CD146-assisted isolation and multi-omics characterisation of human pluripotent stem cell derived cranial Schwann cell precursor-like cells.

Ernst Wolvetang (e.wolvetang@uq.edu.au)
The University of Queensland https://orcid.org/0000-0002-2146-6614

Anushree Balachandran
Children's Cancer Institute

Keyur Dave
QIMR Berghofer Medical Research Institute

Kanupriya Tiwari
The University of Queensland

Cecilia Gomez-Inclan
The University of Queensland

Hannah C. Leeson
Australian Institute for Bioengineering and Nanotechnology, The University of Queensland
https://orcid.org/0000-0002-4694-8513

Shian Su
WEHI

Giovanni Pietrogrande
Australian Institute for Bioengineering and Nanotechnology, The University of Queensland

Dmitry Ovchinnikov
The Ulniversity of Queensland

Owen Hawksworth
The Ulniversity of Queensland

Daniela Amann-Zalcenstein
WEHI

Carola Voss
Helmholtz Center Munich https://orcid.org/0000-0001-5243-4415

Christine Wells
University of Melbourne https://orcid.org/0000-0003-3133-3628

Madeleine. Headlam
QIMR

Justin Cooper-White
The University of Queensland https://orcid.org/0000-0002-1920-8229

Nick Glass
Australian Institute for Bioengineering and Nanotechnology, The University of Queensland
Ilaria Stefani
AIBN
Dedreia Tull
University of Melbourne
Ralf Schittenhelm
Monash University https://orcid.org/0000-0001-8738-1878
Ceilidh Marchant
University of South Australia
Quenten Schwarz
UniSA and SA Pathology https://orcid.org/0000-0002-5958-4181
David De Souza
University of Melbourne https://orcid.org/0000-0002-7225-2632
Malcolm McConville
University of Melbourne https://orcid.org/0000-0002-7107-7887
Lars Nielsen
The University of Queensland https://orcid.org/0000-0001-8191-3511
Jeff Gorman
QIMR Berghofer Medical Research Institute
Shalin Naik
Walter and Eliza Hall Institute of Medical Research https://orcid.org/0000-0003-0299-3301

Article

Keywords: hiPSC, SOX10, cranial Schwann cell precursors, CRISPR/Cas9, proteomics, scRNAseq, metabolomics, splicing, neurocristopathies

Posted Date: September 19th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3324198/v1

Additional Declarations: There is **NO** Competing Interest.

CD146-assisted isolation and multi-omics characterisation of human pluripotent stem cell derived cranial Schwann cell precursor-like cells.

Anushree Balachandran[1], Keyur A. Dave[2], Kanupriya Tiwari[1], Cecilia Gomez-Inclan[1], Hannah Leeson[1], Shian Su[6], Giovanni Pietrogrande[1], Dmitry A. Ovchinnikov[1], Owen Hawksworth[1,3], Daniela Amann-Zalcenstein[6,7], Carola Voss[1], Christine A. Wells[4], Madeleine. J. Headlam[2], Justin Cooper-White[1], Nick Glass[1], Ilaria Stefani[1], Dedreia Tull[8], Ralf B. Schittenhelm[11], Ceilidh Marchant[5], Ines Portella[5], David P. De Souza[8], Malcolm J. McConville[8,9], Lars K. Nielsen[1,10], Jeffrey Gorman[2], Quenten Schwarz[5], Shalin H. Naik[6,7], and Ernst J. Wolvetang[1,*] .

[1]Australian Institute of Bioengineering and Nanotechnology, University of Queensland, Brisbane, Queensland 4072, Australia

[2]QIMR Berghofer Medical Research Institute, Brisbane, Queensland 4006, Australia

[3]School of Biomedical Sciences, University of Queensland, Brisbane, Queensland 4072, Australia

[4]Centre for Stem Cell Systems, MDHS, University of Melbourne, Melbourne, Victoria 3010, Australia

[5] Centre for Cancer Biology, University of South Australia and SA Pathology, Adelaide, South Australia 5001, Australia.

[6] Epigenetics and Development Division (S.S.), Advanced Technology and Biology Division (D.Z.) Immunology Division (S.N.) WEHI, Parkville, Melbourne, Victoria 3052, Australia.

[7] Dept Medical Biology, University of Melbourne, Parkville. Melbourne, Victoria 3010, Australia.

[8]Metabolomics Australia, University of Melbourne, Melbourne, Victoria 3010, Australia.

[9]Department of Biochemistry and Pharmacology, University of Melbourne, Melbourne, Victoria 3010, Australia.

[10]The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark.

[11]Monash Proteomics & Metabolomics Platform, Monash Biomedicine Discovery Institute & Department of Biochemistry and Molecular Biology, Monash University, Clayton, Victoria 3800, Australia.

