reside in the CNH after E9.5^{22,23}.

Tailbud regions exhibit an expression gradient for two developmentally relevant transcription factors, Sox2 and T/Brachyury (encoded by TBXT gene in human)^{20,24,25}. Sox2 expression is higher in rostral parts in contrast to Brachyury which is found to be highly expressed in caudal parts. Based on the expression gradient within the tailbud region, NotoP cells have the highest Brachyury expression, while LPMPs are the cells that least express Brachyury^{20,22}. It can be concluded that axial progenitors can be distinguished by their transcriptional profile in addition to their developmental capacity.

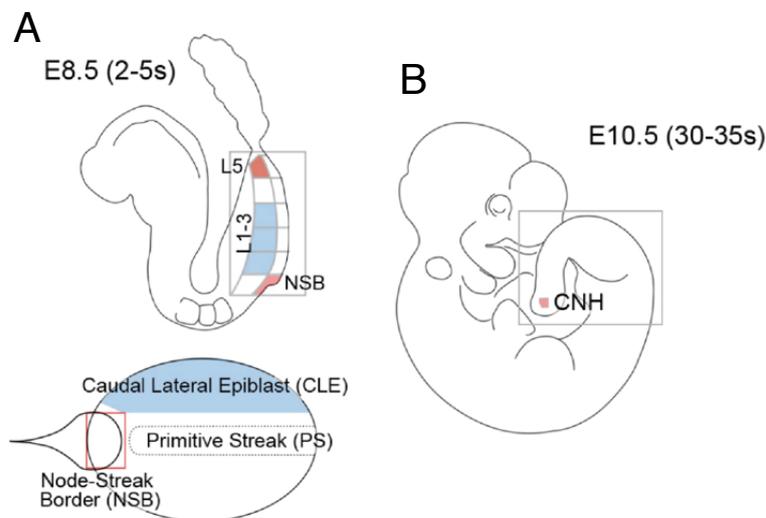


Figure 2: The map of tailbud progenitors

(A) Tailbud of E8.5 mouse embryo corresponding to 2-5 somite stage (s) comprises progenitor cells at caudal lateral epiblast (CLE), node-streak border (NSB) and primitive streak (PS) regions. CLE is subdivided into 5 regions from L1 to L5 based on the developmental capacity of the said cells (adapted from²⁰).

(B) After E10.5 (30-35 somite stage), existing progenitor cells relocate to chordoneural hinge (CNH) region (adapted from²⁰).

1.3. Neuromesodermal progenitors

1.3.1. Signaling pathways involved in NMP maintenance

NMPs are the most striking type of axial progenitors due to their potential to generate both neuroectodermal and mesodermal cells. They are conserved in mouse, human and chick embryos²⁶⁻²⁸. In mouse, they emerge at E7.5, the number of NMPs peaks at E9.5 (22-25 somite stage), they are no longer present after E13.5 (62-63 somite stage) when somitogenesis ceases²². Mouse NMPs are located in NSB and CLE until E9.5 thereafter in CNH. In chick embryos, they arise at HH9 in CLE region and relocate to CNH at HH18^{23,29}.

Steady-state of NMPs is characterized by Brachyury and Sox2 co-expression^{26,27}. Sox2 is a critical transcription factor that plays a role in many developmental processes, such as pluripotency maintenance and neural differentiation. It is highly expressed in ICM cells and its perturbation is lethal at this stage³⁰. Sox2 expression is retained in anterior part of the embryo and decreased in posterior part where NMPs are located, but its high levels can be detected in the differentiating posterior neuroectodermal cells^{20,25,30}. Evidence suggests that Sox2 marks both anterior and posterior neuroectodermal cells. Its expression is driven by N2 enhancer in the anterior embryo and N1 enhancer in posterior embryo^{31,32}. Brachyury is one of the T-box transcription factors. They have a conserved DNA-binding domain in

Metazoa³³. Their absence results in lethality or developmental abnormalities³⁴. Brachyury is first detected in PS cells and plays a role in pluripotency exit together with Eomes, which is also a member of T-box transcription family^{35,36}. Brachyury acts as an upstream factor of TBX genes which are involved in tissue patterning and proper body formation.

FGF and WNT pathways form a gradient through the anterior-posterior axis, but their activity is higher in posterior parts³⁷. Fgf3/4/8/17 are the key FGF ligands expressed in mouse tailbud³⁸. Fgf4/8 play a similar role in chick embryos²⁶. Loss of Fgf4/8 together in mouse embryo causes axial truncation and the downregulation of two WNT ligands: Wnt3a and Wnt5a³⁹. Wnt3a/5a/8a/11 play a role in the axis elongation and proper somite formation in cooperation with FGF signaling in mouse^{40,41}. In addition to the regulatory role of FGF and WNT on axis elongation and its structure formation, these pathways induce Sox2 and Brachyury expression^{27,32,42}. WNT signaling is regulated at the Brachyury level through a positive feedback loop⁴³. Both WNT and FGF pathways are considered as the keys for maintenance and proliferation of undifferentiated and self-renewing NMPs namely steady-state NMPs⁴⁴⁻⁴⁷. NMPs, additionally, have active Notch²² and BMP signaling⁴⁸. Notch is activated via the upstream regulatory role of FGF⁴⁹, while BMP pathway inhibition leads to decreased Brachyury expression in the mouse tail bud⁴⁸. On the other hand, downstream targets of BMP signaling are more enriched in LPMPs compared to NMPs²². It remains unclear whether BMP signaling has a direct effect on NMPs or it is required to orchestrate the formation of the tailbud niche.

Retinoic acid (RA) signaling is the key pathway that negatively regulates the steady-state of NMPs⁵⁰. RA is synthesized by Raldh2 (also known as Aldh1a2) in differentiating paraxial mesoderm cells (PXM) and somites⁵¹, but not in tailbud due to inhibitory activity of Fgf8⁵². Cyp26a1, which is responsible for RA degradation, is expressed in the tailbud^{53,54}. Its expression is gradually downregulated in parallel to the embryo elongation process⁵⁵. Antagonistic expression of Raldh2 and Cyp26a1 generates RA rostrocaudal activity gradient, which enables the tailbud region to be RA-free niche^{56,57}. This gradient is crucial for NMP maintenance and proliferation. It has been shown that perturbation of RA signaling via depletion of RA degradation enzyme in tailbud downregulates Brachyury and Wnt3a expression, leading to the formation of ectopic neural tubes⁵³. RXR γ and RAR γ , which are the isotopes of retinoid X receptor and retinoic acid receptor genes respectively, are expressed within the NMP niche, despite inactive RA signaling^{58,59}.

1.3.2. Molecular characteristics of steady-state NMPs

WNT and FGF pathways promote expression of Cdx genes (Cdx1/2/4) in the caudal embryo⁶⁰⁻⁶². Cdx2 is first expressed in TE cells around E3.5 then in the tailbud region starting from E7.5⁶³. Its absence in TE cells results in embryonic lethality due to abnormalities in extraembryonic development, and lack of Cdx2 expression in tailbud region impairs the axial elongation^{59,63-65}. Its expression has been confirmed in mouse NMPs^{59,66}, and it has been suggested to be one of the prominent factor for developmental potential of NMPs because Cdx2/ Brachyury null mutation exhibits a more severe impact on embryo elongation compared to their individual effects⁶⁵ (Fig. 3). The Cdx gene family is responsible for the induction of HOX genes expression⁶⁷ and Cdx2 alone is sufficient for the regulation of HOX genes^{62,63}. The HOX family comprises 39 transcription factors in four paralogous groups (HOXA-D), which are critical for determination of regional identity along the anteroposterior axis⁶⁸. NMPs are shown to express a plethora of Hox genes^{59,66}. The Hox code does not play a deterministic role in the tail bud NMPs niche. In various transplantation studies, it was showed the NMPs have a high degree of plasticity and change their properties to match the new transplantation site: e.g chick CNH cells, when grafted to caudal node region of early-stage embryos, no longer expressed posterior HOX genes resembling the surrounding host tissue¹⁶.

Nkx1-2 is widely accepted as an NMP marker^{59,69}, although Nkx1-2 expression could not be detected in chick NMPs, similarly to Cdx2⁶⁶. Evx1 and Mnx1 are the genes commonly expressed between mouse and chick NMPs^{22,66}. Evx1 is WNT pathway target and it functions a regulator of anteroposterior patterning⁷⁰. Mnx1 is a well-known marker for motor neurons and pancreatic cells^{71,72}, but its role in NMPs has not been elucidated yet. Oct4, the master regulator of pluripotency, is not expressed in posterior embryo including the NMPs⁷³. It has been found that three of the pluripotency-related factors Sall4 and Lin28a/b are expressed in tailbud region. Their expression has been linked to the increase in numbers of NMPs⁷⁴⁻⁷⁶. NMPs undergo drastic transcriptional changes between E8.5 and E9.5 in mouse^{22,59,66}. Early mouse NMPs have higher expression of Nkx1-2, Wnt8a and Wnt3a while late NMPs are marked by the upregulation of Hox genes (Hox3-Hox13), Sox2, Cdh2, Cyp26a1 and Wnt5a^{59,66}.

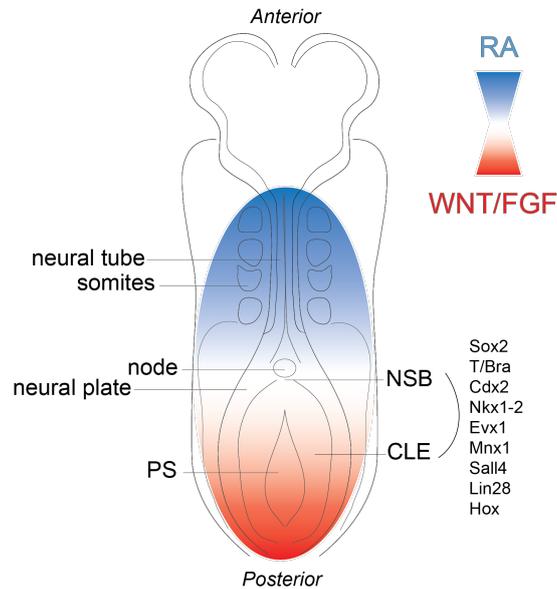


Figure 3: Axial structures and molecular signatures of the bipotent tailbud region.

1.3.3. Lineage commitment of NMPs and development of axial structures

The NMPs are located in the NSB and CLE region of the posterior embryo. It has been demonstrated that they can contribute to both posterior neural and mesodermal lineages^{19-21,45}. It has been suggested that NMP derivatives can have a bias based on their location. The NSB cells have equal potential for neural and mesodermal development. The CLE cells show different propensity to these lineages depending on anterior/posterior or lateral/medial localization within the region²⁰. L1 CLE cells (Fig.2) possess higher tendency for neural contribution than L2-L3 CLE cells which are prone to contribute to somites. Cell fate choice of the NSB or CLE progeny relies on RA, WNT and FGF pathways⁴⁷. Cyp26 null mice have downregulated Brachyury expression in tailbud region resulting in impaired axial elongation and aberrant Sox2 expression, which results in ectopic neural tube formation at trunk level⁵³. These findings highlight that excessive RA steers the cells to the neural fate via modulation of Brachyury and Sox2 expression. Elevated levels of FGF in the tailbud region expands the pool of mesodermal progenitors, while its inhibition leads to ectopic Sox2 expression²⁶. Similarly, continuous WNT activation upregulates Brachyury expression inferring mesodermal commitment, while blocking WNT pathway, as shown for Wnt3a null embryos, drastically decreases the number of mesodermal progenitors and increase neural progenitors compared to WT embryo⁴⁵.

1.3.3.1. Neural fate

Neural fate of NMPs is driven by RA and FGF pathways^{58,77}. RA synthesis is initiated in the differentiating PXM and it is found at high levels further in the somites⁵¹. This behaviour is the opposite of FGF pathway regulation⁷⁸. It has been shown that RA activation alone is able to activate neural differentiation, which is further promoted by FGF inhibition through downregulation of Brachyury and upregulation of early neural progenitor markers such as *Irx3*^{58,79}. Blocking of RA synthesis in *Raldh2* null embryos elicits ectopic *Tbx6* expression resulting in expansion of PXM progenitors into the NMP containing regions⁸⁰. During neural commitment, expression of *Sox2*, which has a role in stimulation of neural control genes, is upregulated therefore it is thought to be the master regulator of differentiation event⁴⁶.

Contribution of NMPs to neural tube has been shown by grafting experiments^{19,20,45}. The neural tube generates the central nervous system (CNS), consisting of brain and spinal cord. The development of the neural tube after formation of neural plate is studied as two consecutive processes; primary and secondary neurulation. It has been shown that these two processes are governed by independent developmental mechanisms as secondary neurulation can occur even when primary neurulation is impaired⁸¹. Primary neurulation, originating from neuroectodermal cells, consists of consecutive folding and closure events of the neural plate leading to formation of the brain and anterior part of the neural tube. The first closure event occurs at E8.5 in mouse embryos at the level of the cervical/hindbrain boundary, followed by the second closure at the forebrain/midbrain boundary and the third closure at the anterior forebrain⁸².

Secondary neurulation involves unique process that gives rise to the posterior neural tube called secondary or elongating neural tube. Tail bud cells were initially designated as the source of secondary neurulation⁸³ and the contribution of NMPs to secondary neural tube has been demonstrated by lineage tracing experiments²⁸. Secondary neurulation can be morphologically discriminated from primary neurulation as this process is driven by cavity formation instead of folding. Catala et al.⁸³ depicted the cascade of secondary neurulation events in chick embryo (Fig. 4A). Tailbud cells (Fig.4A-1) first undergo mesenchymal-to-epithelial transition and condensation at the midline of dorsal region, which gives rise to a structure named as medullary cord (Fig. 4A-2). The medullary cord is comprised of two cell types: peripheral cells with elongated morphology and central cells with heterogeneous morphology. In parallel to the formation of these cells within the dorsal part, the medullary cord extends ventrally (Fig.4A-3) followed by formation of multiple lumens (Fig. 4A-4). The lumens will merge (Fig. A-5) to create the central canal of the neural tube (Fig. A-6). The secondary neural tube is elongated up to the lumbar level and its ventral part meets the dorsal part of primary neural tube in the region named junctional zone⁸⁴⁻⁸⁶. Secondary

neurulation in human embryo has been described as very similar to that of chick embryos, but there is a contradiction regarding the presence of multiple lumens and the junctional zone⁸⁵⁻⁸⁷.

Several morphological and anatomical changes can be observed in mouse embryos compared to the chick embryos⁸⁸ (Fig. 4B). Tail bud cells (Fig. 4B-1) form the medullary rosette (Fig. 4B-2) including only one cell type, morphologically similar to the peripheral cells in chick medullary cord. There are no simultaneous distinct events occurring along the dorsoventral axis in comparison to the chick embryos. The medullary rosette forms the sole lumen structure (Fig. 4B-3-5) as multiple lumens and the junctional zone do not exist in mouse embryos.

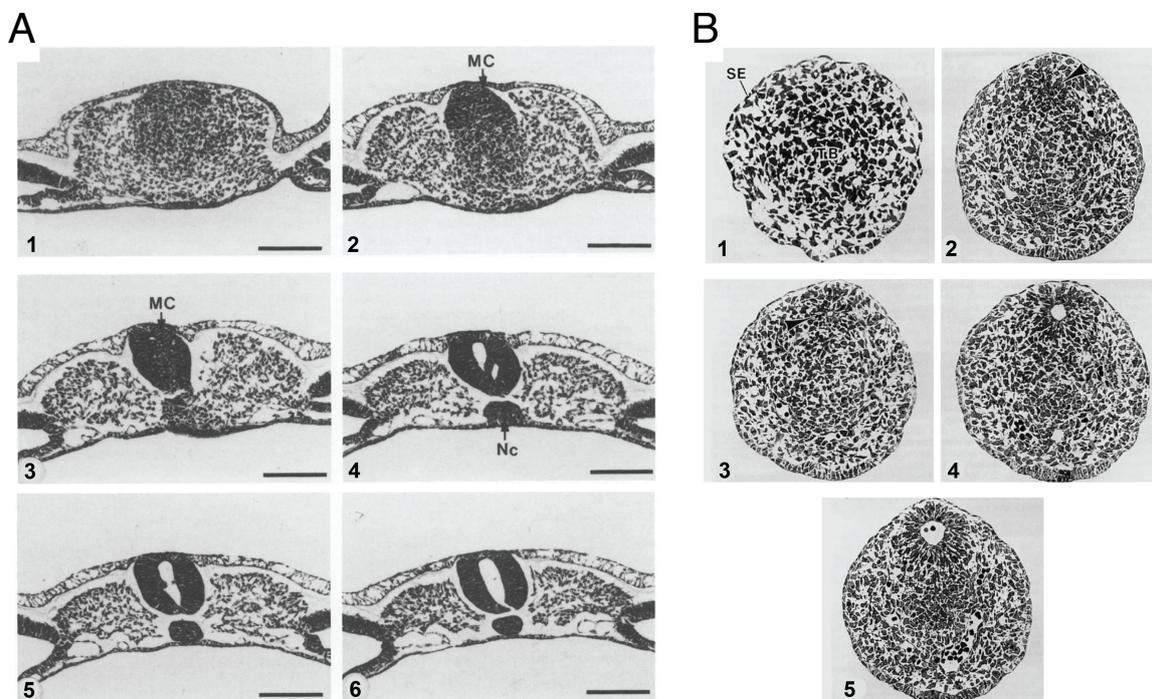


Figure 4: Secondary neural tube formation in chick and mouse embryos

(A) In chick embryo, medullary cord (MC) cells originated from tailbud cells first accumulate at dorsal midline and progress ventrally. Multiple lumens are formed simultaneously and fuse later resulting in a single lumen called central canal (Nc: notochord) (adapted from⁸³).

(B) In mouse embryo, tailbud (TB) cells generate medullary rosette further forming the single lumen structure (adapted from⁸⁸).

A recent study demonstrated the involvement of mouse NMPs in secondary neural tube formation. The study claimed that the developmental timing of NMPs can affect their progeny contribution. They could detect derivatives of early NMPs up to thoracic level while derivatives of late NMPs were present at sacral level²⁸. Additionally, Shaker and colleagues

showed that NMP behavior during the secondary neural tube formation diverges across different species similarly to the morphological differences observed in secondary neurulation between chick and mouse embryos²⁸. Brachyury+ cells are present in the elongating mouse neural tube, unlike the chick, which has Brachyury+ cells only in tailbud region. In both species, Brachyury is downregulated in neural progenitors derived from NMPs followed by the upregulation of Pax6 and Hes5⁶⁶.

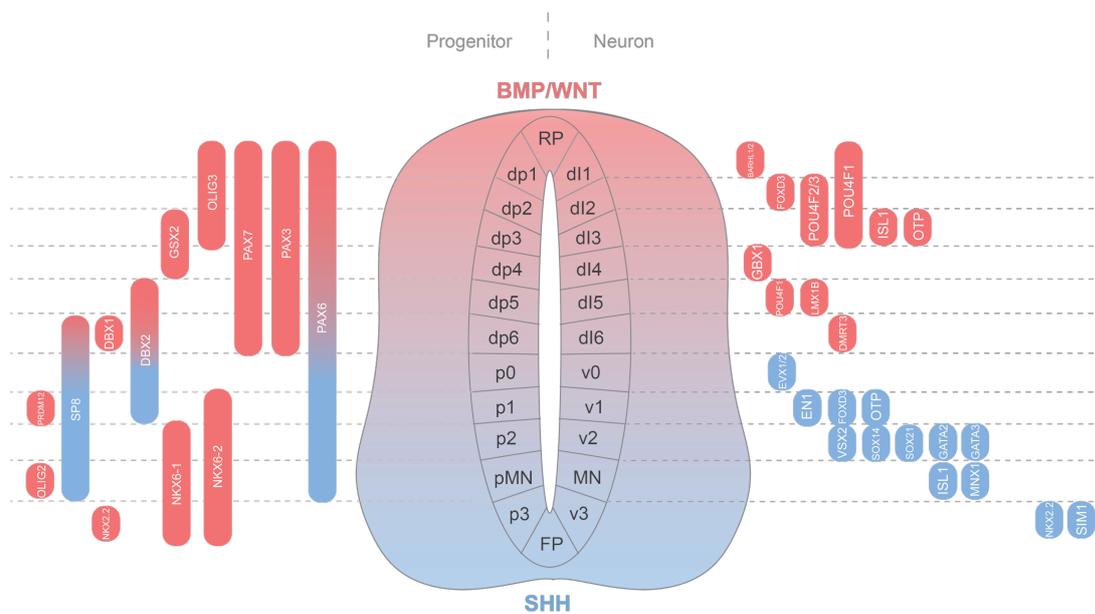


Figure 5: Identification of spinal cord domains

Dorsoventral gradient between BMP/WNT and SHH signaling forms distinguishable spinal cord domains. Spinal cord progenitor cells (left) generate post-mitotic neurons (right) at later developmental stages. Domains are identified based on their respective transcriptional pattern.

Neural progenitors acquire distinct identities within the neural tube that continue to further differentiate into numerous neuronal subtypes. Sequentially secreted SHH from notochord and floor plate, and BMP/WNT ligands secreted from roof plate generate a dorsoventral gradient⁸⁹. These signaling pathways activate Gli repressor proteins (GliR) at dorsal level and Gli activator proteins (GliA) at ventral level thereby generating a Gli gradient, which leads to the formation of subgroups within dorsal and ventral spinal cord regions⁹⁰. Dorsal neural progenitors (DNPs) comprise of 6 subgroups named dp1 to dp6. Each subgroup of DNPs has a specific transcriptional identity (Fig. 5), but they are commonly marked by high IRX3, IRX5, PAX3 and PAX7 expression, in addition to the pan-neural progenitor marker PAX6. Ventral neural progenitors (VNPs) are divided into 5 subgroups: p0-p3 and pMN. The majority of VNPs express either NKX6-1/2 or SP8. DNPs and VNPs differentiate into terminal neurons between E9.0-13.5 in mouse embryos. DNPs (dp1-dp6) and a part of VNPs (p0-p3) mature

into interneurons and are named dl1-dl6 and v0-v3 respectively. Only one of the VNPs subgroups, namely the pMN, is responsible for the formation of motor neurons (MN) which are part of peripheral nervous system (PNS). Each spinal cord domain has a distinct transcriptional pattern as indicated in Fig. 5.

1.3.3.2. Mesodermal fate

Brachyury expression is the initial determining step towards mesodermal specification for NMPs. High levels of Brachyury lead to positive autoregulatory loop within WNT signaling pathway and repression of RA signaling through upregulation of *Cyp26a1* within the NMP niche^{43,91}. The WNT and FGF pathways cumulatively promote mesodermal commitment through consecutive induction of Brachyury and *Tbx6*, both of which are known to inhibit *Sox2* expression^{26,32,45,46}. Disruption of *Tbx6* expression, similarly to disruption in *Cdx2* expression, leads to *Sox2* expression and ectopic neural tube formation in mouse embryos^{32,92}.

It has been shown that the level of BMP signaling influences the submesodermal specification during embryonic development⁹³. Low levels of BMP are required for PXM lineage specification. High and intermediate levels of BMP lead to the formation of lateral and intermediate mesoderm respectively. The contribution of NMPs with regards to somite formation has been demonstrated by numerous studies. The somites are PXM derivatives. A recent finding suggested that *Tbx6* expressing NMP descendants can generate nephric mesenchyme, which is an intermediate mesoderm derivative. The nephric mesenchyme was generated through *Osr1* upregulation from BMP signaling⁹⁴. The derivation mechanism of intermediate mesoderm progeny from NMPs is not well-studied. PXM formation has been shown upon inhibition of neural differentiation by disrupting *Sox2/Sox3* expression in CLE NMPs⁹⁵.

During PXM development, *Tbx6* activates *Hes7* and *Mesp2* which are known to regulate Notch oscillation. The fluctuating activity of Notch is called segmentation clock and it is responsible for somite formation⁹⁶⁻⁹⁸. *Hes7* has an autoregulatory loop that can activate or repress its own expression leading to a negative correlation between *Hes7* transcription and its protein level⁹⁹. The autoregulatory loop initiates *Hes7* expression at posterior PXM and enables its progression anteriorly, functioning as a timing determinant of the segmentation clock¹⁰⁰. *Ripply2* is downstream of *Mesp2* and it negatively regulates *Mesp2* expression, resulting in a negative signaling feedback loop¹⁰¹. *Mesp2* is expressed in the anterior PXM and it acts as a determination factor for PXM-somite boundary or the so-called determination front^{102,103}. *Hes7* and *Mesp2* regulate *Lfng* (lunatic fringe) expression that acts as Notch

signaling inhibitor. The Lfng inhibition is mediated through competitive binding to Notch1 (Notch receptor) instead of Dll1 (Notch ligand), resulting in Notch signaling disruption in the domains where Hes7 and Mesp2 are expressed¹⁰⁴. In the domains with active Notch signaling, successful binding of Notch ligands results in the cleavage of NCID (Notch intracellular domain) that activates its target genes including Hes7 and Mesp2¹⁰⁴. These cycling events result in somite formation and embryo elongation at determination front. New somites are formed every 1.5, 2 and 4-6 hours in chick, mouse, and human embryos respectively^{105,106}. Mature epithelial somites express Meox1/2 and Pax3/7. They give rise to two populations: dermomyotome and sclerotome depending on the signals provided by the surrounding tissues¹⁰⁷⁻¹⁰⁹.

Sclerotome specification, induced in the ventral somites, is orchestrated via a cross play between notochord secreted SHH and BMP inhibition via Noggin expression originating in the notochord and dorsal somites^{110,111} (Fig. 6A). Sclerotome cells undergo epithelial-to-mesenchymal transition (EMT). They are marked by expression of Nkx3.2, Pax1/9 and Sox9, and the downregulation of Pax3/7¹¹²⁻¹¹⁴. After migration towards several routes, Sclerotome cells are divided into four subgroups: dorsal (Msx1/2+), central (Uncx4.1), ventral (Pax1/9+) and lateral (VEGFR2+/Sim1+)^{115,116}. Each sclerotome subgroup presents a distinct profile for their respective progeny. Dorsal cells generate dorsal neural arch and spinous processes. Ventral cells give rise to the vertebral body and intervertebral disks. Central cells derive the neural arch and proximal ribs. Lateral cells produce tendons, endothelial cells and distal ribs^{115,117}.

The dorsal somites do not undergo the EMT process, thus they retain their epithelial morphology. Exposure to WNT signaling secreted from the roof plate induces differentiation of dorsal somite towards the dermomyotome¹¹⁴. Dermomyotome cells are characterized by Pax3 expression (Fig. 6A) and yield dermatome and myotome that further differentiate into dermis and skeletal muscles respectively¹¹⁸. The dermomyotome comprises of three compartments: dorsomedial lip (DML), central dermomyotome (CDM) and ventrolateral lip (VLL)¹¹⁹. DML and CDM contribute to both dorsal dermis and myotome, whereas VLL gives rise to the myotome only^{118,120-122}.

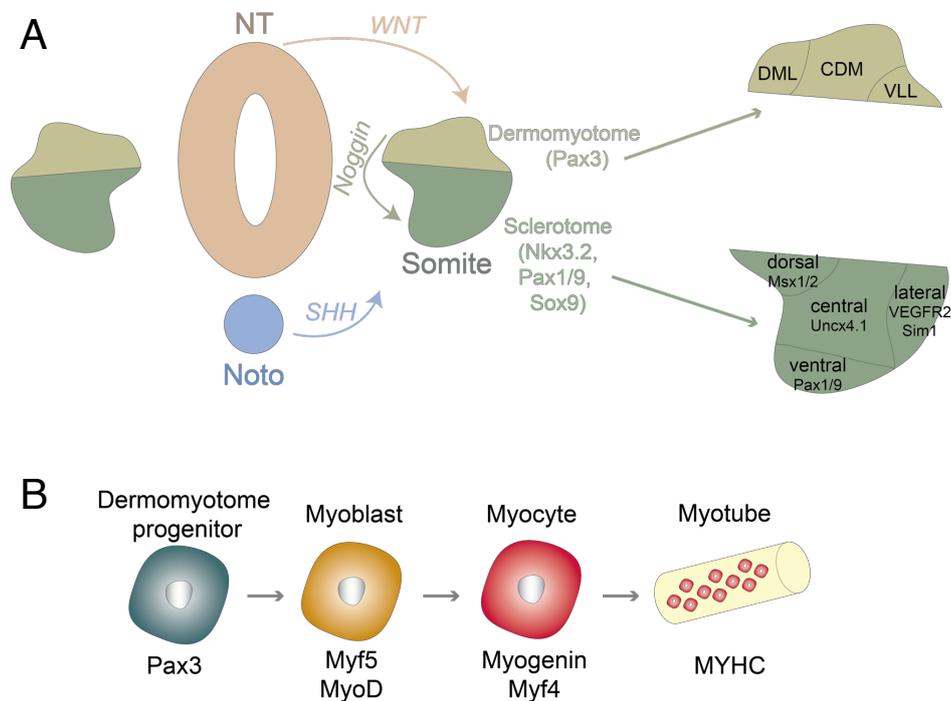


Figure 6: Molecular basis of somite subgroup formation and skeletal muscle development

The myogenesis process is initiated through an interplay of WNT signaling from neural tube, SHH signaling from notochord and Notch signaling activated by neural crest cells, leading to myoblasts formation by upregulating Myf5 and MyoD expression¹²³⁻¹²⁵ (Fig. 6B). Myoblasts further differentiate into myocytes by withdrawing from the cell cycle and activating muscle-specific genes such as Myogenin and Mrf4^{126,127}. This step is followed by myocyte fusion to generate multinucleated myotubes and myofibers that are marked by MYHC expression¹¹⁹. Suppression of FGF signaling is important for the proper skeletal muscle development, as it negatively regulates Myogenin expression, resulting in failure of the myocyte fusion process¹²⁸⁻¹³¹. Disruption of Myf5 and MyoD expression causes an impaired myogenesis¹³², but the absence of MyoD could be compensated by Mrf4 expression¹³³. Taking into account these findings, Myogenin, Myf5, MyoD and Mrf4 were collectively named as myogenic regulatory factors¹³⁴.

1.3.3.3. Neural crest fate

A small fraction of CLE cell derivatives have been found as contributing to trunk neural crest in mouse embryos^{20,69,135}. This finding falls in line with the previous evidences indicating that neural tube and trunk neural crest cells have a common origin^{21,136}. It has been found that loss of Cdx proteins induce a downregulation of neural crest-associated genes at the trunk

level such as Pax3 and Msx1, leading to the developmental impairment of neural crest descendants such as PNS¹³⁷.

A transient population of neural crest precursors namely pre-EMT and pre-migratory neural crest cells (PNCs) is located in dorsal neural tube¹³⁸ and it is characterized by expression of Sox9, Pax3, Zic1/3/5, Msx1, and Gdf7 expression¹³⁹. PNCs undergo Snail mediated EMT to exit the neural tube and migrate dorsally or ventrally thereby leading to their subsequent replacement by roof plate cells¹⁴⁰⁻¹⁴². PNCs downregulate Bmp6 expression and concomitantly with migration Ets1, Snai1, Sox10 are upregulated as common markers for neural crest derivatives¹³⁹.

Spatial organization along anteroposterior axis is a determining factor for the subtype identity of the neural crest progeny¹⁴³ (Fig. 7). Axial identity of NCs is determined by expression of Hox genes. Forebrain, midbrain and the first rhombomere of hindbrain do not express Hox genes¹⁴⁴. The rest of rhombomeres express Hox paralogous group (PG) 1-3, vagal cells Hox PG3-7, trunk cells Hox PG6-10, sacral cells Hox PG11-13¹⁴⁵. Cranial neural crest cells differentiate to craniofacial skeleton (cartilage and bone), cranial ganglia, odontoblasts, pigment and thyroid cells. Vagal neural crest cells generate enteric ganglia, smooth muscle cells and cardiac septa. Trunk neural crest cells give rise through ventral migration between neural tube and somite to the dorsal root ganglia (DRG), Schwann cells and sympathetic ganglia. Through dorsal migration between somites and the ectoderm layer, the trunk neural crest cells give rise to the adrenal medulla and pigment cells¹⁴⁶. Sacral trunk neural crest cells differentiate into enteric ganglia and sympathetic ganglia.

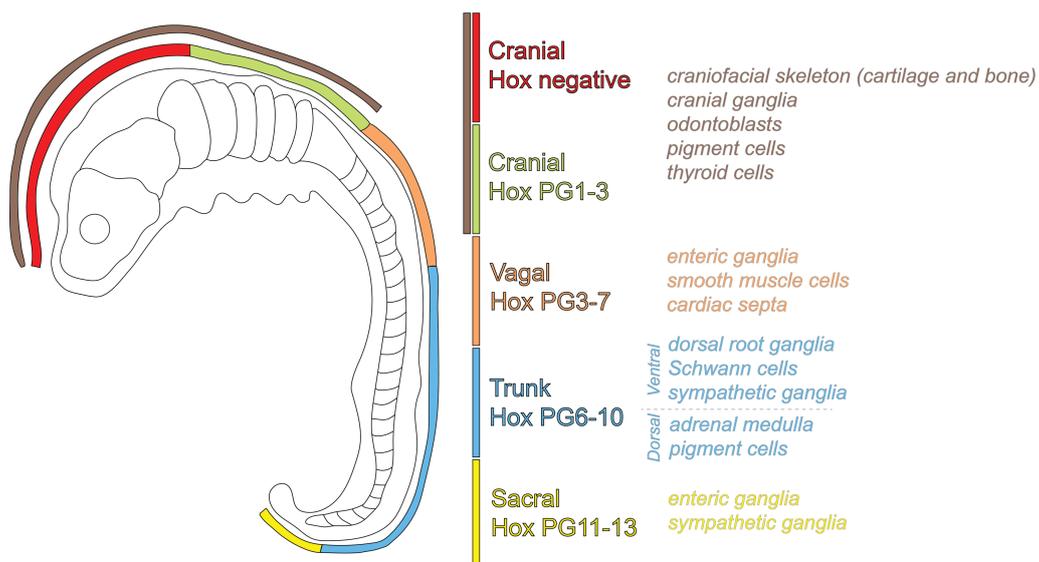


Figure 7: Neural crest subtypes classified by Hox gene expression

Lineage tracing experiments in mouse embryos showed the contribution of NMP descendants to the dorsal root ganglia (DRG)²⁸. DRG cells are categorized into two groups: sensory neurons and non-neural cells. DRG is primarily comprised of the soma of sensory neurons and it is considered a part of peripheral nervous system (PNS) that conveys signals to CNS. Sensory neurons can be identified by *Isl1* and *Pou4f1* (*Brn3a*) expression at E10.5 in addition to *Pou4f2* (*Brn3b*) starting from E11.5^{147,148}. Additional cell types found in the DRG are the satellite glial cells, endothelial and smooth muscle cells¹⁴⁹. Non-neuronal cells are characterized by expression of *Pard3*, *Qki*, *Fkbp5* and *ApoE*¹⁵⁰.

1.4. Stem cells

Stem cells have been widely used to recapitulate the developmental processes in vitro. It is possible to commit stem cells to a specific cell fate by small molecules and cytokine modulations. Stem cells are classified based on their 1) origin such as embryonic and adult stem cells (ESCs and ASCs respectively), 2) region such as ESCs and extraembryonic stem cells (ExESCs), 3) developmental capacity gradually decreasing from totipotent to unipotent stem cells. Totipotent stem cells serve as a snapshot for a specific cell population which emerges in a very short developmental timeline within the first two division cycles^{151,152}. These cells are the origin of both ESCs and ExESCs. Pluripotent stem cells (PSCs) capture the characteristic of the ICM cell population as they can generate all embryonic cell types derived from the three germ layers^{153,154}. Human PSCs are furthermore able to differentiate into extraembryonic cells, e.g. trophoblast cells¹⁵⁵. Multipotent stem cells (MSCs) are capable to produce different types of cells originating within a particular lineage, and unipotent stem cells give rise to only one type of cells¹⁵⁶. Considering their developmental capacity and ease of handling, PSCs and MSCs are extensively used and well-characterized types of stem cells.

Despite each stem cell type having distinct features, their shared trait is the ability to self-renew indefinitely that enables robust in vitro cultures. Self-renewing enables an extended proliferation capacity which is conferred by the maintenance of telomere length. The telomere is a sequence found at the 3' end of each chromosome¹⁵⁷. It is conserved in all eukaryotes, but there are species-specific variations in sequence or size^{158,159}. In somatic cells, each round of DNA replication results in shortening of the telomere length due to end-replication problem that limits the number of division cycles¹⁶⁰. Telomerase, a reverse-transcriptase coded by *TERT* gene and highly conserved in mammals¹⁶¹, plays a role in helping to overcome the end-replication problem. It elongates the DNA strand as a result of binding an RNA sequence known as telomerase RNA component coded by *TERC* gene¹⁶². Shelterin complex accompanies the telomerase activity by forming a T-loop structure to protect the 3' overhang from exonuclease degradation. The shelterin complex consists of

proteins produced by *TRF1*, *TRF2*, *POT1*, *RAP1*, *TIN2*, *TPP1* genes¹⁶³. The maintenance of telomere length by the abovementioned processes permits the indefinite self-renewing capability of stem cells. The telomerase activity correlates with the developmental capacity, meaning that its activity is higher in pluripotent stem cells compared to multipotent stem cells¹⁶⁴.

1.4.1. Pluripotent stem cells

Pluripotent cells were first isolated from mouse preimplantation blastocyst¹⁶⁵ and then from human ICM¹⁶⁶. After prolonged expansion in vitro, these cells exhibited sustained pluripotency by giving rise to all cell types comprising an individual thus, they were named PSCs. It has been shown that there are two states of PSCs with unique properties corresponding to different developmental stages¹⁶⁷⁻¹⁶⁹. These states are termed as naïve and primed PSCs representing ICM of pre and postimplantation blastocyst respectively. In mouse, naïve PSCs are called mouse embryonic stem cells (mESCs) and primed PSCs are denominated as mouse epiblast stem cells (EpiSCs). In human, both states are known as embryonic stem cells however human embryonic stem cell (hESC) term is conventionally referring to the primed state unless it is specified as naïve. Both states can differentiate into the three lineages, but it has been demonstrated that naïve cells are more pluripotent compared to primed cells due to their chimera formation ability and X chromosome state¹⁷⁰. It is possible to convert the states from naïve to primed or vice versa by using transgenes or morphogenes¹⁷¹⁻¹⁷⁴. Naïve and primed state cells show similar gene expression and epigenetic profiles between human and mouse^{167,175}.

Oct4 (*Pou5f1*) and Sox2 have been identified as the core pluripotency factors due to their expression being present in both naïve and primed states^{176,177}. Nanog is also considered as a pluripotency factor, but it has been shown that mouse ESCs can be maintained without Nanog expression¹⁷⁸. The hallmark of naïve PSCs is considered *Klf2/4*, *Fgf4* and *Rex1* expression^{171,179-181}. Expression of *Otx2*, *Zic2/3* and *Lin28a* promotes the exit from naïve state to primed state¹⁸²⁻¹⁸⁴. Essential genes for the human pluripotency network have been described as essentialome which is comprised of *MYBL2*, *SALL4*, *POU5F1*, *PRDM14*, *NANOG*, *FOXB1* and *MYCN*¹⁸⁵. Transcriptome analysis of human naïve and primed state ESCs at the single cell level revealed that *DPPA3/5* and *FGF4* are specific to naïve ESCs, and *POU5F1*, *TDGF1* and *KLF4* are downregulated concomitantly with the transition from naïve to primed. The latter is marked by *ZIC2* and *SOX11* expression¹⁸⁶. In mouse, *Zic2* has been shown as acting downstream of *Zfp281*¹⁸⁷ which is another factor regulating naïve to primed conversion in both human and mouse¹⁸⁸. *Zfp281* interacts with *Tet1*, which is slightly downregulated from naïve to primed state, but not with *Tet2*, which is highly abundant in

naïve state^{188,189}. Tet proteins regulate global DNA methylation together with Dnmt proteins. In the naïve state, Tet1/2 erase the methylation marks and de novo methylation occurs upon primed state transition by increasing Dnmt3a/b occupation on the genome¹⁸⁹⁻¹⁹². Dnmt1, unlike Dnmt3a/b, is responsible for maintenance of the methylation profile and it is expressed at similar levels between naïve and primed state¹⁹².

PSCs are maintained by exogenously modulated signal transduction. It is possible to grow PSCs on feeders or feeder-free culture. Leukemia Inhibitory Factor (LIF) produced by feeder cells induces expression of Oct4, Klf4, Nanog and c-Myc through activation of Stat3¹⁹³⁻¹⁹⁵. It has been shown that LIF is sufficient to maintain pluripotency of mouse ESCs but it is insufficient for the maintenance of human naïve ESCs¹⁹⁶. LIF can be substituted with WNT induction in combination with FGF inhibition¹⁹⁷. FGF pathway activates MEK and ERK signaling, leading to the negative regulation of naïve pluripotent state by inducing lineage commitment¹⁹⁸, thus these pathways need to be blocked by using FGF receptor inhibitor or MEK inhibitor¹⁹⁹. Activation of WNT signaling results in binding of the destruction complex (GSK3 β , Axin, APC) to the WNT receptor, thus preventing β -catenin phosphorylation and allowing nuclear relocation of β -catenin and Tcf3 activation subsequently. Tcf3 acts as a transcriptional repressor of pluripotency-related genes such as Oct4 and Sox2²⁰⁰. However, Xu and colleagues suggested that active WNT signaling supports self-renewal of human naïve ESCs and it does not affect expression of the pluripotency factors²⁰¹.

Smad proteins are important for pluripotency regulation. There are two types of Smad proteins: activatory Smads (Smad1-3/5/8) and inhibitory Smads (Smad 6/7). Smad2/3 and Smad1/5/8 have a role in transduction of Activin/Nodal/TGF β signaling and BMP signaling respectively²⁰². Smad6/7 inhibit BMP signaling, while Smad7 only has an inhibitory effect on Activin/Nodal/TGF β signaling. Expression of Smad7 in mouse ESCs decreases their proliferation rate while Smad6 has no significant effect²⁰³. It has been suggested that Activin/Nodal/TGF β signaling but not BMP signaling is required for mouse ESC proliferation and pluripotency maintenance²⁰³. There are several studies showing Smad1/5 proteins in mESCs have regulatory activity on KLF genes suggesting that BMP signaling is also important for mESC maintenance²⁰⁴. In human naïve ESCs, inhibition of TGF β signaling downregulates expression of pluripotency-related genes (Nanog and Oct4)²⁰⁵, and inhibition of Nodal and BMP leads to differentiation²⁰⁶.

Signaling pathways actively involved in primed pluripotency maintenance are well-characterized. WNT and BMP signaling respectively induce primitive streak and trophoblast differentiation by upregulation of T/Brachyury and APA leading to exit from pluripotency in both EpiSCs and hESCs states¹⁷⁴. In addition to WNT, blocking the FGF pathway via MEK

inhibition promotes differentiation or cell death leading to a disruption of the primed pluripotency state²⁰⁷. Activation of Activin/Nodal and FGF signaling simultaneously supports maintenance of primed PSCs by upregulation of Oct4 expression²⁰⁸. It is possible to convert primed PSCs to naïve PSCs by using transgenes or chemical induction. Exogenous activation of Klf4 and Nanog supported by 2iLIF medium (MEK inhibitor PD0325901, GSK3 β inhibitor CHIR99021, recombinant LIF) reverts EpiSCs to ESCs^{171,172}. Inhibition of MAPK pathway with PD0325901 and SB590885 (B-raf inhibitor), inhibition of LCK/SRC kinases with WH-4-023, inhibition of GSK3 β with IM-12 alongside activation of LIF, Activin A and FGF pathways reverses the primed human PSCs back to naïve state¹⁷³.

Conversion of the mature cells to earlier developmental stages is not utilized only for primed to naïve PSCs transition. Exogenous activation of four transcription factors Oct4, Sox2, Klf4 and c-Myc, namely the Yamanaka factors, under the permissive culture conditions allows for mouse and human somatic cells to be reprogrammed back to the pluripotent state. The cells obtained from a successful reprogramming process were termed induced pluripotent stem cells (iPSCs)²⁰⁹⁻²¹¹. Isolation of human ESCs entails embryonic manipulations that carry ethical and legal issues. The use of iPSCs is a practical alternative for ESCs. iPSCs and ESCs share similarities in their transcriptome, chromatin regulation, morphology and teratoma formation potential^{210,212,213}. Slight changes in the iPSC differentiation propensity compared to ESCs have been demonstrated. Differentiation tendency of iPSCs towards the embryonic lineages is mostly affected by the type of donor cells undergoing reprogramming. It has been shown that iPSCs derived from endothelial cells differentiate more efficiently towards endothelial cell in comparison to fibroblast and cardiac progenitor differentiation²¹⁴.

1.4.2. Multipotent stem cells

Neural progenitor cells (NPCs) are an instance of MSCs with limited developmental capacity in comparison to PSCs. NPCs can be isolated from embryonic or adult tissue²¹⁵⁻²¹⁷. There have been numerous protocols published for in vitro NPCs derivation from hPSCs. The differentiation protocols follow a step-wise procedure starting with dual-SMAD inhibition (BMP and TGF β pathway inhibition)²¹⁸. Upon successful differentiation nestin expressing neuroepithelial cells can be obtained from hPSCs either from 2D or 3D culture. Neuroepithelial cells can be differentiated into neural rosettes which are considered as NPCs²¹⁹. NPCs are marked by ZO1, SOX1/2 and PAX6 expression thus mimicking the embryonic neural tube and they can be further differentiated into different neuronal cell subtypes resembling in vivo PNS/CNS neurons, oligodendrocytes and astrocytes²²⁰⁻²²³.

NPCs differentiate into derivatives of the anterior embryo, but it is possible to posteriorize them by WNT, FGF and RA induction upon differentiation onset²²⁴⁻²²⁷. It has been shown that differentiation modalities employing a posteriorization step produce transient NMP population before NPCs emerge. Continuous activation of FGF and WNT enhances self-renewal of NPCs and their long term maintenance in vitro²²⁸.

1.5. Neuromesodermal progenitors in vitro

It has been demonstrated that NMPs can be derived from human and mouse PSCs in vitro by mimicking the signal transduction at the posterior part of embryo. WNT and FGF signaling are therefore the key components of NMP derivation process. Nevertheless, various protocols of which WNT and FGF are supported by additional morphogenes with the usage of diverse coating materials have been established.

1.5.1. Derivation of NMPs

1.5.1.1. Human NMPs

Gouti and colleagues published in 2014 that SOX2 and TBXT double positive cells, namely NMPs, can be generated within 3 days in vitro by differentiating hPSCs on fibronectin coated plates supplemented with 3 μ M CHIR99021 and 20 ng/ml FGF2, in order to activate WNT and FGF signaling respectively⁴⁴. Pluripotency markers NANOG and OCT4 are downregulated in NMPs compared to hPSCs, however OCT4 expression is retained. The initial establishment protocol resulted in a heterogenous culture containing SOX2⁺/TBXT⁻ cells and TBX6⁺ cells. Several research groups noted that either removal or continuous treatment of CHIR99021 and FGF2 at different concentrations decreases TBXT expression and disrupts the steady-state of NMPs^{44,229}. NMPs generation requires active WNT and FGF signaling in a dose dependent manner, but this modality is inadequate for the prolonged in vitro maintenance of NMPs. Frith et al.^{230,231} utilized a similar protocol as Gouti et al.⁴⁴ for NMP derivation with the same cytokine concentration however, they used either fibronectin or vitronectin as coating solution. Wang et al.²³² established NMPs on Matrigel coating by increasing CHIR99021 concentration to 10 μ M in addition to 20 ng/ml FGF2 with or without TGF β 1 (2-5 ng/ml). The presence of TGF β 1 leads to an increased number of SOX2/TBXT coexpressing cells at day 2 of treatment, SOX2 is downregulated by day 3. Inhibition of TGF β pathway by using 10 μ M SB431542 accompanied with 3 μ M CHIR99021 generates a relatively homogenous SOX2/TBXT coexpressing population on laminin²³³. Verrier et al.²³⁴ obtained NMP population on Geltrex matrix by treating hPSCs with 3 μ M CHIR99021 and 20 ng/ml FGF2 for 2 days and adding 50 ng/ml Noggin and 10 μ M SB431542 for additional 24 hours. Wind and Tsakiridis²³⁵ demonstrated NMPs can be derived on vitronectin coating solution

by inducing hPSCs with 3 μ M CHIR99021, 20 ng/ml FGF2 and 100 nM LDN193189 for 3 days. It has been shown that FGF2 can be replaced with high concentration of FGF8 (200 ng/ml) during NMP derivation²³⁶. Omitting CHIR99021 (3 μ M) on the first day of the establishment process enhances the number of SOX2/TBX1 co-expressing cells at day 3 in comparison to starting with the dual-activation. All of the abovementioned protocols have been summarized in Fig. 8.

1.5.1.2. Mouse NMPs

Treating mESCs on gelatin with either 3 μ M CHIR99021 for 1 day²³⁷ or 10 ng/ml FGF2 for 2 days and subsequently with 10 ng/ml FGF2 and 5 μ M CHIR99021 for 1 day^{44,59} induces Sox2/Brachyury coexpressing cells (Fig. 8). The protocols also result in the upregulation of Tbx6 expression pointing out the heterogeneity of these cultures. Unlike human NMPs, TGF β inhibition by using SB431542 is not able to induce Sox2/Brachyury coexpressing cells^{232,237}. FGF activation alone also does not enable Brachyury induction⁴⁴, yet WNT signaling is essential for Brachyury expression as it was confirmed by its drastic downregulation within 24 hours after CHIR99021 removal²³⁷. Continuous WNT activation with or without FGF signaling upregulates Tbx6, steering the NMPs towards the mesodermal fate^{44,59,237}.

NMPs can be derived from EpiSCs similarly to the NMP derivation from hPSC by treating them with 3 μ M CHIR99021 and 20 ng/ml FGF2 on fibronectin coating⁴⁴. Sox2/Brachyury coexpressing cells are detected on day 3 of the differentiation, albeit Brachyury is downregulated compared to day 2, which is accompanied by upregulation of Tbx6 and Meox1. Tsakiridis and Wilson²³⁸ utilized the same protocol and sorted only Brachyury+ cells by flow cytometry on day 2, which they cultivated further with CHIR and FGF2 for an additional 2 days resulting in Tbx6 positive and Sox2/Brachyury negative cells thus confirming that prolonged WNT and FGF activation leads to the mesodermal commitment of NMPs. Edri and colleagues²³⁹ established a different protocol for NMP derivation by treating EpiSCs on fibronectin with FGF2 (20 ng/ml) for 24 hours, and FGF2 (20 ng/ml) and CHIR99021 (3 μ M) together for further 24 hours. They also iterated mESC-based NMP protocols described above^{44,237} and compared the population yield of the three protocols. The results pointed out that all of three protocols give rise to a heterogenous culture with regards to Sox2/Brachyury coexpression and a great number of Sox2/Brachyury coexpressing cells derived from ESCs but not from EpiSCs retain Oct4 expression.

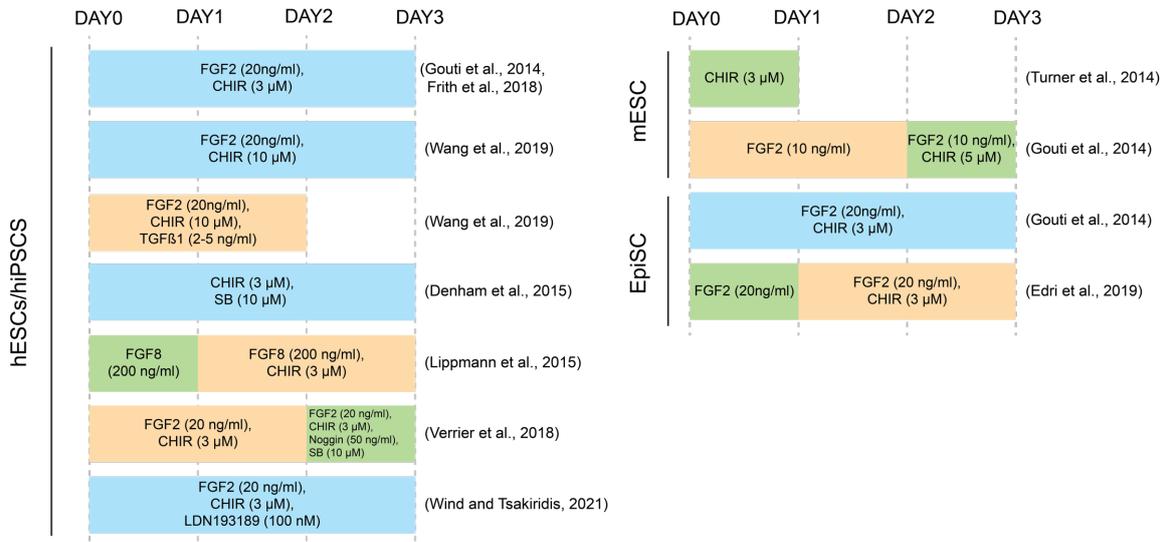


Figure 8: Summary of published human and mouse NMP derivation protocols

1.5.2. Characteristics of NMP derivation cultures

NMPs are characterized by co-expression of SOX2 and TBXT/Brachyury as described in the previous sections. SOX2 is already expressed in human/mouse PSCs, thus it can be suggested that TBXT/Brachyury is the master regulator of NMP establishment and prolonged maintenance. It has been shown that perturbation of the TBXT expression impairs NMP derivation in vitro as well as the anteroposterior segregation in vivo²⁴⁰. Due to its involvement in both pluripotency exit and the subsequent activation of HOX expression, TBXT plays an important role in cell fate determination. In cases where TBXT is not adequately expressed, cells become entrapped in the pluripotent state due to upregulation of pluripotency markers that prevents the cells from advancing to later developmental stages^{36,240}. CDX2 has been defined as an NMP marker due to its coexpression with SOX2 and TBXT/Brachyury^{44,59,229,231,234}. Brachyury expression can be detected in NMPs derived from Cdx null mouse cells, yet it has lower levels compared to WT cells, indicating the regulatory role of Cdx2 in NMP maintenance⁵⁹. NMPs are marked by an extensive repertoire of Hox genes from PG 1 to 13^{44,59,231,236}. Cdx null cells lack posterior Hox genes (Hoxa5, Hoxc6 and Hoxc10) while simultaneously expressing anterior Hox genes (Hoxa1, Hoxb1)⁵⁹. Cdx2 directly regulates Cyp26a1 expression²⁴¹ which is upregulated in NMPs⁵⁹. Nkx1-2 is another known NMP marker present in the in vitro NMPs as well as in their in vivo counterpart^{59,66,229,231,234,239}. Msn is expressed heterogeneously in NMPs in vivo^{59,66} but it can be detected in mouse and human NMPs in vitro^{44,59}. Evx1 which exists in mouse NMPs in vivo⁶⁶ is also shown to be expressed in mouse and human NMP in vitro cultures^{231,239}. Mixl1 and Eomes, which are early primitive streak markers²⁴², are found in NMP in vitro

cultures^{44,234} although they are not homogenously expressed in NMPs in vivo²⁴³. It has been shown that a high number of in vitro NMPs expresses SOX9, MSX1/2, ZIC1/3 and SNAI1/2 which are associated with neural crest development²³¹.

1.5.3. NMP progeny

1.5.3.1. Neural differentiation

NMPs can be differentiated into post-mitotic neurons. In the absence of WNT and FGF activation, either retinoic acid or vitamin A which is a retinoic acid precursor can direct NMPs towards the neural fate^{44,59,239}. The neural commitment is determined by downregulation of Brachyury in Sox2 expressing cells and upregulation of neural progenitor markers Sox1 and Irx3. Wind and Tsakiridis²³⁵ produced neural progenitor cells from NMPs by induction of WNT, FGF, SHH, RA in addition to BMP and TGF β inhibition for 7 days. The resulting neural cultures were characterized by Sox1 and Pax6 expression. Sox1/2 and Pax6 are also expressed in PSC-derived NPCs via dual SMAD (TGF β and BMP) inhibition without WNT and FGF activation, but these cells have unique features such as high Otx2 and low Hox expression in contrast to NMP-derived neural cells⁵⁹. Otx2 is a hallmark of the anterior embryonic development and it is absent in posterior embryonic structures²⁴⁴. Neural progenitors generated from NMPs are named posterior neural progenitors (PNPs). Recent studies suggest that NMP differentiation to PNPs involves a transient population named preneural progenitors which is marked by SOX2 and CDX2 coexpression^{50,245} (see section 1.5.3.3).

It is possible to enhance either ventral or dorsal spinal cord-like PNPs by modulating SHH, BMP and TGF β pathways. SHH is known as a ventralization factor considering the gradient of active pathways in spinal cord in vivo (Fig 5). However, inducing PNPs with SHH agonists (purmorphamine and SAG) does not sufficiently result in ventralization of the differentiating population due to upregulation of both ventral (NKX6-1 and NKX6-2) and dorsal (PAX3 and MSX1) spinal cord progenitor markers²²⁹. Further differentiation of these progenitors into post-mitotic neurons revealed that MNX1 is not expressed, pointing out the absence of motor neurons which are a subtype of the ventral spinal cord. Inhibition of BMP by using LDN193189 with or without DMH-1, and inhibition of TGF β by using SB431542 together with activation of SHH by using purmorphamine and SAG are required for the generation of ventral progenitors as well as motor neurons^{229,246}. Each ventral domain types can be generated in a dose dependent manner of the cytokines used²⁴⁶. Activation of BMP by using BMP4 with or without inhibition of SHH by using cyclopamine induces dorsal patterning^{246,247}. The dorsal-most subdomains (domain 1-3) require prolonged BMP4 exposure²⁴⁶. 24-hour BMP pulse sufficiently induces domain 4 and 5 while domain 5 can also be obtained by SHH inhibition only. Domain 6 can be generated with different modalities: 1)

by a BMP pulse, 2) by SHH inhibition, 3) by BMP and TGF β inhibition. Dorsal and ventral progenitors are differentiated into post-mitotic neurons by using DAPT that inhibits the Notch pathway^{229,246}. Notch inhibition accelerates the differentiation rate by delaying the G1/S phase transition, and active Notch signaling maintains neural progenitor cells in the proliferative state through Hes5 expression^{248,249}. Addition of BMP7 during PNP differentiation enhances the production of dorsal-most neural subtypes (dl1-3)²⁴⁶.

1.5.3.2. Mesodermal differentiation

NMPs are committed to mesodermal fate through 2-day continuous WNT activation by CHIR99021 induction. This step is marked by upregulated PXM markers (TBX6 and MSGN)^{44,59}. Post PXM identity acquisition, CHIR removal and further differentiation steps, the cells in N2B27 medium upregulate MYOD, which is a known marker for skeletal muscle cells⁴⁴. There is a scarcity of studies focusing on the establishment and understanding of the 2D direct differentiation from the NMPs towards skeletal muscle cells. Some preliminary work was published in this direction by Gouti et al.⁴⁴ and more recently Yamanaka and colleagues²⁵⁰ generated human axioids, 3D structures able to recapitulate the segmentation clock, thus demonstrating the contribution of NMPs to the somitogenesis process.