[*]Correspondence: e.wolvetang@uq.edu.au

Summary

Schwann Cell Precursors (SCPs) are multipotent precursor cells and express SOX10, but relatively little is known about the specification and molecular make-up of human SCPs. To address this, we subjected a human *SOX10* knock-in reporter iPSC line to

a one step cranial Schwann cell precursor differentiation protocol. We show that SOX10-expressing cells acquire the morphology, motility and gene expression of SCP-like cells, and we exemplify that these cells migrate to multiple developing craniofacial structures following injection into early mouse embryos. We next defined the bulk and single-cell transcriptomes, metabolomes, and proteomes of SOX10 expressing human SCP-like cells, revealing gene expression heterogeneity, shifts in splicing events, a reduced dependence on glycolysis, and changes in the expression of cell adhesion proteins, miRNAs and LncRNAs that accompany SCP specification. Discovery proteomics identifies MCAM (CD146) as a cell surface marker that permits the isolation of pure SOX10 expressing human cranial SCP-like cells from human pluripotent stem cell lines subjected to SCP differentiation. We further show that this permits isolation of Down syndrome cranial SCP-like cells that display defects in proliferation. Collectively these data provide a detailed map of the molecular make-up of *in vitro* generated human cranial Schwann cell precursor-like cells and exemplify the utility of MCAM-sorted cranial SCP-like cells for modelling human neurocristopathies.

Key Words

hiPSC, SOX10, cranial Schwann cell precursors, CRISPR/Cas9, proteomics, scRNAseq, metabolomics, splicing, neurocristopathies.

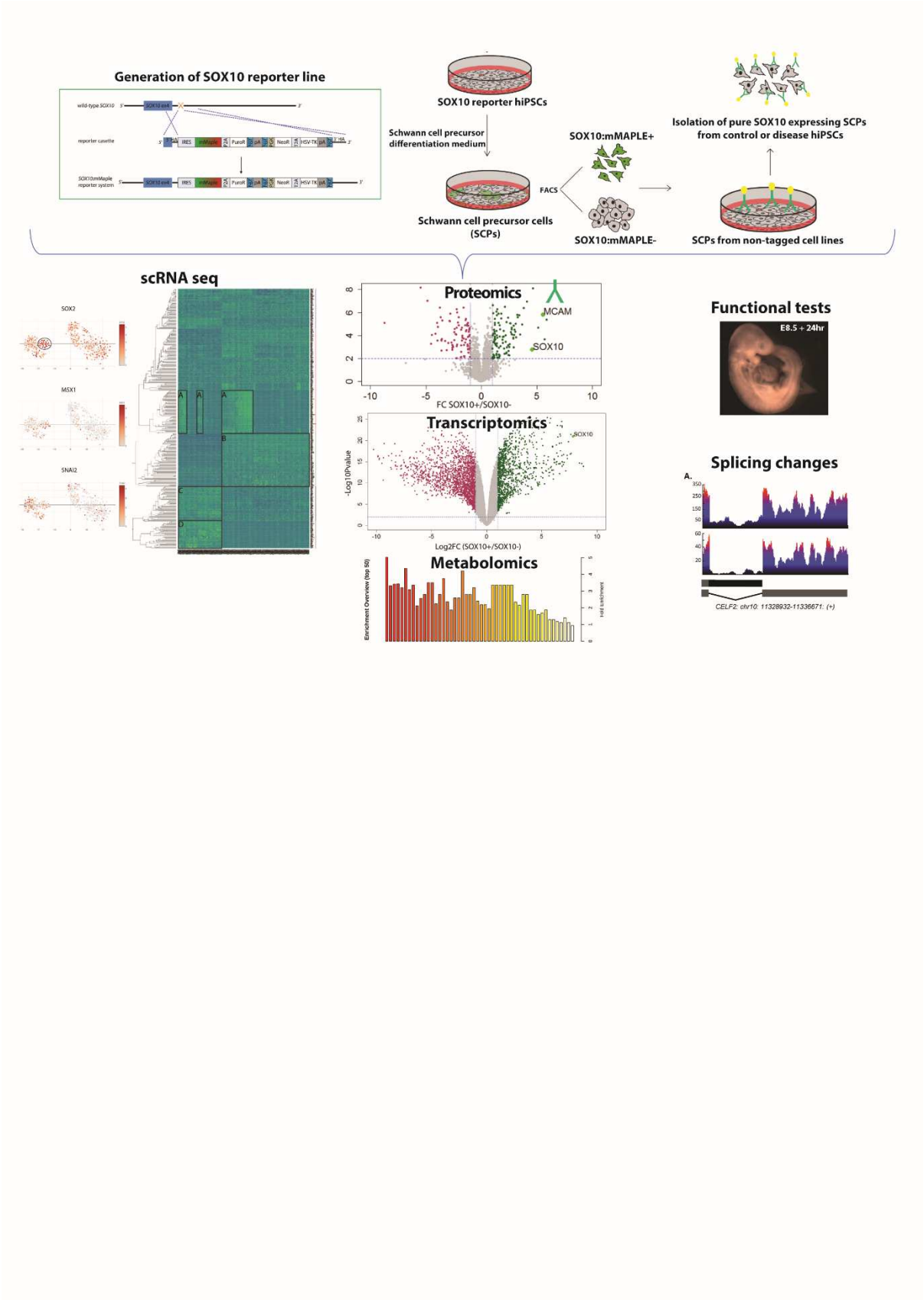
Highlights

- Exemplification of a simple one step cranial Schwann cell precursor differentiation protocol.
- Definition of the proteome, transcriptome, splicing and metabolome changes that accompany *in vitro* generated SOX10 expressing cranial Schwann cell-like precursors.
- Identification of MCAM as a cell surface marker that permits isolation of pure human SOX10 expressing cranial Schwann cell-like precursors that can be used for disease modelling.

In Brief

Balachandran et al. generated a SOX10 reporter hiPSC line and subjected it to a one-step cranial Schwann cell precursor differentiation protocol. Functional assays combined with transcriptomic, metabolomic, splicing, and proteomic analyses were used to define this cranial Schwann cell-like precursor population, and identified MCAM as a specific cell surface marker that permits isolation of these cells from control and disease human pluripotent stem cell lines.

Graphical Abstract



Introduction

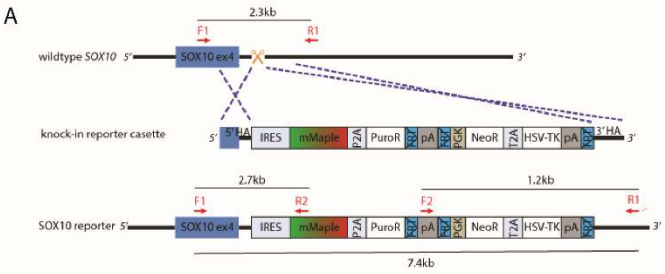
In all chordate embryos, central nervous system development starts with folding and cavitation of the neural plate ectoderm in the midline of the embryo to form the neural tube. In vertebrates, neural crest cells (NCCs) originate from the neural plate border[1], a region of ectoderm at the juncture between neural and non-neural ectoderm. Under the influence of NRG1 and other nerve-derived factors both neural plate border and delaminated NCCs can be specified into multipotent Schwann cell precursors (SCPs) [2,3]. The spatiotemporal evolution of NCCs and the drivers of their cell fate after delamination from the neural tube have been extensively described by Soldatov et al. [4,5]. However, the precise origin of the murine glial SCP population remained unclear, and even less is known about the specification of human SCPs. As their name suggests, SCPs are the progenitors of immature Schwann cells that migrate and proliferate along specified nerve tracts before maturing into Schwann cells that deposit myelin on nerve fibres [6,7]. These multipotent cells also serve as precursors of parasympathetic nerves [8,9]enteric neurons, neuroendocrine cells of the adrenal medulla, the Zuckerkandl organ, endoneurial fibroblasts [10], and various melanocyte populations [11–13]. SCPs are further thought to coordinate neurovascular alignment through their ability to produce CXCL12 and VEGF, and during early craniofacial development can create corridors for migratory nerve precursors that connect the CNS to the developing cranial sensory apparatus and teeth [14–17]. Following detachment from nerve tracts, cranial SCPs retain the ability to generate mesenchymal stem cells that contribute cartilage, bone and fat to various craniofacial structures, including the connective tissue of the ear, eye, and teeth [18–21]. Human SCPs derived from human pluripotent stem cells hold significant potential for regenerative medicine applications given they can promote sciatic nerve regeneration in mice [22]. Furthermore, genetic mutations in SCPs are known to underlie neuroblastoma and neurofibromatosis [23–27], but much remains to be learned about the transcriptional, metabolomic and proteomic make-up of human SCPs. To address this knowledge gap, we isolated pure SOX10 expressing human SCP-like cells [28] by subjecting a Puromycin-selectable SOX10 knock-in reporter hIPSC line to a one-step cranial SCP differentiation protocol. We next used a combination of single-cell and bulk RNAseq analyses to transcriptionally define the SOX10^{pos} SCP-like cells and the SOX10^{neg} neural plate border-like population they emerge from and used discovery proteomics to identify MCAM