Faustino Martins et al.²⁵¹ generated neuromuscular organoids from NMPs established priorly in 2D culture. They showed that FGF2, HGF (hepatocyte growth factor) and IGF (insulin-like growth factor) treatment for 4 days and medium based maturation of organoids induces contracting muscles as well as neural cells by day 50. Transcriptome analysis of day 50 organoids revealed the presence of myogenic progenitors and satellite cells expressing PAX7 and MSC (Musculin), myocytes expressing MYOD, MYOG, CDH15, and skeletal muscle fibers expressing TTN, MYBPH, ACTA1 and ACTN2. Neural and neural crest subtypes at different developmental stages could also be identified. A novel step-wise protocol has been established for 3D differentiation of hiPSCs into both myogenic and trunk neural crest cells²⁵². They observed the presence of NMPs and TBX6 expressing cell after 5 days of WNT/FGF activation in the presence of BMP inhibition. They defined this step as the NMP/PXM stage. Addition of RA for 2 days within the NMP/PXM stage, followed by SHH/WNT activation along BMP inhibition for 4 days resulted in PXM and neural tube cells marked by upregulation of PAX3/7, MEOX2, TFAP2A. Further differentiation by using FGF2 and HGF for 4 days subsequent to SHH/WNT/BMP modulation generated dermomyotome and trunk neural crest cells marked by the expression of PAX3/7, TBX18 and SOX10. Treatment with HGF only of the dermomyotome progenitors resulted in their maturation into skeletal muscle cells marked by MYOD and MYHC. In addition to dermomyotome, NMPs are shown as able to generate chondrocyte and osteocytes which are the derivatives of

sclerotome²³². There has not been a directed differentiation protocol into these cell types, but the study by Wang et al.²³² obtained osteocyte and chondrocyte cells by differentiating mesenchymal stem cells from NMPs.

1.5.3.3. Neural crest differentiation

Continuous activation of WNT and FGF in NMPs has been shown to disrupt the NMP state and decrease TBXT/Brachyury expression (see section 1.5.1.1). A recent paper confirmed this notion by activating WNT and FGF, and blocking RA metabolism⁵⁰. By addition of a Rho-associated kinase inhibitor (Y-27623) to support the cell survival, they managed to culture the cells up to 30 days by regular passaging. By passage 3, TBXT expression was not detectable but the cells remained positive for SOX2 and CDX2. After passage 5, two morphologically distinct cell types were observed that were classified as epithelial and mesenchymal. Epithelial cells were marked by SOX2/CDX2, while mesenchymal cells were distinguished by upregulated NC genes (SNAI1, SOX9/10, and ETS1). The resulting heterogenous culture could be differentiated into neurons by RA addition. They defined the SOX2/CDX2 expressing cells as an intermediate state (preneural progenitors). It is also possible to generate NC cells from NMPs by a 6-day protocol that includes WNT activation (by using CHIR99021), TGF β inhibition (by using SB431542) and moderate level of BMP signaling (by using both BMP4 and DMH-1)²³⁰. NC identity of the cultures was defined by SNAI1/2, SOX5/9/10, TFAP2A/B/C and ETS1 expression. The culture was shown to express HOX6-9 thus, the differentiated cells were identified as trunk NC. These cells were able to contribute to the DRG region upon injection into chick embryos. It was possible to identify during posterior hPSC NC differentiation a transient SOX2/TBXT coexpressing population, whose presence emphasized the NMP role in the derivation of trunk/posterior NC derivatives²⁵³. There have been additional protocols published that yield dual- or multi-fated derivatives such as NC and mesoderm together with or without neural descendants^{251,252}. Faustino Martins et al.²⁵¹ showed the presence of Glia and Schwann cells by GFAP and S100 β expression respectively and Mavrommatis et al.²⁵² showed the presence of neural crest progenitors marked by SOX9/10 and TFAP2A expression. The mesodermal outcomes of these protocols are described in section 1.5.3.2

1.6. Deciphering molecular dynamics during embryonic development

The aim of developmental biology is to reveal key mechanisms throughout the embryonic development. Murine and avian model systems are the primary sources used so far to answer fundamental developmental questions, yet they are not sufficient to unravel mechanism specific to human development. Due to the required embryonic manipulations, investigation of human development raises serious ethical issues. Establishment and differentiation of a

specific stem cell state that captures a snapshot of an embryonic stage of interest and its respective progeny paved the way for major advances in the field of developmental biology. The follow-up challenge in need to be addressed was bridging the gap between molecular dynamics and developmental progression. The standardization of various omics techniques enabled researchers to carry out investigations into the molecular mechanism controlling lineage choice in vitro and adding a time component comparable to developmental timelines in vivo.

RNA sequencing technology was considered a milestone technique for developmental biology²⁵⁴ as it can allow for large scale transcriptome analysis in comparison to more traditional methods like the RT-qPCR, where you are limited to a handful of markers of interest. It was not optimal to solve the puzzle of heterogeneity present in biological samples as it presented as an outcome the average marker gene expression for entire populations. The necessity to understand cellular heterogeneity both in vivo and in vitro led to the breakthrough innovation of single cell RNA sequencing (scRNA sequencing). This method enables transcriptome profiling at single cell resolution. The first readout from this technology was the sequencing of a single blastomere²⁵⁵. Various approaches for scRNA sequencing tools and protocols are continuously being developed based on research demands, such as plate-based or droplet-based, with or without barcoding, low or high transcript coverage. These techniques show differences in terms of efficiency, sensitivity and accuracy^{256,257}. Analysis of scRNA sequencing data sets with bioinformatical methods like RNA velocity and trajectory inference enables traceability of developmental progression in pseudotime^{258,259}.

Transcriptome analysis is frequently substantiated further by proteomic profiling. Mass spectrometry is one of the widely used methods for global proteome profiling²⁶⁰. It aims to elucidate protein-protein interactions, post-translational modifications and primary protein sequence²⁶¹. Over the last decade, a mass spectrometry-based technique aiming to identify chromatin-bound proteins was developed. The name coined for this method is chromatome profiling²⁶². The chromatome analysis opened up an avenue for investigation of chromatin organization and histone modifications leading to identification of epigenetic regulations on a large scale. Ugur et al.¹⁹² further developed the sensitivity of the available chromatome analysis methods and applied their improved methodology to characterize further the chromatome specific to mouse naïve and primed pluripotency and their resemblance to human primed pluripotency state.

1.7. Aim and impact of this study

The theory of how anterior and posterior embryonic development takes place through distinct mechanisms was raised by German scientist David E. Holmdahl in 1925²⁶³. He monitored chick embryos and set forth “primärer Körperentwicklung” and “sekundärer Körperentwicklung” terms. His theory has been confirmed by many other scientists over the decades. Growing evidence is showing that axial structures of the secondary/posterior body originate from a common cellular source namely the NMPs. The machinery underlying NMP emergence has been discovered through in vivo studies. These findings paved the way for generation of NMPs in vitro, which has been the focus of many recently published protocols.

Using state of the art protocols, it is possible to generate NMPs in 2-4 days, but these protocols have considerable drawbacks. Nearly all of them result in very heterogenous cultures that include “off-target” cell types. The obtained NMP progenitor state showcases limited self-renewal capacity and a high propensity towards spontaneous differentiation leading to a high number of mesodermal/neural cellular contaminants. In summary, it is possible to say with confidence that none of the NMP protocols published and reviewed in this introduction allows for an extended in vitro cultivation and passaging of NMPs. Moreover, a stem cell state resembling NMPs or other progenitors giving rise to axial structures has not yet been identified.

Within this study, I aim to prove that I can generate human pluripotent derived axial stem cells (AxSCs) that possess the ability to indefinitely self-renew and give rise to neural and mesodermal progeny, namely CFS AxSCs, thus recapitulating the NMPs as a steady state in vitro. Additionally, I have generated a second axial stem cell state (CS AxSCs) that gives rise to the posterior dorsal neural tube, thus more limited in its progeny compared to CFS cells. I investigated their stem cell identity and determined the molecular hallmarks for each AxSC state. By exploring their potential progeny, I confirmed their developmental correspondence to the in vivo NMPs/axial progenitors. Moreover, neural descendants of both AxSC states were compared to the progeny of neural progenitor cells (NPCs) which have been the widely-used cell state for the in vitro neurogenesis studies. Lastly, I demonstrated that AxSCs can be reproducibly derived between species in line with the species-specific developmental potential of their in vivo correspondence.

In summary, this study is the first of its kind to establish two novel region specific stem cell types representing NMPs and posterior dorsal neural tube progenitors. AxSCs established within this study represent later developmental stages, thus their differentiation into lineage-committed cells requires shorter periods of time as supported by experimental data in the course of this study when compared to hPSC based differentiation. The indefinitely

self-renewing ability of AxSCs opens new avenues for exploring stage specific lineage choice and progeny potential of the posterior embryonic development. Additionally, it presents an outstanding opportunity to explore regenerative medicine applications and patient-targeted therapeutic approaches focusing on stem cell based replacement therapies for neuromuscular and nervous system diseases.

2. MATERIALS AND METHODS

2.1. Materials

2.1.1. Cell lines

H9 (WiCELL Research Institute) and HUES6²⁶⁴ human ESCs, HMGU1²⁶⁵ human iPSCs, E14 mESC-based EpiSCs¹⁷⁴ and Sumatran orangutan iPSCs²⁶⁶ were used in this study.

2.1.2. Cell culture components

All cell culture components used in this study were of research grade and are listed in Table 1.

Table 1: Basal media, supplements, ligands, coating and passaging reagents used in cell culture experiments in this study

Basal Media	Cat no	Manufacturer
DMEM/F12	11320074	Life Technologies
StemMACS iPS Brew XF	130-104-368	Miltenyi Biotech
KnockOut DMEM	10829018	Life Technologies
Neurobasal medium	21103049	Life Technologies
Neurobasal-A medium	10888022	Life Technologies
RPMI-1640 with L-Glutamine	11875093	Life Technologies
StemFit Basic02	Basic02	Nippon Genetics
DMEM low glucose	11885084	Life Technologies
Supplements	Cat no	Manufacturer
B27 supplement	5001207	Life Technologies
B27 supplement without insulin	A1895601	Life Technologies
B27 supplement without vitamin A	12587001	Life Technologies
Insulin-Transferrin-Selenium (ITS - G)	41400045	Life Technologies
Insulin-Transferrin-Selenium-Ethanolamine (ITS-X)	51500056	Life Technologies
Insulin, human recombinant, zinc solution	12585-014	Life Technologies
N2 supplement	17502048	Life Technologies
Ligands	Cat no	Manufacturer
Recombinant Human/Murine/Rat Activin A	120-14E	Peprtech

L-Ascorbic acid 2-phosphate sesquimagnesium salt hydrate	A8960	Sigma
Recombinant Human/Murine/Rat BDNF	450-02	Peprotech
Recombinant Human BMP-4	314-BP	R&D Systems
Adenosine 3',5'-cyclic monophosphate sodium salt monohydrate	A6885	Sigma
CHIR 99021 trihydrochloride	4953	Biotechne
Compound E	6476	Biotechne
DMH-1	4126	Biotechne
Recombinant Human FGF-basic	100-18B	Peprotech
Recombinant Human FGF-8a	100-25A	Peprotech
Recombinant Human GDNF	450-10	Peprotech
Recombinant Murine HGF	315-23	Peprotech
Recombinant Murine IGF-I	250-19	Peprotech
Recombinant Human EGF	E9644	Sigma
IWP2	sc-252928	Santa cruz
LDN-193189 (hydrochloride)	Cay19396	Biomol
Recombinant mouse LIF protein	ESG1106	Millipore
StemMAC PD0325901 in Solution	130-106-541	Miltenyi Biotech
Purmorphamine	4551	R&D Systems
Retinoic acid	R2625	Sigma
SB431542	1614	R&D Systems
Recombinant Human Sonic Hedgehog	100-45	Peprotech
Recombinant Human TGF- β 1	100-21	Peprotech
Y-27632 dihydrochloride	1254	R&D Systems
Recombinant Human FGF-9	100-23	Peprotech
Bovine fetuin	SIALF2379	VWR
Dexamethasone	D4902	Sigma
Coating reagents	Cat no	Manufacturer
Collagen I	354236	BD Corning
Collagen IV	sc-29010	Santa Cruz
Fibronectin human	FALC356008	BD Corning
Laminin	11243217001	Sigma
Matrigel	11543550	BD Corning
Poly-D-lysine	A-003-E	Sigma
Poly-DL-ornithine hydrobromide	P8638	Sigma

Vitronectin	A27940	Thermo Fisher
Passaging reagents	Cat no	Manufacturer
0.05% Trypsin	25300054	Life Technologies
Accutase	A6964	Sigma
Collagenase IV	17104019	Life Technologies
Passaging solution XF	130-104-688	Miltenyi Biotech
TrypLE Express	12605010	Life Technologies
Versene	15040-066	Life Technologies
Papain	P3125	Sigma
Other	Cat No	Manufacturer
Bovine Albumin Fraction V	15260037	Life Technologies
Fetal Bovine Serum	SH30071.03	HyClone
GlutaMAX	35050038	Life Technologies
Knockout Serum Replacement	10828028	Life Technologies
Non-Essential Amino Acid	11140050	Life Technologies
Penicillin/Streptomycin	15070063	Life Technologies
2-Mercaptoethanol	31350-010	Life Technologies

2.1.3. Chemicals and kits

All chemicals and kits used in this study are listed in Table 2.

Table 2: Chemicals and kits used in this study

Chemicals	Cat no	Manufacturer
16% Formaldehyde (w/v), Methanol-free	28908	Thermo Fisher
Ethanol	9065.2	Carl Roth
Hydrochloric acid 32%	100319	Merck Millipore
Isopropanol	6752.2	Carl Roth
Methanol	45631.02	Serva Electrophoresis
NP-40	NP40S	Sigma
cOmplete EDTA-free Protease Inhibitor	11873580001	Roche
Deoxyribonuclease II from bovine spleen	D8764-150KU	Sigma
Trypan Blue Stain (0.4%)	T10282	Life Technologies
Sodium Chloride Solution	59222C	Sigma

Magnesium Chloride Solution	M1028	Sigma
Trizma Hydrochloride Solution pH 7.4	T2194-100ML	Sigma
DAPI	10236276001	Sigma
Triton X-100	X100	Sigma
Tween-20	P9416	Sigma
Glycine	23391.02	Sigma
HEPES	15630056	Thermo Fisher
Urea	U1250	Sigma
EDTA	8043.2	Carl Roth
UltraPure SDS Solution	15553027	Invitrogen
Sera-Mag beads	GE45152105050250	Sigma
Acetonitrile	1000301000	Merck Millipore
Trifluoroacetic acid	85183	Thermo Fisher
Tris(2-carboxyethyl)phosphine	PG82080	Thermo Fisher
Chloroacetamide	C0267	Sigma
Guanidinium chloride	24110	Thermo Fisher
Ammonium hydroxide	1054280250	Supelco
Formic acid	94318	Sigma
LysC Protease	90051	Thermo Fisher
Trypsin Protease	90058	Thermo Fisher
SYBR Green Mastermix	4309155	Applied Biosystems
Taqman Gene Expression Mastermix	4369514	Applied Biosystems
ProLong Glass Antifade Mountant with DAPI	P36935	Life Technologies
Kits	Cat no	Manufacturer
RNeasy Mini Kit	74106	Qiagen
SuperScript III First-Strand Synthesis System	18080051	Life Technologies
Wizard SB Genomic DNA Purification System	A2361	Promega
Pierce™ BCA Protein Assay Kit	23225	Thermo Fisher
Relative Human Telomere Length Quantification qPCR Assay Kit	8908-SC	ScienCell

Chromium Next GEM Single Cell 3' GEM, Library & Gel Bead Kit v3.1	1000121	10X Genomics
Chromium Next GEM Single Cell 3' GEM, Single Index Kit T Set A	1000213	10X Genomics
Chromium Next GEM Single Cell 3' GEM, Dual Index Kit TT Set A, 96 rxns	1000215	10X Genomics
HumanCytoSNP-12 v2.1 BeadChip Kit	WG-320-2101	Illumina
Agilent High Sensitivity kit	5067-4626	Agilent
SP Reagent Kit v1.5/ 100 cycles	20028401	Illumina
S4 Reagent Kit v1.5 /200 cycles	20028313	Illumina

2.1.4. Primers and probes

Probes (Thermo Fisher) and primers (Sigma) used for quantitative real time PCR (qRT-PCR) experiments in this study are listed in Table 3.

Table 3: Probes and primers used in this study

Probes	Cat no	Probes	Cat no
ACTA2	Hs00426835_g1	NGFR	Hs00609976_m1
CDH15	Hs00979297_m1	NKX1-2	Hs01392360_m1
CDH19	Hs00253534_m1	NKX6-1	Hs01055914_m1
CDX2	Hs01078080_m1	NTRK2	Hs00178811_m1
CHAT	Hs00758143_m1	OLIG2	Hs00377820_m1
CHRNA3	Hs01088199_m1	OTP	Hs00259528_m1
DMRT3	Hs00253642_m1	PAX3	Hs00240950_m1
EN1	Hs00154977_m1	PAX6	Hs01088114_m1
ETS1	Hs00428293_m1	PAX7	Hs00242962_m1
FOXD3	Hs00255287_s1	PECAM1	Hs01065279_m1
FOXP1	Hs00908900_m1	PIEZO2	Hs00926218_m1
GAPDH	Hs02758991_g1	POU3F1	Hs00538614_s1
GATA2	Hs00231119_m1	POU4F1	Hs00366711_m1
GBX2	Hs00230965_m1	POU5F1	Hs00999632_g1
GDF7	Hs00766203_m1	PRDM12	Hs00964106_m1
GUCY1A3	Hs01015574_m1	PRPH	Hs00196608_m1

HES5	Hs01387464_g1	S100B	Hs00389217_m1
IRX3	Hs01124217_g1	SALL4	Hs01010838_g1
IRX5	Hs04334749_m1	SIM1	Hs00231914_m1
ISL1	Hs00158126_m1	SOX1	Hs01057642_s1
ISL2	Hs00377575_m1	SOX10	Hs00366918_m1
LHX3	Hs01033412_m1	SOX2	Hs01053049_s1
LHX4	Hs01114435_m1	SOX3	Hs00271627_s1
LIN28A	Hs04189307_g1	SOX9	Hs00165814_m1
LIN28B	Hs01013729_m1	TBXT	Hs00610080_m1
LMX1B	Hs00158750_m1	TBX3	Hs00195612_m1
MNX1	Hs00907365_m1	TBX6	Hs00365539_m1
MSGN	Hs03405514_s1	TLX3	Hs00253271_m1
MSX1	Hs00427183_m1	TNNT2	Hs00943911_m1
MYCN	Hs00232074_m1	UNCX	Hs01394890_g1
MYH1	Hs00947183_g1	VSX1	Hs00232724_m1
MYOD	Hs00159528_m1	WT1	Hs01103751_m1
MYOG	Hs01072232_m1	ZIC1	Hs00602749_m1
NANOG	Hs02387400_g1		
Primers	Sequence	Primers	Sequence
Human ASCL1	F:TTCACCAACTGGTTCTGAG R:TAAAGATGCAGGTTGTGCG	Human TWIST1	F:GAGCTGGACTCCAAGATGG R:TTAAGAAATCTAGGTCTCCGGC
Human CDX2	F:TCGGCAGCCAAGTGAAA R:GATGGTGATGTAGCGACTGTAG	Human ZO1	F:GTCCAGAATCTCGGAAAAGTGCC R:CTTTCAGCGCACCATAACCAACC
Human DCX	F:GCCAGGGAGAACAAGGACTTT R:CACCCCACTGCGGATGA	Mouse Brachyury	F:GAAGGGAGACCCACCGAA R:TTACCTTACAGCACCGGGAAC
Human GAPDH	F:CGCTCTCTGCTCCTCCTGTT R:CCATGGTGTCTGAGCGATGT	Mouse Cdx2	F:CTGGACAAGGACGTGAGCAT R:ACTGCGGAGGACTGACAAAG
Human HES5	F:TCAGCTACCTGAAGCACAG R:AGTAGCCTTCGCTGTAGTC	Mouse Gapdh	F:CATCACTGCCACCCAGAAGACTG R:ATGCCAGTGAGCTTCCCCTTCAG
Human NANOG	F:CCTTCCTCCATGGATCTGCTT R:CTTGACCGGGACCTTGTCTTC	Mouse Hes5	F:AAACACAGCAAAGCCTTCG R:CAGGGTCAGGAACTGTACC
Human PAX6	F:CCCTGGAGAAAGAGTTTGAGAG R:TCCATTTGGCCCTTCGATTAG	Mouse Nanog	F:GAAATCCCTTCCCTCGCCAT R:CAGGCATTGATGAGGCGTTC
Human PHOX2B	F:GCAGATAACAAATTTCCCTCGGT R:GTGAAGAGTTTGTAAAGGAAACCC	Mouse Pax6	F:CTGAGGAACCAGAGAAGACAGG R:CATGGAACCTGATGTGAAGGAGG
Human POU5F1 (OCT4)	F:CAATTTGCCAAGCTCCTGAAG R:AAAGCGGCAGATGGTCGTT	Mouse Pou5f1 (Oct4)	F:GGCTTCAGACTTCGCCTTC R:AGCTTAGCCAGGTTTCGAGGAT
Human SOX1	F:GAGAACCCCAAGATGCACAA R:CCTCGGACATGACCTTCCA	Mouse Sox2	F:CAAAAACCGTGATGCCGACT R:CGCCCTCAGGTTTTCTCTGT

Human SOX2	F:CCTCCGGGACATGATCAGCATGT R:GCAGTGTGCCGTTAATGGCCGTG	Mouse Zic2	F:AAGCCCTTCCCCTGTCCTTT R:TGGAAAGGTTTCTCCCCTGT
Human TBX6	F:CGTGTGAAGAGGAAACTGCG R:GACTACACTCACCTCCGCTC	Orangutan PAX6	F:CAGATCTGCCACTTCCCCTG R:TCACTCCGCTGTGACTGTTC
Human TBXT	F:ATGAGCCTCGAATCCACATAG R:CTGTGATCTCCTCGTTCTGATAA		

2.1.5. Antibodies

Antibodies used in this study are listed in Table 4.

Table 4: Antibodies used in this study

Primary antibodies	Cat no	Manufacturer	Dilution
Rabbit anti-Sox2	2738	Cell Signaling Technology	1:400
Mouse anti-Brachyury	ab209665	Abcam	1:100
Goat anti-Brachyury	AF2085	R&D Systems	1:100
Rabbit anti-CDX2	12306	Cell Signaling Technology	1:200
Rabbit anti-Pax6	901301	Biolegend	1:100
Rabbit anti-Zic2	ab150404	Abcam	1:100
Rabbit anti-Hes5	NBP2-56999	Novus	1:100
Rabbit anti-Islet 1	ab20670	Abcam	1:200
Mouse anti-Brn3a	sc-8429	Santa cruz	1:100
Mouse anti-Mnx1	sc-515769	Santa cruz	1:100
Chicken anti-B3 tubulin	NB100-1612	Novus	1:100
Rabbit anti-Zo1	617300	Life Technologies	1:100
Rabbit anti-Ascl1	ABE1025	Millipore	1:100
Mouse anti-N-cadherin	14215S	Cell Signaling Technology	1:100
Mouse anti-Myod	MA1-41017	LIFE Technologies	1:100
Rabbit anti-M-Cadherin	40491S	Cell Signaling Technology	1:100
Mouse anti-Myogenin	14-5643-82	LIFE Technologies	1:100
Mouse anti-Sox2	MAB2018	R&D Systems	1:100
Rabbit IgG isotype control	GTX35035	GeneTex	1 µg

Mouse IgG isotype control	GTX35009	GeneTex	1 µg
Secondary antibodies	Cat no	Manufacturer	Dilution
Goat anti-Chicken IgY (H L) Secondary Antibody, Alexa Fluor 594	A11042	LIFE Technologies	1:1000
Goat anti-rabbit IgG (H+L), Alexa Fluor 488, highly cross-absorbed	A11034	LIFE Technologies	1:1000
Goat anti-mouse IgG (H+L) Alexa Fluor Plus 647, highly cross-absorbed	A32728	LIFE Technologies	1:1000
Donkey anti-Mouse IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor™ 488	A21202	LIFE Technologies	1:1000
Donkey anti-Goat IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Fluor™ 594	A11058	LIFE Technologies	1:1000
F(ab') ₂ -Goat anti-Rabbit IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Fluor™ 647	A21246	LIFE Technologies	1:1000

2.1.6. Instruments and software

Instruments and software used in this study are listed in Table 5.

Table 5: Instrument and software used in this study

Instruments	Supplier
Megafuge 40R centrifuge	Thermo Fisher
Heraeus PICO 21 centrifuge	Thermo Fisher
SpeedVac centrifuge	Thermo Fisher
Mastercycler® Nexus thermal cycler	Eppendorf
QuantStudio 12K Flex Real-Time PCR System	Thermo Fisher
Axio Observer Z1 Fluorescence microscope	Zeiss
10X Genomics Chromium Controller	10X Genomics
NovaSeq 6000	Illumina
Countess 3 automated cell counter	Thermo Fisher
Biorupter Plus	Diagenode
Easy-nLC 1200	Thermo Fisher
Orbitrap Exploris 480 mass spectrometer	Thermo Fisher
Nanodrop 2000 spectrophotometer	Thermo Fisher
Software	Supplier
Quantstudio	Thermo Fisher
ZEN Microscopy Software	Zeiss
Fiji	National Institutes of Health

Python	Python Software Foundation
Anaconda	Anaconda
GraphPad Prism	GraphPad Software
Adobe Illustrator	Adobe

2.2. Methods

2.2.1. Maintenance of PSCs

2.2.1.1. Human PSCs

Before thawing and passaging, plates were coated with Matrigel diluted 1:100 in DMEM-F12 and incubated for at least 30 minutes in a 37°C incubator. Cells were thawed in a water bath at 37°C and then centrifuged at 300 g for 3 minutes. The supernatant was discarded, and the pellet was resuspended in StemMACS iPS-Brew XF medium including 1X StemMACS iPS-Brew XF supplement and 10 µM Y-27632. Matrigel was removed from the coated plate and cells were seeded. After 24 hours, fresh medium without Y-27632 was added. The medium was changed daily. Cells were maintained at 37°C, 5% CO₂ and 5% O₂.

When cell confluency reached >80%, passaging was performed by washing the cells with PBS and incubating them with StemMACS Passaging solution XF for 5 minutes at room temperature. StemMACS Passaging solution XF was discarded, and cells were harvested in StemMACS iPS-Brew XF medium including 10 µM Y-27632 and plated at 1:10 ratio. After 24 hours, fresh medium without Y-27632 was added.

For freezing, cells were harvested after incubation with StemMACS Passaging solution XF and centrifuged at 300 g for 3 minutes. The supernatant was discarded, and the pellet was resuspended in 1 ml of Cryostor cryopreservation medium and transferred to cryotubes. The cryotubes were stored at -80°C for short-term and in a liquid nitrogen tank for long-term storage.

2.2.1.2. Mouse EpiSCs

Before thawing and passaging, plates were coated with Matrigel diluted 1:100 in DMEM-F12 for at least 30 minutes in a 37°C incubator. Mouse EpiSCs were thawed in a water bath at 37°C and centrifuged at 300 g for 3 minutes. The supernatant was discarded, and the pellet was resuspended in medium including 1:1 DMEM/F12 and Neurobasal medium, 0.5X N2, 0.5X B27, 0.033% BSA, 50 µM β-mercaptoethanol, 1X GlutaMAX, 1X penicillin/streptomycin, 20 ng/ml Activin A, 12 ng/ml FGF2 (bFGF), 2 µM IWP2 (EpiSC medium). Matrigel was removed from the coated plate and cells were seeded in EpiSC medium including 10 µM Y-

27632 for the first 24 hours. Cells were maintained at 37°C, 5% CO₂ and 5% O₂. The medium was changed daily.

When cell confluency reached >70%, passaging was performed by washing the cells with PBS and incubating them with TrypLE Express for 3 minutes at 37°C. Cells were then washed with TrypLE Express and then with EpiSC medium. After centrifugation at 300 g for 3 minutes, the supernatant was discarded, and the pellet was resuspended in EpiSC medium including 10 µM Y-27632 for the first 24 hours. Cells were plated at a ratio between 1:20-1:40.

For freezing, cells were harvested after incubation with TrypLE Express and centrifuged at 300 g for 3 minutes. The supernatant was discarded, and the pellet was resuspended in 1 ml of Cryostor cryopreservation medium and transferred to cryotubes. The cryotubes were stored at -80°C for short-term and in a liquid nitrogen tank for long-term storage.

2.2.1.3. Orangutan iPSCs

Before thawing and passaging, plates were coated with Matrigel diluted 1:100 in DMEM-F12 for at least 30 minutes in a 37°C incubator. Orangutan iPSCs were thawed in water bath at 37°C and centrifuged at 300 g for 3 minutes. The supernatant was discarded, and the pellet was resuspended in medium including StemFit Basic02 medium with 1X penicillin/streptomycin, 100 ng/ml FGF2 (oriPSC medium). Matrigel was removed from the coated plate and cells were seeded in EpiSC medium including 10 µM Y-27632 for the first 24 hours. Cells were maintained at 37°C, 5% CO₂ and 5% O₂. The medium was changed daily.

When cell confluency reached >80%, passaging was performed by washing the cells with PBS and incubating them with Versene for 5 minutes at room temperature. After 5 minutes, Versene was discarded and cells were harvested in oriPSC medium including 10 µM Y-27632 for the first 24 hours. Cells were plated at 1:10 ratio.

For freezing, cells were harvested after incubation with Versene and centrifuged at 300 g for 2 minutes. The supernatant was discarded, and the pellet was resuspended in 1 ml of Cryostor cryopreservation medium and transferred to cryotubes. The cryotubes were stored at -80°C for short-term and in liquid nitrogen tank for long-term storage.

2.2.2. Maintenance of primary human skeletal muscle cells

Primary human skeletal muscle cells were cultured in a medium composed of DMEM low glucose, 0.05 ml/ml FBS, 50 µg/ml Fetuin, 10 ng/ml EGF, 1 ng/ml FGF2, 10 µg/ml Insulin, and 0.4 µg/ml Dexamethasone (SkGM medium), with or without 2% horse serum (HS). The

cells were routinely split at a 1:4 ratio after incubation with 0.25% Trypsin for 5 minutes. 10 μ M Y-27632 was added for the first 24 hours.

2.2.3. Derivation of axial stem cells

2.2.3.1. Derivation of axial stem cells from human ESCs

Before derivation of human AxSCs, H9 cells were adapted to single cell split by passaging with TrypLE Express for at least 3 passages. For this purpose, cells were incubated with TrypLE Express at 37°C for 6 minutes after washing with PBS. Cells were harvested in TrypLE Express and washed in StemMACS iPS-Brew XF medium. After centrifugation at 300 g for 3 minutes, the supernatant was discarded, and the pellet was resuspended in medium including 10 μ M Y-27632 for the first 24 hours.

For human AxSCs derivation, H9 cells (between passage 55-60) were counted after dissociation and plated on a Matrigel-coated plate (1:100) at a density of 71.5×10^3 cells/cm² in StemMACS iPS-Brew XF medium including 10 μ M Y-27632. After 24 hours at 37°C, 5% CO₂ and 5% O₂, cells were rinsed with PBS and the medium was replaced with RPMI-1640 including 1X B27 without insulin and 10 μ M CHIR99021 (induction medium). Cells were cultured in the induction medium for 24 hours. The next day, cells were rinsed with PBS and dissociated with TrypLE Express by incubating for 5 minutes at 37°C. After washing cells in TrypLE Express and then induction medium, the cell suspension was divided into two tubes and centrifuged at 300 g for 3 minutes. The supernatant was removed, and the pellet was resuspended either in CFS medium consisting of RPMI-1640, 1X NEAA, 1X B27 without vitamin A, 100 ng/ml FGF2, 5 μ M CHIR99021 and 10 μ M SB431542, or in CS medium comprised of RPMI-1640, 1X NEAA, 1X B27 without vitamin A, 5 μ M CHIR99021 and 10 μ M SB431542. Cells were plated onto a Matrigel-coated (1:100) 12-well plate at a ratio between 1:5-1:20. 10 μ M Y-27632 was added for the first 24 hours. Fresh medium was applied daily.

Cells were cultured on Matrigel-coated (1:100) 12-well plates until passage 5 and then on Matrigel-coated (1:100) 6-well plates. Cells were rinsed with PBS and then dissociated by incubation with TrypLE Express at 37°C for 4-5 minutes and 4 minutes for CFS and CS respectively. After washing the cells in TrypLE Express and the respective medium, the cell suspension was centrifuged at 300 g for 3 minutes. The supernatant was discarded, and the pellet was resuspended in the respective medium including 10 μ M Y-27632 for the first 24 hours. For the first 7 passages, cells were split twice a week at a ratio between 1:5-1:20. After passage 7, both lines were split at a ratio between 1:10-1:20 however CS cells were split once a week and CFS cells were split twice a week. Cells were maintained at 37°C, 5% CO₂ and 5% O₂.

For freezing, cells were resuspended in 1 ml of Cryostor cryopreservation medium after dissociation and centrifugation. Cryotubes were stored at -80°C for short-term and in liquid nitrogen for long-term storage.

2.2.3.2. Derivation of axial stem cells from EpiSCs

Mouse EpiSCs were rinsed with PBS and incubated with TrypLE Express at 37°C for 4 minutes followed by washing in the same reagent and mouse CFS medium consisting of RPMI-1640, 1X NEAA, 1X B27 without vitamin A, 50 ng/ml FGF2, 3 µM CHIR99021 and 10 µM SB431542. After centrifugation at 300 g for 3 minutes, cells were resuspended in mouse CFS medium including 10 µM Y-27632 only for the first 24 hours and plated at a ratio between 1:20-1:40. Cells were split either two or three times a week by incubating with TrypLE Express for 3 minutes at 37°C, and maintained on either a Matrigel-coated (1:100) plate continuously or on mitotically inactivated fibroblast feeders seeded at a density of 1×10^5 cells/cm² until passage 10 and then on a Matrigel-coated (1:100) plate. For both conditions, cells were cultured in a 37°C incubator with 5% CO₂ and 5% O₂, and 10 µM Y-27632 was added after each split for the first 24 hours. For freezing, cells were resuspended in 1 ml of Cryostor cryopreservation medium after dissociation and centrifugation. Cryotubes were stored at -80°C for short-term and in liquid nitrogen for long-term storage.

2.2.3.3. Derivation of axial stem cells from orangutan iPSCs

When orangutan iPSCs reached >90% confluency, the cells were treated with an induction medium comprising RPMI-1640, 1X B27 without insulin and 5 or 10 µM CHIR99021 for 24 hours at 37°C, 5% CO₂ and 5% O₂. The next day, cells were rinsed with PBS and dissociated with StemMACS Passaging Solution XF by incubating for 5 minutes at room temperature. After 5 minutes, StemMACS Passaging Solution XF was removed, and the cells were harvested in CS medium consisting RPMI-1640, 1X NEAA, 1X B27 without vitamin A, 5 µM CHIR99021 and 10 µM SB431542. Cells were plated onto Matrigel-coated (1:100) plates at a ratio between 1:10-1:20. Cells were split once or twice a week using StemMACS Passaging Solution XF, and 10 µM Y-27632 was added for the first 24 hours after each split. Fresh medium was applied daily and cells were maintained at 37°C, 5% CO₂ and 5% O₂.

2.2.4. Dissection of mouse embryos

Dissection of the stem zone region from CD1 mice at E8.5 was performed by Dr. Silvia Schirge (Institute of Diabetes and Regeneration Research at Helmholtz Centre Munich). Mice were kept and experiments were performed at the central facilities at Helmholtz Munich German

Research Center of Environmental Health in accordance with the German animal welfare legislation and acknowledged guidelines of the Society of Laboratory Animals (GV-SOLAS) and the Federation of Laboratory Animal Science Associations (FELASA). Mice were kept under SPF conditions in animal rooms with a light cycle of 12/12 hours, a temperature of 20-24°C and humidity of 45-65%. Mice received sterile filtered water and a standard diet for rodents ad libitum. For embryo generation mice at the age of ≥ 8 weeks were used. A total of 29 embryos from two mice (#1 and #2) were dissected. They were classified and pooled based on somite numbers as indicated in Table 6. Embryonic tissues were collected in ice-cold PBS.

Table 6: Number of embryos per somite stage and number of embryos pooled

Somite stage	Embryo number from mice #1	Embryo number from mice #2	Number of pooled embryos
2	1	-	1
3	1	1	5
4	3	-	
5	3	-	10
6	6	1	
7	3	2	7
8	-	2	
9	-	4	6
10	-	2	

Pooled embryonic tissues were centrifuged at 400 g for 4 minutes. The supernatant was discarded, and the pellet was dissolved in 0.25% Trypsin and incubated at 37°C for 15 minutes. After the incubation, FBS was added to the samples and each sample was divided into two tubes followed by centrifugation at 400 g for 4 minutes. The pellet was resuspended in 1X penicillin/streptomycin included CFS (RPMI-1640, 1X NEAA, 1X B27 without vitamin A, 100 ng/ml FGF2, 5 μ M CHIR99021 and 10 μ M SB431542) and CS medium (RPMI-1640, 1X NEAA, 1X B27 without vitamin A, 100 ng/ml FGF2, 5 μ M CHIR99021 and 10 μ M SB431542) separately, and 10 μ M Y-27632 was added for the first 24 hours. Cells were plated on 48-well plates. When the cells reached $>80\%$ confluency, they were washed PBS and dissociated with StemMACS Passaging Solution XF for 5 minutes at room temperature. The dissociation reagent was discarded, and cells were collected in medium for respective AxSC state including 10 μ M Y-27632 for the first 24 hours and plated on 12-well plates. After the first passage, cells were dissociated with TrypLE Express as described in section 2.2.2.2 and plated on 12-well plates at a ratio between 1:2 and 1:4. The medium was changed daily, and 1X penicillin/streptomycin was added during the whole cultivation process. Cells were maintained at 37°C, 5% CO₂ and 5% O₂.

2.2.5. Derivation of human neural progenitor cells

H9 cells were rinsed with PBS twice and incubated with Collagenase IV (2 mg/ml) for 30 minutes at 37°C followed by plating 1:1 onto low-attachment 6-well plate in a medium consisting of DMEM/F12, 20% KSR, 1% GlutaMAX, 1% NEAA, 1 µM DMH-1, 10 µM SB431542, 3 µM CHIR99021, 0.5 µM Purmorphamine (PMA), 10 µM Y-27632. The next 2 days, fresh medium was applied without Y-27632. On Day3 and Day4, the medium was replaced with 1:1 DMEM/F12 and Neurobasal-A medium, 1:100 B27 without vitamin A, 1:200 N2, 1% GlutaMAX, 1 µM DMH-1, 10 µM SB431542, 3 µM CHIR99021, 0.5 µM PMA. On Day5, the medium was changed to 1:1 DMEM/F12 and Neurobasal-A medium, 1:100 B27 without vitamin A, 1:200 N2, 1% GlutaMAX, 1 µM DMH-1, 50 µg/ml Ascorbic acid (AA). On Day6, the Day5 medium with 5 ng/ml FGF2 was added. On Day7, cells were plated on a Matrigel-coated (1:100) 6-well plate at 1:1 ratio in a medium consisting of 1:1 DMEM/F12 and Neurobasal-A medium, 1:100 B27 without vitamin A, 1:200 N2, 1% GlutaMAX, 1 µM DMH-1, 50 µg/ml AA, 20 ng/ml FGF2 (NPC medium). NPC medium was changed daily until Day14 and then every other day. On Day14 and Day21, cells were washed with PBS and treated with Collagenase IV (2mg/ml) at 37°C for 15 minutes and TrypLE Express in 37°C for 4 minutes respectively. Cells were plated on Matrigel-coated (1:100) plates at a ratio between 1:4-1:6. After Day21, NPCs were passaged using TrypLE Express when the confluency reached >70%. Cells were maintained at 37°C, 5% CO₂ and 5% O₂.

2.2.6. Neural differentiation from human axial stem cells

The neural differentiation protocol from AxSCs was adapted from Grosch et al²⁶⁷. Plates were coated 20 µg/ml Poly-D-lysine and 20 µg/ml Poly-ornithine in PBS overnight at 37°C. The next day, the coating solution was changed to 20 µg/ml Collagen I and 20 µg/ml Fibronectin in PBS and incubated at 37°C overnight. The following day, the coating solution was replaced with 10 µg/ml Collagen IV in 0.05M HCl and incubated for 2-4 hours at 37°C. For immunostaining experiments, cells were directly differentiated on either coverslips or chamber slides coated as described above.

Human CFS and CS lines were dissociated with TrypLE Express as described in section 2.2.2.1, and plated on previously coated plates in the respective medium including 10 µM Y-27632 at a density of 28.5x10³ cells/cm² for CFS and 71.5x10³ cells/cm² for CS. The next day, the medium was changed to the differentiation medium consisting of 1:1 DMEM/F12 and Neurobasal-A medium, 1X B27, 1X N2, 0.1 µM Retinoic acid (RA), 100 ng/ml SHH, 100 µM cAMP, 10 ng/ml GDNF, 10 ng/ml BDNF, 10 ng/ml IGF, 0.1 µM Compound E. The medium was changed every other day. Cells were maintained at 37°C, 5% CO₂ and 5% O₂.

2.2.7. Neural differentiation from human neural progenitor cells

The protocol from Grosch et al.²⁶⁷ was utilized for neural differentiation from neural progenitor cells. Plates were coated with 10 µg/ml Laminin and 10 µg/ml Poly-ornithine in PBS overnight at 37°C. The next day, the coating solution was changed to 10 µg/ml Collagen I and 10 µg/ml Vitronectin in PBS and incubated at 37°C overnight. The following day, the coating solution was replaced with 10 µg/ml Collagen IV in 0.05M HCl and incubated for 2-4 hours at 37°C. For immunostaining experiments, cells were directly differentiated on either coverslips or chamber slides coated as described above.

NPCs were rinsed with PBS and incubated with TrypLE Express in 37°C for 4 minutes. Cells were harvested in TrypLE Express followed by washing with NPC medium. After centrifugation at 300 g for 3 minutes, cells were plated on previously coated plates in NPC medium including 10 µM Y-27632 at a density of 42.8×10^3 cells/cm². The next day, NPC medium was replaced with differentiation medium consisting of 1:1 DMEM/F12 and Neurobasal-A medium, 1X B27, 1X N2, 0.1 µM RA, 100 ng/ml SHH, 100 µM cAMP, 10 ng/ml GDNF, 10 ng/ml BDNF, 10 ng/ml IGF. Medium was changed every other day until Day14. From Day14 onwards, 0.1 µM Compound E was added to the differentiation medium and fresh medium was applied every other day. Cells were maintained at 37°C, 5% CO₂ and 5% O₂.

2.2.8. Skeletal muscle differentiation from human axial stem cells

Skeletal muscle differentiation was performed using several protocols adapted from Choi et al.²⁶⁸ (SKMD#1) and Mavrommatis et al.²⁵² (SKMD#2). For SKMD#1, 12-well plates were coated with Matrigel (1:100 dilution) at 37°C for 4 hours at least. AxSCs were dissociated into single cells using TrypLE Express as described in section 2.2.2.2, and plated at a density of 5×10^4 (for the condition with DAPT only) and 3×10^5 cells/well (for the condition for DAPT and SHH) in the respective AxSC medium. The next day, the medium was replaced with differentiation medium including DMEM/F12, 1X B27 without vitamin A, 10 µM DAPT with or without 25 ng/ml SHH (medium#1). On Day8, cells were washed with PBS and dissociated using TrypLE Express at 37°C for 6 minutes followed by centrifugation at 400 g for 4 minutes. The supernatant was discarded, and the pellet was resuspended in medium#2 containing DMEM/F12, 1X NEAA, 1X ITS-G, 10 ng/ml FGF2, 100 ng/ml FGF8, with or without 50 µg/ml vitamin C and 5 µM cAMP. Cells were seeded on Matrigel-coated (1:100 dilution, minimum 4 hours at 37°C) plates at ratio of 1:3 and 1:4 per well of 12-well plate for the cells started with 5×10^4 and 3×10^5 cells/well density respectively. 10 µM Y-27632 was added for the first 24 hours. Medium was changed daily during the differentiation process. Cells were split once a week on Matrigel-coated (1:100 dilution, minimum 4 hours at 37°C) 12-well plates by

incubating cells with 0.05% Trypsin at 37°C for 6 minutes followed by centrifugation at 400 g for 4 minutes. The split ratio was 1:3 and 1:4 as described above. 10 µM Y-27632 was added only for the first 24 hours after each split. Cells were either maintained in medium#2 until Day 40, or in medium#3 (DMEM/F12, 1X NEAA, 2% horse serum) between Day26 and 40. Cells were maintained at 37°C, 5% CO₂ and 5% O₂.

For SKMD#2, AxSCs were dissociated with TrypLE Express as described in section 2.2.2.2, and seeded on a Matrigel-coated (1:100 dilution, minimum 4 hours at 37°C) plate in the respective medium including 10 µM Y-27632, at a density of 1.5-2.5x10⁴ cells/cm². The next day, the medium was changed to DMEM/F12, 1X NEAA and 1X ITS-G, 3 µM CHIR99021, 0.5 µM LDN193189, 5 ng/ml FGF2, 10 nM RA. On Day2, FGF2 and RA were replaced with 34 ng/ml SHH. On day 6, the medium was changed to DMEM/F12, 1X NEAA and 1X ITS-G, 10 ng/ml FGF2 and 10 ng/ml HGF. On Day10, cells were split by washing with PBS and incubating with TrypLE Express for 6-7 minutes at 37°C. Cells were harvested and centrifuged at 400 g for 4 minutes. The supernatant was removed, and the pellet was resuspended in DMEM/F12, 1X NEAA and 1X ITS-G, 10 ng/ml FGF2, 10 ng/ml HGF and plated on a Matrigel-coated (1:100 dilution, minimum 4 hours at 37°C) plate at a density of 2-2.8x10⁴ cells/cm². From Day12 onwards, the medium was replaced with DMEM/F12, 1X NEAA and 1X ITS-X, 10 ng/ml HGF. Fresh medium was applied every other day during the differentiation process. Cells were maintained at 37°C, 5% CO₂ and 5% O₂.

2.2.9. Nephron differentiation from human axial stem cells

The nephron differentiation protocol was adapted from Morizane et al.²⁶⁹. Human AxSCs were dissociated with TrypLE Express as described in section 2.2.2.2, and seeded on a Matrigel-coated (1:100) plate in the respective medium including 10 µM Y-27632 at a density of 2.7-3.5x10⁴ cells/cm² for CFS and 5x10⁴ cells/cm² for CS. The medium was changed daily for 4 days. AxSC medium was then replaced with nephron differentiation medium #1 consisting of RPMI-1640, 1X GlutaMAX, 2X B27, 10ng/ml Activin A. On Day3, the medium was changed to nephron differentiation medium #2 (NDM2) consisting of RPMI-1640, 1X GlutaMAX, 2X B27, 10 ng/ml FGF9. On Day5, fresh NDM2 medium with 3 µM CHIR99021 was added. From Day7 onwards, fresh NDM2 medium was applied daily. Cells were maintained at 37°C, 5% CO₂ and 5% O₂.

2.2.10. Cardiomyocyte differentiation from human axial stem cells

The cardiomyocyte differentiation protocol was adapted from Lian et al.²⁷⁰. Human AxSCs were dissociated with TrypLE Express as described in section 2.2.2.2, and seeded on a

Matrigel-coated (1:100) 12-well plate in the respective medium including 10 μ M Y-27632 at a density of 2×10^5 , 5×10^5 and 1×10^6 cells/well. On Day0 and Day1, fresh medium of RPMI-1640, 2% B27 without insulin was added. On Day2, half of the medium was changed to RPMI-1640, 2% B27 without insulin, 10 μ M IWP2. On Day4, the medium was replaced with RPMI-1640, 2% B27 without insulin. From Day6 onwards, the medium was changed every 3 days with RPMI-1640, 2% B27. Cells were maintained at 37°C, 5% CO₂ and 5% O₂.

2.2.11. RNA extraction and cDNA synthesis

RNA extraction was performed using the RNeasy Mini Kit according to the manufacturer's instructions. RNA concentration was measured using the Nanodrop system. The samples with equal RNA concentrations were reverse-transcribed using the SuperScript III First-Strand Synthesis System according to the manufacturer's instructions.

2.2.12. Quantitative RT-PCR

RT-qPCR was performed in 384-well plates using either the Power SYBR Green PCR Mastermix or the Taqman Gene Expression Assay Mastermix. The total reaction volume was 10 μ l for both approaches. 5 μ l of Power SYBR Green PCR Mastermix, 1 μ l of cDNA, 1 μ l of 5 μ M forward and reverse primer and 3 μ l nuclease-free water were mixed in each well of 384-well plate. After centrifugation at 300 g for 2 minutes, the reaction was run under the following cycling conditions: 2 minutes at 50°C, 10 minutes at 95°C, 40 cycles of 15 second at 95°C and 1 minutes at 60°C. 5 μ l of Taqman Gene Expression Assay Mastermix, 1 μ l cDNA, 0.5 μ l Taqman probe and 3.5 μ l nuclease-free water mix was centrifuged at 300 g for 2 minutes and run under following cycling conditions: 2 minutes at 50°, 10 minutes at 95°C, 40 cycles of 15 second at 95°C and 1 minutes at 60°C. Primers and probes are listed in Table 3. For orangutan AxSC analysis, human primers were used except for PAX6. Delta-Delta Ct method was used for calculation of relative expression levels normalized to GAPDH housekeeping gene.

2.2.13. Immunostaining

For the goat anti-Brachyury (#AF2085, R&D Systems) antibody, the following protocol was used. Cells grown on either coverslips or chamber slides were washed with 0.1% BSA in PBS and fixed with 4% paraformaldehyde for 20 minutes at room temperature. After washing twice with 0.1% BSA in PBS, permeabilization was performed using 0.2% Triton X-100 in PBS for 10 minutes at room temperature. Cells were blocked in 5% BSA in PBS for 1 hour at room temperature and incubated with primary antibodies diluted in 0.5% BSA/1% Tween-20 in PBS at 4°C overnight. After washing with 1% Tween-20 in PBS twice for 5 minutes each,

cells were treated with secondary antibodies 1:750 diluted in 0.5% BSA/1% Tween-20 in PBS for 3 hours at room temperature. After washing with 1% Tween-20 in PBS two times for 5 minutes each and once for 10 minutes, coverslips or chamber slides were mounted using Prolong Gold Antifade and incubated overnight at room temperature.

For other antibodies, cells grown on either coverslips or chamber slides were washed with PBS twice followed by fixation with 4% formaldehyde in PBS for 15 minutes at room temperature. After washing with PBS, cells were permeabilized with 0.2% Triton X-100 in PBS for 5 minutes at room temperature. Cells were then blocked in 3% BSA for 30 minutes at room temperature and incubated with primary antibodies diluted in 0.1% Triton X-100/3% BSA in PBS at 4°C overnight. The next day, cells were washed with PBS three times for 5 minutes each and incubated with secondary antibodies 1:1000 diluted in 0.1% Triton X-100/3% BSA in PBS for 1 hour at room temperature. After washing with PBS three times 10 minutes each, coverslips or chamber slides were mounted using Prolong Gold Antifade and incubated overnight at room temperature. Primary and secondary antibodies are listed in Table 4. Imaging was performed using Axio Observer Z1 microscope.

2.2.14. Genomic DNA extraction and telomere length assay

DNA extraction was performed using the Wizard SB Genomic DNA Purification System kit with a minor modification of addition 2 µl of RNase A solution per 250 µl nuclease free water and elution after 10 minutes incubation at room temperature. DNA concentration was measured using the Nanodrop system.

For the telomere length assay, Telomere and single copy reference (SCR) primers from the Relative Human Telomere Length Quantification qPCR kit were prepared according to the manufacturer's instructions. 20 µl total reaction volume consisted of 2 µl of telomere or 2 µl of SCR primers, 1 µl of genomic DNA template (5 ng/µl), 10 µl Power SYBR Green PCR Mastermix and 7 µl nuclease-free water were mixed in one well of a 384-well plate. After centrifugation at 300 g for 2 minutes, the reaction was run under the following cycling conditions: 2 minutes at 50°C, 10 minutes at 95°C, 40 cycles of 15 second at 95°C and 1 minutes at 60°C. Delta-Delta Ct method was used for calculation of the relative expression levels of telomere normalized to SCR.

2.2.15. Single-cell RNA sequencing

2.2.15.1. Preparation of single-cell suspensions

All parental lines (H9 p60, CFS_2 p14, CS_2 p14, NPC p3, CFS_H9/HUES/HMGU p15-18, CS_H9/HUES6/HMGU1 p12-14) and Day2 neural differentiation samples were dissociated into single-cells using TrypLE Express. The harvested cells were centrifuged at 300 g for 3 minutes. The pellet was resuspended in 0.04% BSA in PBS and the cell density was adjusted to 1000 cells/ μ L.

Day 14 and Day28 CFS-derived cells as well as Day28 CS-derived cells were dissociated with 0.05% Trypsin. After centrifugation at 300 g for 3 minutes, the pellet was resuspended in 0.04% BSA in PBS. The cells were filtered using 40 μ m strainers and the cell density was adjusted to 1000 cells/ μ L.

Day 14 CS-derived cells were dissociated using a protocol adapted from Haldane et al.²⁷¹. The cells were rinsed with PBS followed by incubation at 37°C for 25 minutes with a dissociation buffer consisting of 50% Accutase, 10% Papain in PBS. Without removing the dissociation buffer, wash buffer containing DMEM/F12, 33.3 μ g/ml Dnase II and 10 μ M Y-27632 was added. The cell suspension was filtered using a 70 μ m strainer and centrifuged at 300 g for 3 minutes. The pellet was resuspended in neural differentiation medium and filtered using a 40 μ m strainer.

2.2.15.2. Library preparation and sequencing

2.2.15.2.1. Single cell sequencing of parental human AxSC lines

Live cells were counted using trypan blue stain with the Countess 3 automated cell counter. Samples with viability above 85% were used to prepare cell suspensions containing 1000 cells/ μ L, which were subsequently used for the single cell library preparation. The leftover cells were pelleted via centrifugation and collected for DNA isolation for genotyping purposes with the HumanCytoSNP-12 v2.1 BeadChip Kit. The single cell library preparation was done using Chromium Controller instrument, Chromium Next GEM Single Cell 3' GEM, Library & Gel Bead Kit v3.1 and the Single Index Kit T Set A. The final sample loaded for the 10X library prep contained a mixture of 7000 cells/cell line for a total of 21000 cells. Quality control for the library was performed using Agilent High Sensitivity kit following manufacturer instructions. The final pools were sequenced on NovaSeq 6000 platform using SP Reagent Kit v1.5/ 100 cycles with 28x8x91 configuration run leading to 40000 reads/cell.

2.2.15.2.1. Single cell sequencing of neural differentiation time course from human AxSC lines

Counting was performed with Trypan blue for live-dead cell discrimination using Countess automated counter. Samples with viability above 81% were used to prepare cell suspensions containing 500-1000 cells/ μ L. The single cell library preparation was performed using Chromium Controller instrument, Chromium Next GEM Single Cell 3' GEM, Library & Gel Bead Kit v3.1 and Dual Index Kit TT Set A, 96 rxns. Quality control for the library was performed using Agilent High Sensitivity kit following manufacturer instructions. The final pool contained 16 samples each with 10000 cells per time point and they were sequenced on NovaSeq 6000 platform using NovaSeq 6000 S4 Reagent Kit v1.5 /200 cycles with 28x10x10x90 configuration run leading to 70000 reads/cell.

2.2.15.3. Processing of the datasets

For human parental human AxSC lines, SNP sequencing was performed using Illumina GSA v.3 for cells of each cell line. Next, each barcode was assigned to a specific cell line using demuxlet²⁷² by Moritz Thomas (Institute of Computational Biology at Helmholtz Centre Munich). Briefly, demuxlet harnesses genetic variation to identify the genetic origin of individual cells in a mixed sample. Ambiguous cells that could not be confidently matched to one of the three cell lines were removed.

The following process was carried out by Ksenia Arkhipova (Leiden Academic Centre for Drug Research-LACDR, Leiden University). Sequencing reads from parental cell line (H9, CFS and CS AxSCs) dataset and from neural differentiation time course dataset were annotated using CellRanger software (v 6.1.2, 10X Genomics) provided with the custom reference (GRCh38 assembly, NCBI annotation of December 2021). The subsequent analysis of the datasets was performed in the Scanpy²⁷³ environment. Low quality cells (mitochondrial genes > 20%, ribosomal genes > 10% and parameters in Table) were removed. The cell quality filtering process involved applying gene and count thresholds as follows: a minimum of 2000 genes for the differentiating cells, 800 genes for parental CS lines, and 1000 genes for parental CFS lines. Additionally, the total count thresholds were 80000, 60000, and 60000 respectively for the three datasets mentioned. Cell doublets were predicted with Solo software²⁷⁴ and filtered out. Genes present in less than 20 cells were removed. Gene counts were normalized to 10^5 reads per cell and log-transformed. An impact of batch effect on cell clustering was assessed with CarDEC software²⁷⁵. If the general clustering structure was the same after CarDEC, the initial unchanged data and embeddings (UMAP) were used.

Annotated raw count matrices from Rayon et al.²⁷⁶ were downloaded from GEO database (GSE171892), which were filtered in the similar way (minimal number of genes 700, total count 60000) and used for gene expression comparisons.

2.2.15.4. Downstream analysis

Scanpy (1.6.1), Pandas (1.1.3), Numpy (1.19.5) and Louvain (0.7.0) packages were used for downstream analysis. Datasets were visualized by UMAP embedding using the `pl.umap()` function. Differentially expressed gene (DEG) analysis in the H9, CFS and CS included dataset was performed using the `tl.rank_genes_groups()` function with default parameters and the list of the DEGs was extracted using the `get.rank_genes_groups()` function, setting 'reference' as 'rest' and the log2 fold change threshold to 2. The top 50 DEG genes were visualized using the `pl.dotplot()` function. The genes associated with axial development and neural/mesodermal lineages (Fig. 16) were summarized using published studies that carried out scRNA sequencing experiments within in vivo cells^{59,66,243,277}. To assess the expression of dorsoventral spinal cord progenitor genes, the study from Rayon et al.²⁷⁶ was used as a benchmark.

The datasets from the time course neural differentiation experiment were analyzed individually. Cell cycle phases were identified using the `tl.score_genes_cell_cycle()` function with default parameters. The percentage of cells per phase was calculated using the `value_counts()` function. Clustering in each dataset was performed using the `tl.louvain()` function by setting 'resolution' to 2. The marker genes mapping each domain of spinal cord neurons identified by Rayon et al.²⁷⁶ as well as the accompanying database (<https://shiny.crick.ac.uk/scviewer/neuraltube/>), and marker genes for neural crest development identified by Soldatov et al.¹³⁹ were utilized for cluster annotation and summarized in Fig. 5Sa. To evaluate certain cell types, single genes were used for the analysis because these genes have been shown to be specific to only one domain in the database above. To target cell populations which could not be identified by a single gene, coexpression of several genes was assessed. Depending on the expression of marker genes in the datasets, the clusters were manually combined and annotated. Clusters representing the same cellular population across the datasets were shown in the same color.