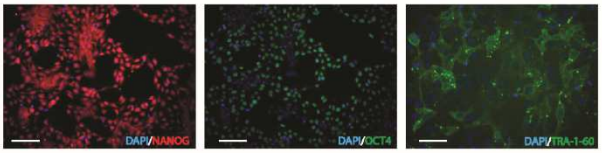
(CD146) as a specific cell surface marker that permits the prospective isolation of pure human cranial SCP-like precursors (CSCP) from genetically unmodified control or disease hiPSC lines. We find that upon acquisition of SOX10 expression, human CSCP adopt a motile phenotype *in vitro* and possess the ability to migrate in the pharyngeal arches following injection into the rhombomere 2 migratory stream of day 8.5 mouse embryos. We demonstrate that the transition into motile SOX10^{pos} SCP involves pronounced shifts in the expression of adhesion molecules, LncRNAs and miRNAs, and is accompanied by alternative splicing events, and a decreased dependence on glycolysis. Collectively this study provides novel insights into human CSCP biology, creates important datasets for further study, and demonstrates MCAM-assisted isolation of these cells as a practical tool that permits investigation of the role(s) of CSCPs in disease conditions.

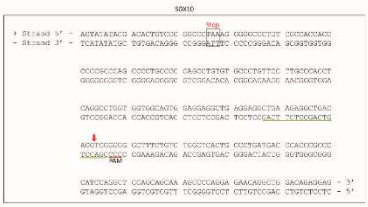
Results

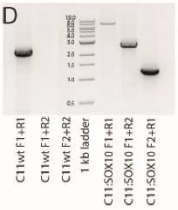
Generation of pure SOX10^{pos} human Schwann cell precursor-like cells.

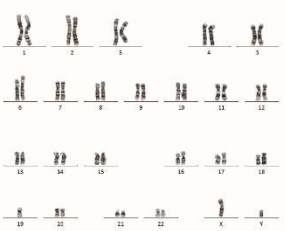
To generate Schwann cell precursor cells (SCPs) under defined culture conditions we modified an existing SCP monolayer differentiation method first described by Kim et al. [22], by adding NRG1 from the very start of differentiation rather than after 6 days of neural differentiation and included the SCP survival and caudalizing factor b-FGF in the medium [29]. This approach allows us to better model the cell-intrinsic driven processes of SCP specification, while recognizing the limitation of lower SCP generation efficiency than that reported by Kim et al. Since SCPs are thought to be derived from and/or co-specified with neural crest cells (NCC) and express SOX10, we generated a clonally derived SOX10 3'UTR knock-in hiPSC reporter line. This reporter line was designed to exhibit locus-appropriate stoichiometric expression between SOX10, mMaple and the puromycin N-acetyl transferase, without affecting the endogenous (haplo-insufficient) SOX10 coding region (for details on the generation and characterisation of the C11:SOX10 reporter line see Suppl. Materials and Methods and Suppl Fig 1).

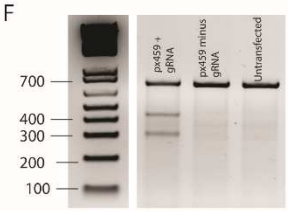

A
2.3kb
F1
R1
wildtype SOX10 5'
SOX10 ex4
3'
knock-in reporter casette
5' HA
IRES
mMaple
P2A
PuroR
pA
NeoR
T2A
HSV-TK
3' HA
2.7kb
1.2kb
F2
R2
SOX10 reporter 5'
7.4kb

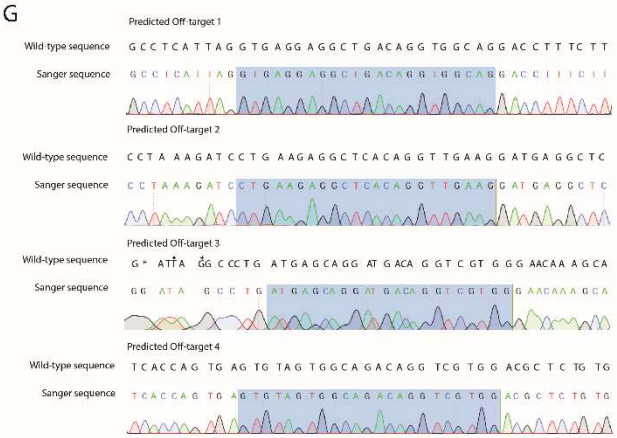

B
DAPI/NANOG
DAPI/OCT4
DAPI/TRA-1-60


C
SOX10
Strn
NdeI


D
C11wt F1+R1
C11wt F1+R2
C11wt F2+R2
1 kb ladder
C11:SOX10 F1+R1
C11:SOX10 F1+R2
C11:SOX10 F2+R1


E


F
700
400
300
200
100
px459 + gRNA
px459 minus gRNA
Untransfected


G
Predicted Off-target 1
Wild-type sequence
Sanger sequence
Predicted Off-target 2
Predicted Off-target 3
Predicted Off-target 4

Suppl Fig 1

Generation of the C11:SOX10 hiPSC line.

(A) Schematic of the knock-in targeting construct and location of the primers used for identification of correctly targeted clones (red arrows). (B) The targeted C11:SOX10 hiPSC line expresses the pluripotency markers NANOG, OCT4 and TRA1-60 (Scale bars = 100µm). (C) Flanking regions of the SOX10 stop codon/upstream 3' UTR were analysed for potential sgRNA target regions. The sgRNA used in this study along with its cutting site and associated PAM sequence is marked. (D) PCR analysis of genomic DNA of the targeted C11:SOX10 hiPSC line with primers anchored within the exogenous cassette sequence within the targeting construct and flanking SOX10 locus sequences exemplifies the correct targeting of the reporter cassette. (E) The correctly-targeted C11:SOX10 hiPSC line displays a normal karyotype (example of 1 of the 15 analyzed metaphase spreads). (F) SURVEYOR nuclease assay demonstrates cutting efficiency of the chosen sgRNA on SOX10 target site in HEK293FT cells. (G) Sanger sequencing cannot detect any mutations or indels in the top four predicted off-target binding sites for the selected gRNA in genomic DNA isolated from the established C11:SOX10 hiPSC line.

Exposure of the C11:SOX10 reporter hIPSC line to our SCP differentiation medium fostered the emergence of ~12% of cells that express mMaple after 21 days (Fig 1A). Photo-switchable mMaple (McEvoy et al. 2012) expression emerged in SOX10 expressing cells confined to small individual islands within the cultures and was readily detectable by fluorescence microscopy (Fig 1B). FACS analysis showed that all mMaple^{pos} cells expressed SOX10 protein, indicating that the reporter faithfully recapitulates endogenous SOX10 expression (Fig 1C). As expected, mMaple expression overlapped with nuclear SOX10 immunoreactivity (Fig 1F-H). Taking advantage of the SOX10-controlled expression of Puromycin N-acetyltransferase we were able to select for cells expressing SOX10:mMaple by exposing day 21 cultures to Puromycin for 4 days (>99% purity) (Fig 1D). Puromycin-selected human SOX10^{pos} SCP-like cells could be expanded in our differentiation medium (without Puromycin) for at least 10 passages (equivalent to a 35-fold expansion over 50 days) without a significant increase in the fraction of SOX10:mMaple negative cells (Fig 1E), highlighting that production of large quantities of cells destined for regenerative medicine applications is feasible.