2.2.16. Chromatome and proteome analysis

The chromatome and proteome experiments were performed in a collaboration with Enes Ugur (Faculty of Biology and Center for Molecular Biosystems at Ludwig Maximilians University).

2.2.16.1. Nuclei isolation and sample preparation for chromatome analysis

AxSCs (p14-15) and H9 (p61) cells were dissociated into single cells using TrypLE Express by incubation for 4-6 minutes at 37°C. Cell suspensions were centrifuged at 300 g for 3 minutes. Cell number was counted. 5×10^6 cells for each derivaton of CS and H9, 3×10^6 cells for each derivation of CFS cells were washed twice with PBS by centrifugation at 2300 g for 5 minutes at room temperature. Supernatant was discarded, and pellet was resuspended in 1 ml of ice-cold lysis buffer consisted of 3 mM MgCl₂, 10 mM NaCl, 10 mM Tris 7.4, 1% NP-40 and freshly added 1X cOmplete protease inhibitor. To dissolve NP-40 thoroughly, lysis buffer was priorly incubated on orbital shaker at mild agitation at 4°C before use for pellet resuspension. Pellet was homogenized by pipetting up and down. Cells were incubated on ice for 20 minutes. Suspension was centrifuged at 2300 g for 5 minutes at 4°C. Supernatant was discarded, and pellet was resuspended and incubated in 3.3 ml PBS with 1% methanol-free formaldehyde for 10 minutes on rotating wheel at mild agitation at room temperature. Reaction was quenched with 125 mM Glycine by incubation for additional 5 minutes on rotating wheel. Nuclei suspension was centrifuged at 2300 g for 5 min at 4°C and washed twice with ice-cold PBS. After additional centrifugation at 2300 g for 5 min at 4°C, supernatant was discarded and pellets were frozen in liquid nitrogen for 15 seconds followed by storage at -80°C. Chromatin was released from crosslinked nuclei by dissolving 300 µl of SDS buffer (50 mM HEPES pH 7.4, 10 mM EDTA pH 8.0, 4% UltraPure™ SDS Solution, along with a newly added 1× cOmplete™ EDTA-free Protease Inhibitor Cocktail) using gentle pipetting. This mixture was left to incubate at room temperature for 10 minutes before adding 900 µl of freshly prepared Urea buffer (10 mM HEPES pH 7.4, 1 mM EDTA pH 8.0, 8 M urea). The solution was then carefully inverted seven times before being centrifuged at room temperature for 30 min at 20000 g. The supernatant was removed, taking care not to disturb the pellet. Two additional wash steps were performed (one wash with SDS and urea, and one wash with only SDS). The final pellet was then dissolved in 300 µl of Sonication buffer (10 mM HEPES pH 7.4, 2 mM MgCl₂, with a freshly added 1× cOmplete™ EDTA-free Protease Inhibitor Cocktail). The chromatin specimens were sonicated using a Bioruptor® Plus at 4°C for 15 cycles (30 s on, 30 s off). The protein concentration was determined using the Pierce™ BCA Protein Assay Kit.

Subsequently, Protein Aggregation Capture (PAC) was performed. In this step, Sera-Mag™ beads (1 mg) were washed three times with 70% acetonitrile for every 100 µg of chromatin solution. After the final wash, 300 µl of the chromatin solution corresponding to 100 µg was added to the beads, followed by 700 µl of 100% acetonitrile. The chromatome-bead mixtures were then vortexed and left to rest on a bench for 10 minutes. The samples were vortexed again and placed into a magnetic rack. The samples were then washed with 700 µl of 100%

acetonitrile, followed by 1 ml of 95% acetonitrile, and finally with 1 ml of 70% ethanol. The remaining ethanol was allowed to evaporate, and the beads were resuspended in 400 μ l of 50 mM HEPES pH 8.5, supplemented with freshly prepared 5 mM Tris(2-carboxyethyl)phosphine (TCEP) and 5.5 mM chloroacetamide (CAA). The samples were then left to incubate at room temperature for half an hour. Protease digestion was initiated by adding LysC (protease to protein ratio of 1:200) and Trypsin (1:100) and allowing the mixture to incubate overnight at 37°C under constant agitation at 1100 rpm. From this point forward, the samples were handled in the same way as the total proteome samples.

2.2.16.2. Sample preparation for total proteome analysis

All proteomic experiments were performed in triplicates. AxSCs (p14-15) and H9 (p61) cells were dissociated into single cells using TrypLE Express by incubation for 4-6 minutes at 37°C. Cell suspensions were centrifuged at 300 g for 3 minutes. Cell number was counted. 2×10^6 cells for each derivation of CS and H9, 1.5×10^6 cells for each derivation of CFS cells were centrifuged at 300 g for 3 minutes. Supernatant was removed, and pellets were snap-frozen in liquid nitrogen for 15 seconds followed by storage at -80°C. Frozen cells were dissolved in 200 μ l of lysis buffer (containing 6 M guanidinium chloride, 100 mM Tris-HCl with a pH of 8.5, and 2 mM DTT) and subjected to heating for 10 minutes at 99°C with a constant agitation rate of 1400 rpm. The sonication of samples was then performed at 4°C using a Bioruptor® Plus sonication device at high-intensity settings, with 30 seconds on/off intervals for 15 rounds. If the sample viscosity was adequately reduced, protein concentrations were determined; if not, sonication was repeated. Protein concentrations were assessed using the Pierce™ BCA Protein Assay Kit in accordance with the manufacturer's instructions. After incubating for at least 20 minutes with 40 mM chloroacetamide, 30 μ g of each proteome sample was diluted in a 30 μ l lysis buffer supplemented with chloroacetamide and DTT. These samples were further diluted in 270 μ l of digestion buffer (containing 10% acetonitrile, 25 mM Tris-HCl at pH 8.5, 0.6 μ g Trypsin/sample, and 0.6 μ g/sample LysC. Proteins were then digested for 16 hours at 37°C with constant shaking at 1100 rpm.

To halt protease activity, 1% (v/v) trifluoroacetic acid (TFA) was added the following day and samples were loaded onto homemade StageTips composed of three layers of SDB-RPS matrix²⁷⁸, previously equilibrated with 0.1% (v/v) TFA. After loading, two washes with 0.1% (v/v) TFA were performed, and peptides were eluted with 80% acetonitrile and 2% ammonium hydroxide. After the eluates were evaporated in a SpeedVac centrifuge, the samples were resuspended in 20 μ l 0.1% TFA and 2% acetonitrile. The peptides were completely solubilized by constant shaking for 10 minutes at 2000 rpm, and peptide concentrations were determined on a Nanodrop™ 2000 spectrophotometer at 280 nm.

2.2.16.2. Nanoflow LC-MS/MS analysis for proteomes and chromatomes

Peptide separation before MS was accomplished using liquid chromatography on an Easy-nLC 1200 with in-house packed 50 cm columns of ReproSilPur C18-AQ 1.9- μm resin. A binary buffer system was used (buffer A: 0.1% formic acid and buffer B: 0.1% formic acid in 80% acetonitrile), with a gradual increase in buffer B concentration (from 5% initially to 95% at the end) to elute the peptides over a 120-minute period at a steady flow rate of 300 nl/min. The peptides were then introduced into an Orbitrap Exploris™ 480 mass spectrometer via a nanoelectrospray source. Each set of triplicates were followed by a washing step while the column temperature was constantly at 55°C.

Data-Dependent Acquisition (DDA) runs used a top12 shotgun proteomics method within a range of 300–1650 m/z, a default charge state of 2, and a maximum injection time of 25 ms. Full scan resolution was set at 60000 and the normalized AGC target at 300%. For MS2 scans, the orbitrap resolution was set at 15000 and the normalized AGC target at 100%, with a maximum injection time of 28 ms.

Data-Independent Acquisition (DIA) runs used an orbitrap resolution of 120000 for full scans in a scan range of 350–1400 m/z, with a maximum injection time of 45 ms. For MS2 acquisitions, the mass range was set to 361–1033 with isolation windows of 22.4 m/z. A default window overlap of 1 m/z was used. The orbitrap resolution for MS2 scans was set at 30000, the normalized AGC target at 1000%, and the maximum injection time at 54 ms.

2.2.16.3. MS data quantification

Raw MS data acquired in DIA mode was analyzed using DIA-NN version 1.8.1²⁷⁹. A hybrid approach utilizing a dedicated DDA and DIA library was employed; the DDA library was generated using SpectroMine and the DIA library using DIA-NN. Cross-run normalization was conducted in an RT-dependent manner. Missed cleavages were set to 1, N-terminal methionine excision was activated, and cysteine carbamidomethylation was set as a fixed modification. Proteins were grouped using the additional command ‘-relaxed-prot-inf’. Match-between runs was enabled, and the precursor FDR was set to 1%.

2.2.16.4. Statistical analyses

Raw data outputs were analyzed downstream with Perseus (version 1.6.0.9)²⁸⁰. CVs were calculated by filtering out proteins or precursors with fewer than 2 out of 3 valid values. Downstream analyses were conducted after imputation of missing values, which was done based on a Gaussian distribution with respect to the standard deviations of measured values

(width of 0.2 and a downshift of 1.8 standard deviations). GO enrichment analyses of differentially enriched proteins were performed against the background of total identified proteins using a Benjamini-Hochberg FDR-corrected Fisher's Exact test of 0.05. These analyses were carried out individually for each cluster. Student's t-tests were calculated with a permutation-based FDR of 0.05 and an s_0 value of 1, unless stated otherwise. For the multiple sample test based on an ANOVA the FDR was set to 0.05 and the s_0 value to 2. ANOVA tests of normalized chromatomes were conducted likewise after calculating pairwise differences of ChAC-DIA and total proteome values.

2.2.16.5. Web application development

Significant changes in chromatome, proteome and relative chromatin binding during AxSCs differentiation were displayed in an interactive heatmap as mean row differences of log₂ intensities.

The web application was programmed using R Shiny with the following libraries besides base R packages for data processing and visualization: shiny (1.7.1), shinydashboard (0.7.2), shinyHeatmaply (0.2.0), plotly (4.10.0), heatmaply (1.3.0) and png (0.1-7). From the tidyverse (1.3.1) family we further utilized tidyr (1.2.0), dplyr (1.0.9), and ggplot2 (3.3.6).

3. RESULTS

3.1. Establishment of human axial stem cells

Generation of an indefinitely self-renewing stem cell state corresponding to that of the *in vivo* axial progenitors requires first and foremost mimicry of the environmental cues resembling those of the posterior embryonic regions, namely the tailbud, where NMPs are located. It is known that active WNT and FGF signaling are the hallmarks of the posterior embryonic regions. FGF signaling is the upstream effector of Notch pathway known for inducing the segmentation clock¹⁰⁰. To avoid commitment of the cells to the mesodermal lineage and simultaneously to obtain bipotent NMP-like cells, I first activated only WNT signaling by using CHIR99021 in H9 human embryonic stem cells (hESC). This step aims to induce posterior identity. H9 cells were routinely split as clumps during their maintenance (**Fig. 9A**). Prior to CHIR99021 induction, I adapted H9 cells to single-cell splitting by performing 3 consecutive passages. The known cytotoxicity of CHIR99021 lead me to plate the H9 cells highly confluent (see Methods section) 24 hours before induction. Next day, the cultures were treated with 10 μ M CHIR99021 for 24 hours which was followed by an additional single-cell split in the two different media compositions comprising of: 1) 10 μ M CHIR99021, 100 ng/ml recombinant FGF2, 10 μ M SB431542 named CFS medium, and 2) 10 μ M CHIR99021, 10 μ M SB431542 named CS medium (**Fig. 9A**). The CFS and CS cells were regularly split as single-cell cultures. Until passage 6-8, I commonly observed heterogeneous cultures for both growth conditions where the stem cell-like colonies and flat differentiated cells were present (**Fig. 9B**). I considered clump passaging as an option to eliminate the differentiated-like cells, but it induced spontaneous differentiation and it further impaired the cultures. As a result, CFS and CS cells were continuously split as single-cell cultures throughout their establishment and long-term maintenance. For establishment period, both types of cells were split twice per week. After 6-8 passages, the flat-shaped cells were not observed in culture and the cultures became considerably more homogeneous (**Fig. 9B**). From this timepoint, I split CFS cells twice a week and CS cells once a week. The cells were kept in culture for more than 30 passages and the stem cell-like morphology of the respective populations did not changed over the time. I derived three CFS and three CS cells from H9 hESC by performing independent derivation experiments (named CFS_1-3, CS_1-3). Both H9 cells and putative axial progenitor cells were continuously cultured in hypoxic conditions as it has been shown that hypoxia enhances SOX2/TBXT colocalization at posterior regions together with CHIR treatment²⁸¹.

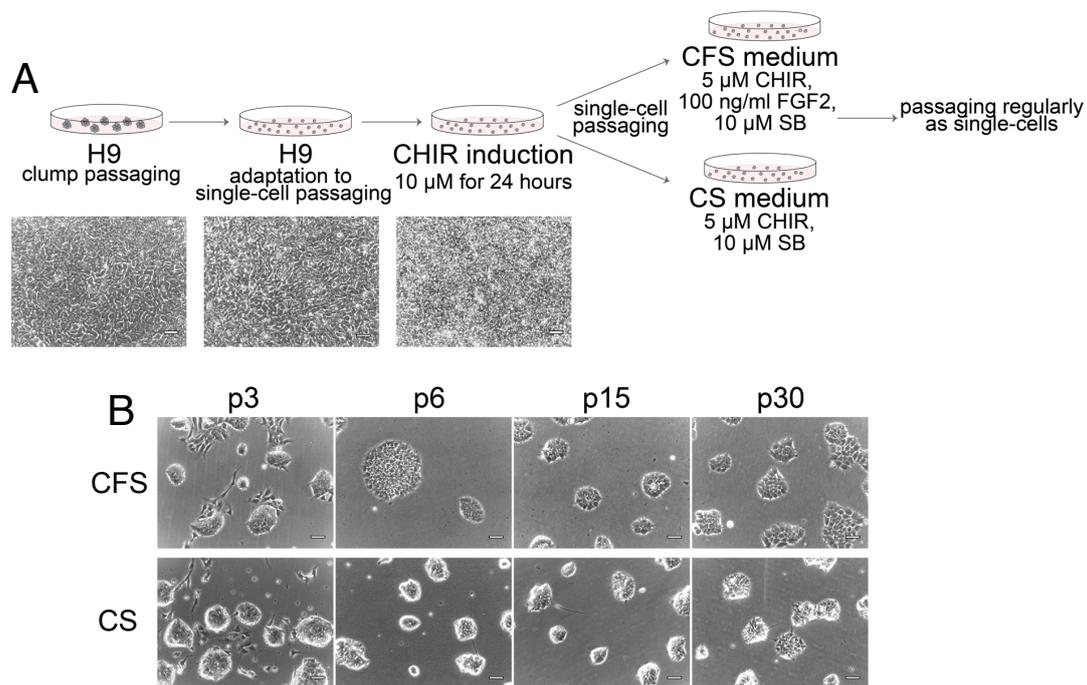


Figure 9: Establishment of human AxSC lines and their long-term morphology

(A) Schematic illustration of the protocol used to generate AxSCs and bright field images from H9 cultures. H9 cells routinely split as clumps were adapted to single-cell splitting over the course of 3 passages (between p55-60). The cultures were used at high confluency supplemented with 10 μ M CHIR99021 for 24 hours. Induced cells were then split into two different AxSC medium which named CFS (5 μ M CHIR99021, 100 ng/ml FGF2 and 10 μ M SB431542) and CS (5 μ M CHIR99021 and 10 μ M SB431542). CFS and CS cells were regularly split as single-cells (scale bars: 50 μ M).

(B) Brightfield images of human AxSCs at different time points (p: passage number, scale bars: 50 μ M).

After establishment of stable AxSCs cultures, I evaluated expression of the key developmental markers associated with pluripotent cells, NMPs, and neural tube progenitors. I performed RT-qPCR experiment by using 24-hour induced cells with CHIR99021, CFS_1-3 and CS_1-3 between passage 1 to 27 (**Fig. 10**). The expression values were normalized to H9 hESCs. *OCT4* and *NANOG* the well-known pluripotency markers²¹⁰ were downregulated in both CFS and CS cultures. *SOX2*, which is expressed in both pluripotent cells and NMPs, was found in CFS and CS cells at similar level compared to H9. Expression of *TBXT* (also known as *T* in human and *Brachyury* in mouse), which is the NMP/axial progenitor marker²⁷, was detected as the highest in the 24-hour induced cells, then gradually decreased in both cell types. CFS cells from passage 1 to 27 express similar levels of *TBXT*, especially starting from p10 coinciding with the timepoint when CFS cultures showcase homogeneous stem cell like phenotype with minor spontaneous differentiation based on the morphological observations. *TBXT* expression was sharply downregulated in CS cells after p1 and it did not show an upregulation over time. CS culture exhibited high *PAX6* expression, which is known as a

neural tube marker²⁸², while CFS cells did not. Additionally, I tested expression of *CDX2* because NMPs are shown to be *CDX2* positive. I found that it was gradually upregulated in CFS cells while CS cells did not have *CDX2* expression after p1.

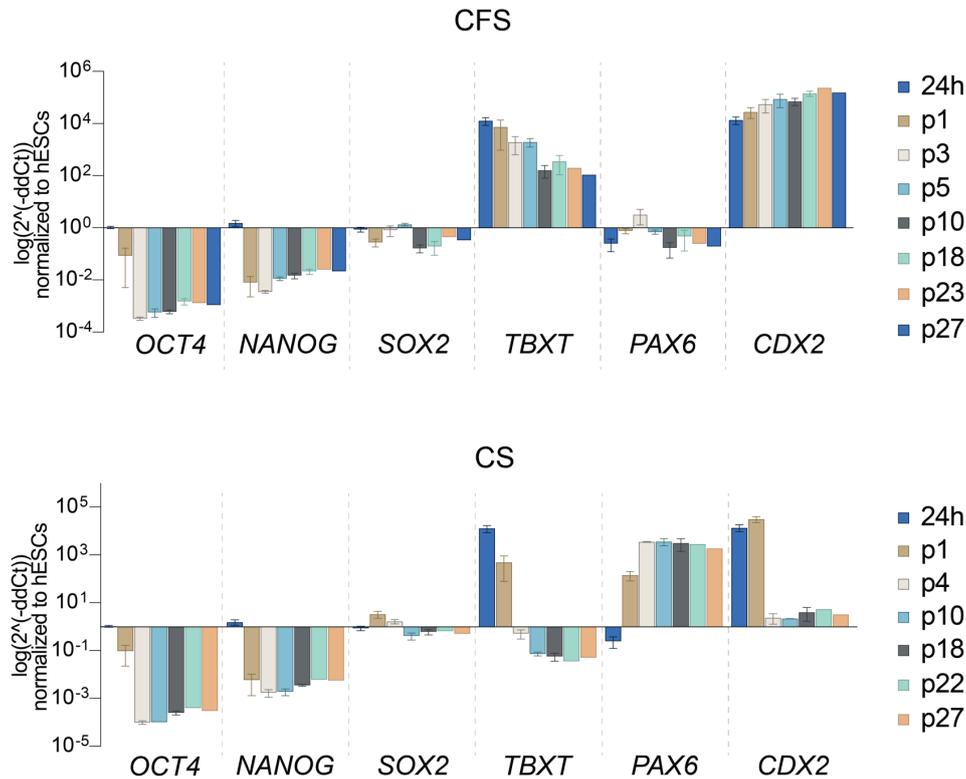


Figure 10: Transcriptional characterization of human AxSCs by the markers of different developmental stages

Expression analysis of pluripotency- (*OCT4*, *NANOG*), NMP- (*SOX2*, *TBXT*, *CDX2*) and neural progenitor (*PAX6*)-related transcription factors by RT-qPCR in three independent derivations per CFS (named CFS_1-3) and CS (named CS_1-3) lines generated from H9 cells (p55-60). The Ct values were normalized to undifferentiated H9 cells. Error bars represent SEM (p: passage number, N=2 or 3 up to p10 and N=1 for p22-27.)

Expression of *SOX2*/*TBXT*/*CDX2* in CFS cells and *SOX2*/*PAX6* in CS cells were also confirmed at the protein level by performing immunostaining experiments (**Fig. 11**). I used H9 cells as control and I did not observe an antibody stain for *CDX2*, *TBXT* and *PAX6* but I could obtain a *SOX2* positive stain (Supp. Fig. 1A). Furthermore, staining with IgG controls was applied for each cell line (Supp. Fig. 1B) that validated the results shown in **Fig. 11**. A slight signal was detected in mouse IgG controls, but it was much lower than the *SOX2* signal and not nuclear-specific. Taken together, *SOX2*/*TBXT*/*CDX2* and *SOX2*/*PAX6* are defined as the marker signature of CFS and CS cells respectively. Marker expression on the protein level did not change over the time implying the stem cell-like characteristics of CFS and CS cells.

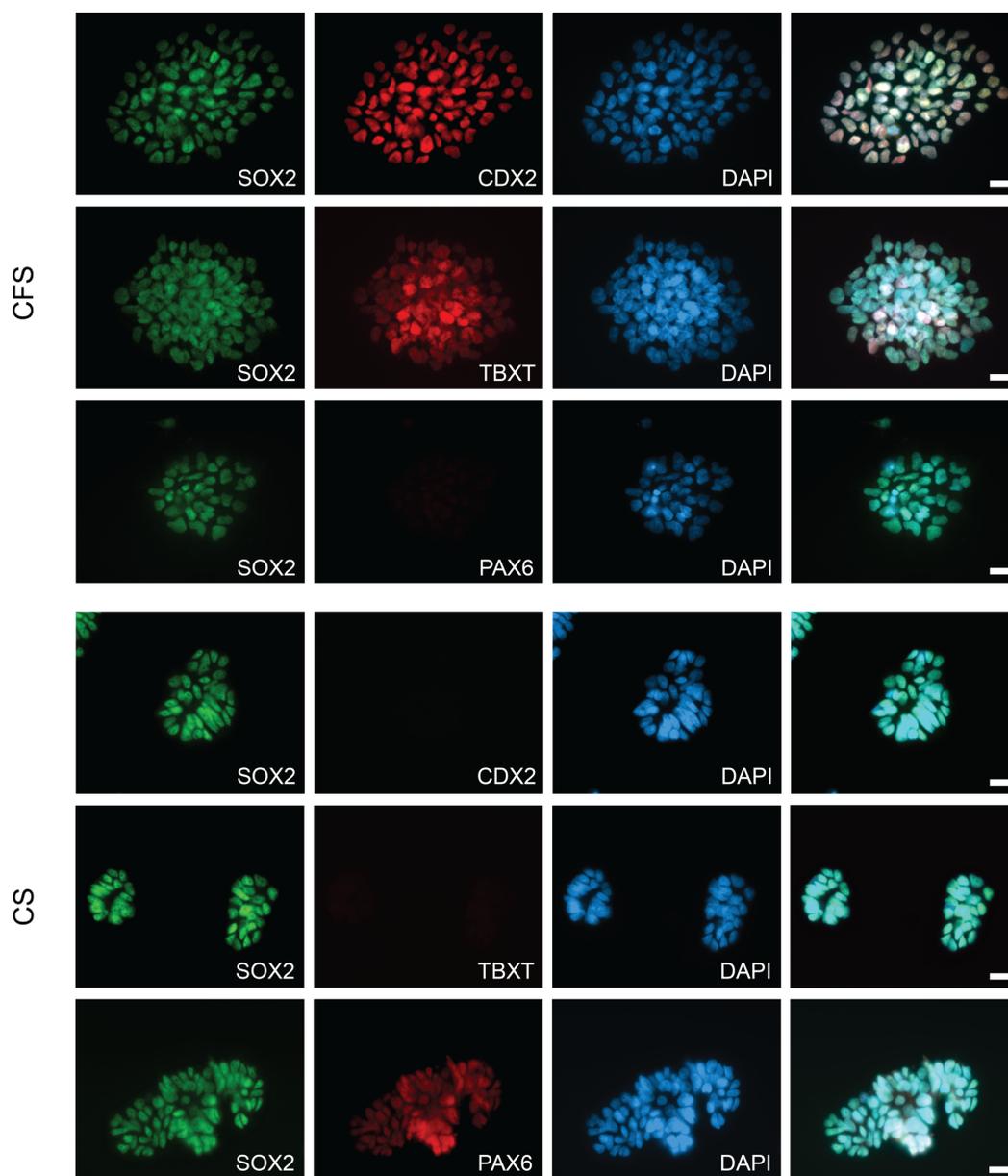


Figure 11: Expression of key transcription factors in human AxSCs on the protein level

Representative images from immunostaining with SOX2, CDX2, TBXT, PAX6 and DAPI in CFS_2 line at passage 30 and CS_2 line at passage 30 (scale bars: 20 μ M).

One of the key features characterizing stem cells is the maintenance of telomere length that enables their indefinite self-renewing capabilities²⁸³. I assessed the relative telomere length via a time course experiment for the 3 independently derived CFS (each at p15, p25 and p35) and CS lines (each at p10, p20 and p30) (**Fig. 12**) that were transcriptionally analyzed in **Fig.**

10A. Two timepoints from H9 cells (p54 and p64) were used as positive control for the assay. I performed RT-qPCR by using telomere-specific primers and normalized all values to H9 cells at p45 (internal control). The results pointed out that the telomere length does not decrease with increased passaging neither for CFS or CS lines (**Fig. 12**). One of the CS lines showed an increased telomere length over time however this phenomenon was similarly detected for H9 cells. In conclusion, the maintenance of telomere length proved the self-renewal activity in CFS and CS cells therefore they are collectively named as axial stem cells (AxSCs).

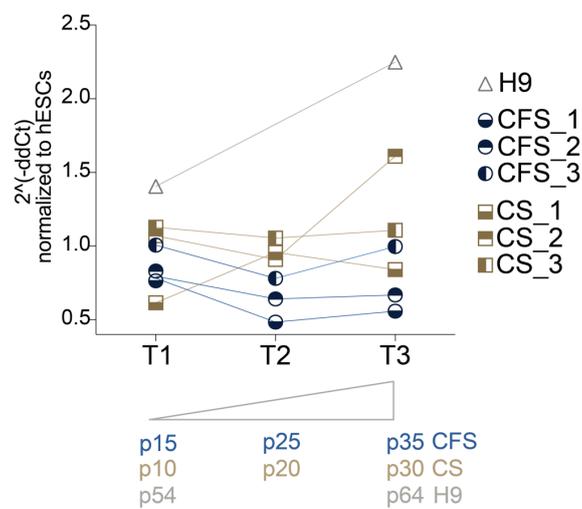


Figure 12: Relative quantification of telomere length in H9 and human AxSC lines

Telomere length was assessed in human AxSC lines as well as H9 by RT-qPCR. AxSCs (CFS_1-3 and CS_1-3 lines, Fig. 10) were analyzed at three different time points (CFS lines at p15, p25 and p35; CS lines at p10, p20 and 30). H9 cells were analyzed at p54 and p64. The values were normalized to H9 cells at p45 as an internal control (T: Time point, p: passage number).

3.2. Transcriptomic characterization of axial stem cells

To better understand molecular signatures in both AxSC states, single-cell RNA (scRNA) sequencing was performed by using CFS_2 and CS_2 (both at p14) as well as H9 cells (p60). The preprocessing of the dataset was performed by Ksenia Arkhipova (Leiden University, the Netherlands). Expression of selected pluripotency-related and AxSC marker genes from the dataset are shown in **Fig. 13**. Pluripotency markers *POU5F1* and *NANOG* were expressed only in H9 cells. *EOMES*, which plays a role in pluripotency exit and mesodermal commitment³⁶, was abundantly expressed in CFS cells but not neither in CS nor H9 cells. Expression of *SOX2* was detected in both AxSCs and H9, while *TBXT* and *CDX2* were specific to CFS cells (**Fig. 13**). *PAX6* was found as highly expressed only in CS cells.

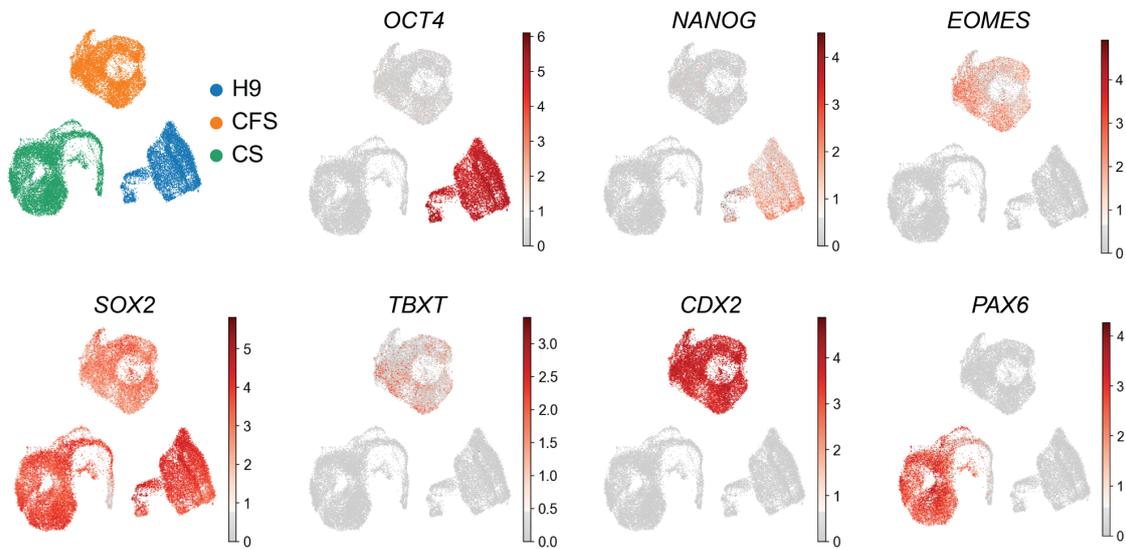


Figure 13: Expression of pluripotency and AxSC marker genes in single-cell resolution

Representative UMAP plots indicating the expression of pluripotency and AxSC markers in scRNA sequencing dataset comprising of CFS_2 (p14), CS_2 (p14) and H9 (p60) cells (10504, 12603 and 8671 single-cells respectively).

Next, I sought to identify the transcriptomic differences between AxSC states by performing differentially expressed gene (DEG) analysis in the presented dataset by comparing CFS and CS cells. The top 50 genes for CFS and CS were plotted in **Fig. 14A** and **14B** respectively. Multiple *HOX* genes were found in DEG list from CFS cells. *URAD*, *WNT5A*, *DUSP6* and *FGF17* which are expressed in NMPs/tailbud region^{37,59,66,245} were detected in CFS cells. *DUSP6* was also expressed in a relatively low number of CS cells. *GBX2* was one of the DEGs in CFS cells however CS cells also showed *GBX2* partial expression. *MLLT3*, which acts as a hematopoietic regulator and downstream effector of *TBXT*²⁸⁴, exhibited high and homogeneous expression in CFS cells (**Fig. 14A**). Remarkably, *SPRY4* and *SPRY2* were found in the DEG list. In the literature, *SPRY4* has been suggested as supportive of ESCs stemness and its levels decline during spontaneous differentiation^{285,286}. *SPRY4* was abundantly expressed in CFS cells, but only 20% of H9 cells expressed it. *SPRY2* has been associated with both mesodermal and neural development²⁸⁷. CFS cells exhibited high *SPRY2* expression while it was also found in low percentage in CS cells. Neural tube-associated genes *IRX3*, *IRX5*, *HES5*, *SOX1* and *SOX3* were detected in the list of DEGs for CS cells (**Fig. 14B**). I could determine that CS cells express *HES4*, *FABP7* and *PRTG*, which are known for playing a role in neural development^{66,288-290}. The expression of *ZFH3* and *ZFH4*. *ZFH3* regulates muscle and neural development²⁹¹⁻²⁹³ and it was expressed in both CFS and CS. Expression of *ZFH4*, which has shown to be involved in neural differentiation and to be suppressed during muscle differentiation²⁹⁴, was only detected in CS cells.

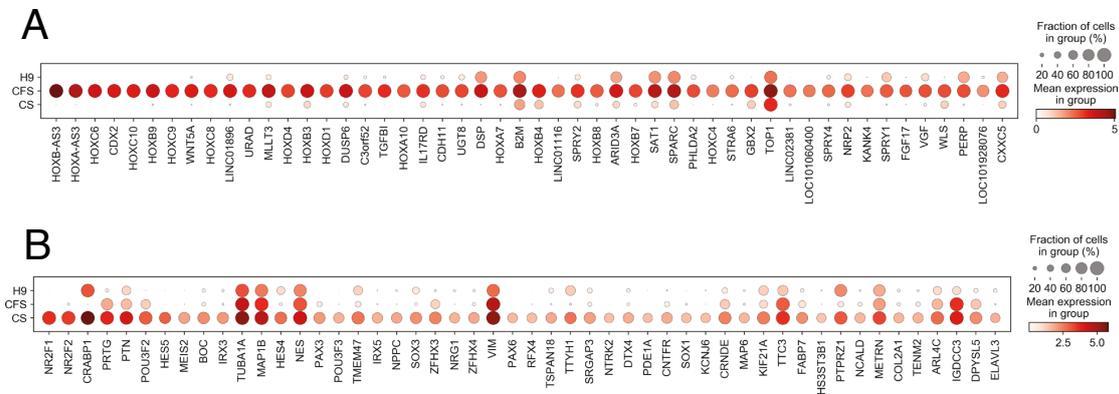


Figure 14: Differentially expressed genes in human CFS and CS cells

Dot plots showing the top 50 DEGs in CFS_2 (A) and CS_2 (B) lines in the scRNA sequencing dataset shown in Fig. 13. The threshold was determined as $>\log_2$ fold change. Dot size represents percentage of the cellular population expressing the gene of interest and color scale displays the mean expression.

Through DEG analysis, I sought to find novel AxSC marker genes. Several transcription factors were detected in either state, but they also exhibited partial expression in the other axial state, except for *HES5*. It was found highly and specifically expressed in CS cells (**Fig. 14B**). Its expression was also confirmed by UMAP plots (**Fig. 15A**). I, also, sought to identify genes commonly expressed in both states. Similarly to *SOX2*, *ZIC2* exhibited quite homogeneous expression in both AxSCs as well as H9. To validate *HES5* and *ZIC2* expression on the protein level, I performed immunostaining experiments. In line with the transcriptome results, both states were marked as *ZIC2* positive while only CS cells were *HES5* positive (**Fig. 15B**). IgG controls are shown in Supp. Fig. 2B and *ZIC2* stain detected in H9 cells is shown in Supp. Fig. 2A.

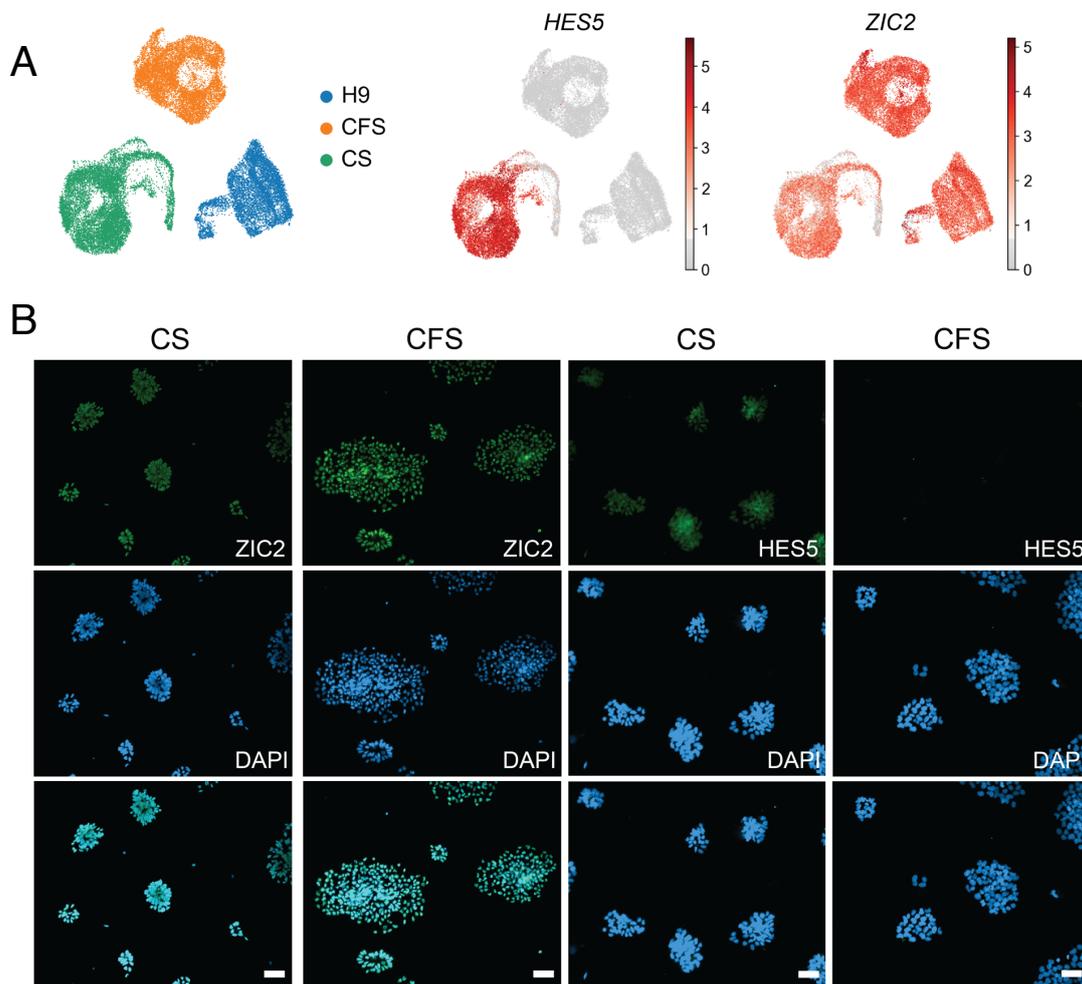


Figure 15: Expression of HES5 and ZIC2 in human AxSCs

(A) UMAP plots showing *HES5* and *ZIC2* expression in CFS_2, CS_2 and H9 cells determined by scRNA sequencing (same as in Fig. 13).

(B) Representative images from immunostaining of AxSC lines (CFS_2 p18 and CS_2 p18) for HES5 and ZIC2 (green) and DAPI (blue), (scale bar: 50 μ M).

The expression pattern characterizing CS cells implied a neural bias while CFS exhibited both neural and mesodermal potential based on the transcriptome analysis. To investigate the developmental bias of AxSCs in further detail and to detect further indications for potential lineage commitment, I summarized the marker genes associated with NMPs, neural and mesodermal development at different stages (**Fig. 16**) by reviewing published scRNA sequencing papers containing in vivo datasets^{59,66,243,277}. The summary of literature review is presented in **Fig 16**.

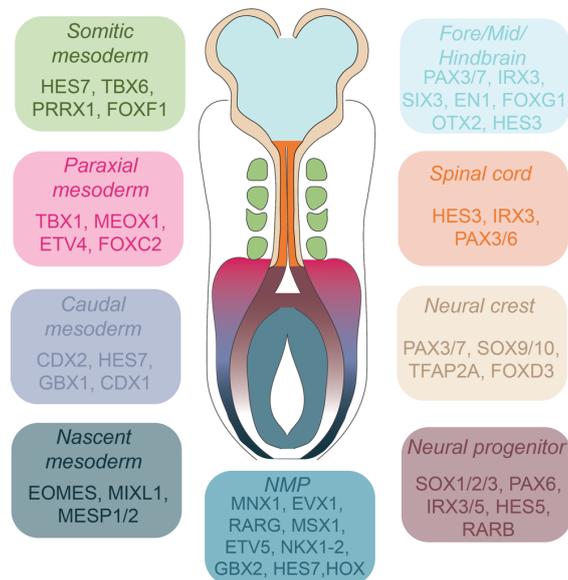


Figure 16: Embryo illustration showing the key developmental markers for different stages of axial development

The genes depicted in illustration were summarized based on scRNA sequencing results from embryonic tissues^{59,66,243,277}.

SOX2, *TBXT* and *CDX2* are commonly defined as the signature genes for NMPs, various genes have been additionally suggested as candidate NMP markers (**Fig. 16**). The analysis of candidate NMP genes in my scRNA dataset comprised of H9 cells and AxSCs showed that *MNX1*, *EVX1*, *NKX1-2* and *HES7* were found as highly and specifically expressed in CFS cells (**Fig. 17A**). Among these genes, *MNX1* and *NKX1-2* showed a relatively homogeneous pattern. *RARG* expression was high in CFS cells, but it was also partially detected in H9 cells. CFS cells were marked by abundant and homogeneous *MSX1* and *GBX2* expression, however these markers were detected also in CS cells, particularly *GBX2*. Finally, high *ETV5* expression was identified in CFS cells, but low levels were detected in CS and H9 cells. A hallmark of both in vivo and in vitro NMPs is the expression of multiple *HOX* genes, which are not detected in pluripotent cells⁶⁸. I analyzed all *HOX* genes present in the dataset (**Fig. 17B**) and found that the *HOX* gene expression spans paralogous group (PG) 1 to 13 in CFS cells. CS cells displayed a *HOX* profile containing PG 1 to 9. This profile can be considered relatively more anterior in comparison to CFS cells. *HOXC* and *HOXD* expression was not detected in CS cells. *HOX* genes were not identified in H9 cells, similarly to what has been reported in the literature (Steens and Klein, 2022).

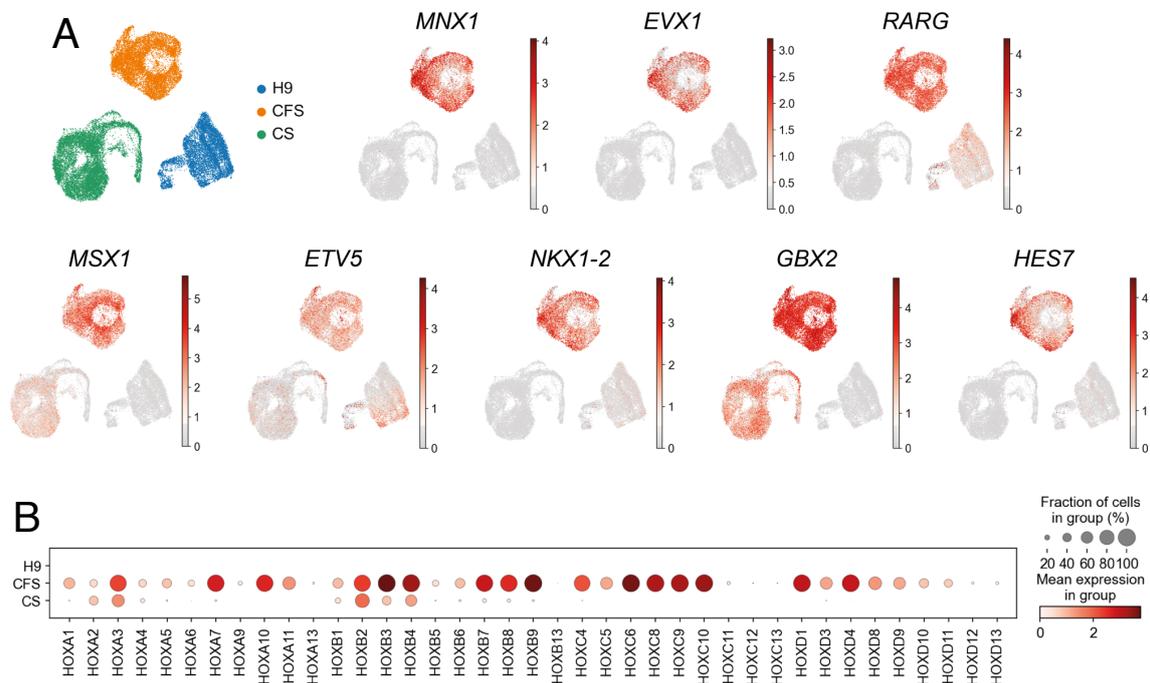


Figure 17: Expression of NMP-associated genes and HOX profile in human AxSC

(A) Representative UMAP plots showing the expression of candidate NMP markers (summarized in Fig. 16) in scRNA sequencing dataset including human AxSCs and H9 cells.

(B) Dotplot displaying HOX gene expression in scRNA sequencing dataset (A). Dot size represents percentage of the cellular population expressing the gene of interest and color scale displays the mean expression.

Mesoderm-associated genes were classified into four groups as nascent mesoderm, caudal mesoderm, paraxial mesoderm and somitic mesoderm (Fig. 16). *EOMES*, one of the nascent mesoderm markers, was previously shown as expressed only in CFS cells (Fig. 13). Other markers *MIXL1*, *MESP1* and *MESP2* displayed differential pattern among the three cell lines. *MIXL1* expression was found higher in CFS cells albeit its heterogeneity (Fig. 18). Most of the CS cells were marked with *MESP1* expression. *MESP2* showed higher levels in H9 cells compared to both AxSCs, where it was barely expressed. *CDX2* and *HES7*, which have been shown as NMP markers, are also expressed in caudal mesoderm cells (Fig. 16). Their cumulative expression alongside markers such as *GBX1* and *CDX1* can be used as a caudal mesoderm signature. *CDX2* and *HES7* were previously identified in CFS cells (Fig. 13 and Fig. 18 respectively), *CDX1* expression was slightly detected in CFS cells and it was not present in neither CS nor H9 (Fig. 18). CS cells were marked by *GBX1* expression, but they were negative for *CDX2* and *HES7*. (Fig. 13 and Fig. 18 respectively). Two of the paraxial mesoderm markers *TBX1* and *FOXC2* were expressed at very low levels in the three assessed cell lines, while all of them showed abundant *ETV4* expression (Fig. 18). CS cells were partly marked by elevated *MEOX1* expression. The cells expressing *MEOX1* were previously detected as low *PAX6* and

HES5 expressing cells in comparison to the CS transcriptome overall (**Fig. 13** and **Fig. 15** respectively). *TBX6*, *PRRX1* and *FOXF1* have been defined as somitic mesoderm markers based on the literature (**Fig. 16**). *TBX6* has also been shown as highly upregulated in NMP derivation cultures in vitro and heterogeneously expressed in NMPs in vivo^{44,59,237,238}. Its expression was barely detected in AxSCs, although a portion of H9 cells showed *TBX6* expression. (**Fig. 18**). A group of CFS cells were marked by *PRRX1* expression and none of the assessed cell lines was identified to be *FOXF1* positive.

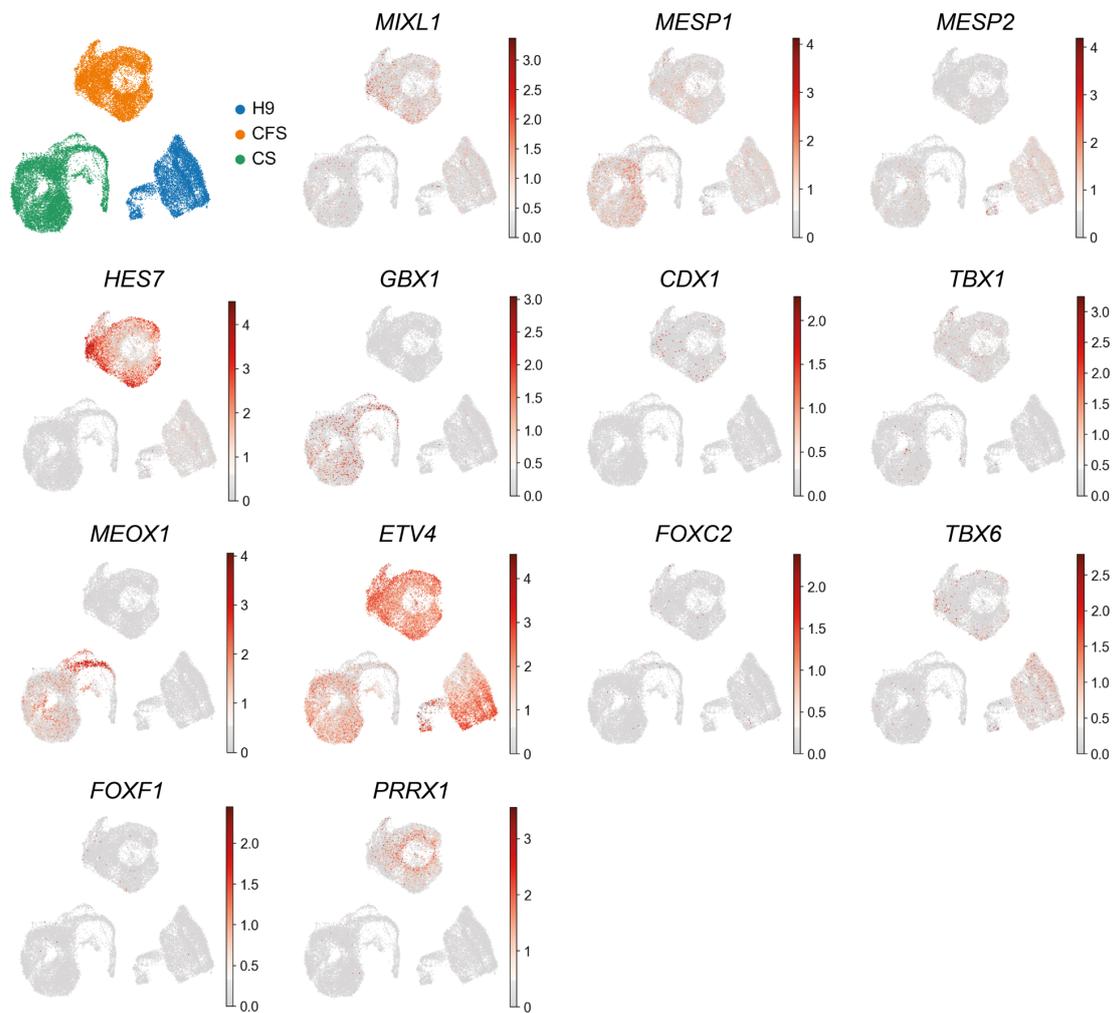


Figure 18: Expression of mesoderm development-associated genes in human AxSCs and H9 cells

Representative UMAP plots from the scRNA sequencing dataset indicating the expression of nascent, caudal, paraxial and somatic mesoderm-related genes as summarized in Fig. 16.

Neural development-associated genes were categorized into four groups, neural progenitors, neural crest, spinal cord and forebrain/midbrain/hindbrain (anterior neural tube) (**Fig. 16**). Neural progenitor markers *SOX1*, *SOX3*, *IRX3* and *IRX5* showed a clear expression pattern as they were homogeneously expressed in CS cells (**Fig. 19**). Small portion of CFS cells exhibited *SOX1* expression. *IRX3* and *IRX5* were expressed in CFS and H9 at similar levels. H9 cells were marked by more abundantly by *SOX3* expression compared to CFS. In addition to these four transcription factors, *RARB* which is one of the retinoic acid receptors, was detected only in CS cells. Neural crest-related markers pointed out a more variable pattern among the analysed cell lines. Both AxSCs were marked by *PAX3* expression albeit its lower levels in CFS cells. *PAX7* expression was detected only in CS cells similarly to *TFAP2A* expression. The CS subpopulation marked by *TFAP2A* was previously found as low *PAX6/HES5* expressing cells (**Fig. 13** and **Fig. 15**). *MEOX1* was partially expressed in *TFAP2A*⁺ population. *SOX10* and *FOXD3* expression was determined to be H9 specific, while *SOX9* expression was identified in all three cell lines (**Fig. 19**). In contrast to *IRX3*, *PAX3* and *PAX6*, the spinal cord marker *HES3* was barely detected. The expression of forebrain, midbrain and hindbrain markers *SIX3*, *FOXP1* and *OTX2* was not found in either AxSCs, but *EN1* expression was detected in CFS cells. *OTX2* expression was exclusively detected in H9 cells.

CS cells indicated to have a neural lineage identity based on transcriptome analysis, I questioned whether they harbor dorsoventral neural tube/spinal cord bias. I summarized the genes widely expressed in either dorsal or ventral progenitors²⁷⁶ (**Fig. 20A**). The expression of *MSX1* and *PAX3/PAX7* was previously shown in **Fig. 17** and **Fig. 19**, respectively. *MSX1* was detected at higher levels in CFS cells, while all other dorsal progenitor markers exhibited elevated levels in CS cells (**Fig. 20B**). The dorsal markers *PAX7* and *GSX1* were identified as specifically expressed in CS cells (**Fig. 19** and **Fig. 20B**). The ventral progenitor markers *NKX6-1* and *SP8* were mainly found in CFS cells (**Fig. 20B**).

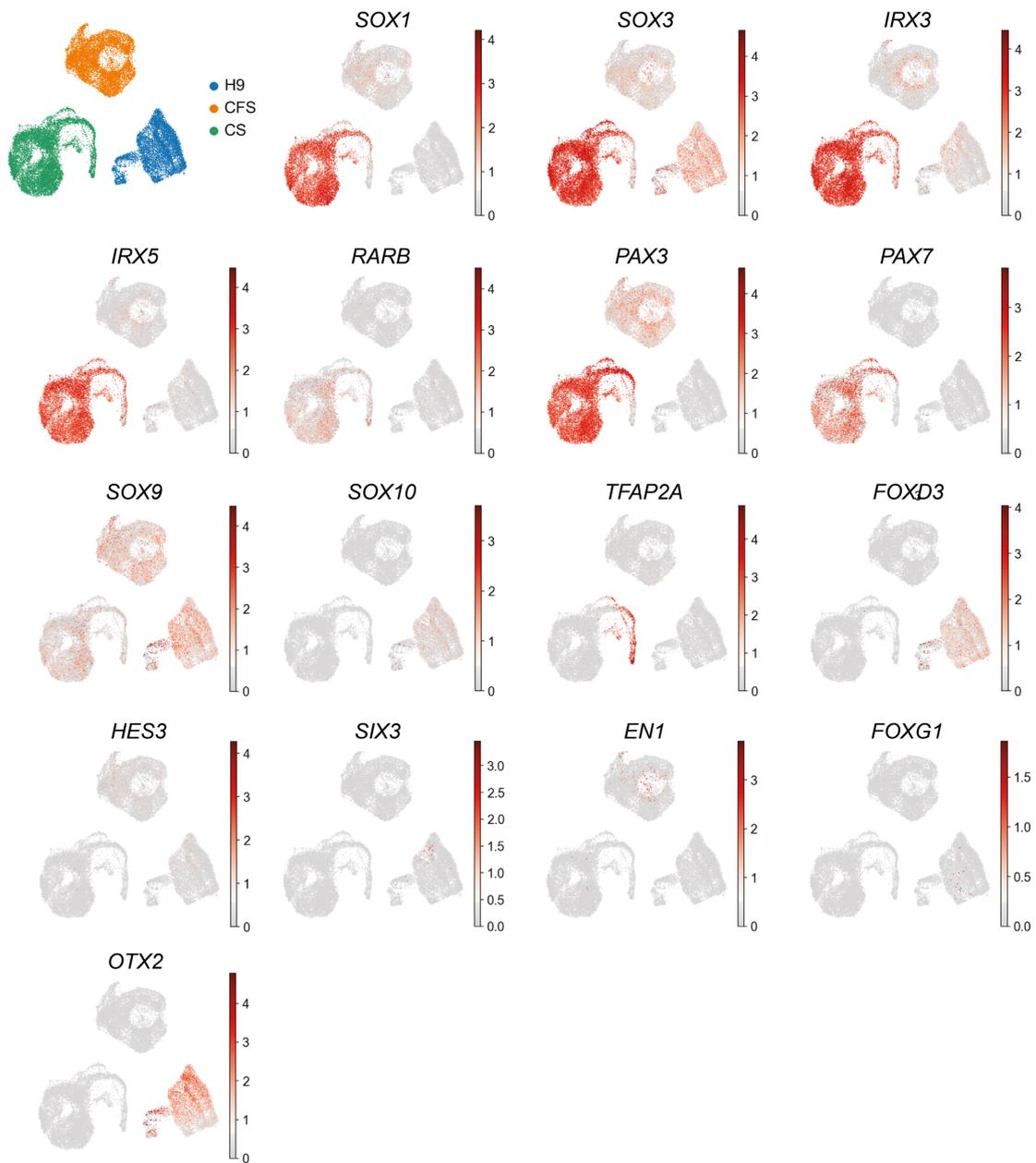


Figure 19: Expression of neural development-associated genes in human AxSCs and H9 cells
 Representative UMAP plots from the scRNA sequencing dataset indicating the expression of neural progenitor, neural crest, spinal cord, fore/mid/hindbrain marker genes as summarized in Fig. 16.

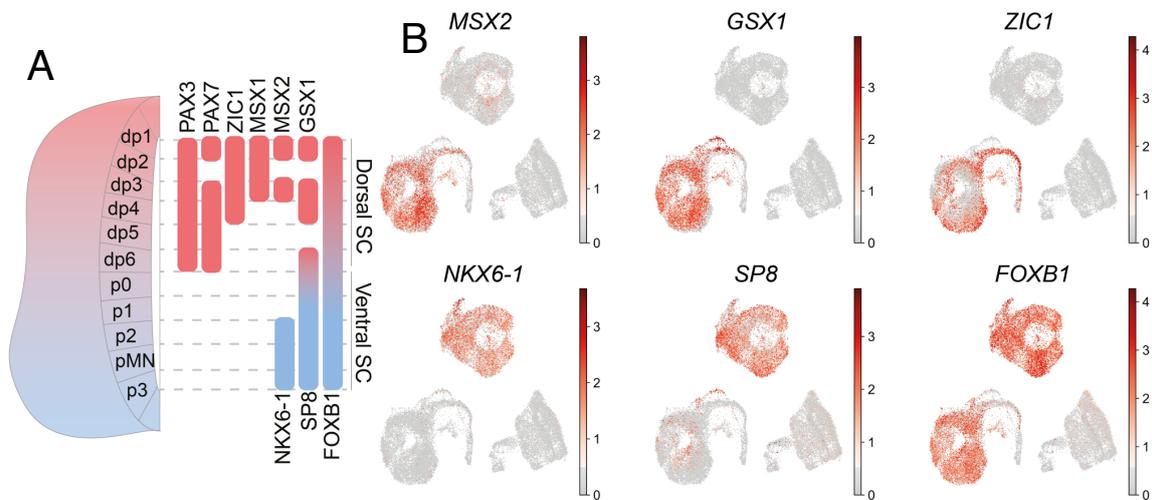


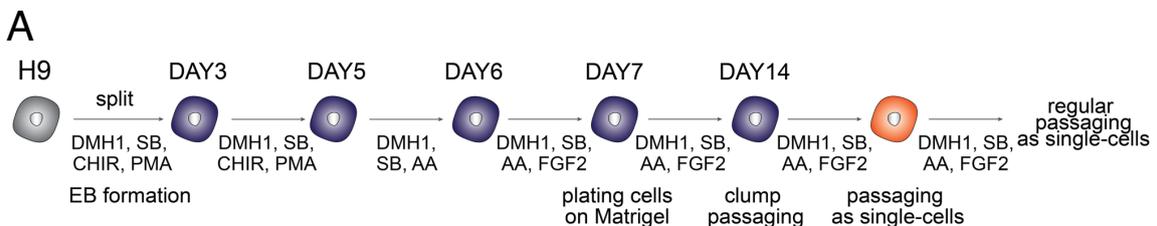
Figure 20: Expression of spinal cord markers in human AxSCs and H9 cells

(A) Spinal cord illustration depicting expression of the selected genes that mark dorsoventral domains of spinal cord/neural tube.