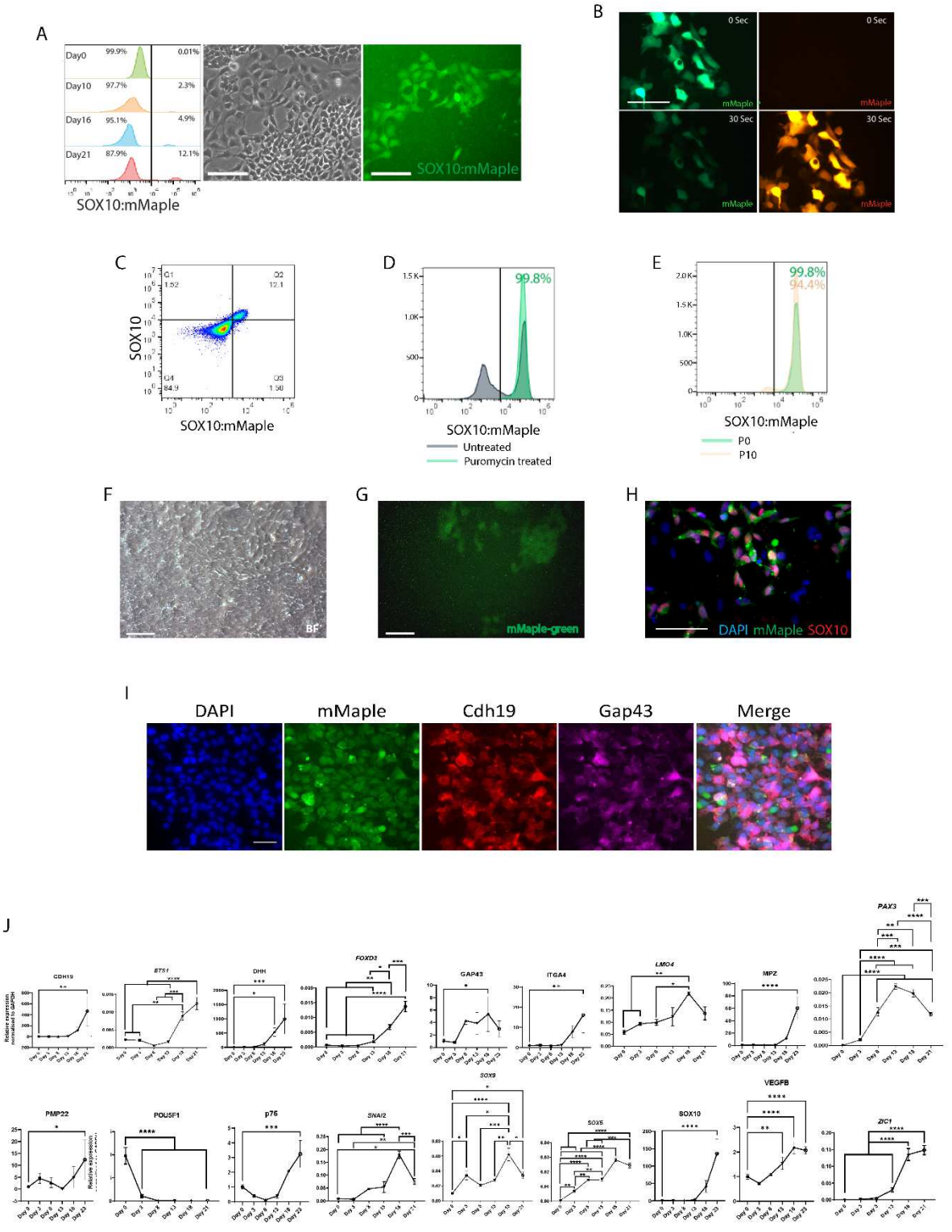
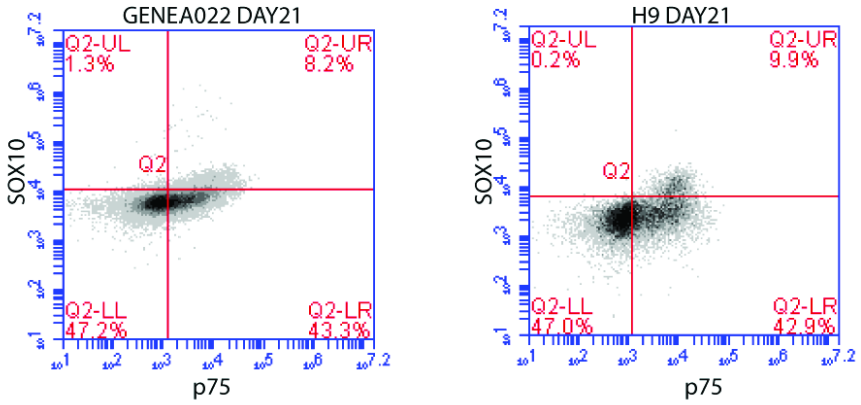

A
Day0
Day10
Day16
Day21
99.9%
0.01%
97.7%
2.3%
95.1%
4.9%
87.9%
12.1%
SOX10:mMaple
SOX10:mMaple
B
0 Sec
0 Sec
30 Sec
30 Sec
mMaple
C
SOX10
SOX10:mMaple
Q1
1.52
Q2
12.1
Q4
84.9
Q3
1.50
D
99.8%
SOX10:mMaple
Untreated
Puromycin treated
E
99.8%
94.4%
SOX10:mMaple
P0
P10
F
BF
G
mMaple-green
H
DAPI mMaple SOX10
I
DAPI
mMaple
Cdh19
Gap43
Merge
J
CDH19
BTG1
DHH
FOXD3
GAP43
ITGA4
LMO4
MPZ
PAX3
PMP22
POU5F1
p75
SNAI2
SOX9
SOX5
SOX10
VEGFB
ZIC1

Figure 1: One-step generation of SCP-like cells from C11:SOX10/mMAPLE reporter hiPSC.

(A) Flow cytometric analysis of mMAPLE expression in C11:SOX10 reporter hIPSC line exposed to our SCP differentiation medium for 0,10,16 and 21 days. (B) mMAPLE fluorescence emerges in small clustered cell populations and is photo-switchable to red fluorescence following 30 sec exposure to 405nm light. (C) Flow cytometric validation of the concordance between mMAPLE expression and SOX10 expression in C11:SOX10 hiPSC subjected to 21 days of SCP differentiation medium. (D) Flow cytometric analysis of mMAPLE expression before and after Puromycin selection. (E) Flow cytometric analysis of mMAPLE expression of sorted SOX10/mMAPLE expressing cells cultured in differentiation medium for 10 passages. (F) Phase contrast (G) mMAPLE fluorescence and (H) Validation of nuclear SOX10 expression via IHC in mMAPLE positive cells counterstained with DAPI (nuclei) Scalebar= 50 µm (I) mMAPLE expressing cells co-express SCP markers CDH19 and Gap43 (J) Q-PCR analysis of the temporal kinetics of SCP and NPB mRNA species in C11:SOX10 hIPSC subjected to SCP-differentiation medium for 21 days (One-way ANOVA. *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001).

To exemplify the general applicability of our one-step SCP-like cell differentiation protocol we next exposed human embryonic stem cell lines H9 [30] and Genea022 [31] as well as the unmodified parental C11 hiPSC line [32] to our differentiation medium for 21 days. This again generated 8-13 % SOX10^{pos} cells (Suppl Fig 2), comparable to the yield of SOX10^{pos} cells obtained with the C11:SOX10 knock-in reporter hiPSC line.

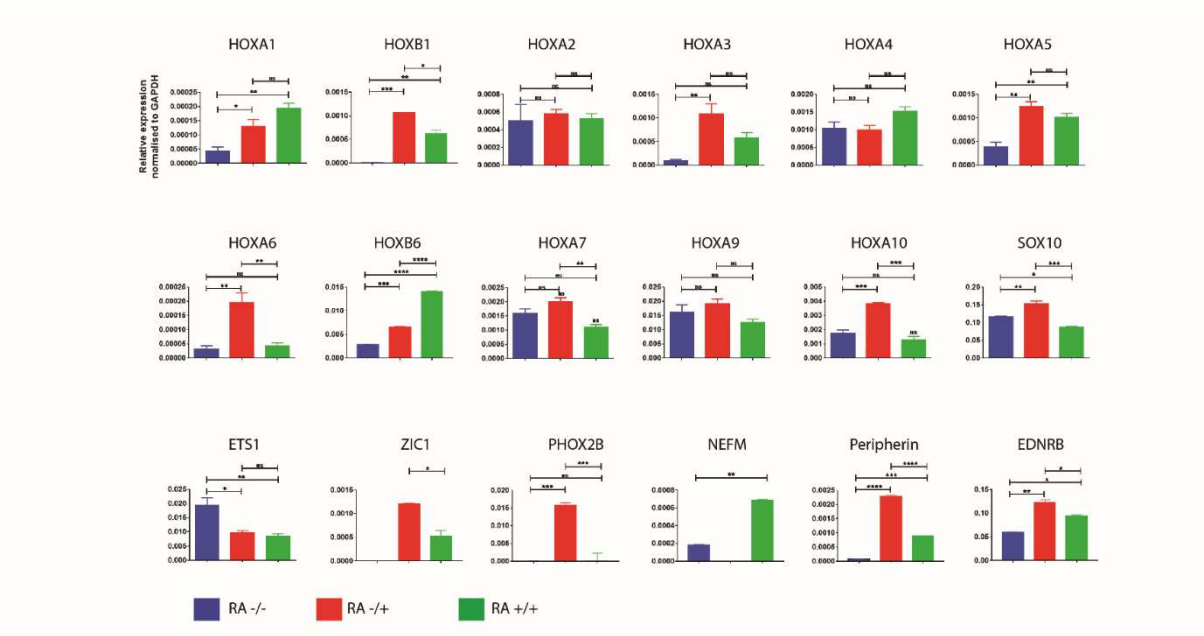


Suppl Fig 2

SCP differentiation protocol applied to other human pluripotent stem cell lines

Flowcytometric analysis of SOX10 expression (using SOX10 antibody labelling) of Gen22 and H9 human ES cells exposed to our one-step SCP differentiation medium for 21 days.

Immunocytochemistry revealed that FACS-sorted mMaple^{pos}/SOX10^{pos} cells expressed the SCP marker proteins GAP43 and CDH19 (Fig 1I). To explore the temporal specification of SOX10^{pos} cells in our cultures we quantified the kinetics of mRNA expression of pluripotency, neural plate border (NPB), neural crest cell (NCC) and SCP genes over the 21-day differentiation time course (Fig 1J). As expected, we observed full extinction of the pluripotency marker OCT4 (POU5F1) by day 8. Expression of ZIC1, a gene expressed in non-neural ectoderm of the NPB, started at day 3 and continued to rise until day 21. Expression of PAX3, a NPB gene, also commenced at day 3, peaked at day 13, and was subsequently downregulated at days 18 and 21. Expression of pre-migratory NCC markers SNAIL2, SOX5 and SOX9 commenced from day 3, peaked at day 18 and subsequently fell at day 21. A similar temporal profile was observed for LMO4, a gene that cooperatively regulates epithelial-mesenchymal transition (EMT) with SNAIL2 [33]. Expression of the cranial NCC marker ETS1 and canonical SCP markers [34] SOX10, CDH19, PMP22, DHH, MPZ and ITGA4, genes that are generally not expressed by neural crest progenitors that colonize ganglia [10,35–37], progressively increased from day 13 and each of these genes peaked at day 21 when mMaple expression became evident (Fig 1A). We did not detect significant expression of posterior HOX genes, as expected for SCP-like cells generated in the absence of retinoic acid, but posterior HOX expression could be upregulated by the inclusion of retinoic acid in the differentiation medium (Suppl Fig 3).