(B) Representative UMAP plots from the scRNA sequencing dataset indicating the expression of dorsoventral markers (A) in human AxSCs and H9 cells. Expression of *MSX1* and *PAX3/7* was previously depicted in Fig. 17 and Fig. 19 respectively.

3.2. Transcriptional differences between neural lineage committed state of axial stem cells and conventional neural progenitor cells

The transcriptome analysis of CS cell showcased a marker signature comparable to that of the neural progenitors, which lead me to question whether they share features with the in vitro derived neural progenitor cells (NPCs) generated by dual SMAD inhibition (BMP and TGF β pathways). To investigate this, I first produced NPCs from H9 by using the protocol in **Fig. 21A**. The NPC identity was validated by RT-qPCR and immunostaining experiments. I detected downregulation of pluripotency factors *POU5F1* and *NANOG* whereby upregulation of NPC markers *PAX6*, *SOX1*, *ASCL1* and *DCX* (**Fig. 21B**). *SOX2* was similarly expressed as in H9 cells. The protein expression of *PAX6*, *ASCL1*, *ZO1* and *CDH2* (N-cadherin) was confirmed by immunostaining experiments (**Fig. 21C**).



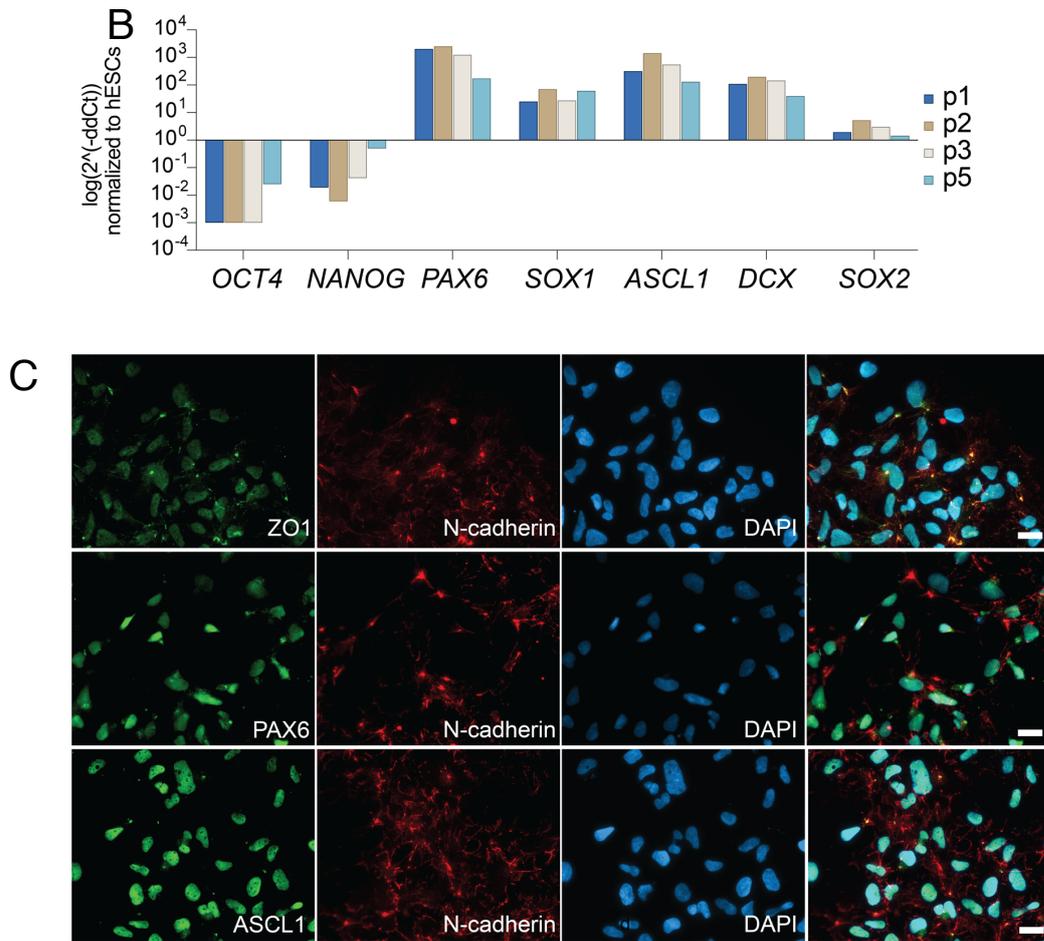


Figure 21: Derivation and characterization of human NPCs from H9 cells

(A) Schematic illustration of the protocol used for human NPC derivation. H9 cells were transferred into ultra-low attachment plates for embryoid body (EB) formation. Dual SMAD inhibition by using dorsomorphin homolog 1 (DMH1) and SB431542 (SB) was applied to the cells along with CHIR99021 (CHIR), purmorphamine (PMA), FGF2, ascorbic acid (AA) at different stages. EBs were transferred to Matrigel-coated plates and split as clumps for the first passage. The cells were routinely split as single-cells upon confluency.

(B) Characterization of human NPCs time course by RT-qPCR. The values were normalized to undifferentiated H9 cells (p: passage number).

(C) Immunostaining of human NPC cells at p5 for ZO1, PAX6, ASCL1, N-cadherin and DAPI (scale bar: 20 μ M).

Next, scRNA sequencing was performed by using NPCs (p3) to assess the differences between the two cell populations. NPC dataset was preprocessed by the same algorithms as in AxSC and H9 dataset (**Fig. 13**) and it was merged with CS population. The transcriptional differences in terms of the expression of neural development-related genes are shown in **Fig. 22A**. NPCs partly displayed high *HES5* expression, but *ZIC2*, *PAX3*, *PAX7*, *MSX1*, *MSX2* and *GBX2* were barely expressed in NPCs compared to CS cells. *NKX6-1* and *SOX9* expression was detected homogeneously in the entire NPC population. A small portion of NPCs were marked by *OTX2* expression. I analyzed the HOX profile in both cell types and found that NPCs do not

exhibit the expression of *HOX* genes. (**Fig. 22B**). These results indicate that CS cells and NPCs have unique features.

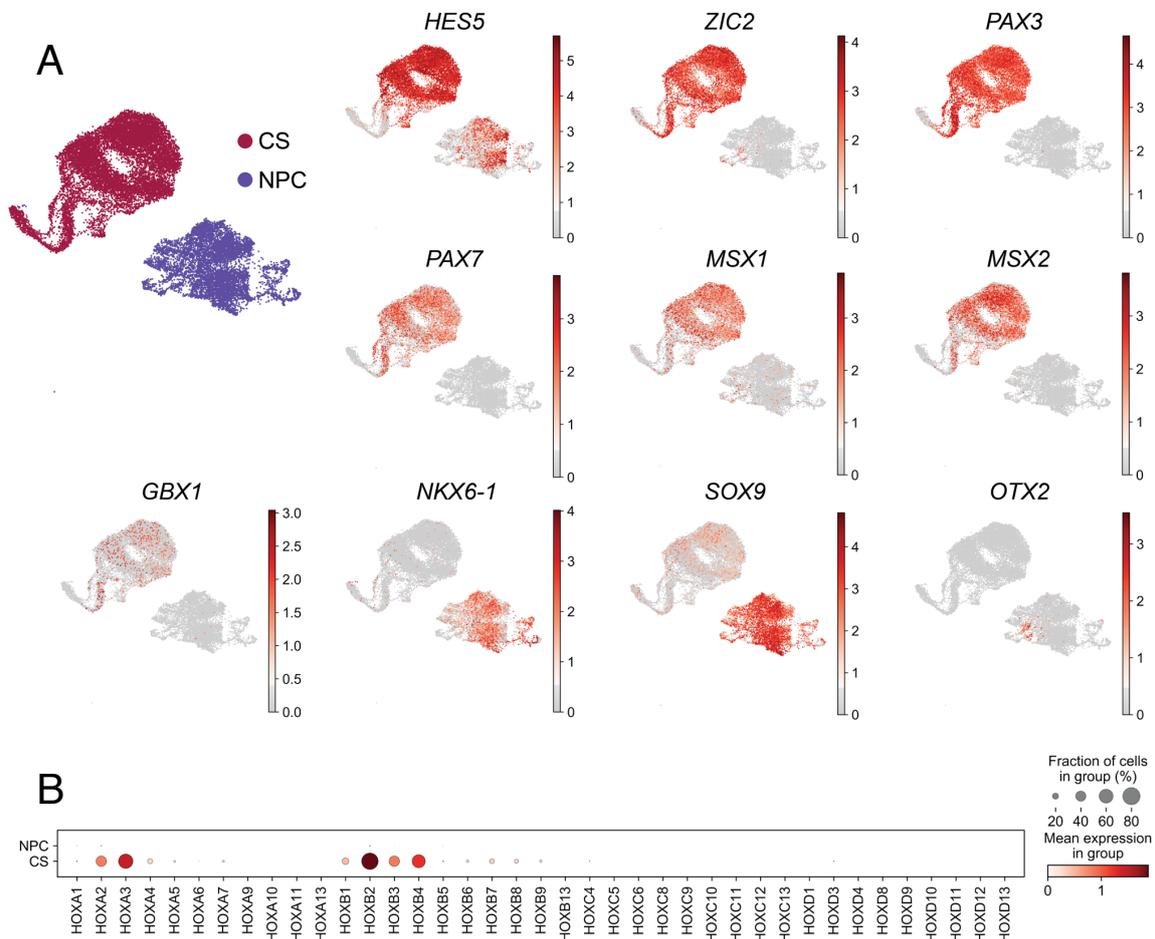


Figure 22: Transcriptional differences between human NPCs and human CS cells

(A) Representative UMAP plots indicating expression of the neural development-associated genes in NPCs (p3) and CS_2 (p14, same as Fig. 13).

(B) Dotplot displaying the expression of *HOX* genes in NPCs and CS_2.

3.3. Proteomic characterization of axial stem cells

To validate scRNA sequencing results in three independent cell lines per AxSC state (CFS_1-3, CS_1-3 in **Fig. 10**) and two technical replicates of H9 cells, global proteome analysis by liquid chromatograph mass spectrometry was performed in collaboration with Enes Ugur and Prof. Dr. Heinrich Leonhardt (Ludwig Maximilian University of Munich, Germany). First, expression of core AxSC markers *SOX2*, *TBXT*, *CDX2*, and *PAX6* was evaluated in the dataset. *TBXT/CDX2* and *PAX6* were elevated in all CFS and CS lines respectively (**Fig. 23A**). *SOX2*

was detected higher in AxSCs compared to H9. Next, expression of the NMP markers (**Fig. 16**) was explored. MNX1, EVX, MSX1 and GBX2 showed a similar expression pattern as in the transcriptome analysis due to their higher expression in CFS cells (**Fig. 23B**). NKX1-2, ETV5 and HES7 were not detected in proteome dataset. Confirming the previous results from AxSCs transcriptome, CFS cells displayed an anterior (PG 1) to posterior HOX profile (PG 11) while CS cells were determined as more anterior due to high expression of HOX PG 1-5 (**Fig. 23C**). CS_2 cells showed HOXC10 expression comparable to all CFS lines which was in contrast to the transcriptomic analysis (**Fig. 17**). I assessed expression of a set of the genes showing cell-specific expression identified by DEG analysis in the scRNA sequencing dataset (**Fig. 14**). CFS_1-3 lines exhibited high expression of WNT5A, FGF17, MLLT3 proteins (**Fig. 23D**). SPRY2 and SPRY4 showed similar expression in CFS and H9 cells unlike their transcript levels (**Fig. 14A**). IRX3, IRX5, SOX1, SOX3, FABP7 and PRTG were elevated in CS_1-3 lines (**Fig. 23D**). CFS_2 line showed slightly increased level of PRTG (**Fig. 23D**) which is in accordance with its transcriptome (**Fig. 14A**). To investigate consistency in each cell line, I conducted correlation analysis (**Fig. 23E**). The results pointed out that the global proteome of all three CFS lines (1-3) and CS lines (1-3) are overall very similar to each other in themselves.

Finally, differentially expressed proteins (DEP) were analyzed and plotted as a heatmap in **Fig. 23F**. The DEP were grouped in seven clusters. Cluster 1 and 2 represent the highly expressed proteins in CFS and CS lines respectively. Cluster 3 contains proteins that were detected as enriched in both AxSC states when compared to H9. Cluster 5 contained low expressed proteins in CS cells but present in elevated levels in H9 and CFS lines. Last two clusters (6 and 7) point out the abundant proteins in H9 cells. The top 40 proteins per cluster are given in Table 7.

Next, GO biological process analysis was conducted for the determined clusters (**Fig. 23F**). GO terms indicated that CFS cells are related both neural and mesodermal development while CS cells are prone to neural commitment. H9-enriched proteins were substantially associated with metabolic processes. The proteins highly expressed in both AxSC lines implied an enhanced ability for oxidative phosphorylation events in AxSCs compared to H9 cells. To investigate this, I stained the mitochondria in the three cell lines. confirmed the CFS and CS cells displayed elevated signal for the mitochondrial staining compared to H9, both as live and fixed cells (**Fig. 24**).

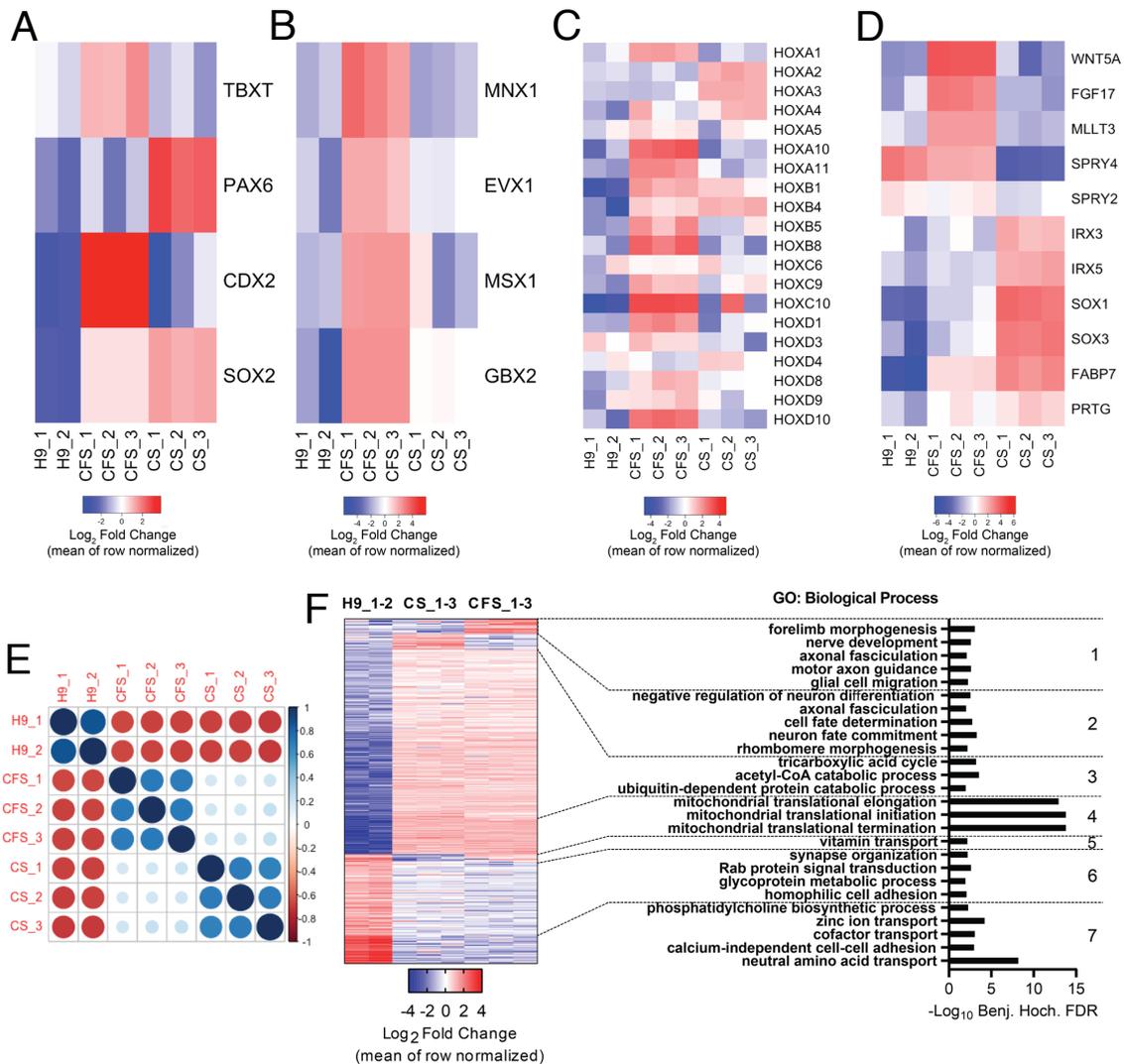


Figure 23: Characterization of AxSCs by global proteome analysis

(A-D) Heatmaps showing expression of the core AxSC markers (A), NMP markers (B), HOX genes (C), the differentially expressed genes (D) determined by scRNA sequencing (Fig. 14). Three CFS_1-3 and three CS_1-3 lines (Fig. 10) were used for the experiment as biological replicates. Two technical replicates of H9 cells were used and named H9_1-2.

(E) Correlation analysis between the samples of human AxSCs and H9 cells

(F) Differentially expressed proteins grouped in seven clusters and GO biological process analysis in the respective clusters

Table 7: Top 40 differentially expressed proteins

Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7
WNT5A	CRABP1	ZBTB39	ODF2L	KRT18	TJP1	CLDN6
KANK4	MEIS1	MRTFB	PHC2	L1TD1	SIDT2	BST2
CDX2	KHDRBS2	RPL26L1	STMN2	LSR	DHCR24	CYP2S1
DNAH1	PDE1A	SELENOH	FABP7	SLC2A3	UTP20	EPCAM
TCEA3	NR2F1	HDCC2	CRIP2	TPM2	EPB41L2	SLC16A3
SCGN	PAX6	MIS12	CKS1B	IFITM2	SPCS_2	TPM2
IL17RD	ELAVL3	MED7	SS18	SPRY4	LARGE1	RAB25
SP6	SOX21	TSC22D1	RBM3	CGNL1	TRAF7	SLC7A3
FGF17	ZDHHC23	EPC1	DACH1	VAMP8	SURF4	SERINC5
EPDR1	ZIM2	WDR76	TAGLN3	IFITM3	SLIT3	CD9
SPARC	GSX1	CPNE2	CSNK1D	SERPINB9	RAB3B	TNFRSF8
HOXA10	LMCD1	PPP1R7	MRPL54	GFPT2	ATG9A	MAL2
MAML3	PAK3	PPIE	CHMP6	SERPINB6	TMEM161B	PODXL
MNX1	TOX	POLA2	MRPL58	RAB15	SEC61A1	CAV1
LGALS8	DCX	PABPC1	COQ7	S100A11	ITGA2	CDH1
RGPD3	PLEKHA7	ACYP2	FUBP1	TAGLN	STX6	HMOX1
HOXB8	PAX6	TSC22D2	GSTA4	STRA6	SEMA4D	LAD1
ZIC5	TFAP2B	AK1	CFL2	ITM2C	HLA-B	CMTM4
EDARADD	PPP1R1A	NCOA1	ETFDH	UGT8	KIF21B	COL1A1
HOXD10	NOS2	PRPF4	PCLAF	HRC	SUN1	CLDN3
COL3A1	TUBB4A	PREX2	DPYSL4	COL6A2	AFAP1L1	STOM
ALX4	PDE1B	PAIP1	GON7	ST6GAL1	MFSD4B	ICAM3
GSN	TOX3	AKR7A2	CKM	PLPP2	SLCO4C1	TPM1
SLC39A8	CRYBA1	TCF25	CCDC134	CLMP	CNNM4	LPCAT3
TPD52L1	RFX4	WIZ	SETMAR	GJA1	ERBB2	IRS4
PLEKHA4	MXRA7	SH3BGRL2	CRB2	MYOF	RAB13	AP1M2
FBLIM1	GDAP1L1	HPF1	PHYKPL	CARMIL1	WIPI1	PFKP
B2M	DACH2	LIN9	POLR2K	GPR19	ORMDL1;ORMDL3	DPAGT1
MLLT3	PBX3	CHTOP	LRRC57	PLEKHG3	SCAP	FADS3
PHLPP1	PBX1	FAM204A	SPATA33	DPP6	ELOVL5	FAM83G
AGMAT	GNAO1	HDAC4	GLUL	TNIK	RAB39A	GALNT3
KLHL14	RFX2	CSTF2	TPPP3	AGTRAP	CLN6	YIPF6
DNM2	SRGAP3	CIAO3	INA	NEBL	RAB3D	KDF1
ARID3A	C4orf46	TRIM36	FKBP9	LGALS3BP	LRRC8B	SPINT2
SP7	PLXNA2	PPIL1	TCF3	ITGB5	TMEM87A	TMEM63A
MSX1	POU3F3	SNX1	SHQ1	DTNA	PARL	DOCK5
CSRP1	LMO3	NDUFV3	ABAT	VGFB	NUP188	METTL7A
EPOP	REM2	OGFOD1	NOVA2	DNMT3B	SEC63	RCE1
EOMES	IRX5	TUBB4B	ZCCHC4	STAT3	DOCK9	ACVR1B
FGF2	GIT2	AK3	GLI3	MBTPS2	KCT2	SLC2A13

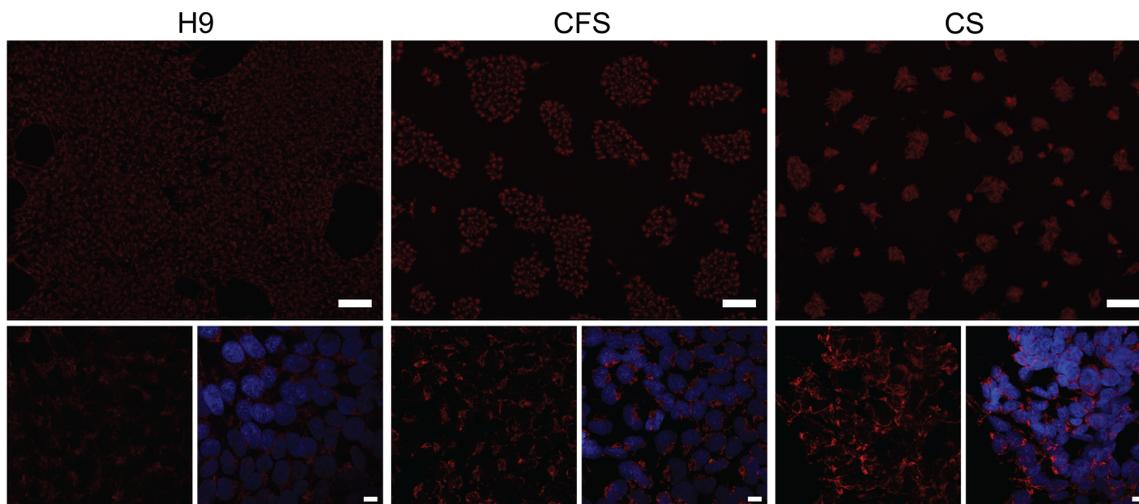


Figure 24: Mitochondrial staining in human AxSCs and H9 cells

Upper panel displays the representative images from the live cells stained for mitochondria (scale bar: 100 μ M). Lower panel shows the mitochondrial staining after fixation of the cells (red: Mitotracker, blue: DAPI, scale bar: 10 μ M). CFS_2 and CS_2 lines were used for the experiment.

Next, I sought to identify the CD markers characterizing the AxSCs. The use of only proteome data as a basis could be misleading due to comprising the normalized values to the mean of each row and then averaged within the replicates. Therefore I first analyzed the CD genes in scRNA sequencing dataset. The CD genes expressed above 40% for all cell population are shown in **Fig. 25A**. The results indicated that there is no CD marker specific to neither AxSC state, however, CD59, CD81, CD151, CD276 and CD320 were expressed slightly higher in CFS cells and CD47 was expressed higher in CS cells. CD46, CD74, CD164, and CD200 were detected with the highest expression in H9 and the lowest expression in CS. The analysis of the same CD marker panel in the proteomic dataset showed that all of the abovementioned CD proteins were downregulated in AxSCs when compared to H9 cells except for CD59 and CD320 (**Fig. 25B**). CD59 was more abundant in CFS lines and CD320 was found at similar levels in both CFS and CS lines. CD24 was not detected in proteome dataset.

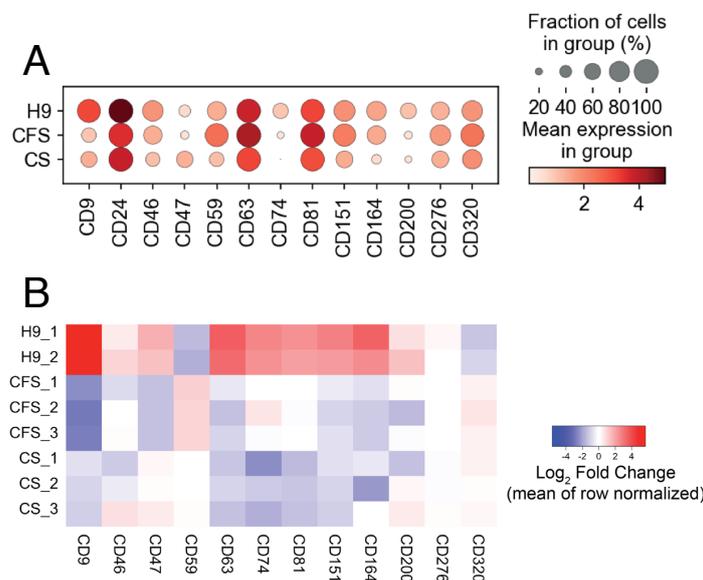


Figure 25: Expression of CD markers at transcript and protein level in human AxSCs and H9 cells included datasets

(A) Dotplot representing the identified CD markers expressed by the analyzed cell populations in the scRNA sequencing dataset. Dot size represents percentage of the cellular population expressing the gene of interest and color scale displays the mean expression.

(B) Heatmap displaying the CD markers in global proteome dataset including 3 cell lines per AxSC states and two technical replicates of H9 cells.

3.4. Chromatome profiling in axial stem cells

For identification of chromatin-bound proteins, chromatome analysis was performed in both AxSC states and H9 in collaboration with Enes Ugur and Prof. Dr. Heinrich Leonhardt (Ludwig Maximilian University of Munich, Germany). To this end, the same samples as used for proteome profiling were used together alongside a third H9 replicate. First, nuclei were isolated from individual cell lines and proteins were fixed by formaldehyde. Next, chromatin was extracted, and proteins were captured after shearing by sonication. Finally, liquid chromatography mass spectrometry was applied. The chromatome analysis showed that TBXT and CDX2 are bound to the CFS chromatin, while PAX6 is bound in CS chromatin (**Fig. 26A**). SOX2 protein was found as more abundant in CFS proteome compared to H9 proteome (**Fig. 23A**), but higher chromatin-bound SOX2 was detected in H9 cells (**Fig. 26A**).

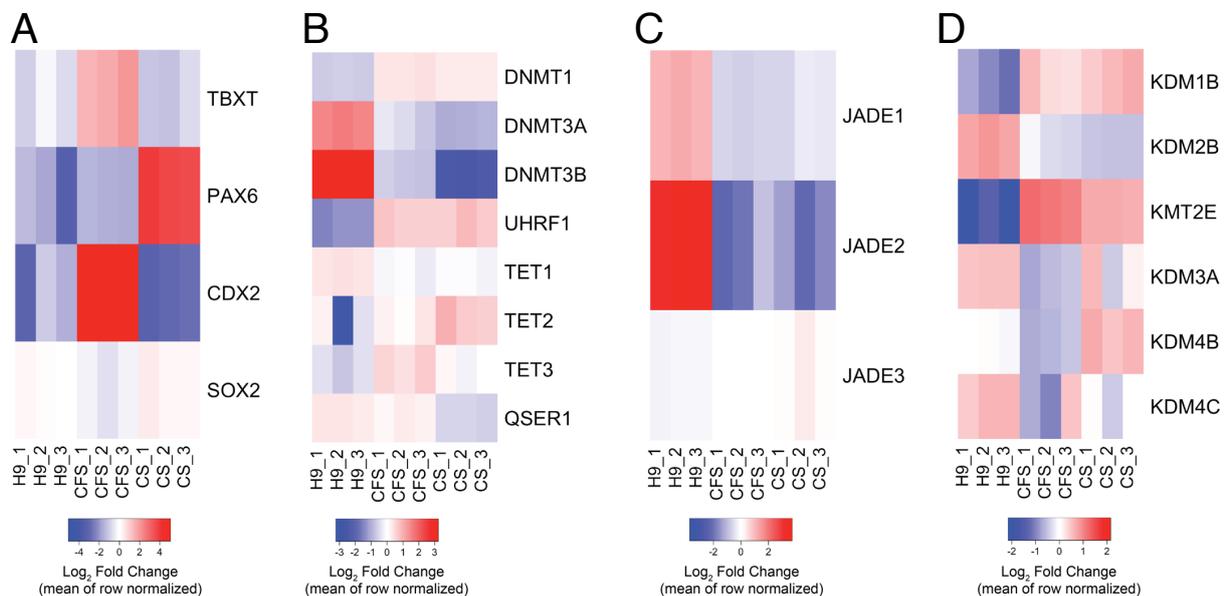


Figure 26: Analysis of chromatin-bound proteins in human AxSCs and H9 cells

Heatmaps displaying the level of AxSC markers (A) and epigenetic regulators for DNA methylation (B), histone acetylation (C) and histone methylation (D) binding to the chromatin (called chromatome) of AxSCs (3 independent cell lines) and H9 cells (2 technical replicates). Same samples in proteome analysis (Fig. 23) were used for chromatome analysis.

To reveal whether there is a difference in epigenetic modifiers between AxSCs and H9 cells, key proteins for DNA methylation including their cofactors/interaction partners (**Fig. 26B**), histone acetylation (**Fig. 26C**) and methylation (**Fig. 26D**) markers were analyzed in the chromatome dataset. The level of de novo methyltransferases DNMT3A and DNMT3B²⁹⁵ was higher in H9 cells while DNMT1 which plays a role in methylation maintenance²⁹⁶ was upregulated in AxSCs, particularly CFS lines (**Fig. 26B**). UHRF1, a DNMT1 cofactor²⁹⁷, was more abundant in AxSCs. DNA demethylases TET1-3 proteins²⁹⁸ showed a distinct pattern between the cell lines. TET1 was detected higher in H9 cells, while TET2 and TET3 were upregulated in CS and CFS lines respectively. In conjunction with the expression of TET proteins, QSER1, which is an interaction partner of TET1²⁹⁹, was downregulated in both AxSC states. Histone acetylases JADE1 and JADE2³⁰⁰ were abundantly found in H9 however JADE3 did not exhibit a notable difference between the populations except for CS_2 line (**Fig. 26C**). Finally, histone methylation regulators KDM1B and KMT2E were upregulated in all AxSC lines, while KDM2B, KDM3A and KDM4C were higher in H9 except for one CS and CFS line for the last two proteins respectively (**Fig. 26D**).

In the literature, the core transcription factors that maintain pluripotency in human cells have been described as the essentialome^{185,301}. It is possible that the essentialome factors

could have a wider role in regulating self-renewal and that this role is not confined to the pluripotency state only. Based on this hypothesis, I analyzed the essentialome factor expression using the chromatome dataset to compare their chromatin-binding level in between H9 cells and AxSCs. The results showed that the core pluripotency factors *POU5F1/NANOG* as well as *PRDM14/TRIM71* were downregulated in AxSCs, *MYCN* was abundantly expressed in both AxSC states (**Fig. 27A**). *SALL4* and *LIN28B* exhibited same levels among the H9 and AxSC lines albeit a slight downregulation of *SALL4* is observed in the CS lines. Expression of these factors was analyzed further at the transcript level. The scRNA sequencing results (**Fig. 13**) confirmed that *POU5F1*, *NANOG* and *PRDM14* were not expressed and *MYCN* was enriched in both CFS and CS (**Fig. 27B**). *TRIM71* was slightly downregulated in AxSCs compared to H9. The level of *SALL4* was same in H9 and CFS, while CS cells showed lower *SALL4* expression. *LIN28B* was expressed at similar levels in all three cell lines.

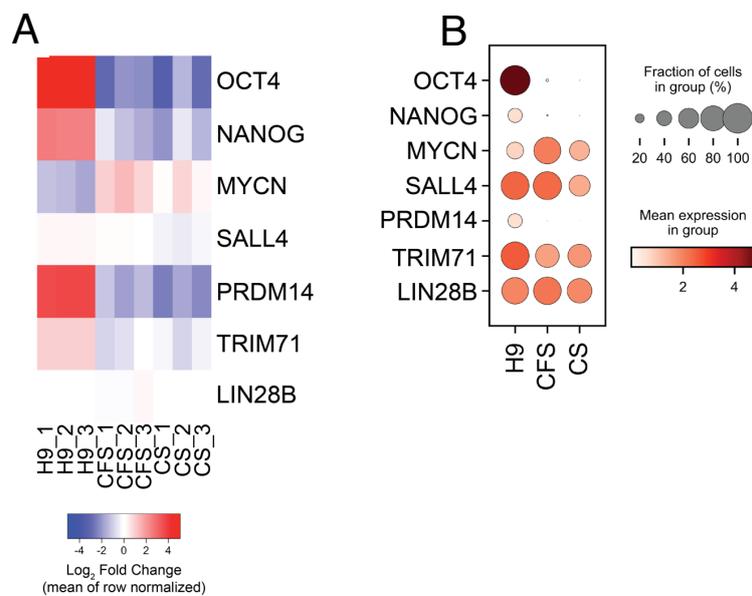


Figure 27: Expression of pluripotency-related genes in human AxSCs and H9 cells

(A) Heatmap displaying chromatin-binding level of the top pluripotency-related (essentialome) proteins in CFS_1-3, CS_1-3 and H9_1-3 (Fig. 26).

(B) Dotplot displaying expression of the panel A genes in CFS_2, CS_2, H9 cells (Fig. 13). Dot size represents percentage of the cellular population expressing the gene of interest and color scale displays the mean expression.

3.5. Hierarchical regulation between axial stem cell states

To investigate whether there is a hierarchical regulation among the derived axial states, we split both AxSCs states and kept them in the opposite maintenance medium to their cell type

(**Fig. 28A**). Three independently derived lines: CFS (1-3) and CS (1-3) (**Fig. 10**) were used for this experiment. At p6 I observed colonies with CS-like morphology in the CFS to CS conversion cultures (**Fig. 28B**). Considerable number of spontaneously differentiated cells as well as the CS-like colonies were present in the CS to CFS conversion cultures. Expression of the core AxSC markers was analyzed as a time course by RT-qPCR experiment throughout the conversion (**Fig. 28C**). The gradual downregulation in *TBXT* and *CDX2* expression but upregulation of *PAX6* was detected during CFS to CS conversion. Neither *TBXT* nor *CDX2* were upregulated in CS to CFS culture. The levels of *PAX6* fluctuated over the time as they were downregulated at p3 compared to the beginning of the conversion, but upregulated again when analyzed at p6. These results indicate that CFS cells are able to generate CS cells but not vice versa.

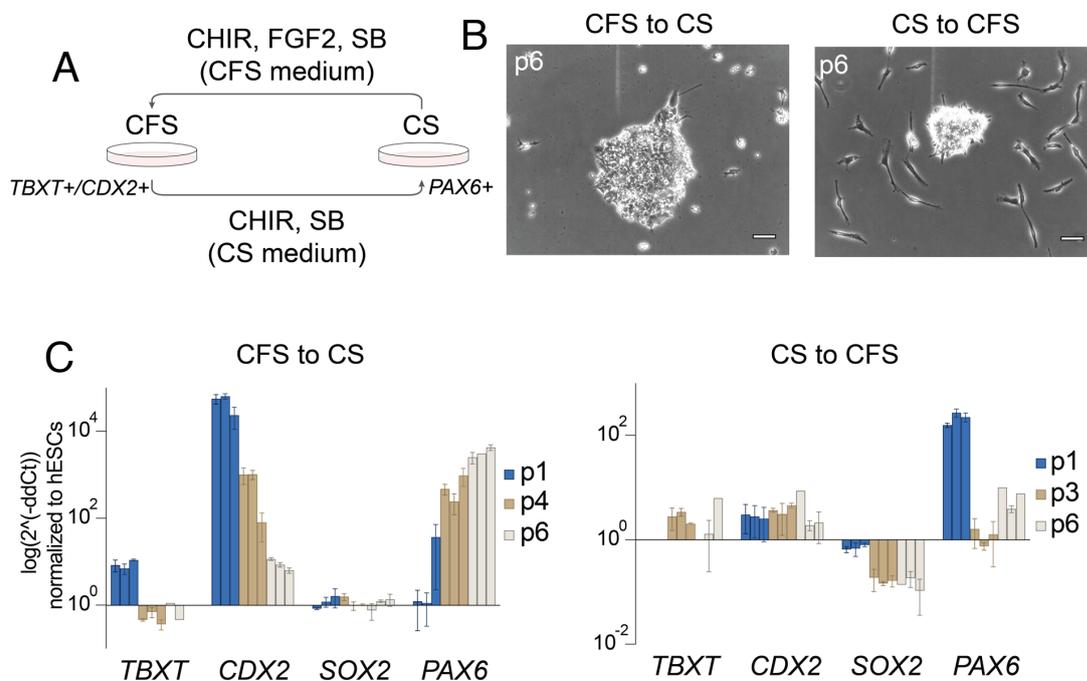


Figure 28: Hierarchical regulation between human AxSC states

(A) Schematic illustration of the medium switching experiment. Each state of AxSCs (CFS_1-3 and CS_1-3 as in Fig. 10) were treated with opponent AxSC state medium.

(B) Brightfield images from CFS to CS and CS to CFS conversion cultures (scale bars: 50 μM, p: passage number).

(C) Expression analysis of AxSC markers by RT-qPCR in the cultures from p1 to p6. The bars with same color display three independent derivation of CFS_1-3 and CS_1-3 lines. Two independent experiments were carried out per CFS to CS and CS to CFS conversion. Error bars represent SEM. The values were normalized to undifferentiated H9 cells (p: passage number).

3.6. Reproducible derivation of axial stem cells from different human pluripotent stem cell lines

The reproducibility of both AxSC states by using H9 hESCs as the starting population was validated by conducting three independent derivations shown in **Fig. 10**. Next, I investigated if the AxSCs can be reproducibly derived from different hPSCs. For this purpose, scRNA sequencing was repeated using CFS_2 and CS_2 lines (H9 background), and AxSC lines derived from HUES6 hESCs and HMGU1 hiPSCs cell lines which were generated by Ejona Rusha and Dr. Dmitry Shaposhnikov (Helmholtz Center Munich) respectively by using the same method as indicated in **Fig. 9A**. The only difference was the passaging modality as iPSC-based lines were split as clumps. The sequencing data was preprocessed by Ksenia Arkhipova (Leiden University, Netherlands). First, I assessed the expression of core AxSC factors in six cell lines with three different backgrounds. *SOX2* was highly expressed in all cell lines as shown in **Fig. 29A**. Potential CFS lines derived from HUES6 and HMGU1 cells were marked by *TBXT* and *CDX2* expression likewise CFS_2 which was characterized previously in this study. CS lines generated from HUES6 and HMGU1 parental cells were marked by *PAX6* expression in consistent manner similarly to the CS_2 line.

NMP and lineage markers (**Fig. 16**) were analyzed in the axial dataset. NMP markers except for *MSX1* and *GBX2* were expressed in all CFS lines albeit their relatively low levels in HUES6-CFS (**Fig. 29B**). Three CS lines were found as abundantly expressing all neural progenitor-related transcription factors, but the expression of retinoic acid receptor *RARB* was very lowly detected (**Fig. 29C**). Neural crest markers presented a very small upregulation in all lines with the exception of *PAX3*, showing high expression. Forebrain, midbrain and hindbrain markers were not expressed, moreover, the anterior development marker *OTX2* was not detected in the dataset. With the exception of *HES3*, the expression of the spinal cord-associated genes *IRX3/PAX3/PAX6* was found to be higher in all CS lines compared to the CFS lines. The majority of markers related to mesoderm development were not identified in AxSC lines except for low and heterogeneous expression of *EOMES*, *MIXL1* and *ETV4* in all three CFS lines, and *PRRX1* in HMGU1-derived CFS cells. Expression of caudal mesoderm markers was not determined in neither state of hESC nor hiPSC-based AxSCs. Due to the previously described dorsal spinal cord/neural tube bias presented by the CS_2 cells in the previous scRNA sequencing experiment (**Fig. 20**), I explored the expression of dorsoventral genes in newly generated AxSC lines (**Fig. 29C**). The results indicated that the CFS lines express both dorsal and ventral markers, while all CS lines exhibit a dorsal bias due to the abundant expression of *PAX3*, *PAX7* and *GSX1*, while lacking the ventral marker *NKX6-1*. Finally, I analyzed the HOX profile. Slight changes were observed in the expression of individual HOX genes such as *HOXC*, *HOXD* among the different hPSC derived AxSCs. The majority of the *HOXA*

genes were attenuated in the HUES6-based CFS cells. The overall pattern was similar between the respective AxSC states. Broad range of HOX expression was found in the three CFS lines, where CS lines were marked by HOX PG 1 to 4, with the exception of *HOXA7*, that was partially detected in CS HMGU1.

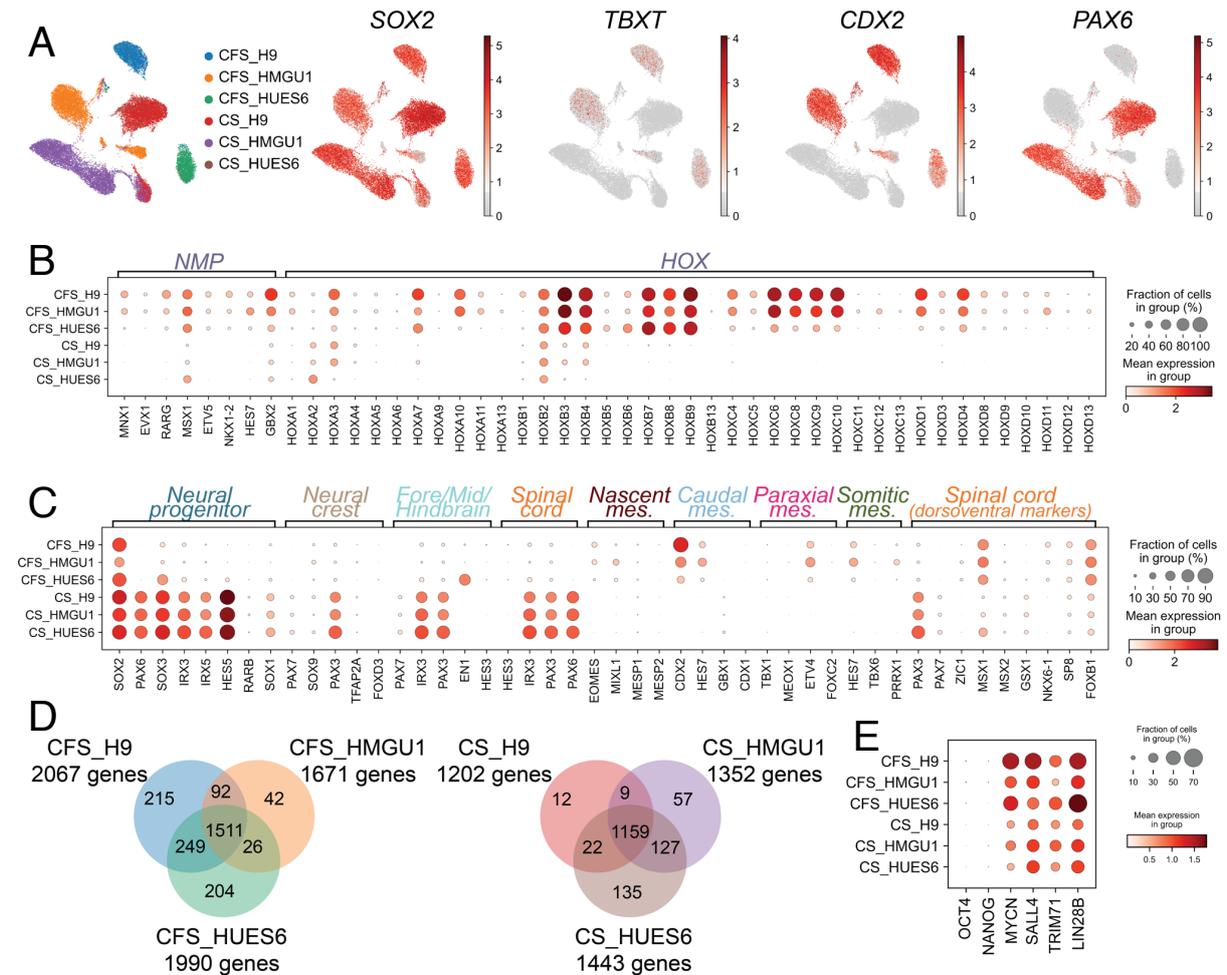


Figure 29: Reproducibility of AxSC states in hESC and hiPSC lines

(A) UMAP plots displaying AxSC markers in CFS and CS lines derived from H9, HUES6 and HMGU1 cells. For H9-derived AxSCs, CFS_2 and CS_2 lines were used at different passages than Fig. 13 (CFS H9: p15, CS H9: p12, CFS HMGU: p18, CS HMGU: p14, CFS HUES6: p14, CS HUES6: p14, p: passage number).

(B-C) Dotplots showing expression of the genes marking different developmental stages as summarized in Fig. 16 and dorsoventral neural tube/spinal cord progenitor markers as depicted in Fig. 20. Dot size represents percentage of the cellular population expressing the gene of interest and color scale displays the mean expression.

(D) Venn diagram indicating the number of the genes that have >1.5 mean expression in each cell line determined in scRNA sequencing dataset (A-B).

(E) Dotplot showing expression of the essentialome pluripotency-related genes in scRNA sequencing dataset (A-D). Dot size represents percentage of the cellular population expressing the gene of interest and color scale displays the mean expression.

Expression of NMP, lineage specific markers and *HOX* genes points out that respective AxSC states share common expression trends irrespective of the background of the starting population. For an unbiased analysis, the genes expressed above 1.5 threshold in each cell line were extracted and the number of the genes is depicted in **Fig. 29D**. It was detected that CFS cells derived from H9 and HUES6 hESCs lines are more similar to each other based on transcriptome analysis among three CFS lines. A contrary result was identified among CS lines as CS HMGU1 and CS HUES6 exhibited a closer profile when compared to CS H9.

Lastly, I analyzed the expression of the essentialome factors (**Fig. 29E**) in the six cell lines. Core pluripotency factors *POU5F1* and *NANOG* were not detected in any lines. *MYCN*, *SALL4*, *TRIM71* and *LIN28B* showed higher expression in CFS lines in comparison to the CS lines, except for *TRIM71* in CFS HMGU1. hESC-derived CFS showed relatively higher level of the essentialome factors among the three CFS lines, while it was opposite for the CS lines due to higher expression observed in CS HMGU1.

3.7. Progeny profiling in axial stem cells

Experimentally, I have collected data supporting that CFS cells exhibit the molecular features of NMPs. I set out to investigate if the CFS state can be considered the in vitro counterpart of the NMP state. I sought to prove this hypothesis by investigating the developmental potential of AxSCs and its similarity to the NMP derivatives. Lineage tracing experiments have reported that NMP descendants are found in neural tube and somites at advanced developmental stages^{15,19,23,28}. I investigated progeny profile of the AxSC lines derived from H9 hESCs (**Fig. 10**). CS cells did not resemble NMPs, but they indicated a neural progenitor-like transcriptome and proteome. I explored CS progeny to understand its differences in comparison to neural progenitor derivatives and to pinpoint the differences between CFS and CS developmental capacity.

3.7.1. Neural differentiation from axial stem cells

I first focused on neural differentiation from AxSCs and applied a previously established differentiation modality with some modifications for motor neuron induction from NPCs²⁶⁷. Differentiation medium comprised RA for neural tube induction, SHH for ventralisation, compound E (γ -secretase inhibitor) and cAMP to accelerate differentiation, neurotrophic factors (BDNF, GDNF and IGF) to support survival and maturation of neurons (**Fig. 30A**, left). The neuronal morphology was observed at different timepoints in both CFS and CS differentiation culture. CS cells exhibited neuronal morphology as early as day 2 of the differentiation and CFS cells from day 7 onward. Two weeks from the differentiation onset,

both cultures presented mature neuron-like morphology. Non-neuronal cell types were observed in CFS culture (**Fig. 30A**, right).

I repeated the neural differentiation from AxSCs using three independently derived CFS (CFS_1-3) and 2 CS (CS_1-2) lines previously characterized in **Fig. 10**. Differentiation cultures were evaluated on day 28 by performing RT-qPCR (**Fig. 30B**). In both CFS- and CS-derived cells, I detected the expression of *CHAT*, responsible for acetylcholine production and used in literature to mark cholinergic neurons. Motor neurons and interneurons can both be cholinergic neurons thus, additional markers are needed to investigate the identification of motor neurons in culture. To this end, I assessed *PRPH* expression which is an intermediate filament and marker of peripheral neurons (motor and sensory neurons). It was abundantly expressed in CFS cultures and slightly upregulated in CS cultures. To determine the differentiation outcome was indeed motor neurons, I tested expression of three transcription factors *ISL1*, *MNX1* and *POU4F1*. *ISL1* has been shown in literature as being expressed in dorsal spinal cord neurons, motor neurons and sensory neurons^{147,148}. *MNX1* is a widely-used marker for motor neurons⁷². *POU4F1* is found in multiple dorsal spinal cord neuron domains and sensory neurons likewise *ISL1*, but unlike *ISL1*, *POU4F1* is not expressed in motor neurons²⁷⁶. I detected the expression of all three transcription factors in CFS derived neuronal cultures indicating the heterogeneity of the assayed cultures represented by motor neurons (*MNX1*+) and sensory or dorsal spinal cord neurons (*ISL1/POU4F1*+). CS-derived cells showed the expression of only *ISL1* and *POU4F1* potentially acquiring either dorsal spinal cord or sensory neuronal fate without motor neurons in culture. There was no variation between technical or biological replicates for this experimental set up. I performed immunostaining experiments on day 28 cultures to validate the expression of *ISL1*, *BRN3A* (produced by *POU4F1*), *MNX1*, β 3-tubulin in CFS culture, and *ISL1*, *BRN3A*, β 3-tubulin in CS culture on the protein level (**Fig. 30C**). The results from IgG control staining are given in Supp. Fig. 3.

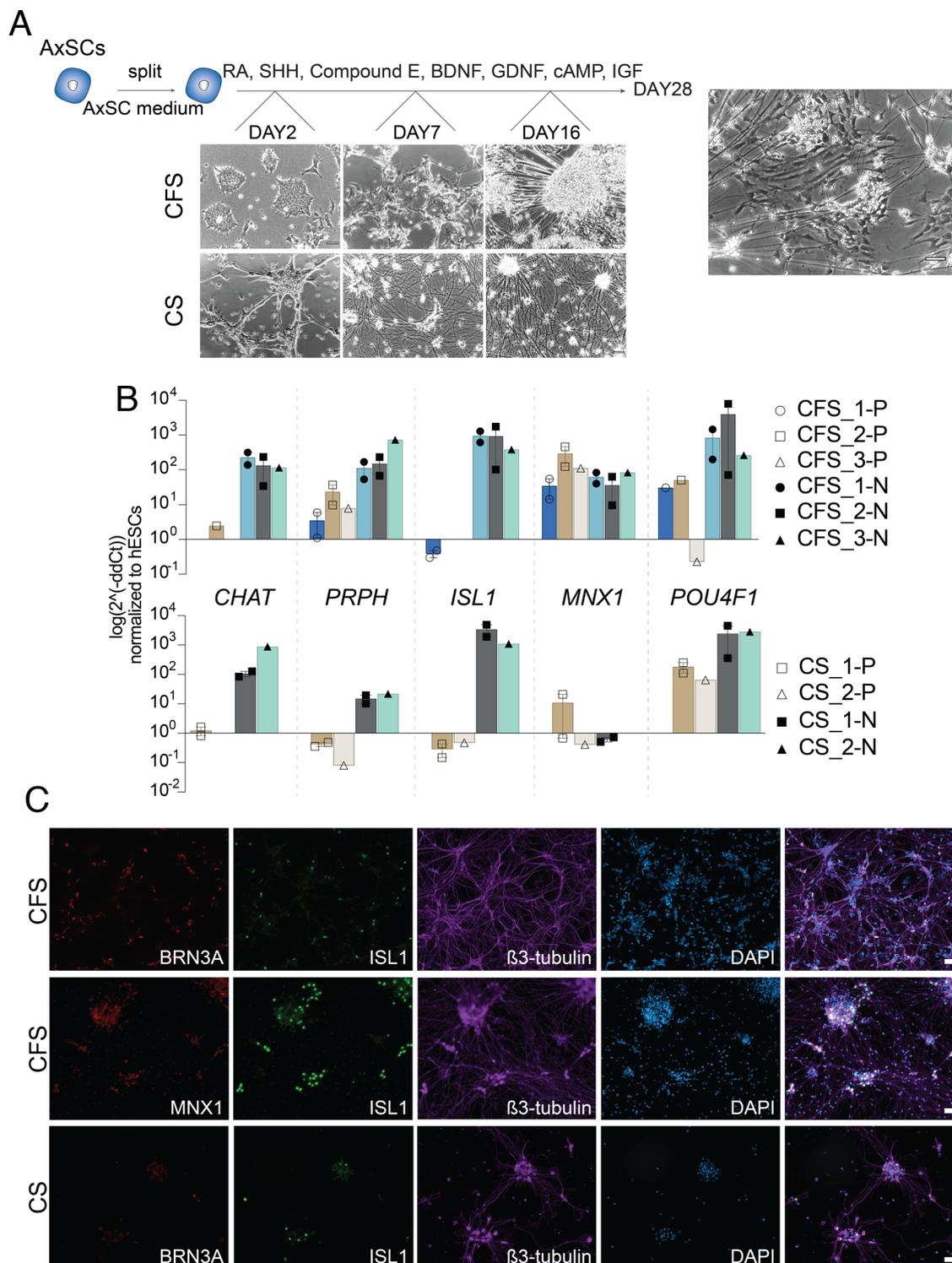


Figure 30: Characterization of neural differentiation cultures from AxSCs

(A) Schematic illustration and morphology of the cells during neural differentiation. H9-derived AxSCs were split in the respective medium and differentiation medium was applied after 24 hours for the next 28 days. Brightfield images taken on day 2, day 7 and day 16 are shown below. The brightfield image on the left side indicate the presence of non-neural cells in CFS culture on day 16 (scale bars: 50 μ M).

(B) Expression analysis by RT-qPCR conducted in day 28 differentiating cultures from each AxSC states. The experiment was carried out by using CFS_1-3 and CS_1-2 lines. Each symbol represents the technical replicates and error bars represent SEM. The Ct values were normalized to undifferentiated H9 cells (P: parental cells, N: neural differentiation culture).

(C) Immunostaining of day 28 differentiating cultures for BRN3A, MNX1, ISL1, β 3-tubulin and DAPI (scale bar: 50 μ M).

To understand heterogeneity of neural subtypes derived from AxSCs, scRNA sequencing time course experiment was performed. The experimental set up for the scRNA time course experiment included CFS, CS-derived neuronal progeny and NPC progeny, in order to determine progeny differences between hESCs-derived NPCs and AxSCs. I first implemented the differentiation from NPCs which were previously characterized in **Fig. 21** and **Fig. 22**. For this purpose, I utilized the original protocol which was modified for AxSC differentiation²⁶⁷. Briefly, 24 hours after plating the cells in NPC maintenance medium, it was switched to differentiation medium comprising RA, SHH, BDNF, GDNF, cAMP and IGF for 14 days (**Fig. 31A**). Subsequently, compound E was added to differentiation medium and the cells were kept in culture for additional 14 days. I observed in day 28 cultures cells with neural morphology and cells resembling neuronal progenitor cells (**Fig. 31B**). Next, I assessed the expression of *CHAT*, *PRPH*, *ISL1* and *MNX1*. Upregulation of all four selected genes indicated the achievement of motor neurons cell fate in my cultures (**Fig. 31C**).

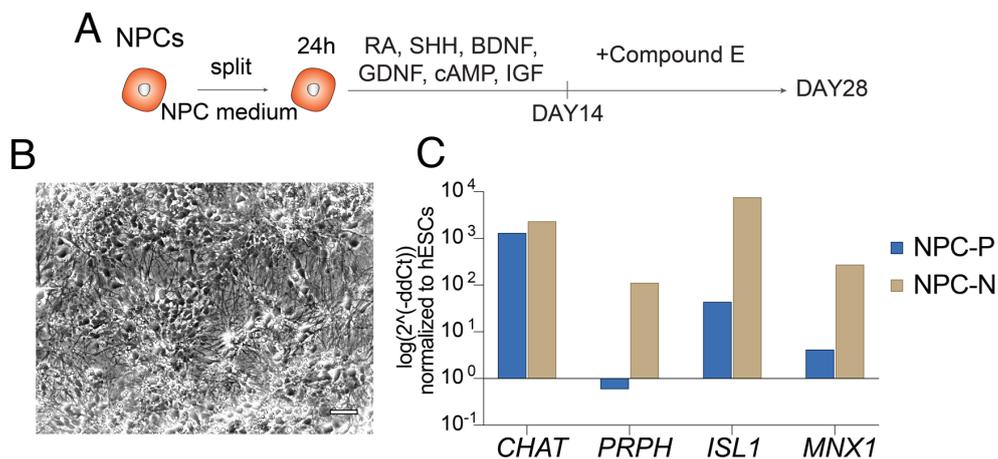


Figure 31: Characterization of neural differentiation from NPCs

(A) Schematic illustration of the neural differentiation protocol from NPCs.

(B) Brightfield image of day 28 differentiation culture (scale bar: 50 μ M).

(C) Expression analysis by RT-qPCR in day 28 differentiating cells. The Ct values were normalized to undifferentiated H9 cells (P: Parental cells, N: Neural differentiation culture).

For scRNA sequencing experiment, the sequencing time points were selected based on the morphological observations made during the AxSCs differentiation (**Fig. 30**): day 0, day 2, day 14 and day 28 from both AxSC states, day 0 and day 28 from NPCs. As a result of sequencing technical issues, day 28 CS sample was excluded from the analysis presented below. First, I analyze the cell-cycle stages to understand the maturation of the differentiation cultures (**Fig. 32A**). This analysis provided an overview of the number of cells present in G1, G2M or S phase in each time point during the differentiation (**Fig. 32B**). Comparison of day 2 cultures to undifferentiated state (day 0) presented large changes in CS but not in CFS cultures. 80% of day2 CS cells were detected in G1 phase which is approximately 65% higher than the day 0 cells. There was no increase found between day 0 and day 2 CFS cells. These results accounted for the morphological changes observed previously (**Fig. 30A**). On day 14 and day 28, a high number of G1 stage cells was detected in CFS culture (~90%) similarly to day 14 CS culture (~80%). NPC cells included at the end timepoint a lower number of G1 phase cells (**Fig. 32B**), which fall in line with the morphological observations shown in **Fig. 31C**. Based on cell cycle activity in the different time points, I proceeded cellular identification analysis with both day 14 and day 28 for CFS, day 14 for CS as well as day 28 NPCs. For the annotation of the various cell populations present in AxSCs- and NPC-derived neuronal cultures, I used the study published by Rayon and her colleagues²⁷⁶ alongside the accompanying database (<https://shiny.crick.ac.uk/scviewer/neuraltube/>) as a guideline to generate a signature marker map for each domain of spinal cord for the Carnegie stages 12, 17 and 19 of human embryonic development. Table 8 summarized the markers used to identify cellular populations throughout the analysis presented below.

Table 8: Marker genes used for annotation of the cellular populations in scRNA sequencing time course datasets

Cell type	Genes	Cell type	Genes
dI1	BARHL1, BARHL2	v2b	GATA2, GATA3
dI2	FOXD3, POU4F1, POU4F2	MN	MNX1
dI3	POU4F1, OTP	aMN	ISL1, UNC5C, PHOX2A, PHOX2B
dI4	GBX1	v3	NKX2-2, SIM1
dI5	LMX1B	SN	ISL1, PIEZO2 POU4F1, DRGX, TAFA1
dI6	DMRT3	preEMT_NC	BMPER, LMX1A, GDF7
v0	EVX1, EVX2	NC_d	ETS1, PRRX1
v1	EN1	Sch	SOX10, MPZ, S100B, CDH19
v2a	VSX2, SOX14, SOX21	CS-like	SOX2, PAX6, HES5, PAX3, PAX7, SOX1, SOX3

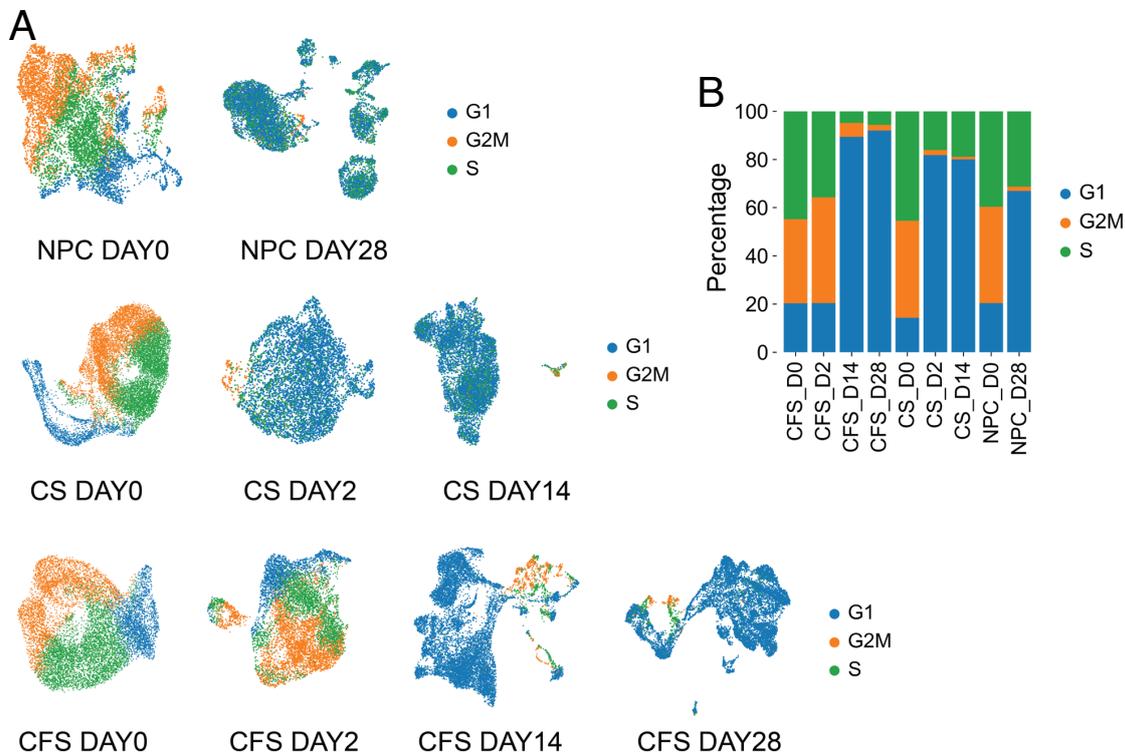


Figure 32: Cell cycle analysis time course during the neural differentiation of AxSCs and NPCs
 (A) Representative UMAP plots for parental CFS_2, CS_2 and NPCs (DAY 0) and their differentiation cultures at several time points
 (B) Stacked bar plot showing the percentage of the cells (A) in respective cell cycle phases

I first analyzed the expression of ventral spinal cord markers as the differentiation protocol results in ventral subtypes, in particular motor neurons, due to the presence of SHH in the differentiation medium. Both CFS day 14 and day 28 dataset contained a small number of cells marked by expression of *EVX1* and *EVX2* similarly to v0-like cells (**Fig. 33A**). A group of *EVX2* expressing cells (cluster 8 in **Fig. 34**) in CFS day 14 dataset displayed expression of sensory neuron (SN) markers thus, this cluster was annotated only as SNs. *EN1* expression marking the v1 neurons was found abundantly and accumulated in cluster-like populations in both datasets. To identify v2a and v2b neurons, co-expression of *VSX2/SOX14/SOX21* and *GATA2/GATA3* respectively was analyzed, as opposed to single gene expression as their individual expression is found in multiple ventral spinal cord domains. Cumulative expression of said gene combinations was not detected indicating that either v2a- or v2b-like cells were not present in the CFS cultures based on the scRNA results. I observed a high number of *MNX1* expressing cells indicating motor neuron presence, which was largely distributed in the datasets. Co-expression of *NKX2-2/SIM1* and *LMX1B/SHH/ARX* that respectively mark v3 neurons and floor plate cells could not be found in either dataset. The

analysis of CS day 14 sample showed that none of the ventral neuron markers are expressed except for *EN1* (**Fig. 33B**). This cluster did not represent post-mitotic neurons based on the cell cycle phase analysis, indicating progenitor-like identity (**Fig. 32A**). Ventral propensity could not be ruled out. Finally, I analyzed the abovementioned markers in NPC day 28 dataset. As shown in **Fig. 33C**, v0 neurons were not detected due to the lack of *EVX1* and *EVX2* expression. *EN1* expression was widely distributed in the dataset representing v1-like neurons. A cluster of cells was marked by co-expression of *VSX2/SOX14/SOX21* indicating v2a neuron-like population. *GATA2/GATA3* co-expression marking the v2b neurons was found in a define cluster alongside two scattered populations. A few cells in the *GATA2/GATA3* positive cluster exhibited a slight *MNX1* expression. The *MNX1* expression was quite homogenous, with the exception of a distinct cluster comprising a few cells. A small population was marked by *NKX2-2/SIM1* expression thus presenting v3 neuron like identity. *LMX1B* alone is the determinant of d5 neurons, its co-expression with *SHH* and *ARX* is an indicator of floor plate cells²⁷⁶. In NPC dataset, I found *LMX1B* expressing cells, displaying *SHH* and *ARX* expression thus, I annotated them as floor plate (FP) cells.

The analysis based on ventral spinal cord markers could not successfully identify all of the cell populations presented in the datasets. I divided the datasets into the clusters (**Fig. 34A**) and proceeded to decipher dorsal spinal cord cells in the datasets. The transcriptome of dorsal spinal cord domains is very similar to each other with the exception of dI5 and dI6 which are respectively marked by *LMX1B* without *SHH* and *ARX*, and *DMRT3* respectively²⁷⁶. To unveil dorsal cell types in the dataset along with dI5 and dI6, I analyzed co-expression of several genes that are simultaneously expressed in certain domains (dI1-4) as shown in Table 8. The results showed that there are no cells expressing *BARHL1* and/or *BARHL2* indicating the lack of dI1-like cells in either AxSCs or NPC differentiation cultures. Expression of *POU4F1/POU4F2/FOXD2*, collectively designating dI2 cells, was slightly upregulated in cluster 3 from CFS day 14 and abundantly expressed in cluster 5 from CFS day 28, but it was not detected in either CS or NPC datasets (**Fig. 34B**). *POU4F1/OTP* co-expressing cells marking dI3 cells were spanning 3 different clusters (2-4) in CFS day 14. Their numbers increased as the differentiation progressed as *POU4F1/OTP* expressing cells were found in several clusters in CFS day 28 dataset, particularly in cluster 4. None of the CS and NPC derivatives were denoted with dI3 identity. Expression of *LMX1B* was detected in all datasets; cluster 7 in CFS day 14, cluster 8 in CFS day 28, and primarily cluster 1, 2 and 4 in CS day 14. *LMX1B* was highly upregulated in three clusters from NPC dataset, corresponding to the cells previously determined as co-expressing *LMX1B/SHH/ARX* (**Fig. 33C**) previously annotated as floor plate ruling out dorsal identity. *DMRT3* expression was found distributed in multiple clusters, yet it was more abundantly present in clusters 1/4/5 in CFS day 14, 3/7 in CFS day 28, 3/4 in CS day 14 datasets (**Fig. 34B**), but it was not detected in NPC progeny.

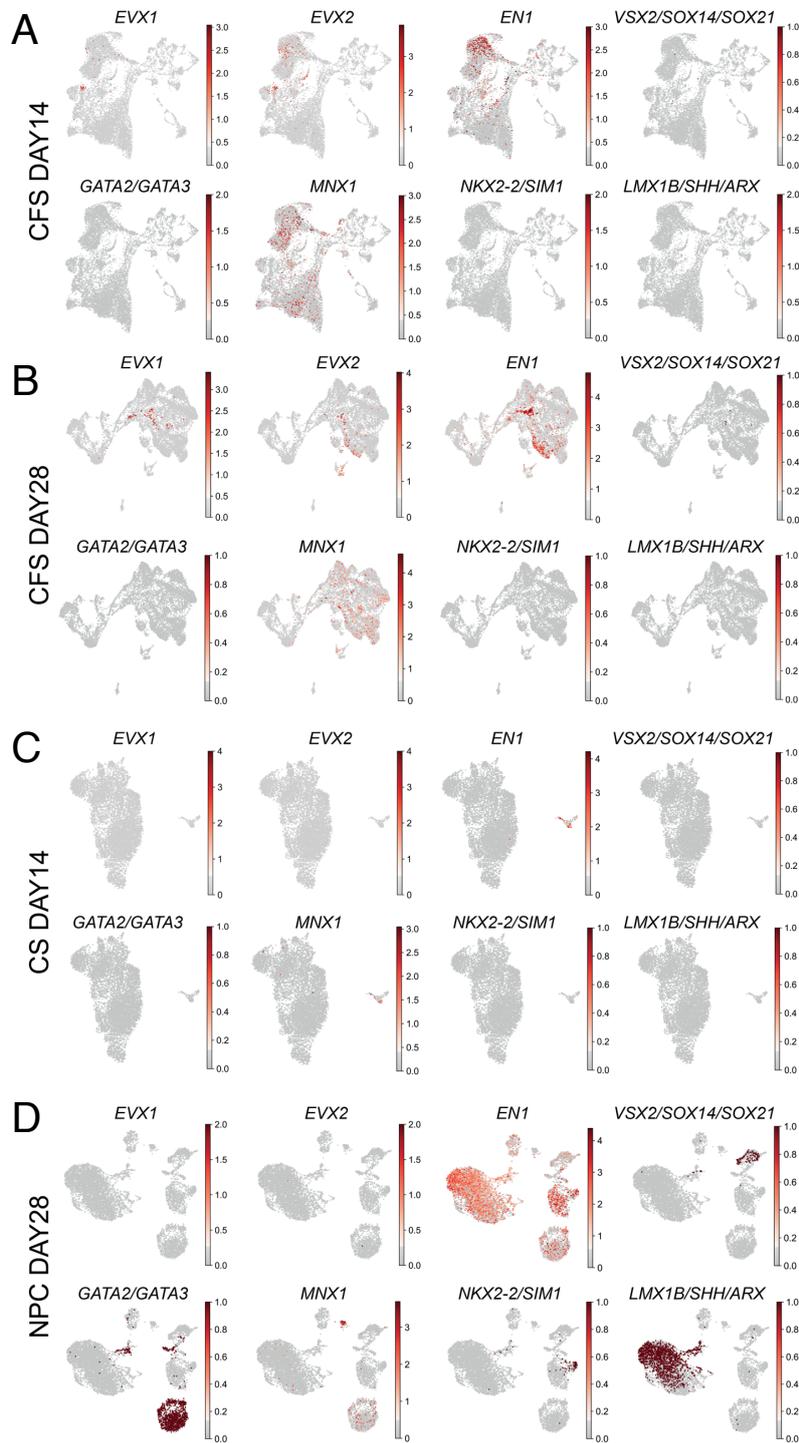


Figure 33: Expression of ventral spinal cord markers in human AxSCs and human NPC differentiation datasets

Representative UMAP plots indicating expression of *EVX1* and *EVX2* (v0 neurons), *EN1* (v1 neurons), *VSX2/SOX14/SOX21* (v2a neurons), *GATA2/GATA3* (v2b neurons), *MNX1* (motor neurons), *NKX2-2/SIM1* (v3

neurons) and *LMX1B/SHH/ARX* (floor plate) in differentiation cultures from CFS_2 on day 14 (A) and 28 (B), CS_2 on day 14 (C) and NPC on day 28 (D).

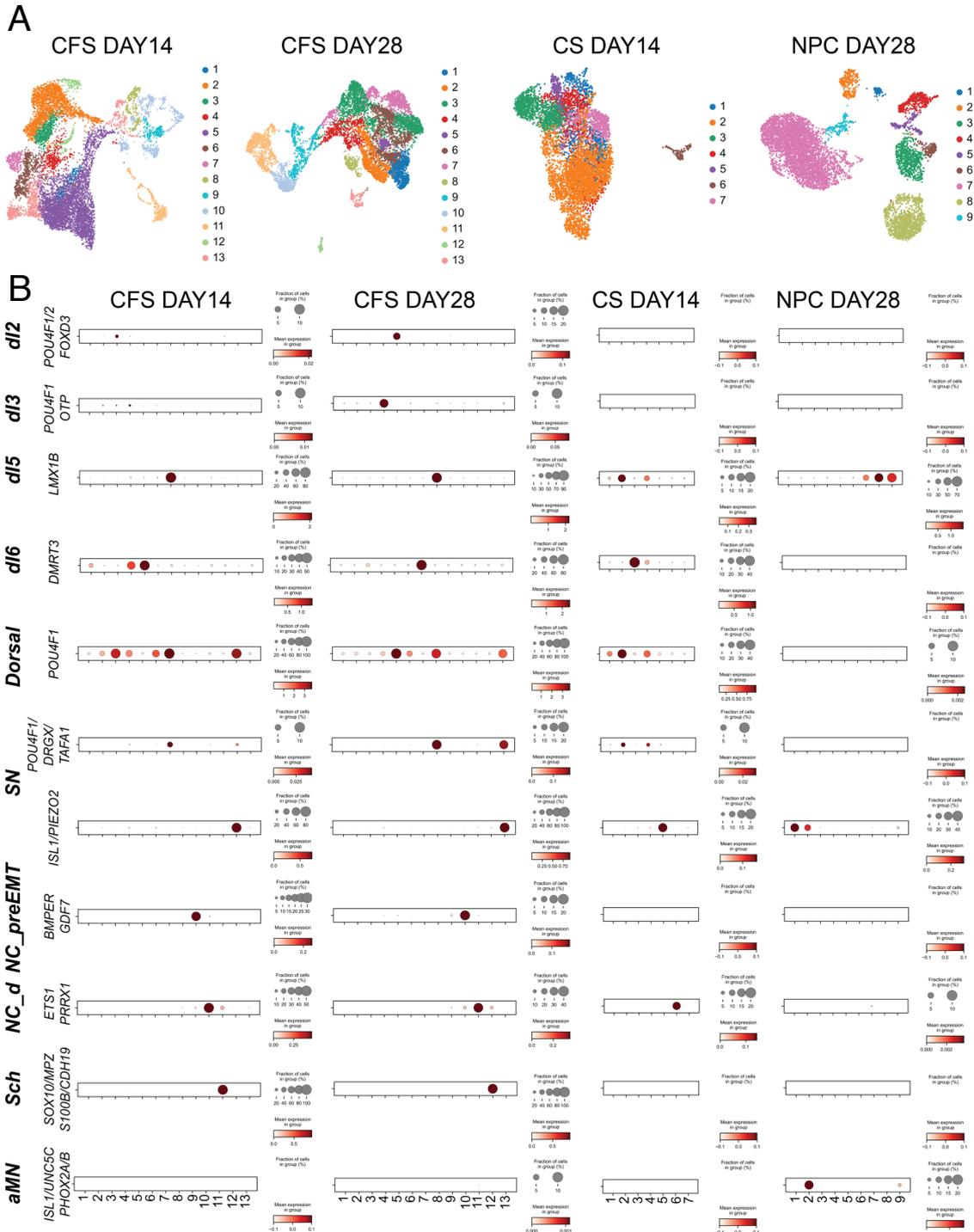


Figure 34: Expression of neural tube and neural crest-development associated genes in human AxSC and human NPC differentiation datasets

(A) Representative UMAP plots indicating the clustering of the cells in each dataset

(B) Dotplots showing the expression of dI2-6 dorsal neural tube markers, neural crest (NC) markers from different stage of cells and accessory motor neuron (aMN) markers. Dot size represents percentage of the cellular population expressing the gene of interest and color scale displays the mean expression (NC_d: neural crest derivatives, NC_preEMT: pre-epithelial to mesenchymal neural crest cells, Sch: Schwann cells, SN: sensory neurons).

Clusters present in day 14 (cluster 9-10, 12-13) and day 28 (cluster 9-11) in CFS samples (**Fig. 34A**) did not align with neither dorsal nor ventral spinal cord identity. High number of dorsal-like cells present in the AxSC datasets suggested that these cells could have acquired neural crest identity, due to the common origin dorsal spinal cord and neural crest cells share in embryonic development²¹. NMPs have been shown to contribute to neural crest development^{20,69,135,231}. To this end, I investigated the expression of genes associated to neural crest development and its progeny by using the stage-specific markers published by Soldatov et al.¹³⁹. The selected genes are listed in Table 8. The results showed that cluster 9 (CFS day 14) and cluster 10 (CFS day 28) were marked by *BMPER/GDF7* co-expression indicating pre-migratory or pre-epithelial to mesenchymal (NC_preEMT) neural crest cells (**Fig. 34B**). The identification of NC_preEMT cell identity at the transcriptomic level is a potential source for the morphological observations made in day 28 CFS differentiating culture (**Fig. 30A**). *ETS1/PRRX1* co-expression has been shown to characterize the majority of neural crest derivatives¹³⁹ and it was found in both CFS and CS datasets (NC_d) (**Fig. 34B**). Sensory neurons (SN), which are a neural crest derivative, were identified in axial and NPC derived progeny. For the classification of SNs, the in vivo spinal cord data from Rayon et al.²⁷⁶ was used. The SN population co-expressing *POU4F1/DRGX/TAF1* was identified in CFS and CS differentiation cultures, while a second SN population co-expressing *ISL1/PIEZO2* was present in both AxSC and NPC derivatives. These two populations might indicate different subtypes of SNs. A group of cells (cluster 11 in CFS DAY14 and cluster 12 in CFS DAY28) displayed *SOX10/MPZ/S100B/CDH19* co-expression indicating the potential presence of Schwann cells (Sch) in the CFS differentiation cultures. A high number of cells showed abundant *POU4F1* but without the co-expression of any additional dorsal spinal cord or neural crest markers in Table 8, thus they were denoted simply as 'Dorsal' cells.

The protocol used for neural differentiation in this study originally targets the NPC differentiation into motor neuron cells. High heterogeneity was detected on the transcriptomic level in both AxSC and NPC differentiation cultures. The number of *MNX1* expressing cells were lower in NPC derivatives when compared to CFS progeny. I questioned whether the more anterior like motor neurons namely accessory motor neurons (aMN) lacking *MNX1* and distinguishable by *ISL1/UNC5C/PHOX2A/PHOX2B* co-expression^{90,302,303} were present in the NPC dataset. Indeed, cluster 2 next to *MNX1* expressing cells and a part

of the cluster 9 exhibited co-expression of these factors while aMNs were not detected in any of the AxSC derivatives (**Fig. 34B**).

The mapping of differentiation cultures to spinal cord and neural crest derivatives revealed the differentiation trajectories of the majority of the cellular populations. Cluster 8 in CFS DAY14 and cluster 9 in CFS DAY28 did not display expression of any of the abovementioned genes. The outcome obtained could be anticipated as based on the cell cycle phase analysis these clusters exhibited immature or progenitor profile (**Fig. 32**). It could be hypothesized that these clusters, in particular cluster 9 in CFS DAY28, were likely to be a representation of a developmental phase or progenitor population which is able to bifurcate towards both spinal cord and neural crest trajectories, especially when taking into account their position on the UMAP (**Fig. 34**). The analysis above showed that CS cells can differentiate into the spinal cord and neural crest derivatives, and taken together with the experimental evidence that CFS cells are able to generate CS cells (**Fig. 28**), it lead me to question if cluster 8 in CFS DAY14 and cluster 9 in CFS DAY28 could have a CS-like cluster identity. I analyzed the markers used to identify the CS cells in the differentiation datasets. The results confirmed this hypothesis due to co-expression of *SOX2*, *PAX6*, *SOX1*, *SOX3* and *PAX3* (**Fig. 35**). The cells present in these clusters were not identical to undifferentiated CS cells due to their low *HES5* and *PAX7* expression indicating a potential transitional state from CS cells, thus named CS-like population. CS DAY14 and NPC DAY28 datasets were found as lacking the so-called CS-like cells.

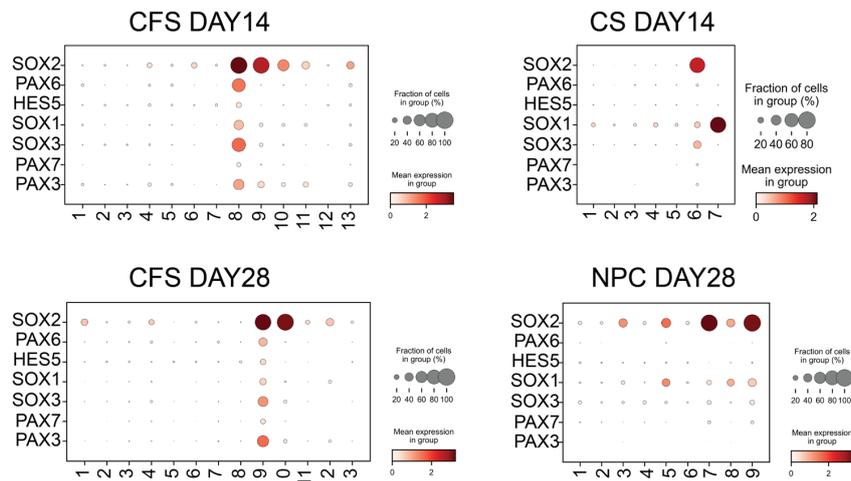


Figure 35: Expression of undifferentiated human CS cell markers in neural differentiation datasets

Dot size represents percentage of the cellular population expressing the gene of interest and color scale displays the mean expression.

Expression of the selected markers used for annotation of the cellular populations in scRNA sequencing dataset was checked experimentally by performing a RT-qPCR experiment (**Fig. 36**). *SOX2* was downregulated in CS day 14 culture, while CFS day 14/28 and NPC day 28 cultures exhibited *SOX2* expression at a similar level compared to undifferentiated H9 cells. *PAX6* and *HES5* expression was upregulated in all samples and CFS differentiating cells in both time points showed the highest levels of *PAX6* and *HES5*. *PAX3* and *PAX7* expression was enriched in both CFS cultures while *PAX7* was also detected in NPC day 28 culture. CS day 14 differentiating cells exhibited low levels of *PAX3* and *PAX7*. *EN1* (v1 neurons) was abundantly expressed in CFS day 14/28 and NPC day 28 cultures, while presenting a slight upregulation in CS day 14 culture. High upregulation of *GATA2* (v2b neurons) was detected in NPC day 28 sample while both CFS cultures showed low *GATA2* expression. All cultures except for CS samples exhibited *MNX1* expression. *SIM1* upregulation was found in all samples, however it was not enough to point out the presence of v3-like cells as *NKX2-2* expression was not examined.

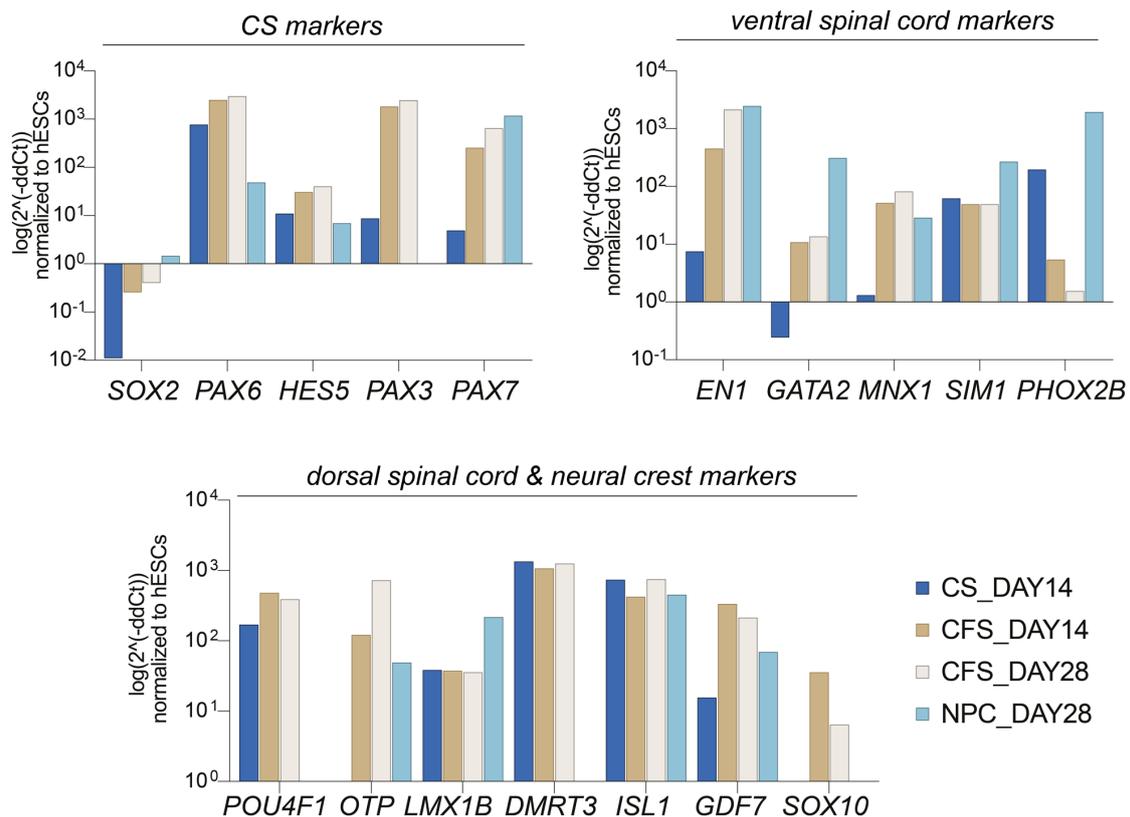


Figure 36: Experimental validation of the markers used to annotate the neural differentiation cultures obtained from human AxSCs and human NPCs

Expression analysis of the markers selected from Table 8 by RT-qPCR in the samples used for time course scRNA sequencing experiment. The Ct values were normalized to undifferentiated H9 cells.

OTP upregulation was detected in both CFS and NPC differentiation cultures. Co-expression of *OTP* with *POU4F1* marks the dI3 cells, while cumulative expression of *OTP* and *EN1* has been shown in v1 cells²⁷⁶. As *POU4F1* was abundantly expressed in CFS- and CS- but not NPC-derived cells, *OTP* expression in NPC derivatives could potentially indicate the presence of v1-like cells, which was then confirmed by *EN1* upregulation. High expression of *LMX1B* and *DMRT3* was detected in all cultures except for the NPC derivatives in accordance with scRNA sequencing results. *ISL1* was abundantly expressed in both AxSC and NPC differentiation cultures as expected based on previous results. Lastly, *GDF7* expression was detected in CFS day 14 and day 28 samples, but also in CS and NPC cultures, unlike the results obtained from the transcriptomic analysis. *SOX10* expression was found only in CFS cultures.

To summarize the results of the scRNA analysis performed thus far I generated representative UMAPs for all the datasets (**Fig. 37**). The v0, v1 and MN-like cells were identified in both CFS samples while for the NPC progeny all the ventral domain-like cells except for v0 could be identified (**Fig. 37**). The aMN and FP-like cells were detected only in NPC progeny. The cells transcriptionally resembling dI5 and dI6 dorsal domains were present in both AxSC derivatives, while CFS samples comprised additionally of dI2 and dI3-like cells. In both CFS cultures, a few clusters were determined as immature/progenitor stage. These clusters implied a different differentiation trajectory other than spinal cord which was more distinguishable for the end timepoint in CFS differentiation. A particular cluster appearing as an intersection of the two trajectories was detected as CS-like cells. A similar population was present in CFS day 14 differentiating cells. Other two clusters, presenting progenitor stage identity were identified as neural crest cells through different developmental stages like preEMT-like neural crest cells and unspecialized neural crest derivatives (NC_d). The latter identity was also found in CS day 14 dataset. Sensory neurons (SN) which are developmentally more mature were identified in both AxSC and NPC samples. Lastly, a population resembling Schwann cells (Sch) was detected in CFS day 14 and day 28 cultures.

Next, I investigated the HOX profile of both parental and differentiating progeny in time course dataset and their similarity to their in vivo spinal cord counterparts. All the time points were integrated in addition to the published data by Rayon et al.²⁷⁶ and they were processed using the same algorithms as for the two previous datasets. Based on the sequencing results, it was noted that *HOX* genes were not expressed in H9 and NPCs (**Fig. 38**).

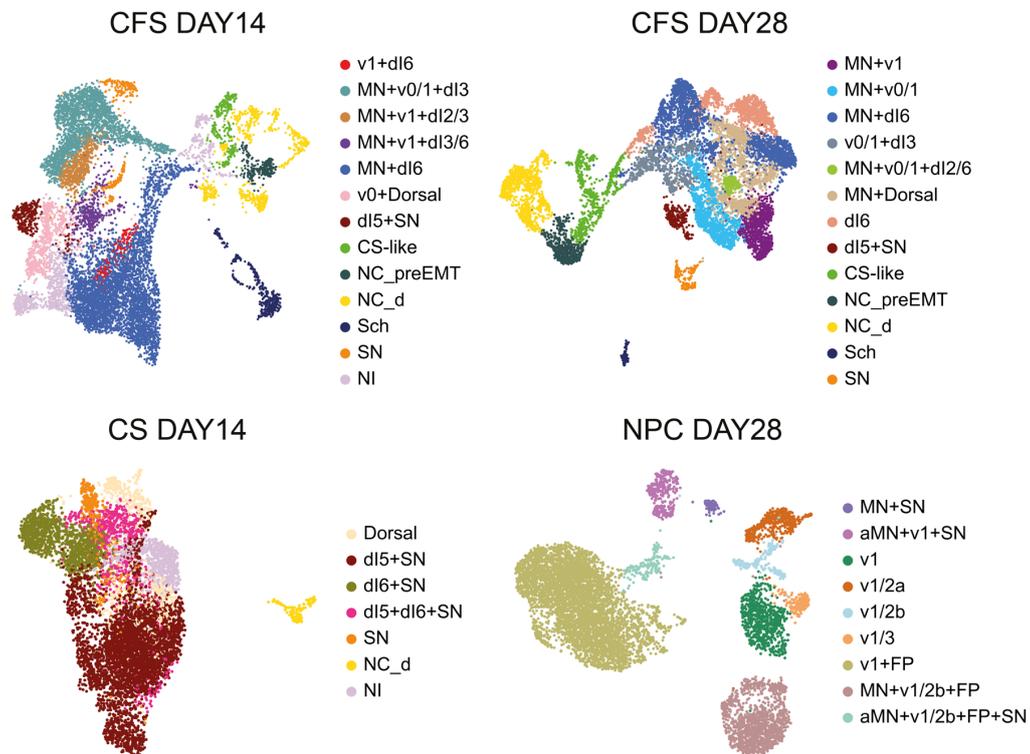


Figure 37: Summary of the progeny analysis in human AxSCs and human NPCs differentiation cultures

Representative UMAP plots showing the annotated cells in differentiating cultures from CFS₂ on day 14 and day 28, CS₂ on day 14 and NPC on day 28. The clusters comprising same cellular populations were marked by same colors (aMN: accessory motor neurons, FP: floor plate, MN: motor neuron, NC_d: neural crest derivatives, NI: not-identified, NC_{preEMT}: pre-epithelial to mesenchymal neural crest cells, Sch: Schwann cells, SN: sensory neurons)

CFS trajectory pointed out a slight downregulation in the anterior *HOX* genes (PG 1-4) and an upregulation of the posterior *HOXA10* and *HOXD9-11*. CS progeny transitioned to a more posterior *HOX* profile due to the upregulation of *HOX* PG 3-9 as they progressed through the later differentiation timepoints. The *HOX* profile (PG 1-8), at the differentiation endpoint was similar between day 28 CS and NPC differentiating cells, however their expression levels were higher in CS progeny. I questioned the expression of multiple *HOX* genes per timepoint and to address this point I compared the *HOX* profile obtained with the *HOX* profile exhibited by the in vivo spinal cord cells, which were marked by similar trend as I could detect anterior to posterior *HOX* gene expression present in both brachial and thoracic spinal cord levels. Comparison of the overall AxSCs progeny profile and in vivo spinal cord cells indicated CFS differentiating cells in day 14 and 28 closely resemble spinal cord cells at Carnegie stage (S) 17 thoracic level. The endpoint of CS cells could very likely represent S12 cells which is an early developmental timepoint and spinal cord identity is not yet clearly distinguished.

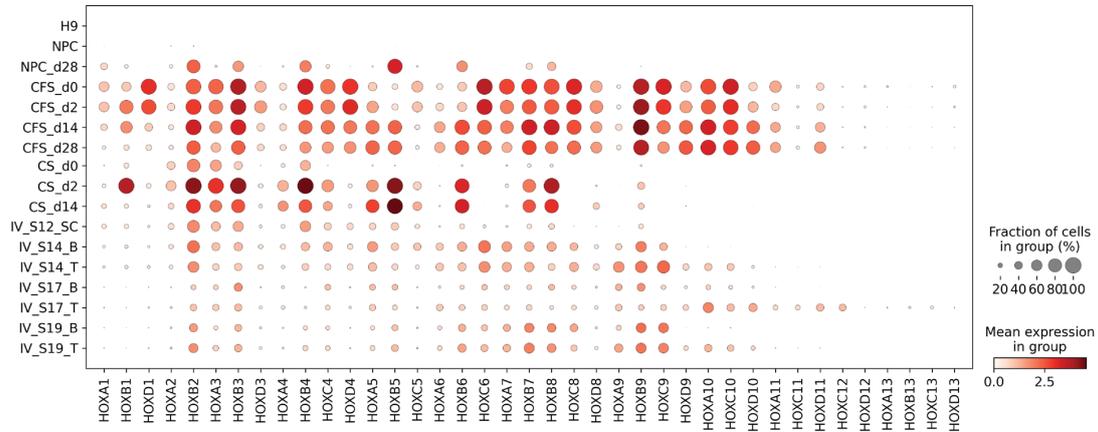


Figure 38: HOX expression profile in neural differentiation trajectory from AxSCs and H9 cells as well as in vivo spinal cord cells

Dotplot indicating the expression of anterior to posterior *HOX* genes in the time course scRNA sequencing dataset. The published data from Rayon et al.²⁷⁶ was preprocessed by the same algorithms used in this study and concatenated with the neural differentiation time course dataset comprising H9 and human AxSCs (CFS_2 and CS_2) (IV: in vivo, S: Carnegie stage, SC: spinal cord, B: brachial, T: thoracic).

3.7.2. Mesodermal differentiation from axial stem cells

Grafting experiments have reported that CLE cells, where NMPs are harbored, contribute to the somites¹⁶. To confirm NMP identity and their mesodermal potential, in vitro derived NMPs were successfully differentiated into skeletal muscle cells⁴⁴. In line with the published studies, I investigated propensity of AxSCs to differentiate into the mesodermal lineage by focusing into the skeletal muscle cell differentiation.

3.7.2.1. Skeletal muscle differentiation targets the segmentation clock

The somites, which are a type of axial structure, give rise to skeletal muscle cells. The somite formation is controlled by the precise cycling event called the segmentation clock. Notch pathway plays a critical role in the regulation of key genes such as *Hes7* and *Mesp2* in the paraxial mesoderm cells. Expression of these genes are mediated by the Notch oscillation through *Tbx6* and *Msgn* expression which are upregulated in the somitic mesoderm^{97,98,304}. The oscillation events lead to the formation of somites which differentiates into skeletal muscle cells. Considering the Notch pathway effect, I identified a skeletal muscle differentiation protocol established from hPSCs that includes Notch inhibition²⁶⁸. I modified the protocol for AxSCs, as from the scRNA results we could show that AxSCs represent a latter development stage compared to hPSCs. I skipped the first three days in the original protocol that involve WNT/FGF pathways activation and started directly with Notch inhibition by using DAPT (**Fig. 39A**). Together with Notch inhibition, I screened several conditions. In the

literature, it has been suggested that SHH activation could improve myogenesis^{305,306}. Inhibition of BMP could potentially enhance paraxial mesoderm formation as BMP is responsible for lateral mesoderm induction³⁰⁷. To this end, I used DMH1 to block BMP signaling. Lastly, Activin A plays a critical role for paraxial mesoderm induction as shown for mouse iPSCs³⁰⁸ therefore I sought to investigate its effect on human axial cultures. Between day 0 and day 8, I applied DAPT with or without SHH, Activin A, DMH1. From day 8, I added medium supplemented only with FGF2 and FGF8. I tested the effect of ascorbic acid (AA) and cAMP because they are shown to regulate skeletal muscle metabolism and to enhance the maturation via further activating ERK1/2 signaling pathway^{309,310}. I treated the cells from selected conditions with horse serum (HS) for 2 weeks as it was shown it plays a role as a maturation factor for myoblasts³¹¹. The abovementioned conditions for skeletal muscle differentiation are summarized in Table 9 and illustrated in **Fig. 39A**.

Table 9: Conditions tested for skeletal muscle differentiation from human CFS lines

Day0-8	Day8-26	Day26-40	#Condition
DAPT	FGF2, FGF8		#1
DAPT	FGF2, FGF8, AA, cAMP		#2
DAPT, SHH	FGF2, FGF8		#3
DAPT, SHH	FGF2, FGF8, AA, cAMP		#4
DAPT, ActA	FGF2, FGF8, AA, cAMP		#5
DAPT, DMH1	FGF2, FGF8, AA, cAMP		#6
DAPT	FGF2, FGF8	HS	#7
DAPT, SHH	FGF2, FGF8	HS	#8

Differentiation cultures were harvested for analysis on day 40. Expression of NMP/early mesoderm marker (*TBXT*), somitic mesoderm markers (*TBX6* and *MSGN*) and skeletal muscle markers (*MYOD* and *MYOG*) was analyzed by performing RT-qPCR experiments (**Fig. 39B** and **Fig. 40**). In all cultures, *TBXT*, *TBX6* and *MSGN* expression were downregulated for the end timepoint of differentiation except for one iteration resulting from the CFS_2 differentiation, where DAPT and FGF2/8 were consecutively applied (#1). The parental lines either express low levels of *MYOD* and *MYOG* or not at all, however these genes were abundantly expressed by day 40 cultures in the #1-4 conditions implying the successful differentiation towards skeletal muscle cells. The morphological observations from #1-4 conditions indicated that the cells were not sufficiently matured and they could resemble myoblast stage considering the upregulation of *MYOD* and *MYOG*. Addition of either SHH (#3) or AA with cAMP (#2 and #4) did not show a significant effect on *MYOD* and *MYOG* expression compared the condition where only DAPT was applied (#1). *MYOG* was upregulated by the Activin A treatment (#5) and DMH1 (#6), but *MYOD* expression was not detected, thus these conditions were omitted from further experiments.

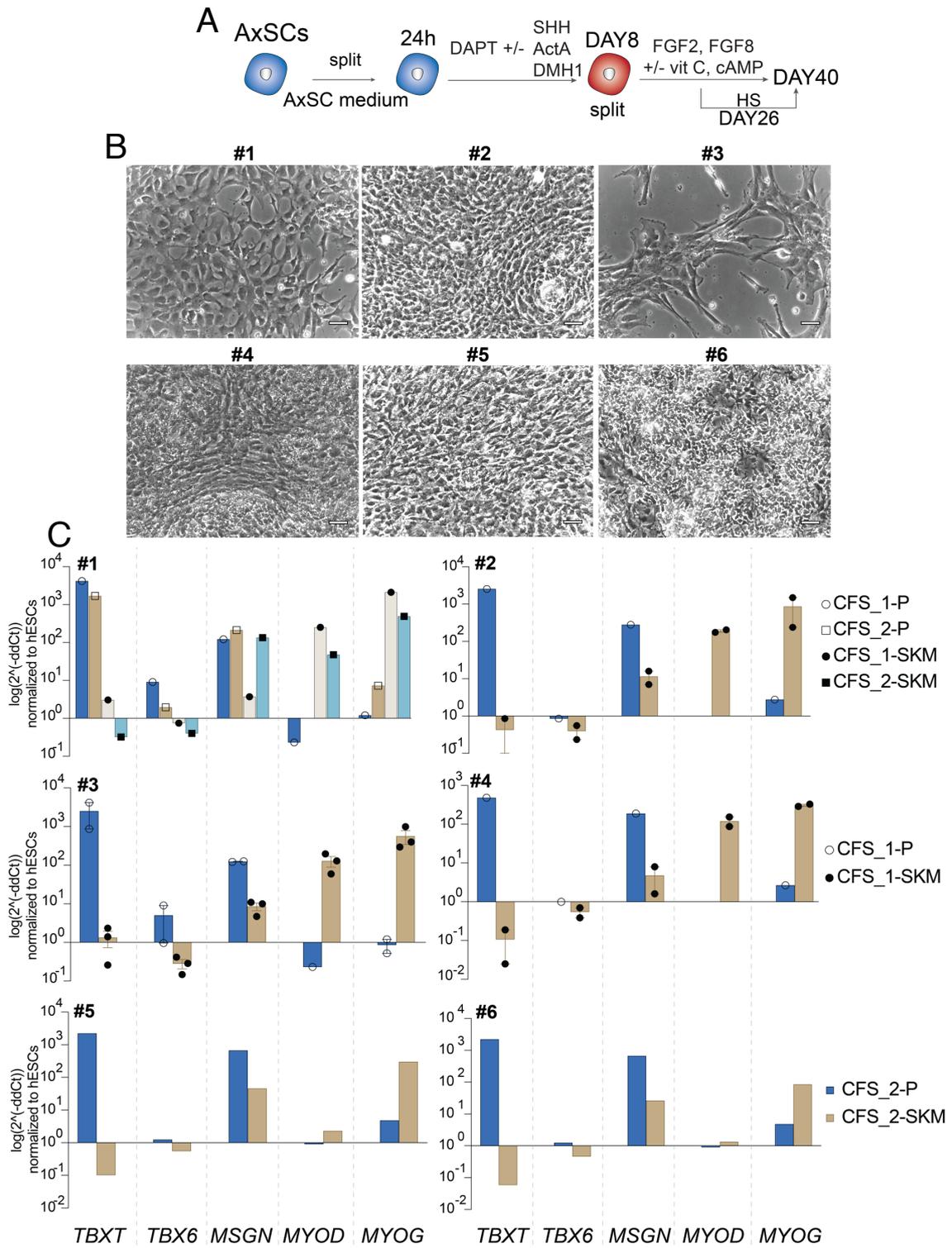


Figure 39: Characterization of skeletal muscle differentiation cultures from human CFS cells by mimicking segmentation clock

(A) Schematic illustration of differentiation protocols

(B-C) Morphology of day 40 differentiating cells and their expression analysis for skeletal muscle-associated markers by RT-qPCR. CFS_1-3 lines in Fig. 10 were used for differentiation experiments. Conditions #1-6 are listed in Table 9. Technical replicates are indicated by the same symbols. The Ct values were normalized to undifferentiated H9 cells. Error bars represent SEM (scale bar: 50 μ M, P: parental, SKM: Skeletal muscle differentiation culture).

Due to the similar expression rate detected for *MYOD* and *MYOG* with or without SHH treatment (#1 and #3 in **Fig. 39B**), I tested the effect of HS on these two cultures conditions to see if the differentiating cells could be fused to form myocytes, thus inducing maturation. For this purpose, I removed FGF2 and FGF8 on day 26 and applied only HS (**Fig. 39A**). The cultures from the conditions #7-8 (Table 9) were assessed on day 40. Mature skeletal muscle-like morphology could not be detected in either of the differentiation cultures as fused cells could not be observed (**Fig. 40A**). Expression of the abovementioned stage-specific markers was examined by performing RT-qPCR experiment (**Fig. 40B**). *TBXT*, *TBX6* and *MSGN* were downregulated in the cells from both conditions (#7 and #8), and the cells were lacking *MYOD* and *MYOG* expression. High *MYOG* expression was detected in parental CFS-2 cells unlike previous experiments (**Fig. 39B**) which could be a technical error. The results indicated that a long term FGF activation concomitant to Notch inhibition can lead to the successful differentiation of skeletal muscle cells from AxSCs and treatment with HS only of human axial differentiating cultures is not sufficient to induce their maturation.

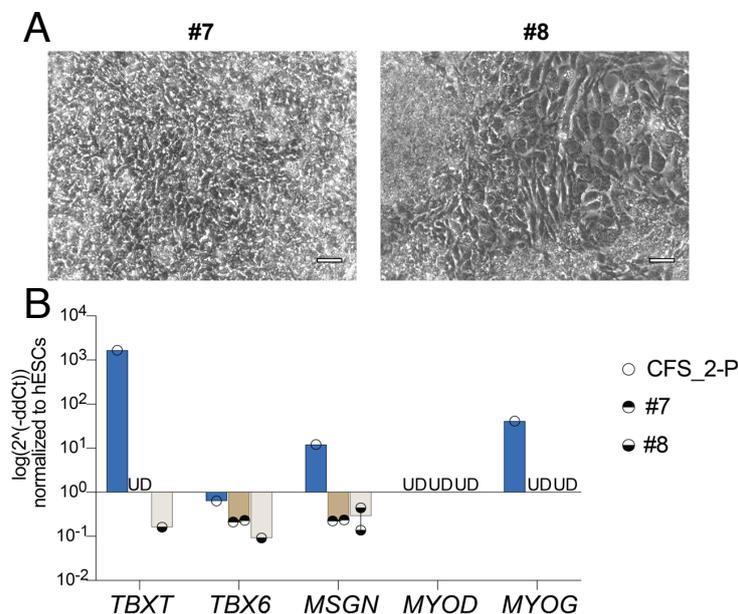


Figure 40: The effect of horse serum on maturation of skeletal muscle cultures from human CFS cells

(A) Morphology of the differentiating cells on day 40. The medium components for the conditions #7 and #8 are listed in Table 9 (scale bar: 50 μ M).

(B) Expression analysis in day 40 differentiation cultures (condition #7 and #8) by RT-qPCR. CFS_2 line was used for the differentiation experiments. Technical replicates are indicated by the same symbols. The Ct values were normalized to undifferentiated H9 cells. Error bars represent SEM (P: parental, SKM: Skeletal muscle differentiation culture).

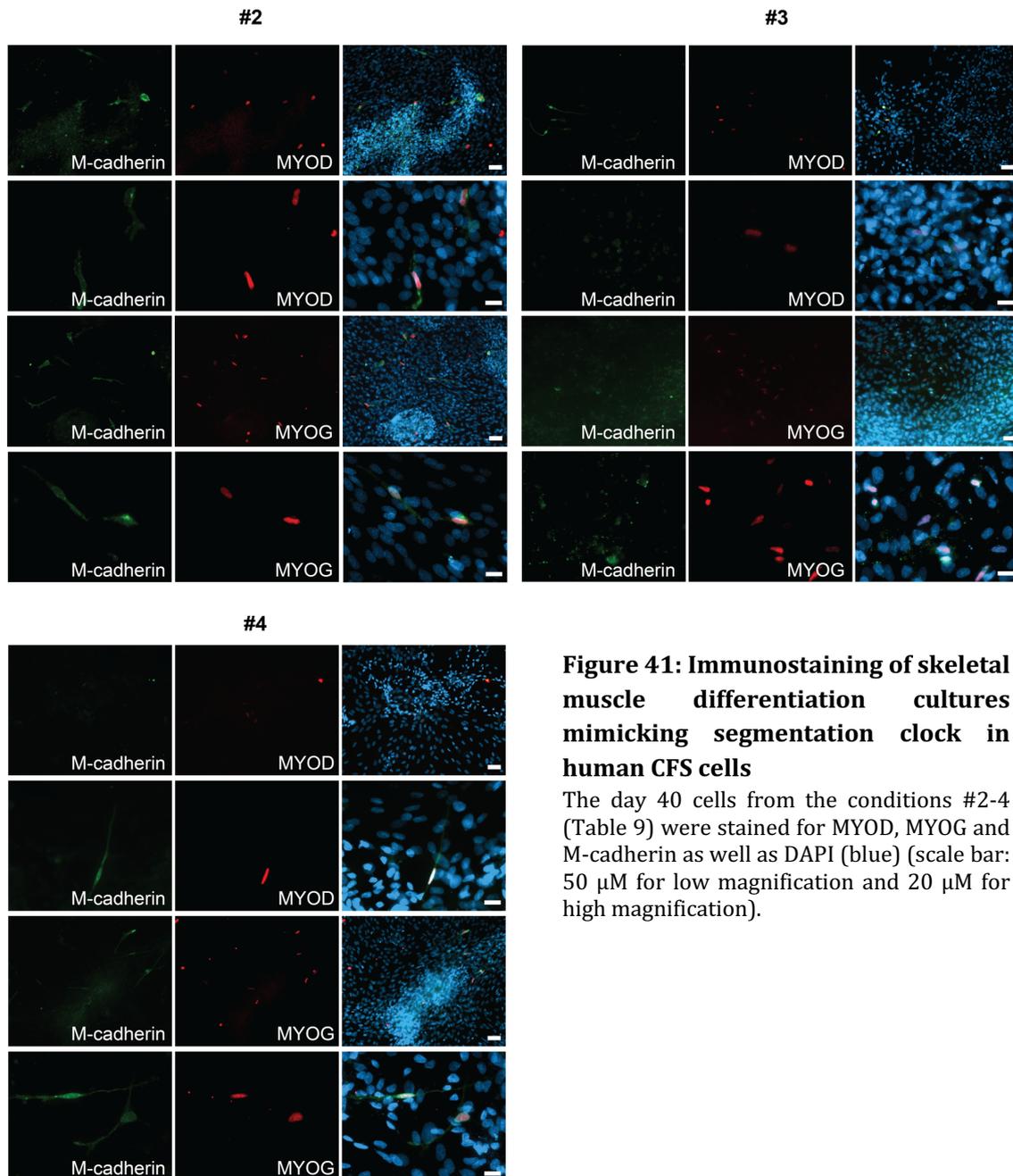


Figure 41: Immunostaining of skeletal muscle differentiation cultures mimicking segmentation clock in human CFS cells

The day 40 cells from the conditions #2-4 (Table 9) were stained for MYOD, MYOG and M-cadherin as well as DAPI (blue) (scale bar: 50 μ M for low magnification and 20 μ M for high magnification).

Next, the cells from the conditions #2-4 were analyzed for MYOD, MYOG and M-cadherin, which is a muscle-specific cytoskeletal marker. The selected markers were detected in all cultures, but the immunostaining results showed that the differentiation efficiency was very low as only a few cells were marked by MYOD or MYOG expression (**Fig. 41**). High background signal in M-cadherin staining was detected. Images of IgG control staining are shown in Supp. Fig. 4.

I investigated differentiation from CS cells by applying DAPT only or DAPT with SHH (**Fig. 39A**) because these conditions led to upregulation of *MYOG* and *MYOD* in CFS differentiation cultures (**Fig. 39B** and **Fig. 41**). The experiment was performed by using two parental CS lines (CS_1-2) and repeated two times independently. The differentiating cells exhibited an explicit neural morphology starting from day 2 and they detached after day 3 (**Fig. 42**).

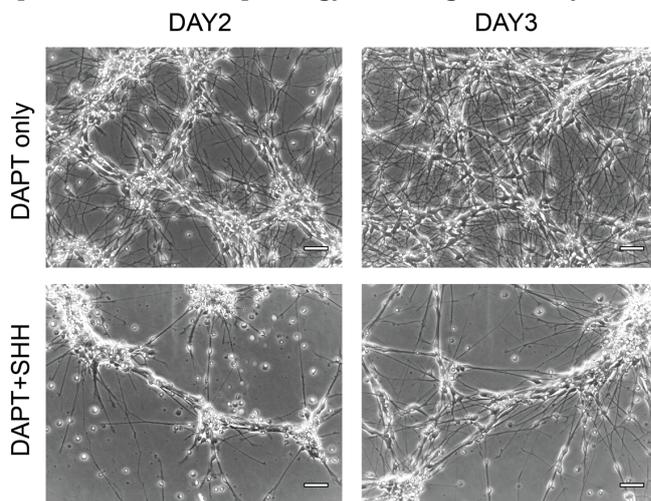


Figure 42: The effect of Notch inhibition with or without SHH activation in human CS cells

Morphology of the cells on day 2 and day 3 after treatment of DAPT (Notch inhibitor), and DAPT with SHH (scale bar: 50 μ M).

3.7.2.2. Skeletal muscle differentiation targeting axial elongation-associated pathways

The previously tested differentiation modalities, mimicking the segmentation clock resulted in low efficiency, I sought to examine different methods for recapitulating step-wise axial elongation. Mavrommatis et al.²⁵² established a hiPSC derived skeletal muscle organoid protocol. They demonstrated a transient NMP population during the differentiation. I applied their model with minor modifications suited to AxSC differentiation on 2D culture. I skipped the first 5 days of the differentiation that involved WNT/FGF activation and BMP inhibition. This is the stage where the NMP population is observed in the original protocol. I treated AxSCs with CHIR/FGF2/LDN/RA for 2 days which was followed by FGF2 removal and SHH addition for the next 4 days (**Fig. 43A**). From day 6 until day 12, the cells were treated with FGF2 and HGF, and split on day 10. Next, only HGF was used until day 35-38 timepoint, when the differentiation was stopped. The experimental set up was repeated by using all three CFS

lines (**Fig. 10**) independently. The differentiation cultures were analyzed by performing RT-qPCR experiments to assess expression of the stage-specific mesodermal markers. The results showed that CFS and early mesoderm markers *TBXT* and *CDX2*, and somatic mesoderm markers *TBX6* and *MSGN* were downregulated at the experimental endpoint with the exception of *MSGN* expression for one of the CFS-2 replicate (**Fig. 43B**). *MYOD* and *MYOG* were abundantly expressed in the differentiating cells except for CFS-3 derivatives that exhibited low *MYOD* and lower *MYOG* expression compared to CFS-1 and CFS-2 derivatives. *CDH15*, encoding for M-cadherin, was highly upregulated in all differentiation cultures. Expression of the selected marker genes was validated by RT-qPCR in primary human skeletal muscle cells which were cultured growth medium with or without HS (Supp. Fig. 5).

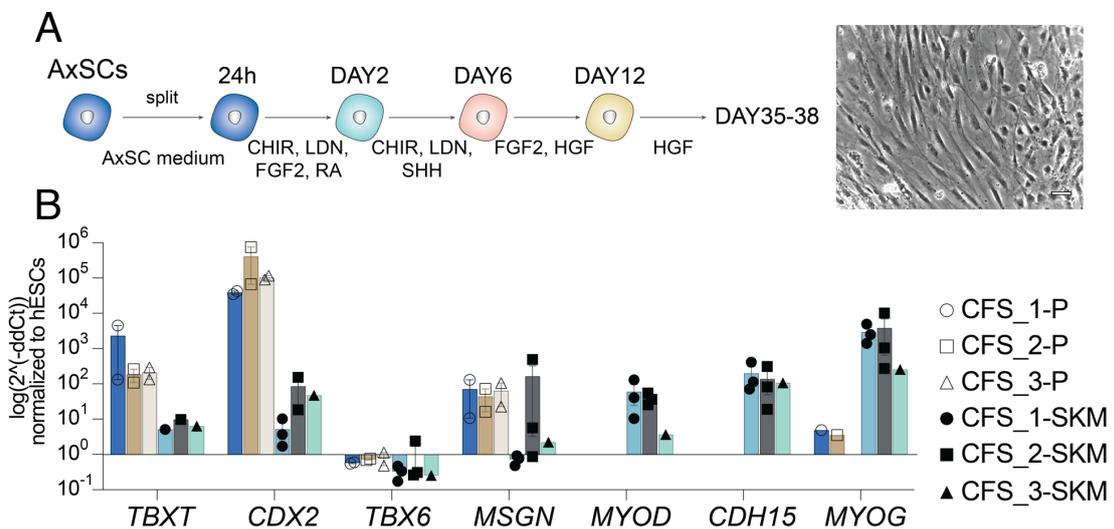


Figure 43: Characterization of skeletal muscle differentiation cultures from human CFS by mimicking axial elongation steps

(A) Schematic illustration of the differentiation protocol (left) and morphology of the CFS differentiating cells on day 35 (right) (scale bar: 50 μ M).

(B) Expression analysis of stage-specific mesodermal markers by RT-qPCR on day 35-38 differentiating cells. The experiment was conducted by using CFS_1-3 lines (Fig. 10). Technical replicates are indicated by the same symbols. The Ct values were normalized to undifferentiated H9 cells. Error bars represent SEM (P: parental, SKM: Skeletal muscle differentiation culture).

The cells were stained for MYOG, MYOD and M-cadherin. The results confirmed the expression of all three markers (**Fig. 44A**). Multinucleated cells that express MYHC and MYOG could be identified in the stained culture (**Fig. 44B**). IgG control staining results are shown in Supp. Fig. 6.

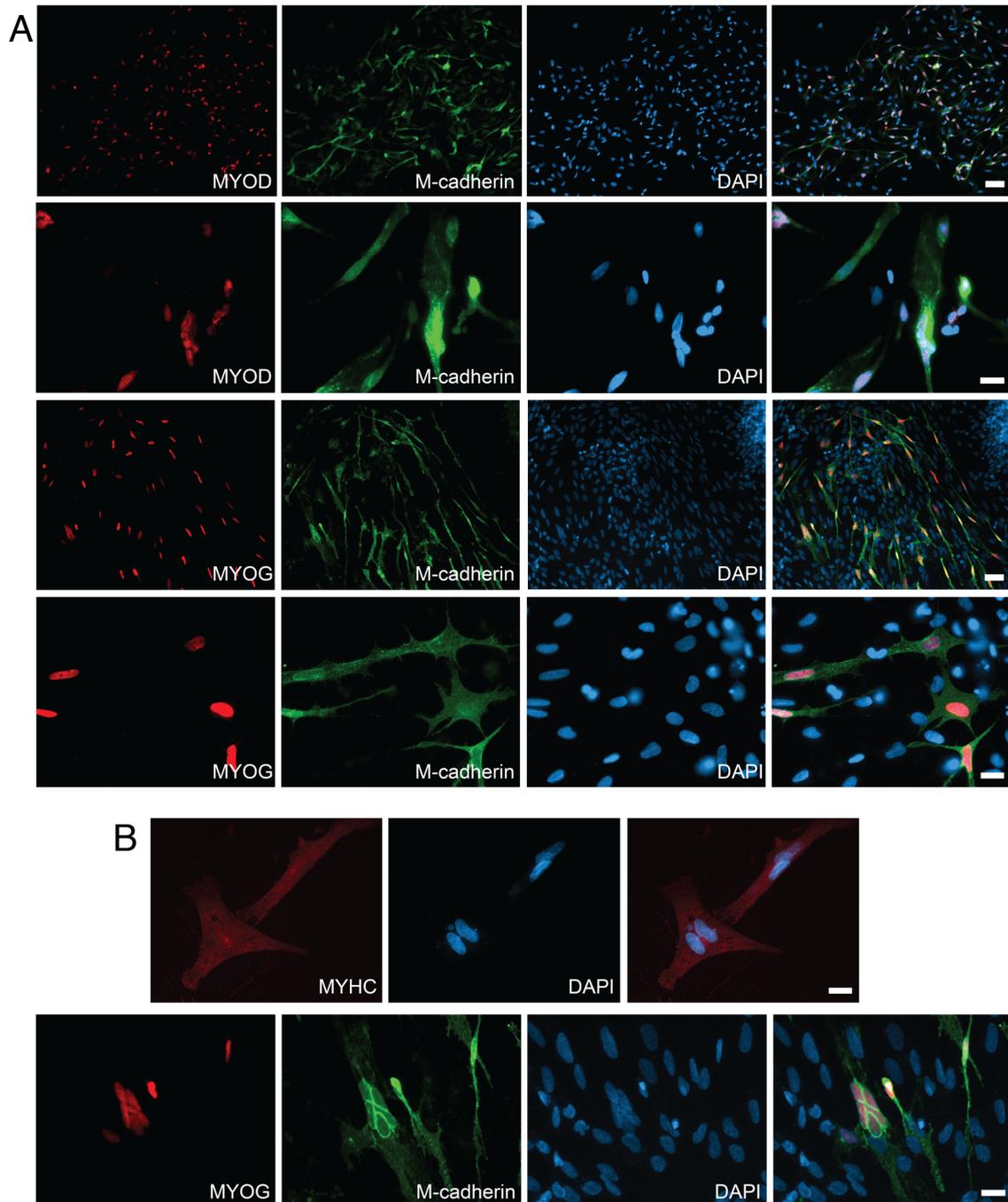


Figure 44: Immunostaining of skeletal muscle differentiation cultures from human CFS cells by mimicking axial elongation

(A) Representative images for MYOD, MYOG and M-cadherin expressing cells on day 35 (scale bar: 50 μM for low magnification and 20 μM for high magnification)

(B) Representative images for multinucleated cells stained for MYHC, MYOG, M-cadherin and DAPI on day 35 (scale bar: 20 μM)

I applied the same differentiation method (**Fig. 43A**) to CS cells. The experiment was conducted by using two CS lines (CS_1-2) and two independent replicates. The cells from all differentiation cultures acquired neural morphology within one week and they detached at various timepoints. Only one of the cultures could be maintained until day 35 (**Fig. 45A**) and evaluated by RT-qPCR experiment. Despite being able to detect MYOG upregulation (**Fig. 45B**), cumulative expression of MYOD/MYOG/M-cadherin was not detected, which was in accordance with the morphological observations. Skeletal muscle cells were obtained from CFS but not CS cells.

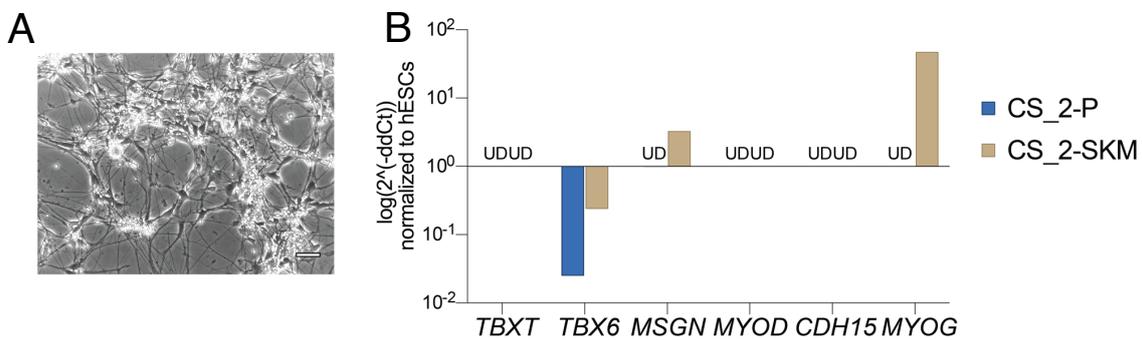


Figure 45: Characterization of skeletal muscle differentiation culture from human CS by mimicking axial elongation steps

(A) Morphology of the CS_2 differentiating cells on day 35 (scale bar: 50 μ M).

(B) Expression analysis in day 35 differentiation culture from CS_2 cells by RT-qPCR targeting mesoderm development-associated genes. The Ct values were normalized to undifferentiated H9 cells (P: Parental, SKM: Skeletal muscle differentiation culture, UD: Undetermined).

3.7.2.3. Differentiation into intermediate and lateral mesoderm derivatives

A recent study suggested that NMPs are able to generate nephric mesenchyme which originates from intermediate mesoderm⁹⁴. I sought to generate nephron cells from both AxSC lines by using a well-established protocol for hPSC differentiation²⁶⁹. First, I applied the original protocol, however the results showed that none of the nephron development-associated genes were upregulated (Supp. Fig. 7A). I adapted the protocol to AxSCs by omitting the 4-day WNT induction and BMP inhibition (**Fig. 46A**). The differentiation cultures were analyzed on day 15 by performing RT-qPCR experiment. Mesoderm commitment (*TBX6*), posterior intermediate mesoderm (*WT1*, *OSR1*, *HOXD11*), metanephric mesenchyme (*SIX2*, *SALL1*), pretubular and renal vesicle (*PAX8*, *LHX1*) markers were highly upregulated in both CFS and CS differentiating cells (**Fig. 46B**). Parental AxSCs exhibited low levels of those markers except for *HOXD11*. To investigate expression trend of the abovementioned genes in detail, I analyzed the differentiation cultures time course (Supp. Fig. 7B). I found that the marker genes were detected at the highest expression on day 11

during CFS differentiation then they were downregulated, while for the CS differentiation the selected markers showed the highest expression on the end time point (day 15).

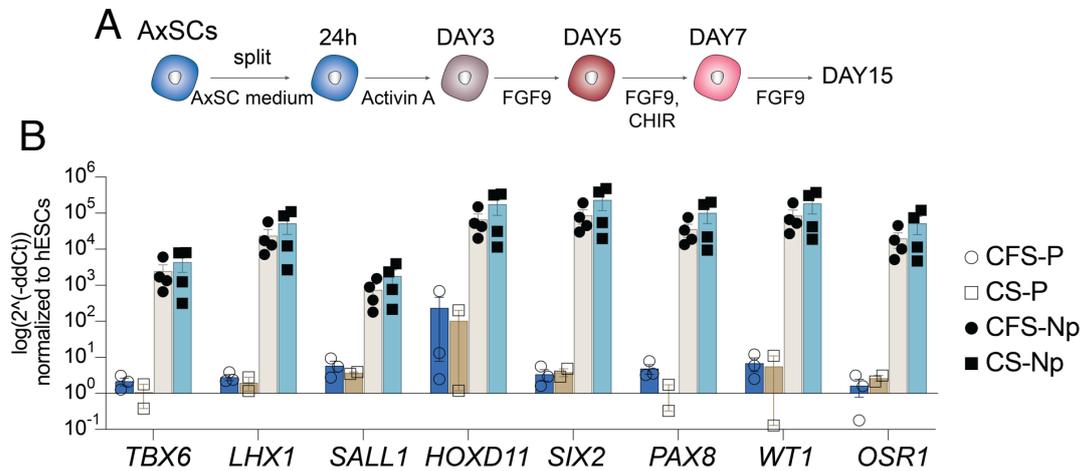


Figure 46: Characterization of nephron differentiation cultures from human CFS and CS cells

(A) Schematic illustration of the nephron differentiation protocol.

(B) Expression analysis of nephron development-associated markers in day 15 differentiation culture from CFS_2 and CS_2 lines (Fig. 10). Technical replicates are indicated by the same symbols. The Ct values were normalized to undifferentiated H9 cells. Error bars represent SEM (P: parental, Np: Nephron differentiation culture).

I also investigated the potential of AxSCs towards lateral mesoderm by performing cardiomyocyte differentiation. For this purpose, I utilized the protocol established by Lian et al.²⁷⁰ (**Fig. 47A**) and tested the effect of cell numbers for the starting parental population. Regardless of the seeding density, contracting cardiomyocytes could not be detected during the differentiation from neither CFS nor CS cells. The cultures were analyzed on day 15 by RT-qPCR for the expression of cardiomyocyte markers *MYOCD* and *TNNT2* (**Fig. 47B**). A slight upregulation of *MYOCD* was observed in CFS differentiating cells, but *TNNT2* levels were downregulated. Neither *MYOCD* nor *TNNT2* did not show elevated expression during CS differentiation. There was no indication based on the expression analysis that the cardiac progeny could be derived from AxSCs.

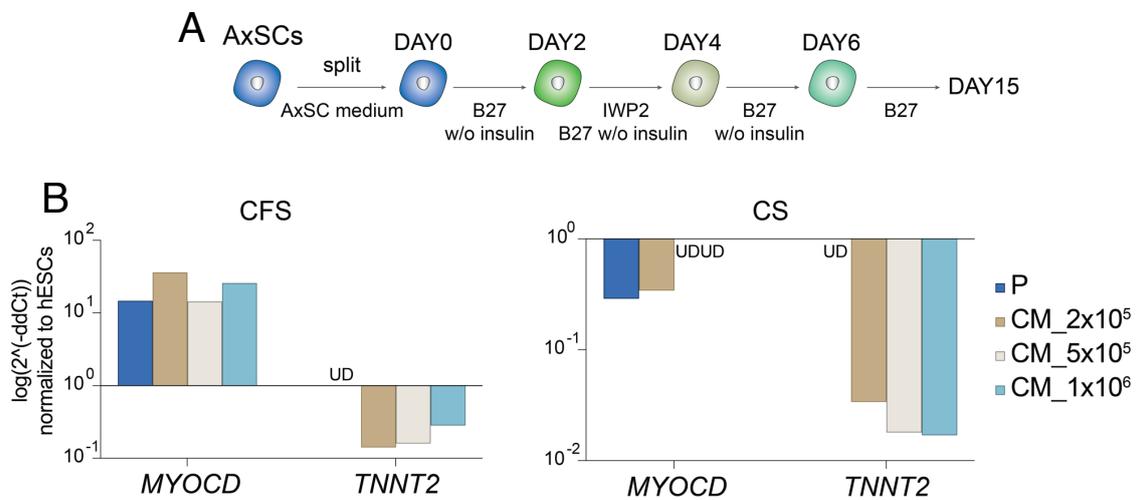


Figure 47: Characterization of cardiomyocyte differentiation from CFS and CS AxSCs

(A) Schematic illustration of the cardiomyocyte differentiation protocol.

(B) Expression analysis of cardiomyocyte markers *MYOCD* and *TNNT2* in day 15 differentiation cultures from CFS_2 and CS_2 lines (Fig. 10). The Ct values were normalized to undifferentiated H9 cells. Three independent experiments were conducted by seeding 2x10⁵, 5x10⁵, 1x10⁶ parental cells (P: parental, CM: Cardiomyocyte differentiation culture).

3.8. Reproducibility of axial stem cells in species

3.8.1. Generation of axial stem cells from mouse embryonic tissues

Throughout this study I have employed human stem cell lines, albeit obtained from different backgrounds, to generate AxSC. Lastly, I sought to generate AxSCs directly from embryonic tissues. For this purpose, 2 CD1 mice were dissected at E8.5 in collaboration with Dr. Silvia Schirge (Helmholtz Center Munich, Germany). The stem zone region from 29 embryos was collected (**Fig. 48A**). The embryos exhibited different somite stages (SS) from 2 to 10 somites, thus the dissected tissue from 3-4, 5-6, 7-8, 9-10 SS embryos were pooled together every other somite. After dissociation, each sample was divided into two in order to establish embryonic outgrowth using the respective AxSC medium (CFS/CS) on Matrigel-coated plates. The cells were split by clump passaging method as they were sensitive to single-cell split. Early 2-6 SS cells could not be long term maintained in neither CFS nor CS medium, similarly to 7-8 SS cells in CS medium.

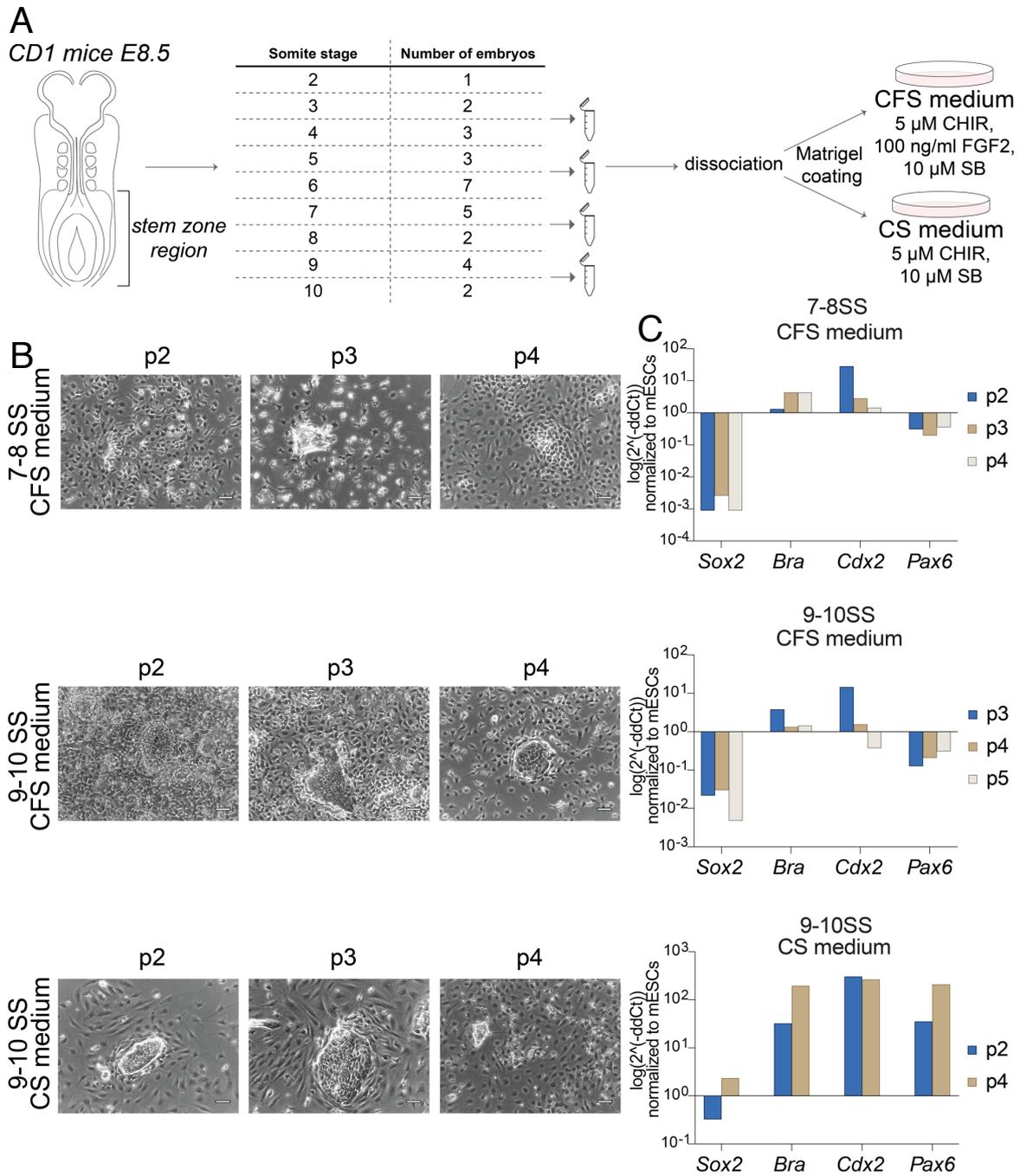


Figure 48: Characterization of the AxSC cultures derived from mouse embryonic tissues

(A) Diagram showing the embryo dissection modality. Stem zone region of 29 CD1 embryos at E8.5 were dissected and embryonic tissues were pooled based on somite stage (SS) of the embryos. After dissociation, the cells were transferred onto Matrigel-coated plates in CFS or CS medium.

(B) Morphology of p2, p3 and p4 cells in CFS medium derived from 7-8 and 9-10 SS cells, and CS medium derived from 9-10 SS cells (scale bar: 50 μ M, p: passage number).

(C) Expression analysis of AxSC markers in the cultures shown in panel B. The Ct values were normalized to undifferentiated mESC cells (p: passage number).

Stem cell-like colonies were observed in 7-8 and 9-10 SS CFS establishment cultures whereas only in 9-10 SS cells for CS condition (**Fig. 48B**). All cultures showed high heterogeneity throughout the establishment process in terms of cellular populations due to presence of morphologically distinguishable flat-shaped mesenchymal-like cells. The number of stem-cell like cells gradually decreased over the time and no stem cell like colonies could not be observed after p4/p5. The expression of AxSC markers was analyzed in a time course experiment from p2 to p4/p5 by performing RT-qPCR (**Fig. 48C**). In 7-8 SS CFS culture, *Pax6* and *Sox2* were downregulated from p2 to p4. *Brachyury* (*Bra* known as *TBXT* ortholog in mouse) was slightly upregulated and *Cdx2* exhibited a gradual downregulation. Similar trends were determined in 9-10 SS CFS culture, yet *Bra* was also downregulated with the increased cell passaging. 9-10 SS CS culture exhibited abundant *Pax6* expression, however, an upregulation in both *Bra* and *Cdx2* was also detected. Neither CFS nor CS establishment cultures exhibited a transcriptional pattern that falls in line with in vitro derived human AxSCs.

3.8.2. Generation of axial stem cells from pluripotent stem cells

Taking into account the results with the mouse embryonic cultures, I sought to optimize the mouse AxSC culture conditions by performing a screening of several small molecules using E14 mESC-based EpiSCs¹⁷⁴. EpiSCs already have heterogeneous *Brachyury* expression as it has been shown in the literature³¹² and in this study (Supp. Fig. 8), the 24-hour CHIR pulse which was previously applied to generate human AxSC lines was not implemented for the establishment of mouse lines. EpiSCs were directly plated in respective AxSC medium on Matrigel-coated plates likewise human cells (**Fig. 49A**). As the human culture conditions did not result in successful derivation of AxSC lines from embryonic tissues, I screened various conditions by increasing or decreasing the cytokine concentrations (**Fig. 49B**). I tested additionally the effect of FGF8, as it has been shown to be an important FGF ligand in mouse development likewise FGF2 for human development³¹³. Among the conditions, only the cells treated with 3 μ M CHIR99021, 50 ng/ml FGF2 and 10 μ M SB431542 (marked by #1) could be maintained in culture for over 10 passages thus designated as a successful derivation. The cells were split two to three times per week depending on their confluency. Colony formation was observed in culture, but flat-shaped cells were detected as shown in **Fig. 49C**.

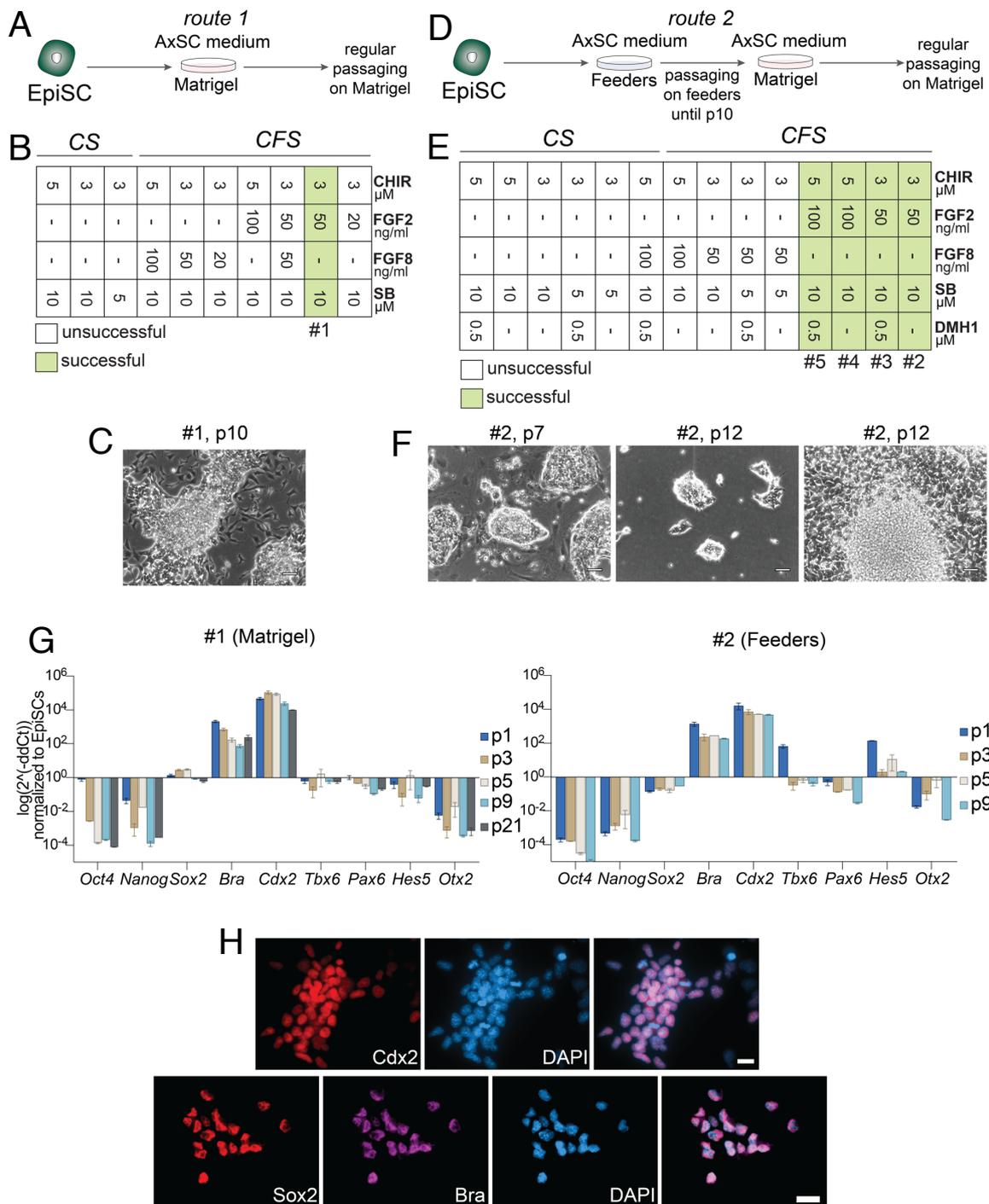


Figure 49: Establishment and characterization of mouse AxSCs from E14 mESC-based EpiSCs
 (A) Schematic illustration of the route 1 establishment protocol.
 (B) The conditions screened by following the protocol in panel A that targets mouse CFS and CS derivation. The successful condition (#1) is shown in green squares.

- (C) Morphology of the p10 cells treated with the cytokines marked by #1 (scale bar: 50 μ M, p: passage number).
(D) Schematic illustration of the route 2 establishment protocol.
(E) The conditions screened by following the protocol in panel D that targets mouse CFS and CS derivation. The successful conditions (#2-5) are shown in green squares.
(F) Morphology of the p7 and p12 cells treated with the cytokines marked by #2. The images from p12 culture were taken 1 day (left) and 3 days (right) after split (scale bars: 50 μ M, p: passage number).
(G) Expression analysis of pluripotency and AxSC markers in the cultures from #1 and #2 conditions by RT-qPCR. The Ct values were normalized to undifferentiated EpiSC cells. Error bars represent SEM (p: passage number).
(H) Immunostaining of the cells from condition #1 for Cdx2, Sox2, Bra and DAPI (scale bar: 20 μ M).

Due to the partial heterogeneity, I employed a different EpiSCs derivation modality by maintaining the cells on mitotically inactivated mouse embryonic fibroblasts (MEFs) until p10 and then switched to feeder free cultures using Matrigel-coating (**Fig. 49D**). I monitored the various conditions and I included DMH1 addition to test the effect of BMP inhibition (**Fig. 49D**). The conditions marked #2-5, that include FGF2 thus potentially resemble CFS state, were considered as successful derivations, as the cells could be kept in culture up to p10 and they were adapted to feeder free conditions. None of the cytokine cocktails without FGF2, presumably designated to obtain CS-like cells, resulted in the generation of stable lines. Cellular morphology generated by applying the same cytokine concentration as for the successful derivation on Matrigel (**Fig. 49C**) is shown in **Fig. 49F**. The AxSC-like colonies were observed on both feeders (p7) and Matrigel (p12). Flat-shaped cells were not detected first day post passaging, but they propagated over time with increasing culture confluency, similarly to the previous derivation attempt on Matrigel only (**Fig. 49C**). Next, I evaluated the derivation cultures for expression of AxSC markers by conducting RT-qPCR experiment (**Fig. 49G**). The cells generated by two modalities either on Matrigel or feeder/Matrigel combination exhibited downregulation of pluripotency markers *Oct4* and *Nanog*. *Sox2* was expressed in a similar level comparing to EpiSCs. *Bra* and *Cdx2* were upregulated while *Tbx6* and *Pax6* were downregulated and *Hes5* did not show notable expression. *Sox2*, *Bra* and *Cdx2* were also tested on the protein level by performing immunostaining experiment using the #1 condition cells (**Fig. 49H**). The results from IgG control staining are given in Supp. Fig. 9.

Next, I examined the mouse axial differentiation cultures from #3-5 conditions which were morphologically indistinguishable from #1 cells due to presence of stem cell-like colonies and flat-shaped cells together (**Fig. 50A**). The RT-qPCR analysis indicated similar transcriptional trend for these conditions. *Oct4* and *Nanog* were downregulated, and *Sox2* expression was not significantly changed compared to EpiSCs (**Fig. 50B**). *Bra* and *Cdx2* were upregulated, but *Tbx6* expression did not increase at any time points. A gradual upregulation in *Pax6* and *Hes5* was detected by the addition of DMH1 to low levels of CHIR99021 (3 μ M) and FGF2 (50 ng/ml) (#3). *Hes5* was also slightly upregulated in the culture which was

treated by high levels of CHIR99021 (5 μ M) and FGF2 (100 ng/ml) with DMH1 (#5), but *Pax6* did not show considerable expression changes for the conditions with or without DMH1 in addition to high CHIR99021 and FGF2 (#5 and #4 respectively). These findings indicated that only CFS like cells could be derived from EpiSCs. BMP inhibition did not result in significant effect on either the morphology or expression of the key AxSC markers and it could not enhance the culture conditions to obtain CS like cells.

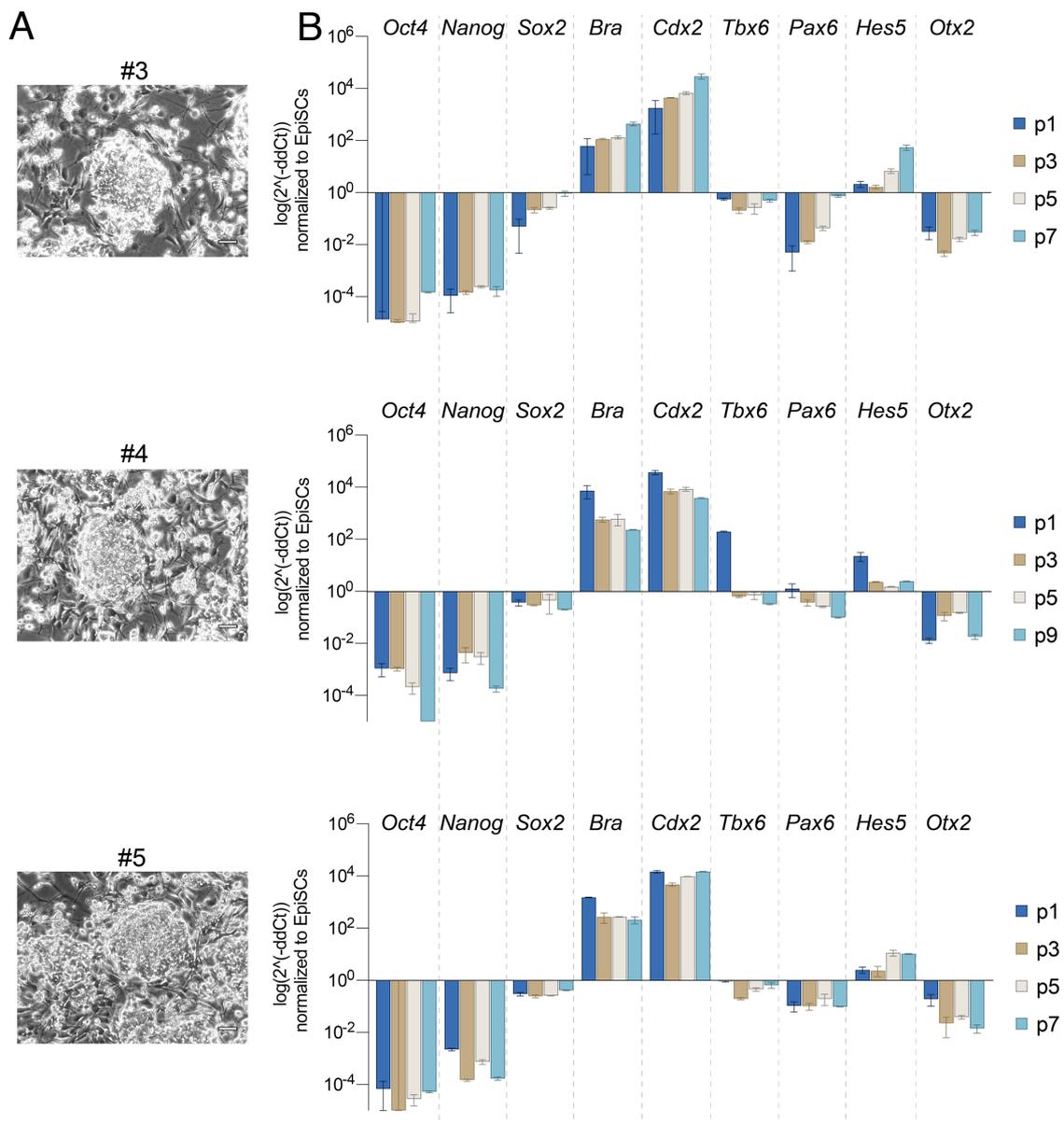


Figure 50: The effect of BMP inhibition and modulation of cytokine concentration on mouse AxSC establishment from E14 mESC-based EpiSCs

(A) Representative brightfield images showing the morphology of the cells (passage 8) from the conditions #3-5 (Fig. 49D-E) (scale bars: 50 μ M).
(B) Expression analysis of pluripotency and AxSC markers in the cultures shown in panel A by RT-qPCR. The Ct values were normalized to undifferentiated EpiSCs. Error bars represent SEM (p: passage number).

The CS AxSC state could not be reproduced in mouse, thus I questioned whether it is due to developmental differences between human and mouse, or failure to identify needed culture conditions for the establishment protocol. To this end, I sought to generate CS state from another species and opted for *Pongo abelii* Sumatran orangutan iPSCs (oriPSCs) clone 70Af1²⁶⁶. The oriPSCs were sensitive to single-cell split, thus I regularly split them as clumps and applied the 24-hour CHIR pulse when the cells reached approximately 90% confluency. I treated the cells with 5 or 10 μ M CHIR for the first 24 hours (**Fig. 51A**). After 24 hours, half of the induced cells were collected to test expression of stem zone/NMP markers and the rest was plated within CS medium on Matrigel. Expression analysis showed that *TBXT* and *CDX2* were highly upregulated regardless of the CHIR concentration, although slightly higher expression levels were detected in the 10 μ M treated cells (**Fig. 51B**). *SOX2* expression exhibited the reverse effect to the different CHIR concentrations. In the light of *TBXT/CDX2* and *SOX2* expression pattern, I carried on the establishment of putative CS lines from both 5 and 10 μ M induced cells by treatment with 5 μ M CHIR and 10 μ M SB, similarly to the human conditions. The cells were regularly split as clumps twice per week and maintained in hypoxic conditions. No morphological differences could be observed between the lines generated from 5 or 10 μ M induced cells (**Fig. 51C**). I conducted two independent derivations per line from 5 or 10 μ M induced cells and analyzed the expression of AxSC markers for p5 and p10 cultures by RT-qPCR (**Fig. 51D**). The results showed that *OCT4*, *NANOG* and *TBX6* were downregulated in both p5 and p10 cells. *SOX2* expression was detected in similar levels as compared to oriPSCs. *TBXT* was slightly upregulated. *CDX2* expression gradually decreased over the time. *PAX6* exhibited low expression at p5 however it was abundantly expressed in the p10 cells similarly to *HES5*. Immunostaining of p10 cells confirmed *SOX2* and *PAX6* expression at protein level while *TBXT* and *CDX2* were not found regardless of the CHIR concentration used for 24-hour induction before the long-term maintenance of stable cell lines (**Fig. 51E**). The results from IgG control staining are given in Supp. Fig. 10.

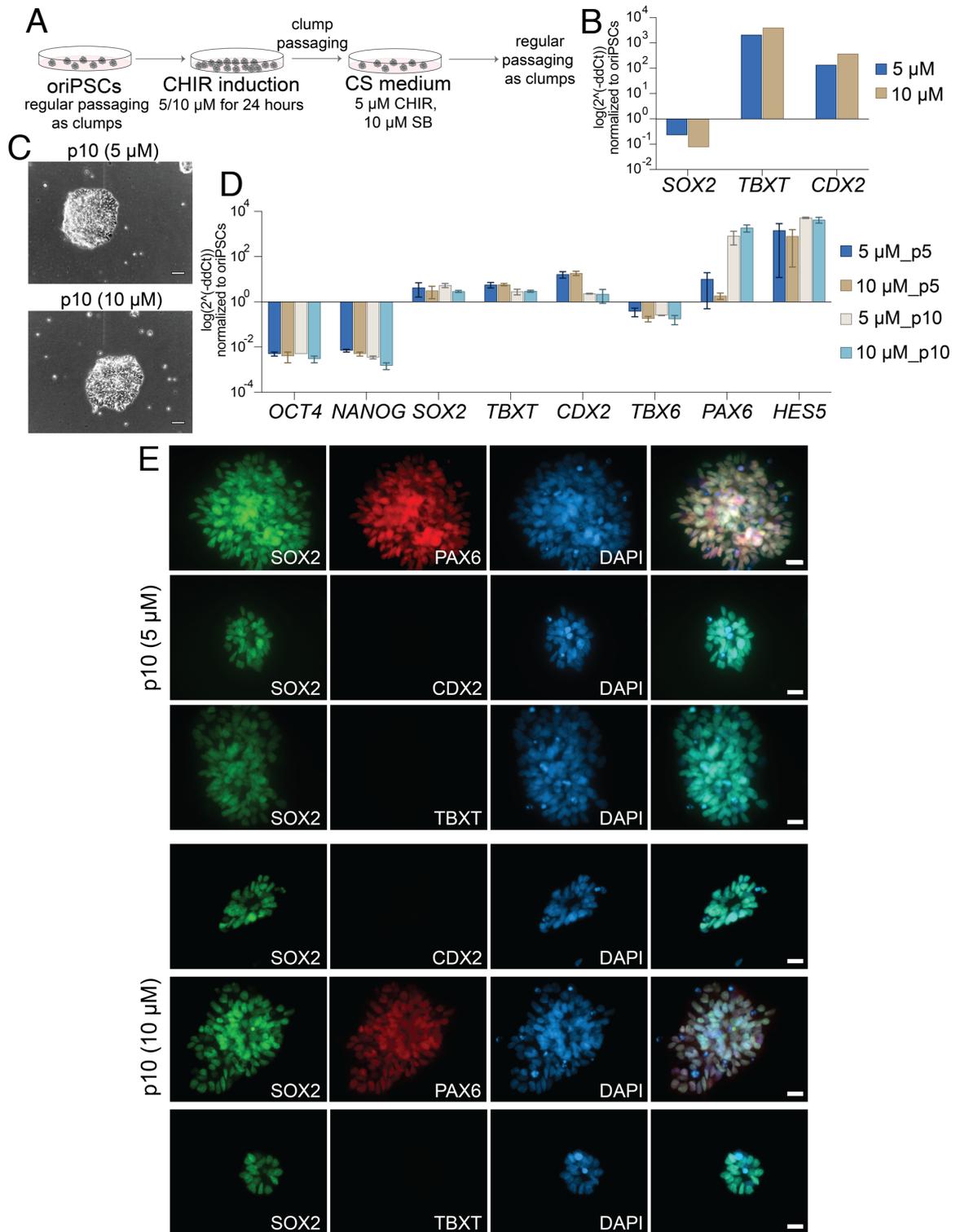


Figure 51: Establishment and characterization of CS lines from orangutan iPSCs

(A) Schematic illustration of the establishment process. Parental Sumatran orangutan iPSCs (oriPSCs) were regularly split as clumps and induced by 5 and 10 μ M CHIR99021 (CHIR) for 24 hours when they reached approximately 90% confluency. The induced cells were split as clumps in CS medium comprising 5 μ M CHIR99021 and 10 μ M SB431542 (SB). Clump passaging was routinely applied to the cells.

(B) Expression analysis of *SOX2*, *TBXT* and *CDX2* in the 24-hour induced cells by 5 or 10 μ M CHIR99021. The Ct values were normalized to uninduced oriPSCs.

(C) Representative brightfield images for the p10 cultures from 5 and 10 μ M induced cells (scale bars: 50 μ M, p: passage number).

(D) Expression analysis of pluripotency and AxSC markers in the 5 and 10 μ M induced cells at p5 and p10 after maintenance within CS medium. The derivation was conducted two times independently per line from either 5 or 10 μ M induced cells. Error bars represent SEM. The Ct values were normalized to undifferentiated oriPSCs (p: passage number).

(E) Representative immunostaining images for the lines at p10 stained for *SOX2*, *TBXT*, *CDX2*, *PAX6* and DAPI (scale bar: 20 μ M, p: passage number).

3.8.3. Generation of CFS cells from mouse embryonic tissue

After successful culture condition optimization for the derivation of mouse axial stem cells by using EpiSCs, I repeated establishment mouse CFS cells from mouse embryos. Dissection of *Foxa2VenusFusion* x *mTmG* embryos was performed in collaboration with Dr. Silvia Schirge (Helmholtz Center Munich, Germany). The stem zone region of E8.5 and E9.0 embryos were collected and seeded on Matrigel-coated plates in mouse CFS medium (3 μ M CHIR, 50 ng/ml FGF2 and 10 μ M SB) after dissociation (**Fig. 52A**). Among the various somite stages, 7/8 and 16/17 SS cells were stabilized, and I was able to culture them for a prolonged time in vitro. I observed heterogeneity in both cultures up to passage 7 (**Fig. 52B**), thus I tested culturing them on Matrigel and MEFs in parallel (**Fig. 52A**). The usage of feeders did not eliminate heterogeneity in culture. The 16/17 SS cells exhibited more homogeneous morphology with continued culturing on Matrigel (**Fig. 52B**). The 7/8 and 16/17 SS cells showcased distinct morphology. I noted that 16/17 cells were more similar to the mouse CFS cells derived from epiblast stem cells, while 7/8 cells were more similar to epiblast stem cel.

To evaluate whether the cells represent CFS-like transcriptional pattern, I analyzed expression of pluripotency, NMP, neuroectoderm and paraxial mesoderm markers by performing RT-qPCR (**Fig. 52C**). The results showed that *Oct4* and *Nanog* pluripotency markers were downregulated in all cultures, while *Sox2* was expressed similarly to EpiSCs. *Bra*, *Cdx2* and *Nkx1-2*, which are NMP/CFS markers, were abundantly expressed in both 7/8 SS and 16/17 SS cells regardless of the coating solution. *Mnx1*, a late NMP and CFS marker, was downregulated in 7/8 SS cells but it showcased a slight upregulation for 16/17 SS cells. Neuroectoderm markers *Pax6* and *Sox1* did not show a considerable expression, but *Hes5* expression was found in 7/8 SS cells. Expression of paraxial mesoderm marker *Tbx6* did not change in any culture conditions when compared to EpiSCs and expression of *Otx2*, which is an anterior development marker, was highly downregulated.

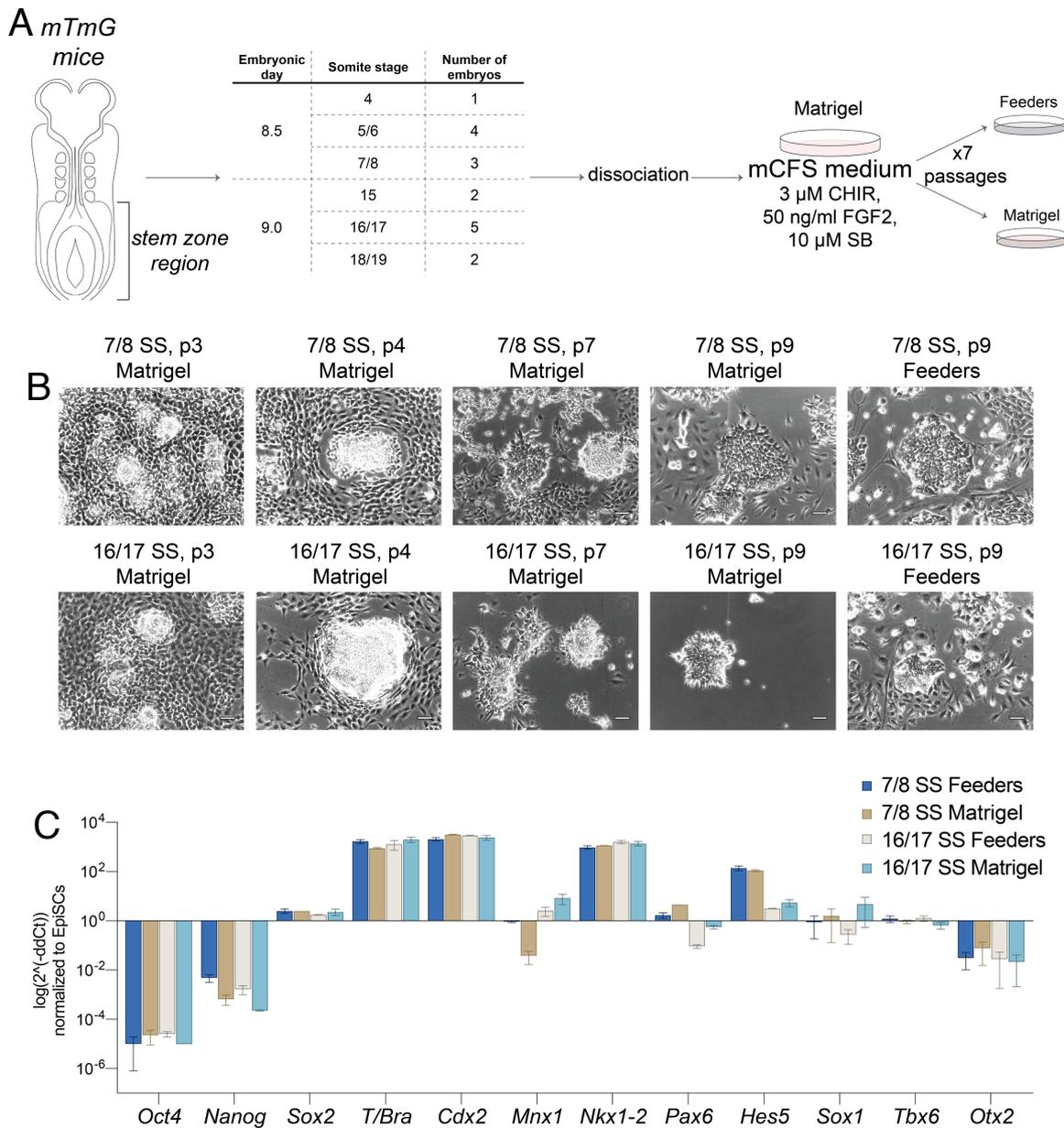


Figure 52: Establishment and characterization of CFS cells from mouse embryos

(A) Schematic illustration of dissection and establishment protocol. Stem zone region of the embryos at E8.5 or E9.0 were dissected. After dissociation, the cells were cultured on Matrigel-coated plates in mouse CFS medium. After passage 7, cells were plated on either Matrigel-coated plates or feeders.

(B) Morphology of the cells at p3, p4, p7 and p9 (scale bar: 50 μ M, p: passage number).

(C) RT-qPCR results showing expression of pluripotency, AxSCs and paraxial mesoderm markers in the p10 cultures derived from 7/8 and 16/17 SS cells grown on either Matrigel or feeders. The Ct values were normalized to EpiSC cells. Error bars represent SEM.

4. DISCUSSION

Posterior embryonic development has been distinguished from anterior regions by means of morphological observations and molecular analysis^{83,263,277}. Unique posterior progenitors populations have been identified through lineage tracing in the embryo and they are involved in posterior/axial elongation, thus named axial progenitors¹⁵. Growing evidence highlighted that there are different types of axial progenitors, but the one captivating the interest of many researchers is the NMP progenitor type. The NMPs harbor the developmental capacity to differentiate into both neural and mesodermal lineage, thus contradicting traditional gastrulation theory^{21,22}. Over the years, many studies have been dedicated to deciphering the molecular hallmarks and progeny profile of *in vivo* and *in vitro* NMPs. The most challenging aspect of NMP-associated studies is the lack of protocols resulting in long term NMP cultures. Prolonged expansion of NMP cells *in vitro* induces spontaneous differentiation. This holdback could be overcome with the generation of an *in vitro* NMP stem cell counterpart, which is what I set out to research during my doctoral studies.

4.1. Establishment of axial stem cells relies on mimicking posterior embryonic development step-wise

The protocols established for NMP generation *in vitro* from either human or mouse PSCs comprise activation of FGF and WNT signaling pathways to imitate the molecular basis of posterior growth zone, or also known as the stem zone where NMPs are located during early embryonic development. I employed this modality to generate long-term equivalent of NMPs from human PSCs. Unlike the published protocols, hESCs (H9) were first induced for differentiation only by using CHIR99021 (WNT pathway activator) to posteriorize and to induce primitive streak formation (**Fig. 9A**). Higher CHIR99021 concentration (10 μ M) was utilized for the initial stimulation compared to maintenance medium (5 μ M), because the highest TBXT and lowest SOX2 expression were determined for the 10 μ M CHIR99021 induction in the range of 0-10 μ M⁵⁰. This expression pattern for TBXT and SOX2 is the molecular feature of most-posterior region of the stem zone²⁰. The CHIR99021 pulse performed in this study might be one of the key factors linked to the achievement of stable long term cultures, as it has been known that FGF and WNT pathway co-activation steers NMPs into mesodermal fate⁴⁴. The resulting mesodermal fate could be attributed to the involvement of the FGF pathway in Notch pathway regulation during the embryo segmentation events³⁸. The second important factor that results in stable cultures could be the applied concentration of the cytokines used in the maintenance medium: 5 μ M CHIR99021 and 100 ng/ml FGF2. Many of the NMP generation protocols utilize 3 μ M CHIR99032 and 20 ng/ml FGF2^{44,234,235,239}. In addition to WNT/FGF activation, TGF β pathway was blocked throughout the cell maintenance by using SB431542, because it has

been shown that TGF β inhibition promotes reprogramming of somatic cells towards the pluripotent state³¹⁴, thus it might have a role in enhancing self-renewing activity. NMPs have also been established by TGF β inhibition and WNT activation²³³, but the effect of TGF β inhibition on the NMP maintenance was not shown. The modulation of abovementioned pathways in hPSCs with or without FGF activation resulted in stable cultures for over 30 passages (**Fig. 9B**). The putative human AxSC populations were named CFS and CS, which was based on the nomenclature of the used cytokines for the line establishment (CHIR99021, FGF2, SB431542 and CHIR99021, SB431542 respectively). Robustness of the protocol and stemness/self-renewing capacity of the cell populations were proven by investigating the transcriptomic profile (**Fig. 10**) and telomere length (**Fig. 12**) over time in three independent derivations per cell line (CFS_1-3, CS_1-3). The researched avenues indicated no change over time, thus providing conclusive evidence towards the ability of the axial stem cells to grow over an extended time in vitro without changing their transcriptional identity or losing their stemness.

4.2. Novel AxSC states recapitulate consecutive phases of axial development

The transcriptional and proteomic analyses of human AxSC lines derived from H9 cells demonstrated that CFS state is the in vitro counterpart of NMPs while CS state is committed to the neural lineage (**Fig. 10-11, 13-19, 23**). This notion was further confirmed by progeny profiling in both states. CFS cells can be converted to CS but not vice versa inferring that CFS is the upstream state of CS (**Fig. 28**).

4.2.1. Human CFS cells exhibit the unique features of NMPs

NMPs are identified by SOX2 and TBXT coexpression. CDX2, which is one of the essential factors for axial elongation, is also expressed in the stem zone region and both in vivo and in vitro NMPs. I confirmed that human CFS cells coexpress *SOX2*, *TBXT* and *CDX2* (**Fig. 11, 13**), as well as *NKX1-2* (**Fig. 17A, 23B**) which was shown as another NMP marker⁵⁹. Another well-known characteristic of NMPs is expression of multiple HOX genes and a similar trend was also detected in CFS cells (**Fig. 17B, 23C**). It has been accepted that HOX genes are not expressed in ESCs and activated progressively in spatial/temporal fashion to confer positional cues³¹⁵. The transcriptomic and proteomic profile of CFS supports a contrary notion suggesting that all of the HOX genes are activated first and subsequently silenced as the cells differentiate further. CDX2 has been shown as a regulatory factor of HOX genes^{62,63} and it might be one of the underlying mechanisms for active HOX panel in CFS state. This needs to be further investigated by expression analysis of HOX and NMP-associated genes in CDX2 knockout lines. In mouse embryos, comparison of early and late NMPs corresponding to the regional transition from CLE/NSB to CNH pointed out that late NMPs exhibit

expression of more posterior HOX genes (PG 9-13) on the transcriptional level. I observed the same pattern in human CFS cells suggesting that these cells are the equivalent of late NMPs (**Fig. 17B**). I confirmed expression of MNX1 in CFS cells (**Fig. 17A, 23B**), which supports this notion as it has been suggested to be specifically expressed in late NMPs²².

In addition to NMP-associated genes, CFS cells express *EOMES* that plays a role in pluripotency exit³⁶, but not *POU5F1* or *NANOG* (**Fig. 13**, Table 7). These findings conclude that CFS cells cannot be considered pluripotent. In vitro NMP derivation cultures show *POU5F1* or *TBX6* expression that indicates the presence of pluripotent-like or mesodermal-committed cells respectively^{44,237,239}. Considerable expression of either transcription factors was not observed in CFS cells highlighting the homogeneity of these cultures and the efficiency of the establishment protocol. The unbiased analysis for differentially expressed proteins within global proteome profile supported equivalence of CFS cells to NMPs by indicating that CFS cells are able to differentiate into both neural and mesodermal lineage, as well as neural crest cells, as shown by GO analysis (**Fig. 23F**).

4.2.2. Human CS cells are committed to neural lineage

The additional AxSC state named CS sustained SOX2 expression, however, these cells did not show neither pluripotency markers (*POU5F1* and *NANOG*) nor NMP markers (*TBXT* or *CDX2*) (**Fig. 10, 13, 23**). These cells were characterized by *PAX6* expression which is an early neuroectodermal marker³¹⁶. This finding suggested a neural bias in CS state. The molecular features detected in human H9 derived CS cells were closely resembling the in vivo neural progenitors originating from NMPs⁶⁶ due to expression of *HES5*, *IRX3*, *IRX5* (**Fig. 15, 19**). These findings question whether CS cells represent the same population as conventional neural progenitor cells (NPCs), established via dual-SMAD ($TGF\beta$ and BMP) inhibition. Unlike NPCs, CS cells showcased distinguishable transcriptomic profiles in terms of key axial developmental genes such as *PAX3* and *PAX7*, ventral spinal cord progenitor marker *NKX6-1*, and anteroposterior demarcation factor *OTX2* (**Fig. 22A**). Another remarkable difference was the HOX profile showcased by CS cells (**Fig. 22B**). NPCs are devoid of both anterior and posterior HOX genes, while CS cells express HOX genes until PG 9. The restricted number of HOX genes expressed in CS cells compared to CFS cells supports the idea of HOX genes being progressively silenced during the developmental timeline, as CS cells have a more restricted developmental capacity as anticipated based on their molecular profile.

The neural bias was also identified by differentially expressed gene (DEG) (**Fig. 14**) and protein (DEP) (**Fig. 23F**) analyses within the transcriptome and proteome analysis of CS cells. In the literature, *ZFH3* has been shown as involved in both neural and mesodermal development²⁹¹⁻²⁹³, while *ZFH4* is specific to neural lineage²⁹⁴. CFS cells showed expression

of both *ZFH3* and *ZFH4*, but only *ZFH4* was expressed in CS cells, hinting neural bias (**Fig. 14**). DEP analysis supported this notion as neural development-associated processes were enriched in CS cells compared to CFS and H9 (**Fig. 23F**).

4.2.3. Human AxSCs share certain traits related to transcriptomic, metabolomic and epigenetic regulations

One of the most important shared features of the human AxSC states is the absence of *OTX2* expression that indicates the posterior/axial identity of these cells^{27,244}. This is also supported by the lack of *HES3* expression (**Fig. 19**), which is a potential indicator of anterior spinal cord as it has been shown to be expressed in forebrain, midbrain and hindbrain regions^{243,277} (**Fig. 16**). *SOX2*, which marks pluripotent cells, NMPs and NPCs based on the combinatorial expression with lineage specific transcription factors, is also a mutually shared characteristic of AxSCs. *FOXB1* and *ETV4* are enriched in both CFS and CS cells (**Fig. 18, 20B**). It has been reported in the literature that *FOXB1* is involved in hindbrain as well as dorsal spinal cord development in mice³¹⁷, and its expression is induced by FGF/WNT pathway co-activation as shown in *Xenopus* embryos³¹⁸. A preprint study highlighted that *FOXB1* upregulates *MSX2*, *PAX7* and *TFAP2B* expression in hESCs after 3-day WNT activation³¹⁹. CS cells, unlike CFS cells, were shown to express *MSX2*, *PAX7* and *TFAP2B* factors (**Fig. 19, 20, Table 7**). Whether *FOXB1* upregulation is a prerequisite to obtain CS cells and what kind of regulatory effect it could play in CFS cells should be investigated in future studies. *ETV4* has been shown as regulated by FGF signaling³²⁰, but its expression in CS cells suggests that WNT signaling could have a role in its induction.

AxSCs cannot be considered as pluripotent as they lack expression of core pluripotency markers *POU5F1* and *NANOG*, but they exhibit enrichment of key pluripotency-related genes namely the essentialome genes^{185,301}. Some of these genes, like *MYC* and *LIN28*, are not exclusively expressed in PSCs, but they have been found to be involved in gaining or sustaining the self-renewal capacity in multipotent stem cells. The self-renewing capacity is possibly obtained by changing chromatin organization or orchestrating a state specific novel gene regulatory network³²¹⁻³²³. *MYC* and *LIN28* have been used to reprogram somatic cells towards the pluripotency^{210,211}. My analysis indicated high expression of *MYCN*, *LIN28B* and *SALL4* in both human CFS and CS cells (**Fig. 27**), therefore I hypothesize that these genes might be involved in the underlying mechanism of indefinite self-renewing activity of AxSCs. NMP pool exhaustion in *Sall4* knockout mouse embryos at earlier timepoints compared to their WT counterparts strongly supports the direct involvement of *Sall4* in self-renewal activity of NMPs⁷⁶. Therefore, whether *SALL4* is also involved in the self-renewal of CFS and CS cells should be investigated by CRISPR knockout studies.

Stem cell methylation status and histone modifications influence the differentiation potential of stem cells due to their chromatin conformation altering function³²⁴. Human AxSCs showed drastic expression changes for known epigenetic regulators when compared to H9 cells (**Fig. 26B-D**). De novo DNA methylation by DNMT3A/B is suppressed in AxSCs while they maintain the global DNA methylation as indicated by chromatin occupancy of DNMT1 and UHRF1 (**Fig. 26B**) which are a part of the methylation maintenance machinery^{325,326}. TET proteins responsible for erasing methylation imprint³²⁷ act differently in H9, CFS and CS lines suggesting a cell-specific demethylation mechanism. This notion could be explained as TET proteins are known for structural differences based on their various target genes³²⁸. QSER1 has been identified as an interaction partner of TET1 in both mouse and human ESCs^{299,329}. Single gene knockouts in hESCs have been reported to reduce the differentiation efficiency towards pancreatic progenitors, while double knockout leads to very inefficient differentiation and to a greatly reduced multipotent pancreatic progenitor population²⁹⁹. The chromatin-binding pattern of TET1 and QSER1 in H9 and AxSC states postulates decreased developmental capacity in AxSCs compared to H9, where CS cells seem to have a lower developmental capacity when compared to CFS cells. This finding is consistent with the molecular features and progeny potential of the respective AxSC states, thus, QSER1/TET1 perturbations in AxSCs and investigation of the progeny differences could elucidate the function of epigenetic regulation with respect to AxSC progeny capacity. Metabolic status has also an impact on self-renewal and potency of stem cells as the shift from glycolysis to oxidative phosphorylation has been confirmed during PSCs differentiation into somatic cells³³⁰. This might explain the indications of higher oxidative phosphorylation in AxSCs compared to H9 (**Fig. 23F, 24**). Lastly, metabolic changes indicating calcium independent processes in H9 cells in contrast to AxSCs were revealed by GO analysis (**Fig. 23F**). This might be because of active WNT pathway in AxSCs as it has been shown that non-canonical WNT signaling increases calcium release and it activates calcium-dependent signalling in cells³³¹.

4.3. Human AxSC states can be distinguished by their lineage potential

4.3.1. Neural progeny of AxSCs

I confirmed that CFS transcriptomic and proteomic profiles are substantially overlapping with NMP characteristics. The CFS state can be considered the stem cell counterpart of NMPs. To further lay more weight to this claim, I investigated the progeny potential of both CFS and CS cells albeit the latter showed a potential biased towards the neural lineage. To this end, I first conducted neural differentiation experiments from AxSCs by applying a previously established protocol for MN differentiation from NPCs²²². The protocol comprises of 14-day

differentiation to the neural progenitor modality, followed by 14-day terminal differentiation by applying Notch inhibition as it has been known that Notch inhibition enhances neural differentiation propensity^{222,249}. Instead of following this trajectory, I inhibited directly Notch signaling at differentiation onset to reveal whether the differentiation could be expedited by using human AxSCs. The morphological observations indicate that CS cells differentiate into neurons faster than CFS cells (**Fig. 30A**) confirming previous findings with regards to CS cells being biased and perhaps limited to the neural lineage and it potentially represents a latter developmental stage compared to the CFS cells. Initial molecular characterization of the differentiation cultures repeated by using CFS_1-3 and CS_1-2 lines highlighted a higher heterogeneity in CFS derivatives compared to the CS ones due to expression of both MN marker (MNX1) and SN/dorsal spinal cord markers POU4F1 alongside ISL1 in CFS culture while CS culture were lacking MNX1 expression. (**Fig. 30B-C**). Previous findings regarding the undifferentiated AxSC state, the presence of non-neural cells in CFS culture (**Fig. 30A**) collectively with the expression of MNX1 and POU4F1 indicate that CFS cells harbor an extended developmental capacity compared to CS cells.

To decipher the cellular subtypes present in differentiation cultures, I performed scRNA sequencing experiment including day 14 CFS and CS, day 28 CFS as well as day 28 differentiating cells from NPCs by following the original protocol to determine differences between AxSC and NPC derivatives (**Fig. 31**). I first confirmed previous morphological observations at the molecular level that CS cells were detected as differentiating faster than CFS cells however AxSCs collectively showed a higher number of G1 phase cells compared to NPCs (**Fig. 32**). These findings suggest that use of AxSCs serves a more efficient way for spinal cord neural differentiation compared to NPCs. The cellular populations were annotated for the differentiation cultures (**Fig. 33-37**) and I concluded that CS cells are involved only in dorsal development as they generate only dl5 and dl6 dorsal spinal cord neurons, SNs and unspecialized NC derivatives (NC_d) (**Fig. 53**). CFS cells showed higher developmental potential as they could differentiate into both dorsal and ventral spinal cord neurons as well as various neural crest derivatives identified as preEMT NC cells, NC_d and SNs. CFS differentiation trajectory towards early neural crest cells (preEMT and NC_d) included an intermediate cellular population closely resembling CS cells (**Fig. 37**) which is in line with the previous findings showing that CFS represents an early developmental time point (**Fig. 28**). The preEMT NC cells were not found in the CS differentiation dataset, I postulate that CS cells can give rise to this population due to the presence of SN and NC_d cells in the dataset. I could have missed this stage during the sampling for the scRNA timecourse considering the higher differentiation rate of CS cells compared to CFS. Lastly, Schwann cells (Sch) were found as CFS but not as CS derivatives that requires further investigation by testing Schwann cell differentiation from both AxSC modalities. I found that NPCs can differentiate into SNs

and ventral spinal cord neurons. Strikingly, I determined a population called accessory motor neurons (aMN) in NPC but not in AxSC derivatives. It has been shown that aMNs are located at cervical level³³². These findings highlight the anterior propensity of NPCs generated by dual SMAD inhibition in contrary to AxSCs generated by WNT/FGF activation and TGF β inhibition. *OTX2* expression in NPCs further supports this notion (**Fig. 22**).

The presence of dorsal development-associated descendants in AxSC differentiation cultures can be questioned as SHH, which has been determined as a ventralisation factor³³³, was used during in the differentiation protocol. A similar outcome has been identified in NMP derivatives upon SHH activation. Wind et al.²²⁹ showed that neural differentiation from NMPs by SHH activation results in generation of both dorsal and ventral neurons, while inhibition of BMP and TGF β alongside SHH activation suppresses the dorsal propensity to a great extent. Iyer et al.²⁴⁶ also proved that NMPs are highly versatile and they can be directed towards each spinal cord domain by using various cytokine cocktail combinations. Similar strategies can be applied to AxSCs to test if they can generate individual spinal cord domains of interest. This could be achieved by using CFS cells, but I hypothesize that CS cells will not be able to generate ventral spinal cord domains considering the molecular profile of undifferentiated CS cells. The restricted capacity of CS cells towards only dorsal derivatives was already implied in the undifferentiated cells due to expression of *PAX3/6*, *MSX1/2*, *ZIC1* and *GSX2*, while CFS did not show any dorsal or ventral bias (**Fig. 17, 19-20**). In parallel to screening various cytokine cocktails to further improve the posterior neural differentiation, it needs to be investigated whether another neural-committed population showing a ventral propensity can be acquired, which expresses core CS markers (*SOX2* and *PAX6*) and ventral neural tube markers rather than dorsal markers like the CS cells established in this study (**Fig. 54**).

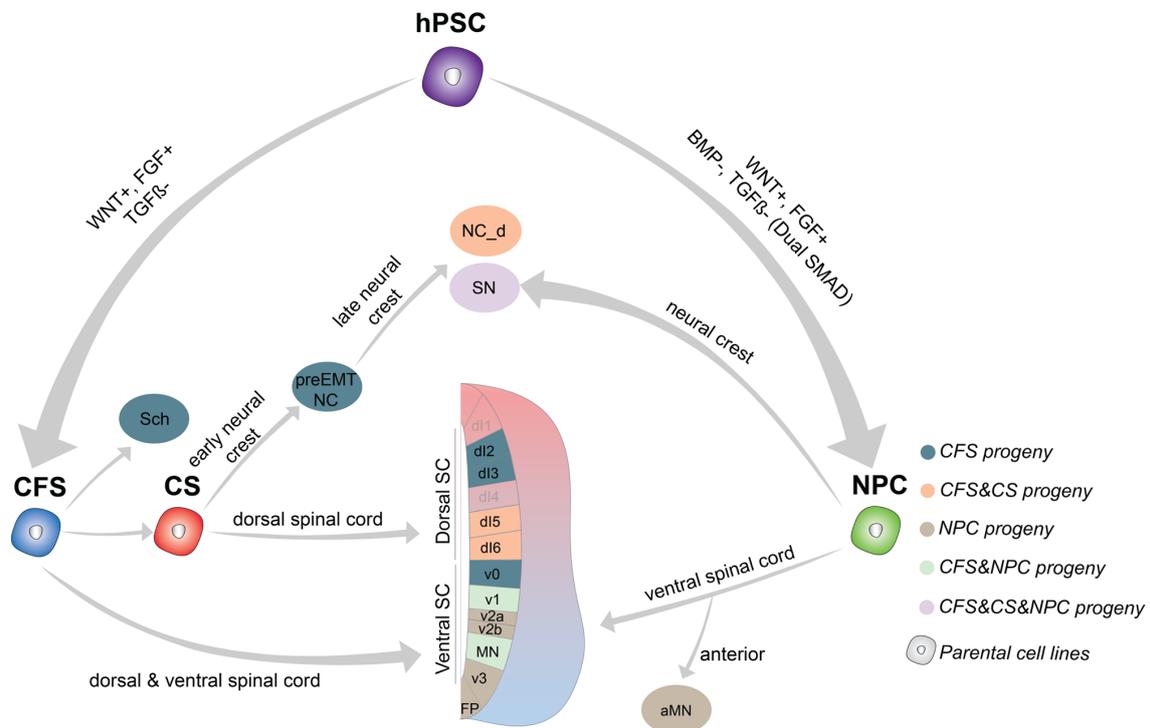


Figure 53: Summary of AxSC and NPC differentiation trajectory towards neural lineage

WNT/FGF activation with TGFβ inhibition generates CFS cells while addition of BMP inhibition generates NPCs from hPSCs. CS cells can be acquired from either directly hPSCs or CFS cells by WNT activation and TGFβ inhibition. Various cell types from dorsal and ventral spinal cord domains as well as neural crest derivatives are obtained from CFS cells. CS cells exhibit limited developmental potential because they can differentiate into only two domains of dorsal spinal cord and neural crest derivatives. Progeny profile of NPCs is not identical with either CFS or CS cells because they can give rise to only ventral spinal cord cells in addition to accessory motor neurons (aMN). Nevertheless, a small population of sensory neurons (SN) can be obtained from NPCs (Sch: Schwann cells, preEMT NC: pre-epithelial to mesenchymal transition neural crest cells), NC_d: unspecialized neural crest derivatives, MN: motor neurons, FP: floor plate)

4.3.2. Mesodermal progeny of AxSCs

4.3.2.1. Skeletal muscle differentiation

One of the striking features of NMPs *in vivo* is their contribution to somites^{19,20}. Skeletal muscle cells are a somite derivative, *in vitro* derived NMPs are characterized by their ability to differentiate into skeletal muscle cells and neurons⁴⁴. It has been shown that NMPs transiently appear during the differentiation of hPSCs towards somitic mesoderm^{250,252} emphasizing the NMPs involvement in somite formation. I decided to test the AxSC mesodermal potential by targeting differentiation to skeletal muscle cells.

Somite formation is achieved by a biological process called the segmentation clock. Notch signaling oscillations are the key factor that governs this process³³⁴. Each oscillation event

from the posterior embryo towards the determination front produces a single somite at the border region, which enables axial elongation^{105,335}. Taking into account the Notch effect, I first conducted differentiation experiments by applying a previously established protocol²⁶⁸ to AxSCs which includes Notch inhibition followed by FGF activation. It has been suggested that SHH and Activin A enhance the myogenic potential and inhibition of BMP signaling increases its efficiency as it suppresses lateral mesoderm induction³⁰⁵⁻³⁰⁷. I screened Activin A and SHH activation as well as BMP inhibition in parallel to the Notch inhibition (Table 9). cAMP and Ascorbic Acid (AA) have been shown to stimulate in vitro skeletal muscle maturation^{309,310}, thus, I applied FGF activation with or without cAMP and AA to test their effect. Based on the analysis of CFS_1-2 differentiation cultures on day 40 (**Fig. 39C-41**), I confirmed that CFS cells were able to generate skeletal muscle cells via Notch inhibition with or without SHH activation. Expression profile and morphological observations suggested that the differentiated cells were at myoblast stage due to the presence of MYOD/MYOG expressing cells and the absence of multinucleated cells. BMP inhibition and Activin induction in parallel to cAMP and AA treatment did not enhance the yield of CFS differentiation into skeletal muscle cells. For this reason, I applied HS, which is well known maturation factor³¹¹, to the cells after Notch inhibition with or without SHH activation followed by FGF activation (Table 9). Strikingly, neither *MYOD* nor *MYOG* were not expressed in day 40 differentiation cultures, pointing out the adverse effect of HS on CFS differentiation.

I detected skeletal muscle cells in the CFS differentiation cultures, but the efficiency in these attempts was very low. The protocol from Choi and colleagues²⁶⁸ was established for hPSCs differentiation, thus I presumed the differentiation timing of stimulatory factors used was most likely incompatible with the developmental stage of NMP-like cells. I found a recent study demonstrating the acquisition of skeletal muscle cell fate from hPSCs by recapitulating the step-wise axial elongation, which includes a NMP-like population²⁵². I adapted this modality to AxSCs, thus shortening the protocol. I confirmed the presence of skeletal muscle cells in CFS_1-3 differentiation cultures on day 35-38 (**Fig. 43B-44**) and I obtained higher differentiation efficiency compared to the outcome of previous strategies (**Fig. 39-41**). I detected multinucleated cells (**Fig. 44B**), thus inferring the existence of myocytes which represent a latter stage of muscle development compared to myoblasts.

I applied both types of modalities, which target Notch oscillation and step-wise axial elongation, to CS cells. I performed the experiments by using independently derived CS lines and noted that the CS differentiating cells regardless of the parental line showed neural morphology within one week and detached from culture. One iteration could be maintained until day 35. The neural morphology observed in culture alongside the negative transcriptional analysis which pointed out the absence of MYOD and MYOG, classical skeletal

muscle markers, one can conclude that CS cells lack the developmental capacity to give rise to skeletal muscle cells. Taken together, I concluded that CFS cells possess the developmental potential capable of differentiating into both mesodermal and neural lineage therefore they are long-term in vitro equivalent of NMPs, however, CS cells harbor relatively restricted capacity and represent a different type of axial progenitor. (**Fig. 53, 54**).

4.3.2.1. Intermediate and lateral mesoderm differentiation

Tbx6 is a well-known marker for somatic mesoderm development. A recent study suggested that Tbx6 expressing NMP progeny is versatile and it can be directed to intermediate mesoderm as a result of Osr1 upregulation⁹⁴. These cells are reported to contribute to the nephric mesenchyme as shown by Sox2-N1 enhancer labeled cells used for lineage tracing experiments in mouse embryos. In the light of this finding, I examined the potential of AxSCs towards intermediate mesoderm by conducting nephron differentiation experiments. I utilized a previously established protocol for this purpose²⁶⁹. I did not detect considerable upregulation of nephron markers but their expression, albeit in low levels, implicated the potential of AxSCs. I modified and shortened the protocol to potentially fit the developmental stage of AxSCs. Strikingly, expression analysis of day 15 differentiation cultures from both CFS_2 and CS_2 lines indicated the presence of nephron like cells due to the abundant expression of all nephron related markers (**Fig. 46B**). I carried out a time course experiment and found that the nephron markers reached the highest expression rate on day 11 in CFS and day 13-15 in CS differentiation cultures (Supp. Fig. 7B). I was unable to further confirm the expression of those markers by immunostaining experiments as nephron differentiation cultures require high confluency leading to technical difficulties in performing specific nuclear staining. Conducting AxSCs differentiation in 3D might facilitate immunochemical analysis through cryosectioning and elucidate the nephric potential more accurately. Preliminary findings in my study suggest that CFS cells can potentially differentiate into nephron like cells more rapidly than CS cells. The derivation of nephron cells from CFS could be expected as it is an NMP counterpart⁹⁴, but it was surprising to find evidence that CS state hold this capability as throughout this study CS cells have shown a bias to dorsal neural development (**Fig. 53, 54**). Close relation between caudal neural crest and embryonic kidney has been suggested in the literature^{336,337}. Neural crest cells are not essential for nephrogenesis, but it has been shown that they are involved in mouse kidney development, by orchestrating environmental signaling³³⁷. These findings could explain the nephric potential of CS cells considering they represent an earlier phase in the development compared to neural crest cells as CS cells can contribute to both neural crest derivatives and spinal cord. Their transcriptional signature also indicates a very distinguishable profile from neural crest progenitors.

Based on my literature search, there has not been a well-established study showing the NMP contribution to lateral mesoderm. I investigated if CFS cells also possess this developmental potential. I carried out cardiomyocyte differentiation experiments. Due to the previously identified implication for derivation of intermediate mesoderm from CS cells, I applied the same modality to both CFS and CS cells. The protocol is straightforward²⁷⁰, thus I did not modulate the cytokine cocktail but tested different seeding cell densities. It has been shown that beating cardiomyocytes can be acquired by day 10 when applying the original protocol to hPSCs. Regardless of the seeding density, I did not observe contraction in the differentiating cultures or upregulation of cardiomyocyte markers (**Fig. 47B**). These findings suggested that neither CFS nor CS cells are able to generate lateral mesoderm.

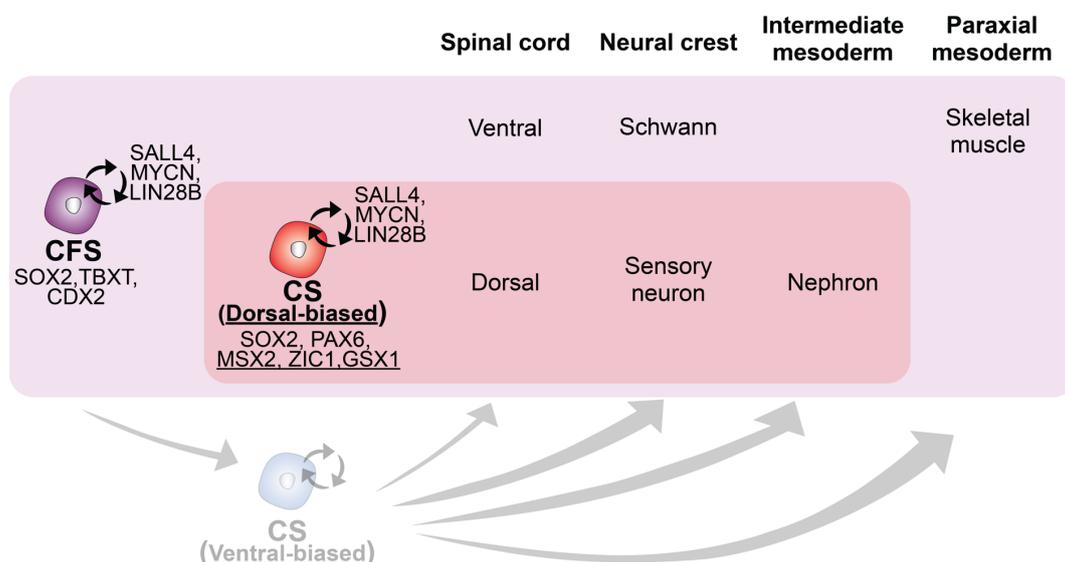


Figure 54: Proposed model for developmental capacity of human AxSCs

Two types of human AxSCs named CFS and CS have been identified in this study. CFS cells can be considered the developmental parental state of CS cells as they can be converted to both undifferentiated CS cells and their progenies. CFS and CS cells are respectively characterized by SOX2/TBXT/CDX2 and SOX2/PAX6 expression. SALL4, MYCN and LIN28B are commonly expressed in both AxSC states and they could be involved in the self-renewing mechanism governing axial stem cells. CS cells show dorsal development-bias due to the expression of MSX2, ZIC1, GSX1 which are not detected in CFS cells. The CS state dorsal-bias is proven by progeny profiling through the limited acquisition of dorsal spinal cord and sensory neurons (neural crest derivative). The differentiation capacity of CS state towards nephron cells is suggested within this study. In addition to CS related descendants, CFS cells generate ventral spinal cord, Schwann and skeletal muscle cells. As the CFS state covers dorsal-biased CS state and its derivatives while giving rise to ventral development-associated cell types raises the possibility of acquisition of ventral neural tube-committed (ventral-biased) state expressing core CS markers (SOX2 and PAX6) and ventral neural tube markers that needs to be confirmed by future studies including progeny analysis.

4.4. AxSCs are reproducible, depending on their in vivo counterparts and the different mechanisms of development in different species

I confirmed accuracy of the AxSC establishment protocol by generating CFS and CS cells from H9 cells three times independently. Transcriptome, proteome and chromatome analysis, as well as progeny investigation, confirmed that the unique AxSC features are the same across the lines. I validated CFS and CS identity by deriving axial stem cell lines from HUES6 hESCs and HMGU1 hiPSCs (**Fig. 29**). Along with the expression of core AxSC markers, I demonstrated that the overall transcriptomic profile is very similar in each AxSC state regardless of the hPSC background.

Next, I aimed to generate AxSCs directly from the mouse stem zone region where NMPs and axial progenitors are located in. After embryo dissection, I applied human culture conditions to the cells, however, clean and stable cultures could not be achieved as the stem cell-like colonies were present only for a few passages in both CFS and CS conditions (**Fig. 48**). These results suggested that the derivation protocol needs to be adapted for mouse cells. For optimization purposes, I used EpiSCs as the starting population because it has been shown that EpiSCs rather than mESCs are closely resembling hESCs¹⁶⁷. I skipped the 24-hour WNT induction pulse as undifferentiated EpiSCs already express Brachyury heterogeneously³¹² (Supp. Fig. 8). I optimized the culture conditions by screening a range of cytokine concentrations, as well as tested two different growth matrices: Matrigel and mitotically inactivated feeders (**Fig. 49**). Only the cells grown with FGF2, similar to the CFS derivation, could be sustained long-term in culture. I demonstrated that one condition which comprises of lower CHIR99021 and FGF2 concentration compared to human AxSCs modality is reproducible regardless of the ECM, while the high CHIR99021 and FGF2 condition work only in presence of feeders. I verified the CFS identity for the cells grown in both matrices and the modulated cytokine concentrations did not change the expression pattern of core CFS markers Sox2, Brachyury and Cdx2 (**Fig. 49G**). BMP inhibition added to CFS condition resulted in a slight upregulation of CS markers, but Brachyury and Cdx2 levels were not downregulated, thus concluding that BMP modulation is not sufficient for the acquisition of the CS identity (**Fig. 50**). None of the cultures exhibited Tbx6 expression which is frequently detected in NMP derivation cultures in vitro indicating that AxSC establishment protocol results in a more efficient induction of NMP-like cells. Taken together, these findings concluded that CFS but not CS state is reproducible between human and mouse species.

I hypothesize that the CS state, established from hPSC, cannot be derived for EpiSC cultures as there is no equivalent for CS population in mouse. This could be explained by the developmental differences between human and mouse as in shown in the literature. AxSCs, which lack OTX2 expression, are involved in the posterior development including the

formation of the secondary neural tube (SNT). SNT is formed by sequential events which are morphologically distinguishable in mouse and chick embryos⁸³. It has been postulated that human SNT development is closely resembling the chick rather than mouse SNT development³³⁸. Molecular basis underlying these divergent mechanism has not been elucidated yet, lineage tracing experiments conducted by Shaker et al.²⁸ highlighted a species-specific difference in the NMP mode of action during SNT formation. In the chick embryos, *Sox2/Brachyury* coexpressing cells are confined to tailbud region only, while they are also present in elongating SNT in mouse embryos. This observation could hint at the possibility that neural progenitors generated from mouse and chicken NMPs represent different developmental stages, with the mouse being further ahead in development and the chick neural progenitors require an interim step from NMPs to neural progenitors. Considering human and chicken SNT development are more similar to each other among the three species, it can be suggested that CS state represents the intermediate phase in chicken development that does not have a corresponding state in mouse development. Abundant *HES4* expression detected in human CS cells (**Fig. 14**) strongly supports this notion as it does not have an ortholog in mouse³³⁹. Lastly, my hypothesis proposing that CS reproducibility depends on the species-specific development is further supported by the successful generation of CS cells from orangutan iPSCs similar to human cells (**Fig. 51**).

Optimization of the conditions for the derivation of mouse axial stem cells using EpiSCs, lead to the establishment of mouse CFS state and the finding that CS state cannot be reproduced in mouse. I repeated CFS derivation from mouse embryonic tissues. Stem zone regions of the embryos at E8.5 and E9.0 corresponding to various somite stages (SS) from 4 to 19 were dissected. Two cultures, including 7/8 and 16/17 SS embryonic primary cells, were pooled individually and I was able to maintain them for an extended time in vitro (**Fig. 52B**). The 7/8 SS cells resulted in heterogeneous cultures regardless of the coating material tested, while the 16/17 SS cells grown on Matrigel were stabilized after passage 7. I observed morphological differences between 7/8 and 16/17 SS derived cells and detected that 16/17 cells were similar to mouse CFS cells generated from EpiSCs (**Fig. 49C&F, 52B**). I confirmed the CFS identity for all 7/8 and 16/17 SS cultures by analyzing *Sox2*, *Brachyury*, *Cdx2* and *Nkx1-2* expression (**Fig. 52C**). Differences in *Hes5* and *Mnx1* expression for the derived cultures suggested that they do not represent identical populations. Transcriptional pattern of 16/17 SS cells was more similar to both mouse and human CFS cells supporting the idea that CFS cells represent late NMPs (section 4.2.1). These results suggested that two different states of CFS cells can be captured in vitro that resemble early and late NMPs.

4.5. The promise of AxSCs for clinical translation

The self-renewing ability of stem cells paved the way for their use in the development of stem cell based replacement therapies in the field of regenerative medicine. The safety aspect of hPSC based therapies is a limiting factor as it requires extensive animal studies to determine the teratoma formation potential of a stem cell derived drug substance³⁴⁰. AxSCs do not exhibit the broad-spectrum developmental capacity unlike PSCs yet they still have self-renewing ability and generate a spectacular number of diverse lineage derivatives. It is important to study if AxSCs and their progeny can cause teratoma formation as this would provide a basis to overcome the safety concerns regarding usage of AxSCs in cell therapies focusing on neuromuscular diseases.

Somatic cell reprogramming developed in the recent years, has enabled patient-specific therapies while simultaneously overcoming ethical issues surrounding the usage of hESC, although hiPSC usage does not reduce the teratoma risk³⁴¹. Applying a similar strategy it can be possible to generate induced AxSCs (iAxSCs). This would be an important milestone as it can offer advantages such as shortening the timeline for obtaining AxSCs. Reprogramming somatic tissues to PSCs requires usage of pioneer transcription factors such as Yamanaka factors which are OCT4, KLF4, SOX2, MYC^{210,342}. Different factor combinations and delivery methods have been identified³⁴³. SOX2/TBXT/CDX2 and SOX2/PAX6 might be the potential pioneer factors to generate induced CFS and CS cells respectively. Usage of SALL4, MYCN, LIN28B can improve reprogramming efficiency considering these factors possibly are the regulators of the self-renewing AxSC network as discussed in section 4.2.3. Identification of AxSC essential genes following a strategy similar to that previously described in hPSCs¹⁸⁵ would decipher the AxSC gene regulatory network and serve as a map for identifying pioneer factors needed for the iAxSCs derivation.

LITERATURE

1. Streeter, G. L. Developmental horizons in human embryos. Description of age group XI, 13 to 20 somites, and age group XII, 21 to 29 somites. *Contrib. Embryol. Carnegie Inst.* **30**, 211–245 (1942).
2. Streeter, G. L. Developmental horizons in human embryos: description of age group XIII, embryos of 4 or 5 millimeters long, and age group XIV, period of identification of the lens vesicle. *Contrib Embryol Carnegie Inst* **31**, 27 (1945).
3. Streeter, G. L. Developmental horizons in human embryos. Description of age groups XV, XVI, XVII and XVIII. [Carnegie Institute, Washington; pub. no. 575]. *Contrib Embryol* **133**, 162 (1948).
4. GL, S. Developmental horizons in human embryos, description of age groups XIX, XX, XXI, XXII and XXIII. *Contrib Embryol* **230**, 167–196 (1951).
5. Theiler, K. *The house mouse: atlas of embryonic development*. (Springer Science & Business Media, 2013).
6. Hamburger, V. & Hamilton, H. L. A series of normal stages in the development of the chick embryo. *Developmental dynamics* **195**, 231–272 (1992).
7. Stephenson, R. O., Yamanaka, Y. & Rossant, J. Disorganized epithelial polarity and excess trophectoderm cell fate in preimplantation embryos lacking E-cadherin. *Development* **137**, (2010).
8. Morris, S. A. *et al.* Origin and formation of the first two distinct cell types of the inner cell mass in the mouse embryo. *Proc Natl Acad Sci U S A* **107**, (2010).
9. Murry, C. E. & Keller, G. Differentiation of embryonic stem cells to clinically relevant populations: lessons from embryonic development. *Cell* **132**, 661–680 (2008).
10. Handrigan, G. R. Concordia discors: Duality in the origin of the vertebrate tail. *Journal of Anatomy* vol. 202 Preprint at <https://doi.org/10.1046/j.1469-7580.2003.00163.x> (2003).
11. Herrmann, B. G., Labeit, S., Poustka, A., King, T. R. & Lehrach, H. Cloning of the T gene required in mesoderm formation in the mouse. *Nature* **343**, (1990).
12. Takada, S. *et al.* Wnt-3a regulates somite and tailbud formation in the mouse embryo. *Genes Dev* **8**, (1994).
13. Greco, T. L. *et al.* Analysis of the vestigial tail mutation demonstrates that Wnt-3a gene dosage regulates mouse axial development. *Genes Dev* **10**, (1996).
14. Gont, L. K., Steinbeisser, H., Blumberg, B. & De Robertis, E. M. Tail formation as a continuation of gastrulation: The multiple cell populations of the *Xenopus* tailbud derive from the late blastopore lip. *Development* **119**, (1993).
15. Cambray, N. & Wilson, V. Axial progenitors with extensive potency are localised to the mouse chordoneural hinge. *Development* **129**, (2002).
16. McGrew, M. J. *et al.* Localised axial progenitor cell populations in the avian tail bud are not committed to a posterior Hox identity. *Development* **135**, 2289–2299 (2008).
17. Neijts, R., Simmini, S., Giuliani, F., van Rooijen, C. & Deschamps, J. Region-specific regulation of posterior axial elongation during vertebrate embryogenesis. *Developmental Dynamics* vol. 243 Preprint at <https://doi.org/10.1002/dvdy.24027> (2014).
18. Zhai, J., Xiao, Z., Wang, Y. & Wang, H. Human embryonic development: from peri-implantation to gastrulation. *Trends in Cell Biology* vol. 32 Preprint at <https://doi.org/10.1016/j.tcb.2021.07.008> (2022).
19. Cambray, N. & Wilson, V. Two distinct sources for a population of maturing axial progenitors. *Development* **134**, (2007).
20. Wymeersch, F. J. *et al.* Position-dependent plasticity of distinct progenitor types in the primitive streak. *Elife* **5**, (2016).
21. Tzouanacou, E., Wegener, A., Wymeersch, F. J., Wilson, V. & Nicolas, J. F. Redefining the Progression of Lineage Segregations during Mammalian Embryogenesis by Clonal Analysis. *Dev Cell* **17**, (2009).
22. Wymeersch, F. J. *et al.* Transcriptionally dynamic progenitor populations organised around a stable niche drive axial patterning. *Development (Cambridge)* **146**, (2019).
23. Wymeersch, F. J., Wilson, V. & Tsakiridis, A. Understanding axial progenitor biology in vivo and in vitro. *Development (Cambridge)* vol. 148 Preprint at <https://doi.org/10.1242/dev.180612> (2021).

24. Zhu, J., Kwan, K. M. & Mackem, S. Putative oncogene Brachyury (T) is essential to specify cell fate but dispensable for notochord progenitor proliferation and EMT. *Proc Natl Acad Sci U S A* **113**, (2016).
25. Kawachi, T., Shimokita, E., Kudo, R., Tadokoro, R. & Takahashi, Y. Neural-fated self-renewing cells regulated by Sox2 during secondary neurulation in chicken tail bud. *Dev Biol* **461**, (2020).
26. Olivera-Martinez, I., Harada, H., Halley, P. A. & Storey, K. G. Loss of FGF-Dependent Mesoderm Identity and Rise of Endogenous Retinoid Signalling Determine Cessation of Body Axis Elongation. *PLoS Biol* **10**, (2012).
27. Henrique, D., Abranches, E., Verrier, L. & Storey, K. G. Neuromesodermal progenitors and the making of the spinal cord. *Development (Cambridge)* **142**, (2015).
28. Shaker, M. R. *et al.* Spatiotemporal contribution of neuromesodermal progenitor-derived neural cells in the elongation of developing mouse spinal cord. *Life Sci* **282**, (2021).
29. Wilson, V., Olivera-Martinez, I. & Storey, K. G. Stem cells, signals and vertebrate body axis extension. *Development* vol. 136 Preprint at <https://doi.org/10.1242/dev.021246> (2009).
30. Avilion, A. A. *et al.* Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes Dev* **17**, (2003).
31. Iwafuchi-Doi, M. *et al.* The Pou5f1/Pou3f-dependent but SoxB-independent regulation of conserved enhancer N2 initiates Sox2 expression during epiblast to neural plate stages in vertebrates. *Dev Biol* **352**, (2011).
32. Takemoto, T. *et al.* Tbx6-dependent Sox2 regulation determines neural or mesodermal fate in axial stem cells. *Nature* **470**, (2011).
33. Seb e-Pedr s, A. *et al.* Early evolution of the T-box transcription factor family. *Proc Natl Acad Sci U S A* **110**, (2013).
34. Papaioannou, V. E. The t-box gene family: Emerging roles in development, Stem cells and cancer. *Development (Cambridge)* **141**, (2014).
35. Inman, K. E. & Downs, K. M. Localization of Brachyury (T) in embryonic and extraembryonic tissues during mouse gastrulation. *Gene Expression Patterns* **6**, (2006).
36. Tomic, J. *et al.* Eomes and Brachyury control pluripotency exit and germ-layer segregation by changing the chromatin state. *Nat Cell Biol* **21**, (2019).
37. Aulehla, A. & Pourqui , O. Signaling gradients during paraxial mesoderm development. *Cold Spring Harbor perspectives in biology* vol. 2 Preprint at <https://doi.org/10.1101/cshperspect.a000869> (2010).
38. Wahl, M. B., Deng, C., Lewandowski, M. & Pourqui , O. FGF signaling acts upstream of the NOTCH and WNT signaling pathways to control segmentation clock oscillations in mouse somitogenesis. *Development* **134**, (2007).
39. Naiche, L. A., Holder, N. & Lewandoski, M. FGF4 and FGF8 comprise the wavefront activity that controls somitogenesis. *Proc Natl Acad Sci U S A* **108**, (2011).
40. Andre, P., Song, H., Kim, W., Kispert, A. & Yang, Y. Wnt5a and Wnt11 regulate mammalian anterior-posterior axis elongation. *Development (Cambridge)* **142**, (2015).
41. Cunningham, T. J., Kumar, S., Yamaguchi, T. P. & Ducrest, G. Wnt8a and Wnt3a cooperate in the axial stem cell niche to promote mammalian body axis extension. *Developmental Dynamics* **244**, (2015).
42. Yamaguchi, T. P., Takada, S., Yoshikawa, Y., Wu, N. & McMahon, A. P. T (Brachyury) is a direct target of Wnt3a during paraxial mesoderm specification. *Genes Dev* **13**, (1999).
43. Schwaiger, M. *et al.* An ancestral Wnt–Brachyury feedback loop in axial patterning and recruitment of mesoderm-determining target genes. *Nat Ecol Evol* **6**, (2022).
44. 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**, (2014).
45. Garriock, R. J. *et al.* Lineage tracing of neuromesodermal progenitors reveals novel wnt-dependent roles in trunk progenitor cell maintenance and differentiation. *Development (Cambridge)* **142**, (2015).
46. Koch, F. *et al.* Antagonistic Activities of Sox2 and Brachyury Control the Fate Choice of Neuro-Mesodermal Progenitors. *Dev Cell* **42**, (2017).
47. Sambasivan, R. & Steventon, B. Neuromesodermal Progenitors: A Basis for Robust Axial Patterning in Development and Evolution. *Frontiers in Cell and Developmental Biology* vol. 8 Preprint at <https://doi.org/10.3389/fcell.2020.607516> (2021).

48. Sharma, R. *et al.* Bmp signaling maintains a mesoderm progenitor cell state in the mouse tailbud. *J Cell Sci* **130**, (2017).
49. Akai, J., Halley, P. A. & Storey, K. G. FGF-dependent Notch signaling maintains the spinal cord stem zone. *Genes Dev* **19**, (2005).
50. Cooper, F. *et al.* Rostrocaudal patterning and neural crest differentiation of human pre-neural spinal cord progenitors in vitro. *Stem Cell Reports* **17**, (2022).
51. Swindell, E. C. *et al.* Complementary domains of retinoic acid production and degradation in the early chick embryo. *Dev Biol* **216**, (1999).
52. Patel, N. S. *et al.* FGF Signalling Regulates Chromatin Organisation during Neural Differentiation via Mechanisms that Can Be Uncoupled from Transcription. *PLoS Genet* **9**, (2013).
53. Sakai, Y. *et al.* The retinoic acid-inactivating enzyme CYP26 is essential for establishing an uneven distribution of retinoic acid along the antero-posterior axis within the mouse embryo. *Genes Dev* **15**, (2001).
54. Cunningham, T. J. & Duester, G. Mechanisms of retinoic acid signalling and its roles in organ and limb development. *Nature Reviews Molecular Cell Biology* vol. 16 Preprint at <https://doi.org/10.1038/nrm3932> (2015).
55. Diez del Corral, R. & Morales, A. V. Retinoic acid signaling during early spinal cord development. *Journal of Developmental Biology* vol. 2 Preprint at <https://doi.org/10.3390/jdb2030174> (2014).
56. Rhinn, M. & Dollé, P. Retinoic acid signalling during development. *Development* **139**, (2012).
57. Roberts, C. Regulating retinoic acid availability during development and regeneration: the role of the CYP26 enzymes. *J Dev Biol* **8**, 6 (2020).
58. Corral, R. D. Del *et al.* Opposing FGF and retinoid pathways control ventral neural pattern, neuronal differentiation, and segmentation during body axis extension. *Neuron* **40**, (2003).
59. Gouti, M. *et al.* A Gene Regulatory Network Balances Neural and Mesoderm Specification during Vertebrate Trunk Development. *Dev Cell* **41**, (2017).
60. Pilon, N. *et al.* Cdx4 is a direct target of the canonical Wnt pathway. *Dev Biol* **289**, (2006).
61. Keenan, I. D., Sharrard, R. M. & Isaacs, H. V. FGF signal transduction and the regulation of Cdx gene expression. *Dev Biol* **299**, 478–488 (2006).
62. Mazzoni, E. O. *et al.* Saltatory remodeling of Hox chromatin in response to rostrocaudal patterning signals. *Nat Neurosci* **16**, (2013).
63. Chawengsaksophak, K., De Graaff, W., Rossant, J., Deschamps, J. & Beck, F. Cdx2 is essential for axial elongation in mouse development. *Proc Natl Acad Sci U S A* **101**, (2004).
64. Jedrusik, A., Cox, A., Wicher, K., Glover, D. M. & Zernicka-Goetz, M. Maternal-zygotic knockout reveals a critical role of Cdx2 in the morula to blastocyst transition. *Dev Biol* **398**, (2015).
65. Amin, S. *et al.* Cdx and T Brachyury Co-activate Growth Signaling in the Embryonic Axial Progenitor Niche. *Cell Rep* **17**, (2016).
66. Guillot, C., Djeflal, Y., Michaut, A., Rabe, B. & Pourquié, O. Dynamics of primitive streak regression controls the fate of neuromesodermal progenitors in the chicken embryo. *Elife* **10**, (2021).
67. van Rooijen, C. *et al.* Evolutionarily conserved requirement of Cdx for post-occipital tissue emergence. *Development (Cambridge)* **139**, (2012).
68. Seifert, A. Role of Hox genes in stem cell differentiation. *World J Stem Cells* **7**, (2015).
69. Albors, A. R., Halley, P. A. & Storey, K. G. Lineage tracing of axial progenitors using nkx1-2creert2 mice defines their trunk and tail contributions. *Development (Cambridge)* **145**, (2018).
70. Bell, C. C. *et al.* The Evx1/Evx1as gene locus regulates anterior-posterior patterning during gastrulation. *Sci Rep* **6**, (2016).
71. Harrison, K. A., Druey, K. M., Deguchi, Y., Tuscano, J. M. & Kehrl, J. H. A novel human homeobox gene distantly related to proboscipedia is expressed in lymphoid and pancreatic tissues. *Journal of Biological Chemistry* **269**, (1994).
72. Arber, S. *et al.* Requirement for the homeobox gene Hb9 in the consolidation of motor neuron identity. *Neuron* **23**, (1999).
73. Aires, R. *et al.* Oct4 Is a Key Regulator of Vertebrate Trunk Length Diversity. *Dev Cell* **38**, (2016).
74. Aires, R. *et al.* Tail Bud Progenitor Activity Relies on a Network Comprising Gdf11, Lin28, and Hox13 Genes. *Dev Cell* **48**, (2019).

75. Robinton, D. A. *et al.* The Lin28/let-7 Pathway Regulates the Mammalian Caudal Body Axis Elongation Program. *Dev Cell* **48**, (2019).
76. Tahara, N. *et al.* Sall4 regulates neuromesodermal progenitors and their descendants during body elongation in mouse embryos. *Development (Cambridge)* **146**, (2019).
77. Cunningham, T. J., Colas, A. & Duester, G. Early molecular events during Retinoic acid induced differentiation of neuromesodermal progenitors. *Biol Open* **5**, (2016).
78. Dubrulle, J., McGrew, M. J. & Pourquié, O. FGF signaling controls somite boundary position and regulates segmentation clock control of spatiotemporal Hox gene activation. *Cell* **106**, (2001).
79. Diez del Corral, R., Breitzkreuz, D. N. & Storey, K. G. Onset of neuronal differentiation is regulated by paraxial mesoderm and requires attenuation of FGF signalling. *Development* vol. 129 Preprint at <https://doi.org/10.1242/dev.129.7.1681> (2002).
80. Cunningham, T. J. *et al.* Retinoic acid activity in undifferentiated neural progenitors is sufficient to fulfill its role in restricting Fgf8 expression for somitogenesis. *PLoS One* **10**, (2015).
81. Costanzo, R., Watterson, R. L. & Schoenwolf, G. C. Evidence that secondary neurulation occurs autonomously in the chick embryo. *Journal of Experimental Zoology* **219**, (1982).
82. Copp, A. J., Greene, N. D. E. & Murdoch, J. N. The genetic basis of mammalian neurulation. *Nature Reviews Genetics* vol. 4 Preprint at <https://doi.org/10.1038/nrg1181> (2003).
83. Catala, M., Teillet, M. A. & Douarin, N. M. Le. Organization and development of the tail bud analyzed with the quail-chick chimaera system. *Mech Dev* **51**, (1995).
84. Le Douarin, N. M., Teillet, M. A. & Catala, M. Neurulation in amniote vertebrates: A novel view deduced from the use of quail-chick chimeras. *International Journal of Developmental Biology* **42**, (1998).
85. Saitsu, H., Yamada, S., Uwabe, C., Ishibashi, M. & Shiota, K. Development of the posterior neural tube in human embryos. *Anat Embryol (Berl)* **209**, (2004).
86. Dady, A., Havis, E., Escriou, V., Catala, M. & Duband, J. L. Junctional Neurulation: A unique developmental program shaping a discrete region of the spinal cord highly susceptible to neural tube defects. *Journal of Neuroscience* **34**, (2014).
87. Müller, F. & O'Rahilly, R. The development of the human brain from a closed neural tube at stage 13. *Anat Embryol (Berl)* **177**, (1988).
88. Schoenwolf, G. C. Histological and ultrastructural studies of secondary neurulation in mouse embryos. *American journal of anatomy* **169**, 361–376 (1984).
89. Ulloa, F. & Martí, E. Wnt won the war: Antagonistic role of Wnt over Shh controls dorso-ventral patterning of the vertebrate neural tube. *Developmental Dynamics* vol. 239 Preprint at <https://doi.org/10.1002/dvdy.22058> (2010).
90. Stifani, N. Motor neurons and the generation of spinal motor neuron diversity. *Frontiers in Cellular Neuroscience* vol. 8 Preprint at <https://doi.org/10.3389/fncel.2014.00293> (2014).
91. Martin, B. L. & Kimelman, D. Regulation of Canonical Wnt Signaling by Brachyury Is Essential for Posterior Mesoderm Formation. *Dev Cell* **15**, (2008).
92. van de Ven, C. *et al.* Concerted involvement of Cdx/Hox genes and Wnt signaling in morphogenesis of the caudal neural tube and cloacal derivatives from the posterior growth zone. *Development* **138**, (2011).
93. James, R. G. & Schultheiss, T. M. Bmp signaling promotes intermediate mesoderm gene expression in a dose-dependent, cell-autonomous and translation-dependent manner. *Dev Biol* **288**, (2005).
94. Hayashi, S., Suzuki, H. & Takemoto, T. The nephric mesenchyme lineage of intermediate mesoderm is derived from Tbx6-expressing derivatives of neuro-mesodermal progenitors via BMP-dependent Osr1 function. *Dev Biol* **478**, (2021).
95. Yoshida, M. *et al.* Regulation of mesodermal precursor production by low-level expression of B1 Sox genes in the caudal lateral epiblast. *Mech Dev* **132**, (2014).
96. Hofmann, M. *et al.* WNT signaling, in synergy with T/TBX6, controls Notch signaling by regulating Dll1 expression in the presomitic mesoderm of mouse embryos. *Genes Dev* **18**, (2004).
97. Yasuhiko, Y. *et al.* Tbx6-mediated Notch signaling controls somite-specific Mesp2 expression. *Proc Natl Acad Sci U S A* **103**, (2006).
98. González, A., Manosalva, I., Liu, T. & Kageyama, R. Control of Hes7 Expression by Tbx6, the Wnt Pathway and the Chemical Gsk3 Inhibitor LiCl in the Mouse Segmentation Clock. *PLoS One* **8**, (2013).

99. Bessho, Y., Hirata, H., Masamizu, Y. & Kageyama, R. Periodic repression by the bHLH factor Hes7 is an essential mechanism for the somite segmentation clock. *Genes Dev* **17**, (2003).
100. Niwa, Y. *et al.* Different types of oscillations in notch and Fgf signaling regulate the spatiotemporal periodicity of somitogenesis. *Genes Dev* **25**, (2011).
101. Morimoto, M. *et al.* The negative regulation of Mesp2 by mouse Ripply2 is required to establish the rostro-caudal patterning within a somite. *Development* **134**, (2007).
102. Saga, Y. Segmental border is defined by the key transcription factor Mesp2, by means of the suppression of Notch activity. *Dev Dyn* **236**, 1450–1455 (2007).
103. Oginuma, M., Niwa, Y., Chapman, D. L. & Saga, Y. Mesp2 and Tbx6 cooperatively create periodic patterns coupled with the clock machinery during mouse somitogenesis. *Development* **135**, (2008).
104. Morimoto, M., Takahashi, Y., Endo, M. & Saga, Y. The Mesp2 transcription factor establishes segmental borders by suppressing Notch activity. *Nature* **435**, (2005).
105. Palmeirim, I., Henrique, D., Ish-Horowitz, D. & Pourquié, O. Avian hairy gene expression identifies a molecular clock linked to vertebrate segmentation and somitogenesis. *Cell* **91**, (1997).
106. Turnpenny, P. D. *et al.* Abnormal vertebral segmentation and the notch signaling pathway in man. *Developmental Dynamics* vol. 236 Preprint at <https://doi.org/10.1002/dvdy.21182> (2007).
107. Mankoo, B. S. *et al.* The concerted action of Meox homeobox genes is required upstream of genetic pathways essential for the formation, patterning and differentiation of somites. *Development* **130**, (2003).
108. Relaix, F., Rocancourt, D., Mansouri, A. & Buckingham, M. Divergent functions of murine Pax3 and Pax7 in limb muscle development. *Genes Dev* **18**, (2004).
109. Tani, S., Chung, U. il, Ohba, S. & Hojo, H. Understanding paraxial mesoderm development and sclerotome specification for skeletal repair. *Experimental and Molecular Medicine* vol. 52 Preprint at <https://doi.org/10.1038/s12276-020-0482-1> (2020).
110. McMahon, J. A. *et al.* Noggin-mediated antagonism of BMP signaling is required for growth and patterning of the neural tube and somite. *Genes Dev* **12**, (1998).
111. Murtaugh, L. C., Chyung, J. H. & Lassar, A. B. Sonic hedgehog promotes somitic chondrogenesis by altering the cellular response to BMP signaling. *Genes Dev* **13**, (1999).
112. Christ, B., Jacob, H. J. & Jacob, M. On the formation of the myotomes in avian embryos. An experimental and scanning electron microscope study. *Experientia* **34**, (1978).
113. Rodrigo, I., Hill, R. E., Balling, R., Münsterberg, A. & Imai, K. Pax1 and Pax9 activate Bapx1 to induce chondrogenic differentiation in the sclerotome. *Development* vol. 130 Preprint at <https://doi.org/10.1242/dev.00240> (2003).
114. Cairns, D. M., Sato, M. E., Lee, P. G., Lassar, A. B. & Zeng, L. A gradient of Shh establishes mutually repressing somitic cell fates induced by Nkx3.2 and Pax3. *Dev Biol* **323**, (2008).
115. Chal, J. & Pourquié, O. *Patterning and Differentiation of the Vertebrate Spine*. CSHL Press (2009).
116. Hughes, D. S., Keynes, R. J. & Tannahill, D. Extensive molecular differences between anterior- and posterior-half- sclerotomes underlie somite polarity and spinal nerve segmentation. *BMC Dev Biol* **9**, (2009).
117. Brent, A. E. & Tabin, C. J. Developmental regulation of somite derivatives: Muscle, cartilage and tendon. *Current Opinion in Genetics and Development* vol. 12 Preprint at [https://doi.org/10.1016/S0959-437X\(02\)00339-8](https://doi.org/10.1016/S0959-437X(02)00339-8) (2002).
118. Ben-Yair, R. & Kalcheim, C. Lineage analysis of the avian dermomyotome sheet reveals the existence of single cells with both dermal and muscle progenitor fates. *Development* **132**, (2005).
119. Chal, J. & Pourquié, O. Making muscle: Skeletal myogenesis in vivo and in vitro. *Development (Cambridge)* vol. 144 Preprint at <https://doi.org/10.1242/dev.151035> (2017).
120. Olivera-Martinez, I., Coltey, M., Dhouailly, D. & Pourquie, O. Mediolateral somitic origin of ribs and dermis determined by quail-chick chimeras. *Development* **127**, (2000).
121. Ordahl, C. P., Berdugo, E., Venters, S. J. & Denetclaw, J. The dermomyotome dorsomedial lip drives growth and morphogenesis of both the primary myotome and dermomyotome epithelium. *Development* **128**, (2001).
122. Pu, Q. *et al.* Correction: The dermomyotome ventrolateral lip is essential for the hypaxial myotome formation. *BMC Dev Biol* **13**, (2013).

123. Schuster-Gossler, K., Cordes, R. & Gossler, A. Premature myogenic differentiation and depletion of progenitor cells cause severe muscle hypotrophy in Delta1 mutants. *Proc Natl Acad Sci U S A* **104**, (2007).
124. Ustanina, S., Carvajal, J., Rigby, P. & Braun, T. The Myogenic Factor Myf5 Supports Efficient Skeletal Muscle Regeneration by Enabling Transient Myoblast Amplification. *Stem Cells* **25**, (2007).
125. Rios, A. C., Serralbo, O., Salgado, D. & Marcelle, C. Neural crest regulates myogenesis through the transient activation of NOTCH. *Nature* **473**, (2011).
126. Compton, R. S. & Konigsberg, I. R. Cell cycle withdrawal without concomitant differentiation: Analysis of a G1-specific temperature-sensitive murine myoblast cell line. *Dev Biol* **129**, (1988).
127. Jang, Y.-N. & Baik, E. J. JAK-STAT pathway and myogenic differentiation. *JAKSTAT* **2**, (2013).
128. Fox, J. C. & Swain, J. L. Auto and transactivation of FGF expression: Potential mechanism for regulation of myogenic differentiation. *In Vitro Cellular & Developmental Biology - Animal: Journal of the Society for In Vitro Biology* **29**, (1993).
129. Hasty, P. *et al.* Muscle deficiency and neonatal death in mice with a targeted mutation in the myogenin gene. *Nature* **364**, (1993).
130. Nabeshima, Y. *et al.* Myogenin gene disruption results in perinatal lethality because of severe muscle defect. *Nature* **364**, (1993).
131. Itoh, N., Mima, T. & Mikawa, T. Loss of fibroblast growth factor receptors is necessary for terminal differentiation of embryonic limb muscle. *Development* **122**, (1996).
132. Kablar, B., Krastel, K., Tajbakhsh, S. & Rudnicki, M. A. Myf5 and MyoD activation define independent myogenic compartments during embryonic development. *Dev Biol* **258**, (2003).
133. Kassar-Duchossoy, L. *et al.* Mrf4 determines skeletal muscle identity in Myf5:MyoD double-mutant mice. *Nature* **431**, (2004).
134. Moncaut, N., Rigby, P. W. J. & Carvajal, J. J. Dial M(RF) for myogenesis. *FEBS Journal* vol. 280 Preprint at <https://doi.org/10.1111/febs.12379> (2013).
135. Javali, A. *et al.* Co-expression of Tbx6 and Sox2 identifies a novel transient neuromesoderm progenitor cell state. *Development (Cambridge)* **144**, (2017).
136. Serbedzija, G. N., Bronner-Fraser, M. & Fraser, S. E. Developmental potential of trunk neural crest cells in the mouse. *Development* **120**, (1994).
137. Sanchez-Ferras, O. *et al.* A direct role for murine cdx proteins in the trunk neural crest gene regulatory network. *Development (Cambridge)* **143**, (2016).
138. Theveneau, E. & Mayor, R. Neural crest delamination and migration: From epithelium-to-mesenchyme transition to collective cell migration. *Developmental Biology* vol. 366 Preprint at <https://doi.org/10.1016/j.ydbio.2011.12.041> (2012).
139. Soldatov, R. *et al.* Spatiotemporal structure of cell fate decisions in murine neural crest. *Science (1979)* **364**, (2019).
140. Cano, A. *et al.* The transcription factor Snail controls epithelial-mesenchymal transitions by repressing E-cadherin expression. *Nat Cell Biol* **2**, (2000).
141. Ahlstrom, J. D. & Erickson, C. A. The neural crest epithelial-mesenchymal transition in 4D: A 'tail' of multiple non-obligatory cellular mechanisms. *Development* **136**, (2009).
142. Krispin, S., Nitzan, E., Kassem, Y. & Kalcheim, C. Evidence for a dynamic spatiotemporal fate map and early fate restrictions of premigratory avian neural crest. *Development* **137**, (2010).
143. Rothstein, M., Bhattacharya, D. & Simoes-Costa, M. The molecular basis of neural crest axial identity. *Developmental Biology* vol. 444 Preprint at <https://doi.org/10.1016/j.ydbio.2018.07.026> (2018).
144. Creuzet, S., Couly, G., Vincent, C. & Le Douarin, N. M. Negative effect of Hox gene expression on the development of the neural crest-derived facial skeleton. *Development* vol. 129 Preprint at <https://doi.org/10.1242/dev.129.18.4301> (2002).
145. Cooper, F. & Tsakiridis, A. Shaping axial identity during human pluripotent stem cell differentiation to neural crest cells. *Biochemical Society Transactions* vol. 50 Preprint at <https://doi.org/10.1042/BST20211152> (2022).
146. Rocha, M. *et al.* From head to tail: Regionalization of the neural crest. *Development (Cambridge)* vol. 147 Preprint at <https://doi.org/10.1242/dev.193888> (2021).

147. Dykes, I. M., Tempest, L., Lee, S. I. & Turner, E. E. Brn3a and Islet1 act epistatically to regulate the gene expression program of sensory differentiation. *Journal of Neuroscience* **31**, (2011).
148. Zou, M., Li, S., Klein, W. H. & Xiang, M. Brn3a/Pou4f1 regulates dorsal root ganglion sensory neuron specification and axonal projection into the spinal cord. *Dev Biol* **364**, (2012).
149. Haberberger, R. V., Barry, C., Dominguez, N. & Matusica, D. Human dorsal root ganglia. *Frontiers in Cellular Neuroscience* vol. 13 Preprint at <https://doi.org/10.3389/fncel.2019.00271> (2019).
150. Nguyen, M. Q., von Buchholtz, L. J., Reker, A. N., Ryba, N. J. P. & Davidson, S. Single-nucleus transcriptomic analysis of human dorsal root ganglion neurons. *Elife* **10**, (2021).
151. Tarkowski, A. K. Experiments on the development of isolated blastomeres of mouse eggs. *Nature* **184**, (1959).
152. Rossant, J. Postimplantation development of blastomeres isolated from 4 and 8 cell mouse eggs. *J Embryol Exp Morphol* **36**, (1976).
153. Gardner, R. L. Clonal analysis of early mammalian development. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* vol. 312 Preprint at <https://doi.org/10.1098/rstb.1985.0186> (1985).
154. Beddington, R. S. P. & Robertson, E. J. An assessment of the developmental potential of embryonic stem cells in the midgestation mouse embryo. *Development* **105**, (1989).
155. Viukov, S. *et al.* Human primed and naïve PSCs are both able to differentiate into trophoblast stem cells. *Stem Cell Reports* **17**, (2022).
156. Zakrzewski, W., Dobrzyński, M., Szymonowicz, M. & Rybak, Z. Stem cells: Past, present, and future. *Stem Cell Research and Therapy* vol. 10 Preprint at <https://doi.org/10.1186/s13287-019-1165-5> (2019).
157. Blackburn, E. H. Switching and signaling at the telomere. *Cell* vol. 106 Preprint at [https://doi.org/10.1016/S0092-8674\(01\)00492-5](https://doi.org/10.1016/S0092-8674(01)00492-5) (2001).
158. Chen, J. L., Blasco, M. A. & Greider, C. W. Secondary structure of vertebrate telomerase RNA. *Cell* **100**, (2000).
159. Gomes, N. M. V. *et al.* Comparative biology of mammalian telomeres: Hypotheses on ancestral states and the roles of telomeres in longevity determination. *Aging Cell* **10**, (2011).
160. Levy, M. Z., Allsopp, R. C., Futcher, A. B., Greider, C. W. & Harley, C. B. Telomere end-replication problem and cell aging. *J Mol Biol* **225**, (1992).
161. Kim, W. *et al.* Regulation of the Human Telomerase Gene TERT by Telomere Position Effect—Over Long Distances (TPE-OLD): Implications for Aging and Cancer. *PLoS Biol* **14**, (2016).
162. Greider, C. W. & Blackburn, E. H. The telomere terminal transferase of tetrahymena is a ribonucleoprotein enzyme with two kinds of primer specificity. *Cell* **51**, (1987).
163. De Lange, T. Shelterin: The protein complex that shapes and safeguards human telomeres. *Genes and Development* vol. 19 Preprint at <https://doi.org/10.1101/gad.1346005> (2005).
164. Cheung, H.-H. *et al.* Telomerase Protects Werner Syndrome Lineage-Specific Stem Cells from Premature Aging. *Stem Cell Reports* **2**, 534–546 (2014).
165. Martin, G. R. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc Natl Acad Sci U S A* **78**, (1981).
166. Thomson, J. A. Embryonic stem cell lines derived from human blastocysts. *Science (1979)* **282**, (1998).
167. Tesar, P. J. *et al.* New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature* **448**, (2007).
168. Brons, I. G. M. *et al.* Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature* **448**, (2007).
169. Guo, G. *et al.* Naive Pluripotent Stem Cells Derived Directly from Isolated Cells of the Human Inner Cell Mass. *Stem Cell Reports* **6**, (2016).
170. Nichols, J. & Smith, A. Naive and Primed Pluripotent States. *Cell Stem Cell* vol. 4 Preprint at <https://doi.org/10.1016/j.stem.2009.05.015> (2009).
171. Guo, G. *et al.* Klf4 reverts developmentally programmed restriction of ground state pluripotency. *Development* **136**, (2009).
172. Silva, J. *et al.* Nanog Is the Gateway to the Pluripotent Ground State. *Cell* **138**, (2009).

173. Theunissen, T. W. *et al.* Systematic identification of culture conditions for induction and maintenance of naive human pluripotency. *Cell Stem Cell* **15**, (2014).
174. Kurek, D. *et al.* Endogenous WNT signals mediate BMP-induced and spontaneous differentiation of epiblast stem cells and human embryonic stem cells. *Stem Cell Reports* **4**, (2015).
175. Hanna, J. *et al.* Human embryonic stem cells with biological and epigenetic characteristics similar to those of mouse ESCs. *Proc Natl Acad Sci U S A* **107**, (2010).
176. Yeo, J. C. & Ng, H. H. The transcriptional regulation of pluripotency. *Cell Research* vol. 23 Preprint at <https://doi.org/10.1038/cr.2012.172> (2013).
177. Weinberger, L., Ayyash, M., Novershtern, N. & Hanna, J. H. Dynamic stem cell states: naive to primed pluripotency in rodents and humans. *Nat Rev Mol Cell Biol* **17**, 155–169 (2016).
178. Chambers, I. *et al.* Nanog safeguards pluripotency and mediates germline development. *Nature* **450**, (2007).
179. Boroviak, T., Loos, R., Bertone, P., Smith, A. & Nichols, J. The ability of inner-cell-mass cells to self-renew as embryonic stem cells is acquired following epiblast specification. *Nat Cell Biol* **16**, (2014).
180. Kalkan, T. *et al.* Tracking the embryonic stem cell transition from ground state pluripotency. *Development (Cambridge)* **144**, (2017).
181. Messmer, T. *et al.* Transcriptional Heterogeneity in Naive and Primed Human Pluripotent Stem Cells at Single-Cell Resolution. *Cell Rep* **26**, (2019).
182. Buecker, C. *et al.* Reorganization of enhancer patterns in transition from naive to primed pluripotency. *Cell Stem Cell* **14**, (2014).
183. Sang, H. *et al.* Dppa3 is critical for Lin28a-regulated ES cells naïve-primed state conversion. *J Mol Cell Biol* **11**, (2019).
184. Yang, S. H. *et al.* ZIC3 Controls the Transition from Naive to Primed Pluripotency. *Cell Rep* **27**, (2019).
185. Yilmaz, A., Peretz, M., Aharony, A., Sagi, I. & Benvenisty, N. Defining essential genes for human pluripotent stem cells by CRISPR-Cas9 screening in haploid cells. *Nat Cell Biol* **20**, (2018).
186. Bi, Y. *et al.* Cell fate roadmap of human primed-to-naive transition reveals preimplantation cell lineage signatures. *Nat Commun* **13**, (2022).
187. Mayer, D. *et al.* Zfp281 orchestrates interconversion of pluripotent states by engaging Ehmt1 and Zic2. *EMBO J* **39**, (2020).
188. Fidalgo, M. *et al.* Zfp281 Coordinates Opposing Functions of Tet1 and Tet2 in Pluripotent States. *Cell Stem Cell* **19**, (2016).
189. Mulholland, C. B. *et al.* Distinct and stage-specific contributions of TET1 and TET2 to stepwise cytosine oxidation in the transition from naive to primed pluripotency. *Sci Rep* **10**, (2020).
190. Bagci, H. & Fisher, A. G. Dna demethylation in pluripotency and reprogramming: The role of Tet proteins and cell division. *Cell Stem Cell* vol. 13 Preprint at <https://doi.org/10.1016/j.stem.2013.08.005> (2013).
191. Hill, P. W. S., Amouroux, R. & Hajkova, P. DNA demethylation, Tet proteins and 5-hydroxymethylcytosine in epigenetic reprogramming: An emerging complex story. *Genomics* vol. 104 Preprint at <https://doi.org/10.1016/j.ygeno.2014.08.012> (2014).
192. Ugur, E. *et al.* Comprehensive chromatin proteomics resolves functional phases of pluripotency and identifies changes in regulatory components. *Nucleic Acids Res* **51**, (2023).
193. Cartwright, P. *et al.* LIF/STAT3 controls ES cell self-renewal and pluripotency by a Myc-dependent mechanism. *Development* **132**, (2005).
194. Niwa, H., Ogawa, K., Shimosato, D. & Adachi, K. A parallel circuit of LIF signalling pathways maintains pluripotency of mouse ES cells. *Nature* **460**, (2009).
195. Do, D. V. *et al.* A genetic and developmental pathway from STAT3 to the OCT4-NANOG circuit is essential for maintenance of ICM lineages in vivo. *Genes Dev* **27**, (2013).
196. Dahéron, L. *et al.* LIF/STAT3 Signaling Fails to Maintain Self-Renewal of Human Embryonic Stem Cells. *Stem Cells* **22**, (2004).
197. Ying, Q. L. *et al.* The ground state of embryonic stem cell self-renewal. *Nature* **453**, (2008).
198. Furlan, G., Huyghe, A., Combémourel, N. & Laval, F. Molecular versatility during pluripotency progression. *Nat Commun* **14**, (2023).

199. Khan, S. A. *et al.* Probing the signaling requirements for naive human pluripotency by high-throughput chemical screening. *Cell Rep* **35**, (2021).
200. Tam, W.-L. *et al.* T-Cell Factor 3 Regulates Embryonic Stem Cell Pluripotency and Self-Renewal by the Transcriptional Control of Multiple Lineage Pathways. *Stem Cells* **26**, (2008).
201. Xu, Z. *et al.* Wnt/ β -catenin signaling promotes self-renewal and inhibits the primed state transition in naïve human embryonic stem cells. *Proc Natl Acad Sci U S A* **113**, (2016).
202. Massagué, J., Seoane, J. & Wotton, D. Smad transcription factors. *Genes and Development* vol. 19 Preprint at <https://doi.org/10.1101/gad.1350705> (2005).
203. Ogawa, K. *et al.* Activin-Nodal signaling is involved in propagation of mouse embryonic stem cells. *J Cell Sci* **120**, (2007).
204. Morikawa, M. *et al.* BMP Sustains Embryonic Stem Cell Self-Renewal through Distinct Functions of Different Krüppel-like Factors. *Stem Cell Reports* **6**, (2016).
205. Osnato, A. *et al.* Tgf β signalling is required to maintain pluripotency of human naïve pluripotent stem cells. *Elife* **10**, (2021).
206. Guo, G. *et al.* Human naive epiblast cells possess unrestricted lineage potential. *Cell Stem Cell* **28**, (2021).
207. Kim, H. *et al.* Modulation of β -catenin function maintains mouse epiblast stem cell and human embryonic stem cell self-renewal. *Nat Commun* **4**, (2013).
208. Vallier, L., Alexander, M. & Pedersen, R. A. Activin/Nodal and FGF pathways cooperate to maintain pluripotency of human embryonic stem cells. *J Cell Sci* **118**, (2005).
209. Takahashi, K. & Yamanaka, S. Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell* **126**, (2006).
210. Takahashi, K. *et al.* Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. *Cell* **131**, (2007).
211. Yu, J. *et al.* Induced pluripotent stem cell lines derived from human somatic cells. *Science (1979)* **318**, (2007).
212. Guenther, M. G. *et al.* Chromatin structure and gene expression programs of human embryonic and induced pluripotent stem cells. *Cell Stem Cell* **7**, (2010).
213. Bock, C. *et al.* Reference maps of human es and ips cell variation enable high-throughput characterization of pluripotent cell lines. *Cell* **144**, (2011).
214. Hu, S. *et al.* Effects of cellular origin on differentiation of human induced pluripotent stem cell-derived endothelial cells. *Journal of Clinical Investigation* **1**, (2016).
215. Uchida, N. *et al.* Direct isolation of human central nervous system stem cells. *Proc Natl Acad Sci U S A* **97**, (2000).
216. Rybachuk, O., Kopach, O., Pivneva, T. & Kyryk, V. Isolation of Neural Stem Cells from the Embryonic Mouse Hippocampus for in vitro Growth or Engraftment into a Host Tissue. *Bio Protoc* **9**, (2019).
217. Ahmed, A. K. M. A., Isaksen, T. J. & Yamashita, T. Protocol for mouse adult neural stem cell isolation and culture. *STAR Protoc* **2**, (2021).
218. Galiakberova, A. A. & Dashinimaev, E. B. Neural Stem Cells and Methods for Their Generation From Induced Pluripotent Stem Cells in vitro. *Frontiers in Cell and Developmental Biology* vol. 8 Preprint at <https://doi.org/10.3389/fcell.2020.00815> (2020).
219. Medelnik, J. P. *et al.* Signaling-Dependent Control of Apical Membrane Size and Self-Renewal in Rosette-Stage Human Neuroepithelial Stem Cells. *Stem Cell Reports* **10**, (2018).
220. Koch, P., Opitz, T., Steinbeck, J. A., Ladewig, J. & Brüstle, O. A rosette-type, self-renewing human ES cell-derived neural stem cell with potential for in vitro instruction and synaptic integration. *Proc Natl Acad Sci U S A* **106**, (2009).
221. Obayashi, S., Tabunoki, H., Kim, S. U. & Satoh, J. I. Gene expression profiling of human neural progenitor cells following the serum-induced astrocyte differentiation. *Cell Mol Neurobiol* **29**, (2009).
222. Qu, Q. *et al.* High-efficiency motor neuron differentiation from human pluripotent stem cells and the function of Islet-1. *Nat Commun* **5**, (2014).
223. Bian, J., Zheng, J., Li, S., Luo, L. & Ding, F. Sequential differentiation of embryonic stem cells into neural epithelial-like stem cells and oligodendrocyte progenitor cells. *PLoS One* **11**, (2016).

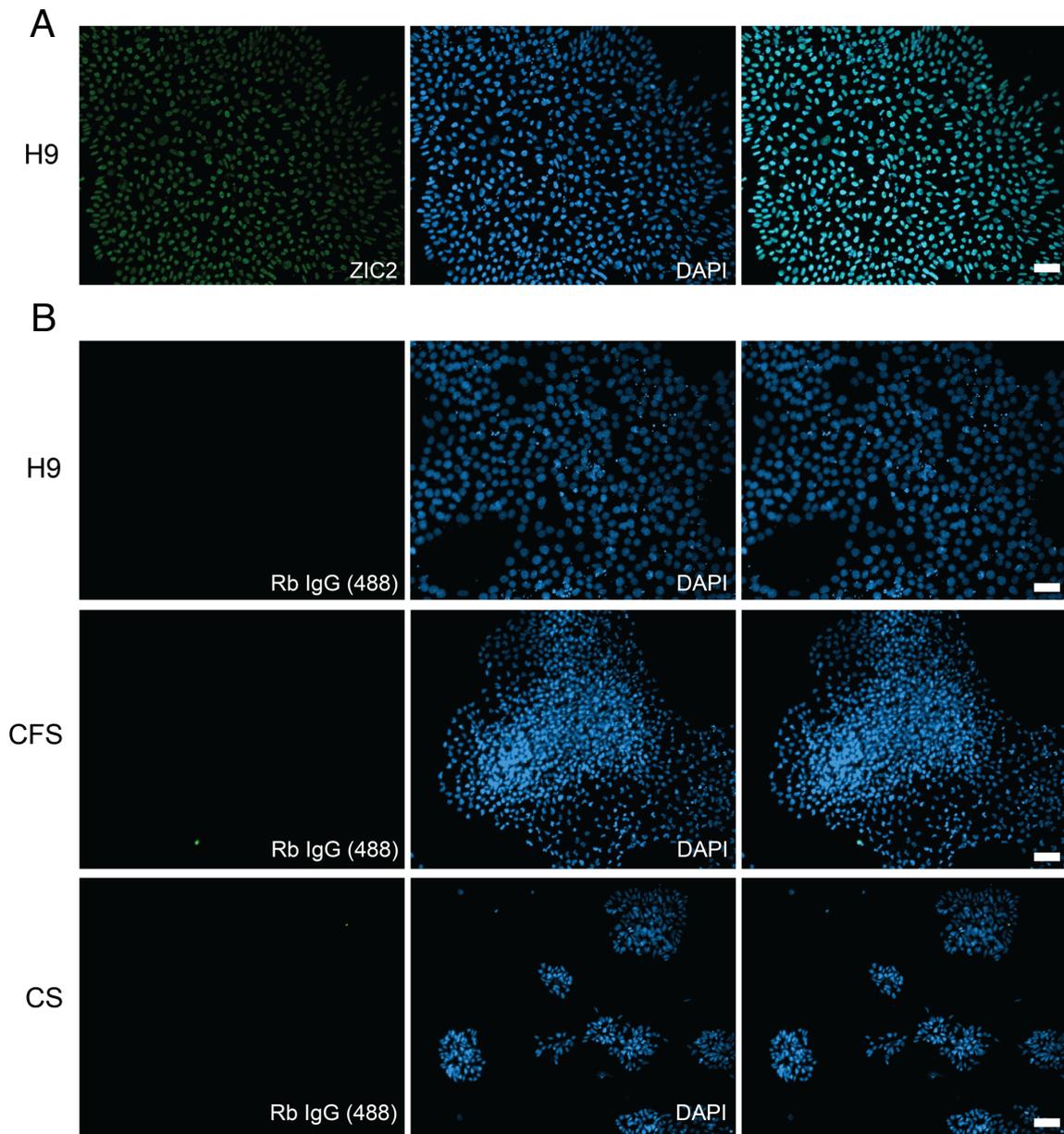
224. Elkabetz, Y. & Studer, L. Human ESC-derived neural rosettes and neural stem cell progression. in *Cold Spring Harbor Symposia on Quantitative Biology* vol. 73 (2008).
225. Elkabetz, Y. *et al.* Human ES cell-derived neural rosettes reveal a functionally distinct early neural stem cell stage (vol 22, pg 152, 2008). *Genes Dev* **22**, 1257 (2008).
226. Metzis, V. *et al.* Nervous System Regionalization Entails Axial Allocation before Neural Differentiation. *Cell* **175**, (2018).
227. Dady, A., Davidson, L., Halley, P. A. & Storey, K. G. Human spinal cord in vitro differentiation pace is initially maintained in heterologous embryonic environments. *Elife* **11**, (2022).
228. Israsena, N., Hu, M., Fu, W., Kan, L. & Kessler, J. A. The presence of FGF2 signaling determines whether β -catenin exerts effects on proliferation or neuronal differentiation of neural stem cells. *Dev Biol* **268**, (2004).
229. Wind, M. *et al.* Defining the signalling determinants of a posterior ventral spinal cord identity in human neuromesodermal progenitor derivatives. *Development (Cambridge)* **148**, (2021).
230. Frith, T. J. R. & Tsakiridis, A. Efficient Generation of Trunk Neural Crest and Sympathetic Neurons from Human Pluripotent Stem Cells Via a Neuromesodermal Axial Progenitor Intermediate. *Curr Protoc Stem Cell Biol* **49**, (2019).
231. Frith, T. J. R. *et al.* Human axial progenitors generate trunk neural crest cells in vitro. *Elife* **7**, (2018).
232. Wang, H. *et al.* Characterization and therapeutic application of mesenchymal stem cells with neuromesodermal origin from human pluripotent stem cells. *Theranostics* **9**, (2019).
233. Denham, M. *et al.* Multipotent caudal neural progenitors derived from human pluripotent stem cells that give rise to lineages of the central and peripheral nervous system. *Stem Cells* **33**, (2015).
234. Verrier, L., Davidson, L., Gierlinński, M., Dady, A. & Storey, K. G. Neural differentiation, selection and transcriptomic profiling of human neuromesodermal progenitor-like cells in vitro. *Development (Cambridge)* **145**, (2018).
235. Wind, M. & Tsakiridis, A. In Vitro Generation of Posterior Motor Neurons from Human Pluripotent Stem Cells. *Curr Protoc* **1**, (2021).
236. Lippmann, E. S. *et al.* Deterministic HOX patterning in human pluripotent stem cell-derived neuroectoderm. *Stem Cell Reports* **4**, (2015).
237. Turner, D. A. *et al.* Wnt/ β -catenin and FGF signalling direct the specification and maintenance of a neuromesodermal axial progenitor in ensembles of mouse embryonic stem cells. *Development (Cambridge)* **141**, (2014).
238. Tsakiridis, A. & Wilson, V. Assessing the bipotency of in vitro-derived neuromesodermal progenitors. *F1000Res* **4**, (2015).
239. Edri, S., Hayward, P., Baillie-Johnson, P., Steventon, B. J. & Arias, A. M. An epiblast stem cell-derived multipotent progenitor population for axial extension. *Development (Cambridge)* **146**, (2019).
240. Gogolou, A. *et al.* Early anteroposterior regionalisation of human neural crest is shaped by a pro-mesodermal factor. *Elife* **11**, (2022).
241. Savory, J. G. A. *et al.* Cdx2 regulation of posterior development through non-Hox targets. *Development* **136**, (2009).
242. Blassberg, R. *et al.* Sox2 levels regulate the chromatin occupancy of WNT mediators in epiblast progenitors responsible for vertebrate body formation. *Nat Cell Biol* **24**, (2022).
243. Pijuan-Sala, B. *et al.* A single-cell molecular map of mouse gastrulation and early organogenesis. *Nature* **566**, (2019).
244. Iwafuchi-Doi, M. *et al.* Transcriptional regulatory networks in epiblast cells and during anterior neural plate development as modeled in epiblast stem cells. *Development (Cambridge)* **139**, (2012).
245. Anand, G. M. *et al.* Controlling organoid symmetry breaking uncovers an excitable system underlying human axial elongation. *Cell* **186**, (2023).
246. Iyer, N. R. *et al.* Modular derivation of diverse, regionally discrete human posterior CNS neurons enables discovery of transcriptomic patterns. *Sci Adv* **8**, (2022).
247. Gupta, S. *et al.* In vitro atlas of dorsal spinal interneurons reveals Wnt signaling as a critical regulator of progenitor expansion. *Cell Rep* **40**, (2022).
248. Bansod, S., Kageyama, R. & Ohtsuka, T. Hes5 regulates the transition timing of neurogenesis and gliogenesis in mammalian neocortical development. *Development (Cambridge)* **144**, (2017).

249. Borghese, L. *et al.* Inhibition of notch signaling in human embryonic stem cell-derived neural stem cells delays G1/S phase transition and accelerates neuronal differentiation in vitro and in vivo. *Stem Cells* **28**, (2010).
250. Yamanaka, Y. *et al.* Reconstituting human somitogenesis in vitro. *Nature* **614**, (2023).
251. Faustino Martins, J. M. *et al.* Self-Organizing 3D Human Trunk Neuromuscular Organoids. *Cell Stem Cell* **26**, (2020).
252. Mavrommatis, L. *et al.* Human skeletal muscle organoids model fetal myogenesis and sustain uncommitted PAX7 myogenic progenitors. (2023) doi:10.7554/elife.87081.1.
253. Gomez, G. A. *et al.* WNT/ β -catenin modulates the axial identity of embryonic stem cell-derived human neural crest. *Development (Cambridge)* **146**, (2019).
254. Emrich, S. J., Barbazuk, W. B., Li, L. & Schnable, P. S. Gene discovery and annotation using LCM-454 transcriptome sequencing. *Genome Res* **17**, (2007).
255. Tang, F. *et al.* mRNA-Seq whole-transcriptome analysis of a single cell. *Nat Methods* **6**, (2009).
256. Ziegenhain, C. *et al.* Comparative Analysis of Single-Cell RNA Sequencing Methods. *Mol Cell* **65**, (2017).
257. Ding, J. *et al.* Author Correction: Systematic comparison of single-cell and single-nucleus RNA-sequencing methods (Nature Biotechnology, (2020), 38, 6, (737-746), 10.1038/s41587-020-0465-8). *Nature Biotechnology* vol. 38 Preprint at <https://doi.org/10.1038/s41587-020-0534-z> (2020).
258. La Manno, G. *et al.* RNA velocity of single cells. *Nature* **560**, (2018).
259. Saelens, W., Cannoodt, R., Todorov, H. & Saeys, Y. A comparison of single-cell trajectory inference methods. *Nat Biotechnol* **37**, (2019).
260. Cui, M., Cheng, C. & Zhang, L. High-throughput proteomics: a methodological mini-review. *Laboratory Investigation* vol. 102 Preprint at <https://doi.org/10.1038/s41374-022-00830-7> (2022).
261. Aebersold, R. & Mann, M. Mass spectrometry-based proteomics. *Nature* vol. 422 Preprint at <https://doi.org/10.1038/nature01511> (2003).
262. Sigismondo, G., Papageorgiou, D. N. & Krijgsveld, J. Cracking chromatin with proteomics: From chromatome to histone modifications. *Proteomics* vol. 22 Preprint at <https://doi.org/10.1002/pmic.202100206> (2022).
263. Holmdahl, D. E. Experimentelle Untersuchungen über die Lage der Grenze zwischen primärer und sekundärer Körperentwicklung beim Huhn. *Anat. Anz* **59**, 6 (1925).
264. Rao, J. *et al.* Erratum: Stepwise Clearance of Repressive Roadblocks Drives Cardiac Induction in Human ESCs (Cell Stem Cell (2016) 18 (341-353)). *Cell Stem Cell* vol. 18 Preprint at <https://doi.org/10.1016/j.stem.2016.03.008> (2016).
265. Kunze, C. *et al.* Synthetic AAV/CRISPR vectors for blocking HIV-1 expression in persistently infected astrocytes. *Glia* **66**, (2018).
266. Geuder, J. *et al.* A non-invasive method to generate induced pluripotent stem cells from primate urine. *Sci Rep* **11**, (2021).
267. Grosch, M. *et al.* Nucleus size and DNA accessibility are linked to the regulation of paraspeckle formation in cellular differentiation. *BMC Biol* **18**, (2020).
268. Choi, I. Y. *et al.* Concordant but Varied Phenotypes among Duchenne Muscular Dystrophy Patient-Specific Myoblasts Derived using a Human iPSC-Based Model. *Cell Rep* **15**, (2016).
269. Morizane, R. & Bonventre, J. V. Generation of nephron progenitor cells and kidney organoids from human pluripotent stem cells. *Nat Protoc* **12**, (2017).
270. Lian, X. *et al.* Directed cardiomyocyte differentiation from human pluripotent stem cells by modulating Wnt/ β -catenin signaling under fully defined conditions. *Nat Protoc* **8**, (2013).
271. Haldane, J. & Julie Jerber James Haldane, J. S. D. P. M. P. Dissociation of neuronal culture to single cells for scRNA-seq (10x Genomics) . *protocols.io* (2020).
272. Kang, H. M. *et al.* Multiplexed droplet single-cell RNA-sequencing using natural genetic variation. *Nat Biotechnol* **36**, (2018).
273. Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: Large-scale single-cell gene expression data analysis. *Genome Biol* **19**, (2018).
274. Bernstein, N. J. *et al.* Solo: Doublet Identification in Single-Cell RNA-Seq via Semi-Supervised Deep Learning. *Cell Syst* **11**, (2020).

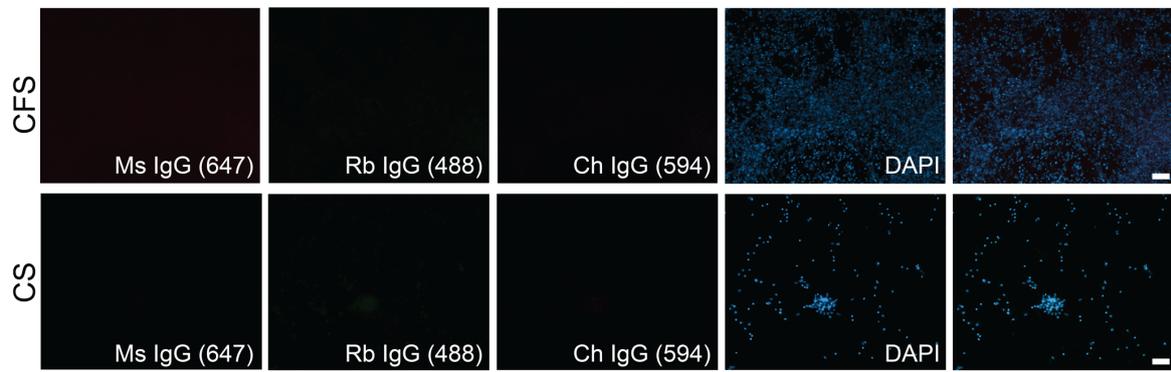
275. Lakkis, J. *et al.* A joint deep learning model enables simultaneous batch effect correction, denoising, and clustering in single-cell transcriptomics. *Genome Res* **31**, (2021).
276. Rayon, T., Maizels, R. J., Barrington, C. & Briscoe, J. Single-cell transcriptome profiling of the human developing spinal cord reveals a conserved genetic programme with human-specific features. *Development (Cambridge)* **148**, (2021).
277. Qiu, C. *et al.* Systematic reconstruction of cellular trajectories across mouse embryogenesis. *Nat Genet* **54**, (2022).
278. Rappsilber, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat Protoc* **2**, (2007).
279. Demichev, V., Messner, C. B., Vernardis, S. I., Lilley, K. S. & Ralser, M. DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. *Nat Methods* **17**, (2020).
280. Tyanova, S. *et al.* The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nature Methods* vol. 13 Preprint at <https://doi.org/10.1038/nmeth.3901> (2016).
281. López-Anguita, N. *et al.* Hypoxia induces an early primitive streak signature, enhancing spontaneous elongation and lineage representation in gastruloids. *Development (Cambridge)* **149**, (2022).
282. Bertrand, N., Medevielle, F. & Pituello, F. FGF signalling controls the timing of Pax6 activation in the neural tube. *Development* **127**, (2000).
283. Armstrong, L. *et al.* Overexpression of Telomerase Confers Growth Advantage, Stress Resistance, and Enhanced Differentiation of ESCs Toward the Hematopoietic Lineage. *Stem Cells* **23**, (2005).
284. Lolas, M., Valenzuela, P. D. T., Tjian, R. & Liu, Z. Charting Brachyury-mediated developmental pathways during early mouse embryogenesis. *Proc Natl Acad Sci U S A* **111**, (2014).
285. Felfly, H. & Klein, O. D. Sprouty genes regulate proliferation and survival of human embryonic stem cells. *Sci Rep* **3**, (2013).
286. Lee, J.-Y. *et al.* Novel Function of Sprouty4 as a Regulator of Stemness and Differentiation of Embryonic Stem Cells. *Dev Reprod* **20**, (2016).
287. Chambers, D. & Mason, I. Expression of sprouty2 during early development of the chick embryo is coincident with known sites of FGF signalling. *Mech Dev* **91**, (2000).
288. Arai, Y. *et al.* Role of Fabbp7, a downstream gene of Pax6, in the maintenance of neuroepithelial cells during early embryonic development of the rat cortex. *Journal of Neuroscience* **25**, (2005).
289. Vesque, C., Anselme, I., Couvé, E., Charnay, P. & Schneider-Maunoury, S. Cloning of vertebrate Protogenin (Prtg) and comparative expression analysis during axis elongation. *Developmental Dynamics* **235**, (2006).
290. El Yakoubi, W. *et al.* Hes4 controls proliferative properties of neural stem cells during retinal ontogenesis. *Stem Cells* **30**, (2012).
291. Berry, F. B. *et al.* Positive and Negative Regulation of Myogenic Differentiation of C2C12 Cells by Isoforms of the Multiple Homeodomain Zinc Finger Transcription Factor ATBF1. *Journal of Biological Chemistry* **276**, (2001).
292. Jung, C. G. *et al.* Homeotic factor ATBF1 induces the cell cycle arrest associated with neuronal differentiation. *Development* **132**, (2005).
293. Russ, D. E. *et al.* A harmonized atlas of mouse spinal cord cell types and their spatial organization. *Nat Commun* **12**, (2021).
294. Hemmi, K. *et al.* A homeodomain-zinc finger protein, ZFHx4, is expressed in neuronal differentiation manner and suppressed in muscle differentiation manner. *Biol Pharm Bull* **29**, (2006).
295. Okano, M., Bell, D. W., Haber, D. A. & Li, E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99**, (1999).
296. Leonhardt, H., Page, A. W., Weier, H. U. & Bestor, T. H. A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. *Cell* **71**, (1992).
297. Ren, Y. Regulatory mechanism and biological function of UHRF1-DNMT1-mediated DNA methylation. *Functional and Integrative Genomics* vol. 22 Preprint at <https://doi.org/10.1007/s10142-022-00918-9> (2022).
298. Wu, X. & Zhang, Y. TET-mediated active DNA demethylation: Mechanism, function and beyond. *Nature Reviews Genetics* vol. 18 Preprint at <https://doi.org/10.1038/nrg.2017.33> (2017).

299. Dixon, G. *et al.* QSER1 protects DNA methylation valleys from de novo methylation. *Science (1979)* **372**, (2021).
300. Foy, R. L. *et al.* Role of Jade-1 in the histone acetyltransferase (HAT) HBO1 complex. *Journal of Biological Chemistry* **283**, (2008).
301. Yilmaz, A. & Benvenisty, N. Defining Human Pluripotency. *Cell Stem Cell* vol. 25 Preprint at <https://doi.org/10.1016/j.stem.2019.06.010> (2019).
302. Dillon, A. K. *et al.* UNC5C is required for spinal accessory motor neuron development. *Molecular and Cellular Neuroscience* **35**, (2007).
303. Roome, R. B. *et al.* Phox2a Defines a Developmental Origin of the Anterolateral System in Mice and Humans. *Cell Rep* **33**, (2020).
304. Chalamalasetty, R. B. *et al.* The Wnt3a/ β -catenin target gene Mesogenin1 controls the segmentation clock by activating a Notch signalling program. *Nat Commun* **2**, (2011).
305. Straface, G. *et al.* Sonic hedgehog regulates angiogenesis and myogenesis during post-natal skeletal muscle regeneration. *J Cell Mol Med* **13**, (2009).
306. Lee, H. *et al.* Identification of Small Molecules Which Induce Skeletal Muscle Differentiation in Embryonic Stem Cells via Activation of the Wnt and Inhibition of Smad2/3 and Sonic Hedgehog Pathways. *Stem Cells* **34**, (2016).
307. Row, R. H. *et al.* BMP and FGF signaling interact to pattern mesoderm by controlling basic helix-loop-helix transcription factor activity. *Elife* **7**, (2018).
308. Sakurai, H. *et al.* In vitro modeling of paraxial mesodermal progenitors derived from induced pluripotent stem cells. *PLoS One* **7**, e47078 (2012).
309. Fu, X., Matsui, T. & Funaba, M. Enhancement of vitamin C-induced myogenesis by inhibition of extracellular signal-regulated kinase (ERK) 1/2 pathway. *Biochem Biophys Res Commun* **612**, (2022).
310. Lautherbach, N. *et al.* Urocortin 2 promotes hypertrophy and enhances skeletal muscle function through cAMP and insulin/IGF-1 signaling pathways. *Mol Metab* **60**, (2022).
311. Lawson, M. A. & Purslow, P. P. Differentiation of myoblasts in serum-free media: Effects of modified media are cell line-specific. *Cells Tissues Organs* **167**, (2000).
312. Song, L., Chen, J., Peng, G., Tang, K. & Jing, N. Dynamic heterogeneity of brachyury in mouse epiblast stem cells mediates distinct response to extrinsic bone morphogenetic protein (bmp) signaling. *Journal of Biological Chemistry* **291**, (2016).
313. Tyser, R. C. V *et al.* Single-cell transcriptomic characterization of a gastrulating human embryo. *Nature* **600**, (2021).
314. Maherali, N. & Hochedlinger, K. Tgf β Signal Inhibition Cooperates in the Induction of iPSCs and Replaces Sox2 and cMyc. *Current Biology* **19**, (2009).
315. Steens, J. & Klein, D. HOX genes in stem cells: Maintaining cellular identity and regulation of differentiation. *Frontiers in Cell and Developmental Biology* vol. 10 Preprint at <https://doi.org/10.3389/fcell.2022.1002909> (2022).
316. Zhang, X. *et al.* Pax6 is a human neuroectoderm cell fate determinant. *Cell Stem Cell* **7**, (2010).
317. Cui, G. *et al.* Mouse gastrulation: Attributes of transcription factor regulatory network for epiblast patterning. *Dev Growth Differ* **60**, (2018).
318. Takebayashi-Suzuki, K., Kitayama, A., Terasaka-Iioka, C., Ueno, N. & Suzuki, A. The forkhead transcription factor FoxB1 regulates the dorsal-ventral and anterior-posterior patterning of the ectoderm during early *Xenopus* embryogenesis. *Dev Biol* **360**, (2011).
319. Leung, A. W. *et al.* Pre-Border Gene Foxb1 Regulates the Differentiation Timing and Autonomic Neuronal Potential of Human Neural Crest Cells. *bioRxiv* (2019) doi:10.1101/646026.
320. Mao, J., McGlinn, E., Huang, P., Tabin, C. J. & McMahon, A. P. Fgf-Dependent Etv4/5 Activity Is Required for Posterior Restriction of Sonic hedgehog and Promoting Outgrowth of the Vertebrate Limb. *Dev Cell* **16**, (2009).
321. Kerosuo, L. *et al.* Myc increases self-renewal in neural progenitor cells through Miz-1. *J Cell Sci* **121**, (2008).
322. Yang, M. *et al.* Lin28 promotes the proliferative capacity of neural progenitor cells in brain development. *Development (Cambridge)* **142**, (2015).

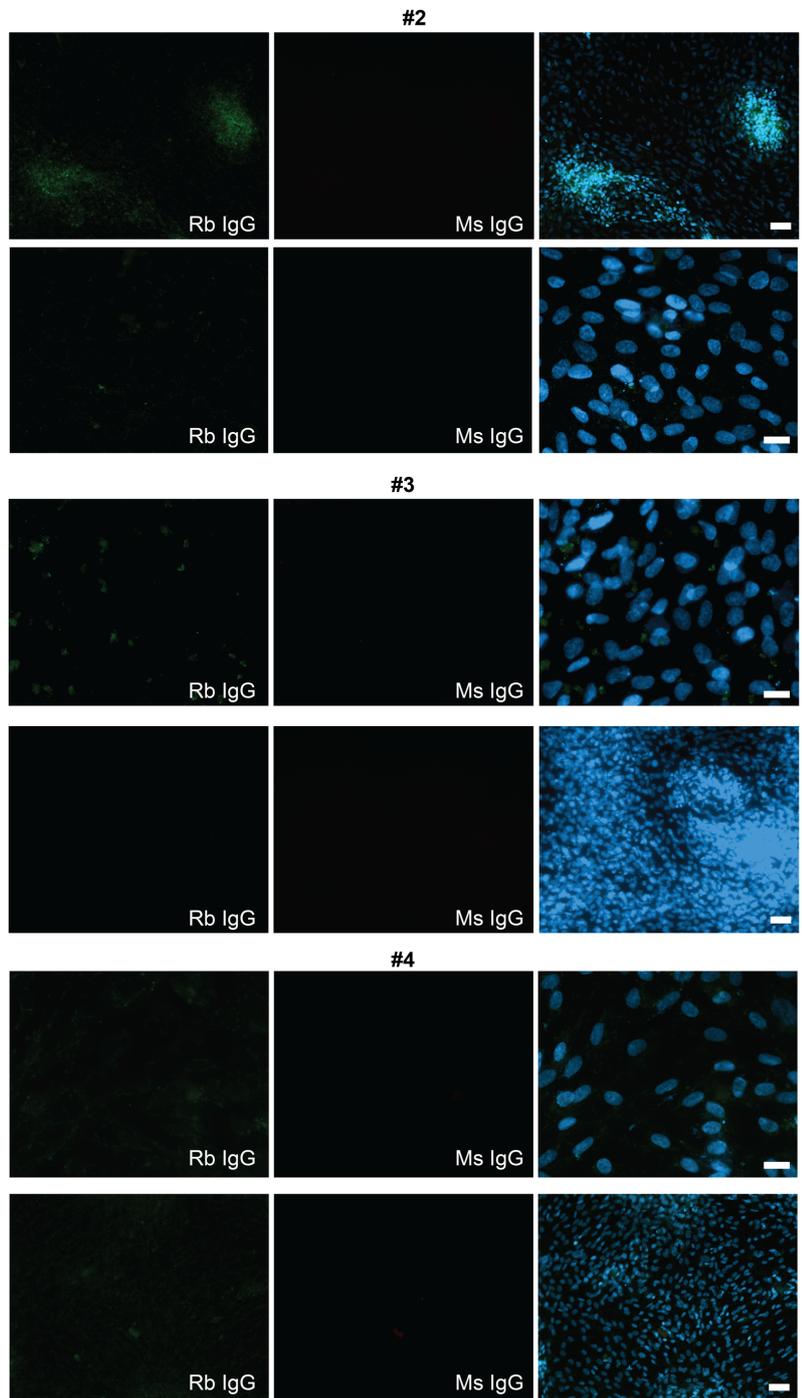
323. Fagnocchi, L. *et al.* A Myc-driven self-reinforcing regulatory network maintains mouse embryonic stem cell identity. *Nat Commun* **7**, (2016).
324. Berdasco, M. & Esteller, M. DNA methylation in stem cell renewal and multipotency. *Stem Cell Research and Therapy* vol. 2 Preprint at <https://doi.org/10.1186/scrt83> (2011).
325. Bostick, M. *et al.* UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science (1979)* **317**, (2007).
326. Moore, L. D., Le, T. & Fan, G. DNA methylation and its basic function. *Neuropsychopharmacology* vol. 38 Preprint at <https://doi.org/10.1038/npp.2012.112> (2013).
327. Pastor, W. A., Aravind, L. & Rao, A. TETonic shift: Biological roles of TET proteins in DNA demethylation and transcription. *Nature Reviews Molecular Cell Biology* vol. 14 Preprint at <https://doi.org/10.1038/nrm3589> (2013).
328. Yang, J., Bashkenova, N., Zang, R., Huang, X. & Wang, J. The roles of TET family proteins in development and stem cells. *Development (Cambridge)* **147**, (2020).
329. Costa, Y. *et al.* NANOG-dependent function of TET1 and TET2 in establishment of pluripotency. *Nature* **495**, (2013).
330. Hu, C. *et al.* Energy metabolism plays a critical role in stem cell maintenance and differentiation. *International Journal of Molecular Sciences* vol. 17 Preprint at <https://doi.org/10.3390/ijms17020253> (2016).
331. Rasmussen, M. L., Ortolano, N. A., Romero-Morales, A. I. & Gama, V. Wnt signaling and its impact on mitochondrial and cell cycle dynamics in pluripotent stem cells. *Genes* vol. 9 Preprint at <https://doi.org/10.3390/genes9020109> (2018).
332. Dillon, A. K. *et al.* Molecular control of spinal accessory motor neuron/axon development in the mouse spinal cord. *Journal of Neuroscience* **25**, (2005).
333. Litingtung, Y. & Chiang, C. Control of Shh activity and signaling in the neural tube. *Developmental Dynamics* vol. 219 Preprint at [https://doi.org/10.1002/1097-0177\(2000\)9999:9999::AID-DVDY1050>3.3.CO;2-H](https://doi.org/10.1002/1097-0177(2000)9999:9999::AID-DVDY1050>3.3.CO;2-H) (2000).
334. Pourquié, O. The segmentation clock: Converting embryonic time into spatial pattern. *Science* vol. 301 Preprint at <https://doi.org/10.1126/science.1085887> (2003).
335. Diaz-Cuadros, M. & Pourquie, O. In vitro systems: A new window to the segmentation clock. *Development Growth and Differentiation* vol. 63 Preprint at <https://doi.org/10.1111/dgd.12710> (2021).
336. Wiese, C. B., Deal, K. K., Ireland, S. J., Cantrell, V. A. & Southard-Smith, E. M. Migration pathways of sacral neural crest during development of lower urogenital tract innervation. *Dev Biol* **429**, (2017).
337. Itäranta, P., Viiri, K., Kaartinen, V. & Vainio, S. Lumbo-sacral neural crest derivatives fate mapped with the aid of Wnt-1 promoter integrate but are not essential to kidney development. *Differentiation* **77**, (2009).
338. Shaker, M. R., Lee, J. H. & Sun, W. Embryonal neuromesodermal progenitors for caudal central nervous system and tissue development. *J Korean Neurosurg Soc* **64**, (2021).
339. Kageyama, R., Ohtsuka, T. & Kobayashi, T. The Hes gene family: Repressors and oscillators that orchestrate embryogenesis. *Development* **134**, (2007).
340. Lee, A. S., Tang, C., Rao, M. S., Weissman, I. L. & Wu, J. C. Tumorigenicity as a clinical hurdle for pluripotent stem cell therapies. *Nature Medicine* vol. 19 Preprint at <https://doi.org/10.1038/nm.3267> (2013).
341. Gutierrez-Aranda, I. *et al.* Human induced pluripotent stem cells develop teratoma more efficiently and faster than human embryonic stem cells regardless the site of injection. *Stem Cells* vol. 28 Preprint at <https://doi.org/10.1002/stem.471> (2010).
342. Soufi, A. *et al.* Pioneer transcription factors target partial DNA motifs on nucleosomes to initiate reprogramming. *Cell* **161**, (2015).
343. González, F., Boué, S. & Belmonte, J. C. I. Methods for making induced pluripotent stem cells: Reprogramming à la carte. *Nature Reviews Genetics* vol. 12 Preprint at <https://doi.org/10.1038/nrg2937> (2011).



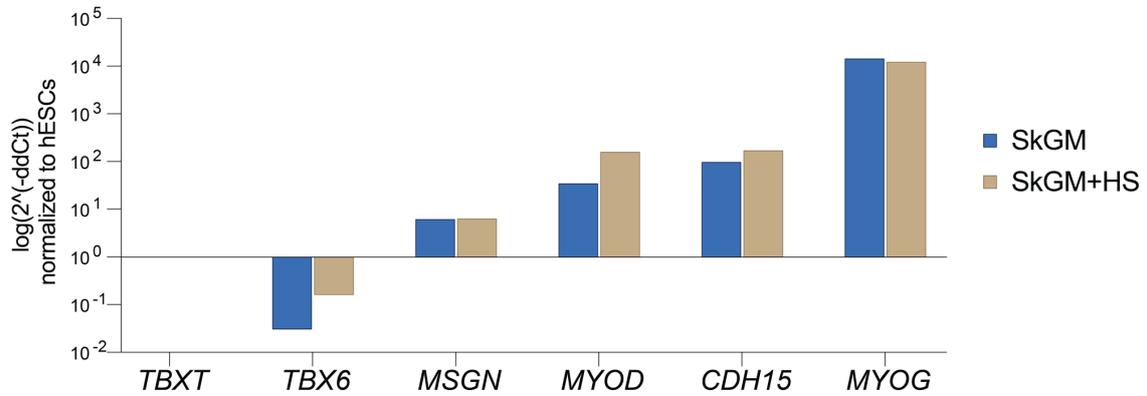
Supplementary Figure 2: Representative immunostaining images from undifferentiated H9 cells stained for ZIC2 and DAPI (A), and isotype controls from undifferentiated H9, human CFS and human CS cells stained with rabbit (488) IgG (B) (Rb: rabbit, scale bars: 50 μ M).



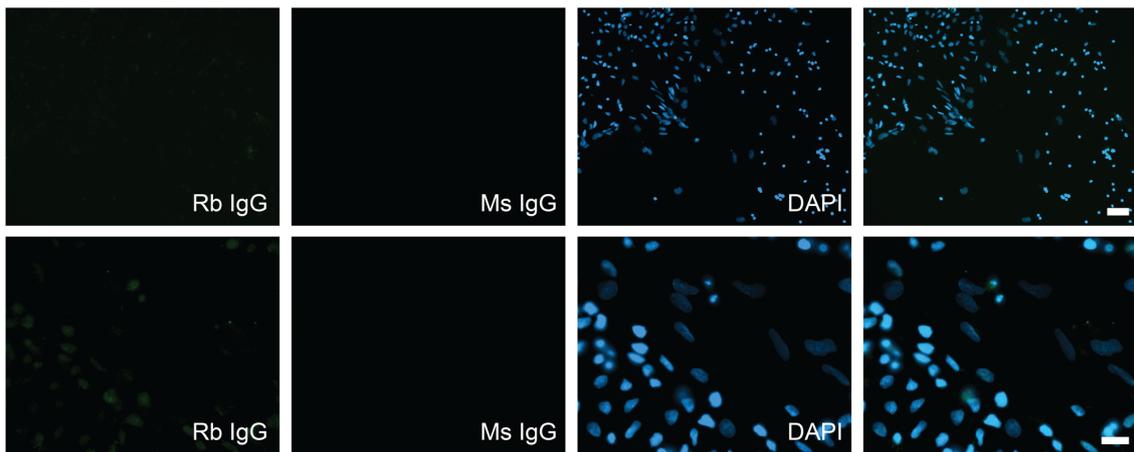
Supplementary Figure 3: Representative isotype control images from day 28 human CFS and human CS neural differentiating cultures stained with mouse (647), rabbit (488) and chicken (594) IgG (Ms: mouse, Rb: rabbit, Ch: Chicken, scale bars: 50 μ M).



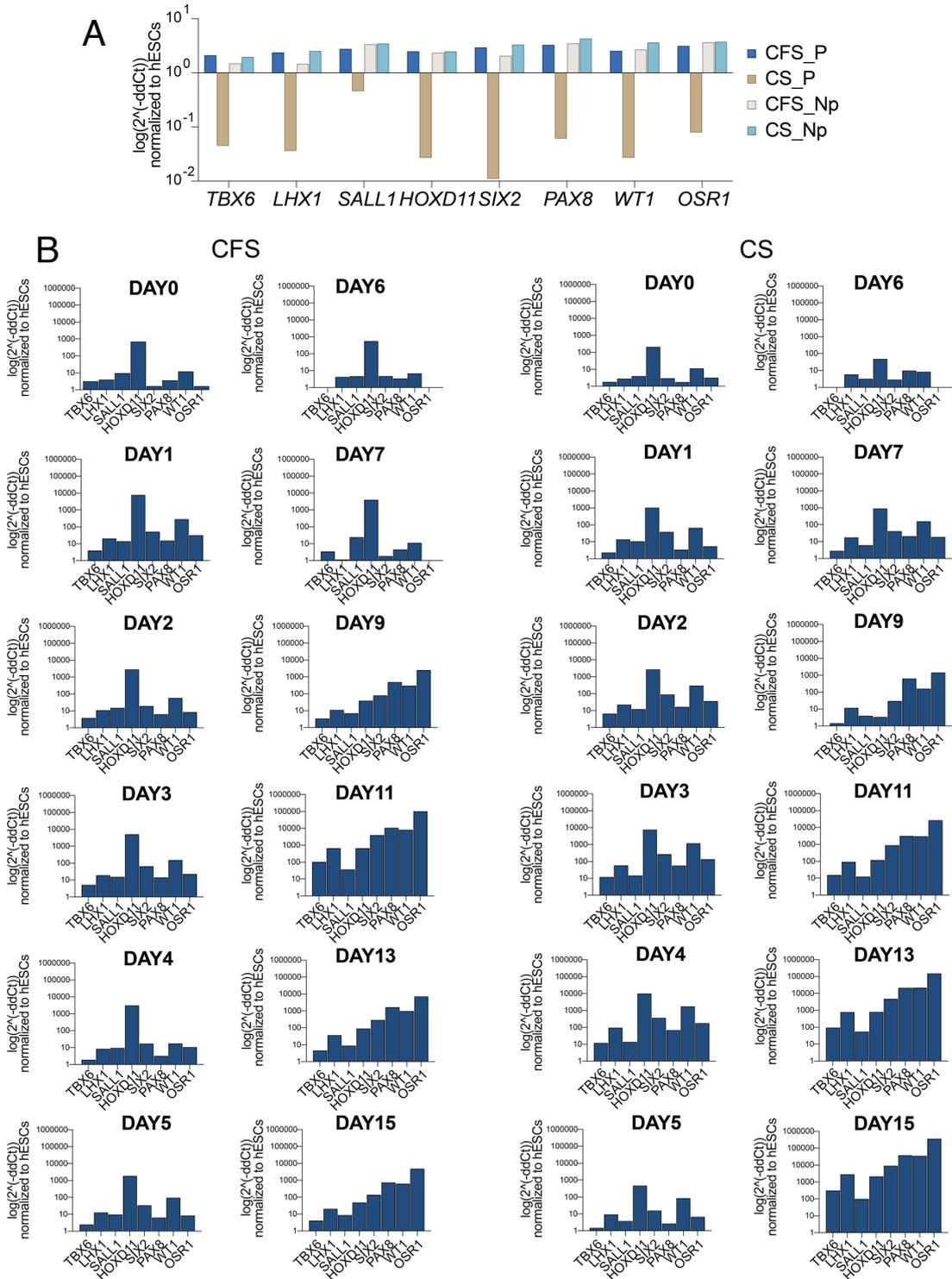
Supplementary Figure 4: Representative isotype control images from day 40 human CFS differentiating cultures into skeletal muscle cells (condition #2-4 in Table 9) stained with rabbit (488) and mouse (647) IgG (Rb: rabbit, Ms: mouse, scale bars: 50 μ M for lower magnification and 20 μ M for higher magnification).



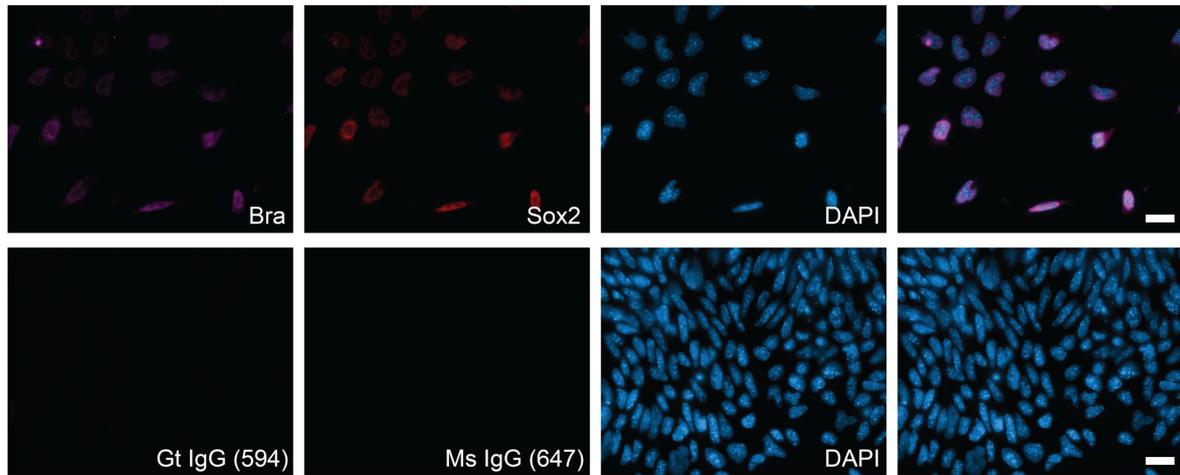
Supplementary Figure 5: Expression analysis of the selected stage-specific markers for mesodermal development in human primary skeletal muscle culture maintained in skeletal muscle growth medium (SkGM) with or without horse serum (HS). The Ct values were normalized to undifferentiated H9 cells.



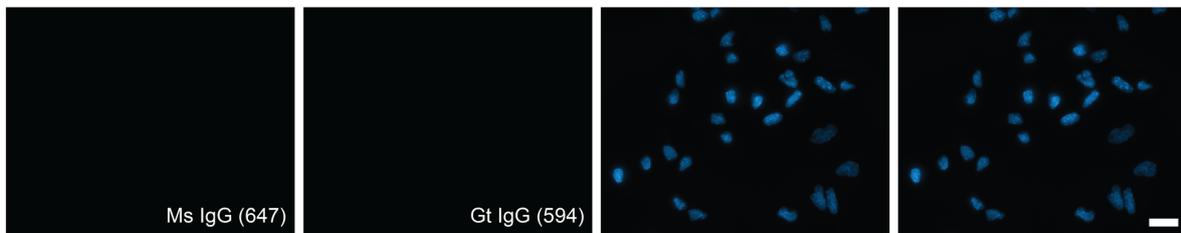
Supplementary Figure 6: Representative isotype control images from day 35 human CFS differentiating cultures into skeletal muscle cells stained with rabbit (488) and mouse (647) IgG (Rb: rabbit, Ms: mouse, scale bars: 50 μ M for lower magnification and 20 μ M for higher magnification).



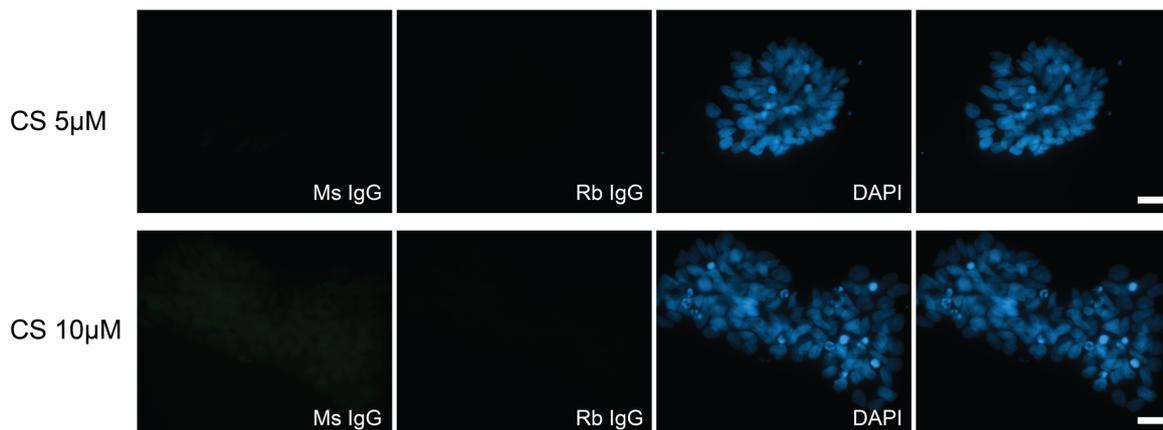
Supplementary Figure 7: Expression analysis in human CFS and human CS nephron differentiation cultures on day 15 by using the original protocol (A) and time course by using the modified protocol in Fig. 36 (B). The Ct values were normalized to undifferentiated H9 cells (P: Parental, Np: Nephron differentiation culture).



Supplementary Figure 8: Representative immunostaining images from undifferentiated EpiSCs stained for Bra, Sox2, and DAPI (upper panel), and isotype controls from undifferentiated EpiSCs stained with goat (594) and mouse (647) IgG (lower panel) (Gt: Goat, Ms: mouse, scale bars: 20 μ M).



Supplementary Figure 9: Representative isotype control images from mouse CFS (#1) cells stained with goat (594) and mouse (647) IgG (Gt: Goat, Ms: mouse, scale bar: 20 μ M).



Supplementary Figure 10: Representative isotype control images from undifferentiated orangutan CS cells stained with mouse (488) and rabbit (647) IgG (Ms: mouse, Rb: rabbit). Upper and lower panels represent CS cells induced by 5 μ M and 10 μ M CHIR99021 respectively (scale bars: 20 μ M).

ACKNOWLEDGEMENTS

I would like to sincerely thank my supervisor, Prof. Dr. Micha Drukker, for granting me the opportunity to carry out my research under his guidance. His perspective and provision of freedom enabled me to pursue my scientific curiosity while shaping the direction of this work. I would like to extend my sincere gratitude to my thesis committee members, Prof. Dr. Magdalena Götz and Prof. Dr. Dominik Paquet, for their great support, valuable feedback, and constant inspiration. I am also deeply thankful to German Academic Exchange Service-DAAD for their support throughout my PhD studies, which not only alleviated the financial challenges but also paved the way for me to continue my career in Germany.

My gratitude extends to all members of ISF-P, whose endless support and motivation created an exceptional working atmosphere that I will always cherish. I would like to express my deepest appreciation to Ejona Rusha. This study would not have been possible without her companionship and brainstorming sessions about the project. Her patience in motivating me and helping me navigate through challenges was immeasurable. Beyond scientific discussions, her friendship, which I hope will be everlasting, was truly priceless. I am extremely grateful to Polyxeni Nteli, my dear Xenia. No one can deny that we were a moral support duo! I will miss all the chat -you name it, but I prefer calling it information transfer- that we had during the breaks. I am furthermore thankful to Alexandra de la Porte for her friendship during our PhD journey and her assistance with all the German documents; who knows, she may have even saved my life! I give a big thank to Dr. Anna Pertek for her encouragement all the time, and always making me laugh. Even though we spent a short time together, it was an absolute pleasure to work with Sebastian Ittermann, Chaido Ori, Markus Grosch and Dr. Friederike Matheus. I am grateful to them for sharing their knowledge and offering guidance when I was new in the lab. I must also express my gratitude to all the collaborators, especially Enes Uğur, who made invaluable contributions to the success of this project. Moreover, I very much appreciate Moritz Thomas who helped me a lot to get started on data analysis.

Beyond academic guidance, the completion of this thesis hinged on immense emotional support. I could not have undertaken this journey without Onur Özdemir. I am wholeheartedly grateful for his boundless encouragement, patience and understanding. I believe he also deserves a Ph.D. degree, perhaps not in science but definitely in psychology. I am deeply thankful to my parents, Emel Kelle and Hüseyin Kelle, for their sacrifices and unwavering support at every step of my life. I am truly privileged to have a family that not only encourages me in my both scientific and personal life, but also raised me to be an open-minded and independent woman. Special thanks go to my sister, Damla Kelle, for making life endurable and for always being with me no matter where she is. I will keep our motto forever, 'we are together, this is special'. I would also like to thank and commemorate my late grandfather, whose encouragement propelled me forward on my scientific journey.

Many thanks to my friends, Burcu Sucu, Yağız Akdaş, Melike Korkmaz and Öztaç Korkmaz, for their moral support and always making time for long discussions. Furthermore, I am profoundly indebted to Barış Kalaycı, who has been my life coach from time to time, guiding me along this path.

Last but not least, I want to thank me for doing all this hard work!

AFFIDAVIT



LUDWIG-
MAXIMILIANS-
UNIVERSITÄT
MÜNCHEN

Dekanat Medizinische Fakultät
Promotionsbüro



Affidavit

Kelle Özdemir, Dolunay

Surname, first name

I hereby declare, that the submitted thesis entitled

Establishment of novel axial stem cell states recapitulating neuromesodermal and dorsal neural tube progenitor identities

is my own work. I have only used the sources indicated and have not made unauthorised use of services of a third party. Where the work of others has been quoted or reproduced, the source is always given.

I further declare that the dissertation presented here has not been submitted in the same or similar form to any other institution for the purpose of obtaining an academic degree.

Neuherberg,
10.08.2023

Place, Date

Dolunay Kelle Özdemir

Signature doctoral candidate

Affidavit

Date: 07.08.2023