Suppl Fig 3

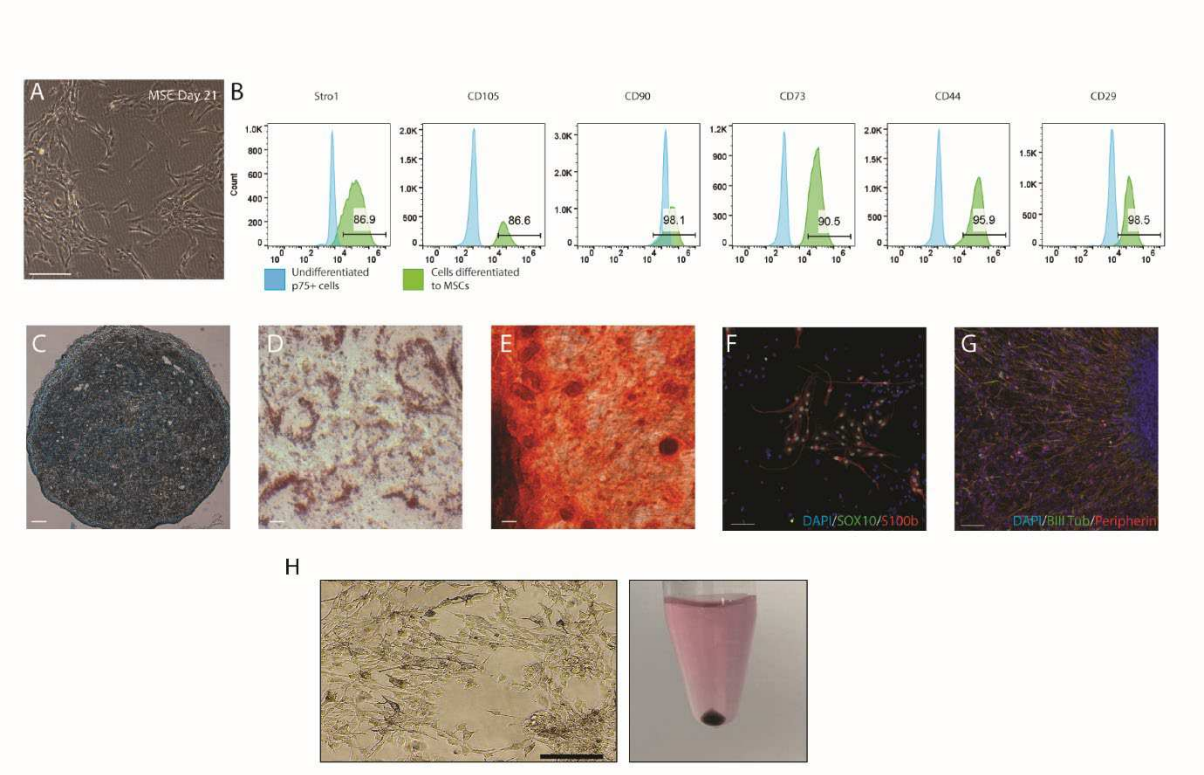
Hox gene expression in SOX10pos-SCP-like cells generated in the absence or presence of retinoic acid.

qPCR analysis of HOX gene expression in reporter line derived Puromycin selected SOX10pos-SCP-like cells maintained in differentiation medium either with or without RA. Data are shown as mean of relative expression normalized to GAPDH ± SEM of three independent experiments (N=3). Unpaired t-test with Welch's correlation not assuming equal standard deviation (ns not significant, * $p<0.05$, ** $p<0.01$, *** $p<0.001$).

We concluded that under our differentiation conditions human iPSC sequentially exit pluripotency, progress through NPB- and early NCC-like stages, before upregulating SCP genes.

Differentiation potential and *in vitro* migration properties of human Schwann cell precursor-like cells.

Exposure of FACS-sorted SOX10^{pos} cells to foetal calf serum containing medium for 3 weeks fostered the efficient generation of adherent spindle-shaped cells that expressed the cell surface marker profile (Supp Fig 4A) and tri-lineage differentiation potential of mesenchymal stromal cells (osteogenic, chondrogenic and adipogenic) (Suppl Fig 4C-E), a cell type that is known to be the progeny of SCPs with a cranial identity. When subjected to directed differentiation conditions the SOX10^{pos} SCP-like cells also generated S100B-expressing Schwann cells (Suppl Fig 4F) and melanocytes (Suppl Fig 4H). SOX10^{pos} SCP-like cells however repeatedly failed to generate peripheral neurons. In contrast, p75^{pos} MACS sorted NCC generated from the same cell line with the Menendez protocol [38] readily generated both tri-potential MSC and peripheral neurons (Suppl Fig 4G), indicating the SOX10^{pos} SCP-like cells were not functionally equivalent to NCC.

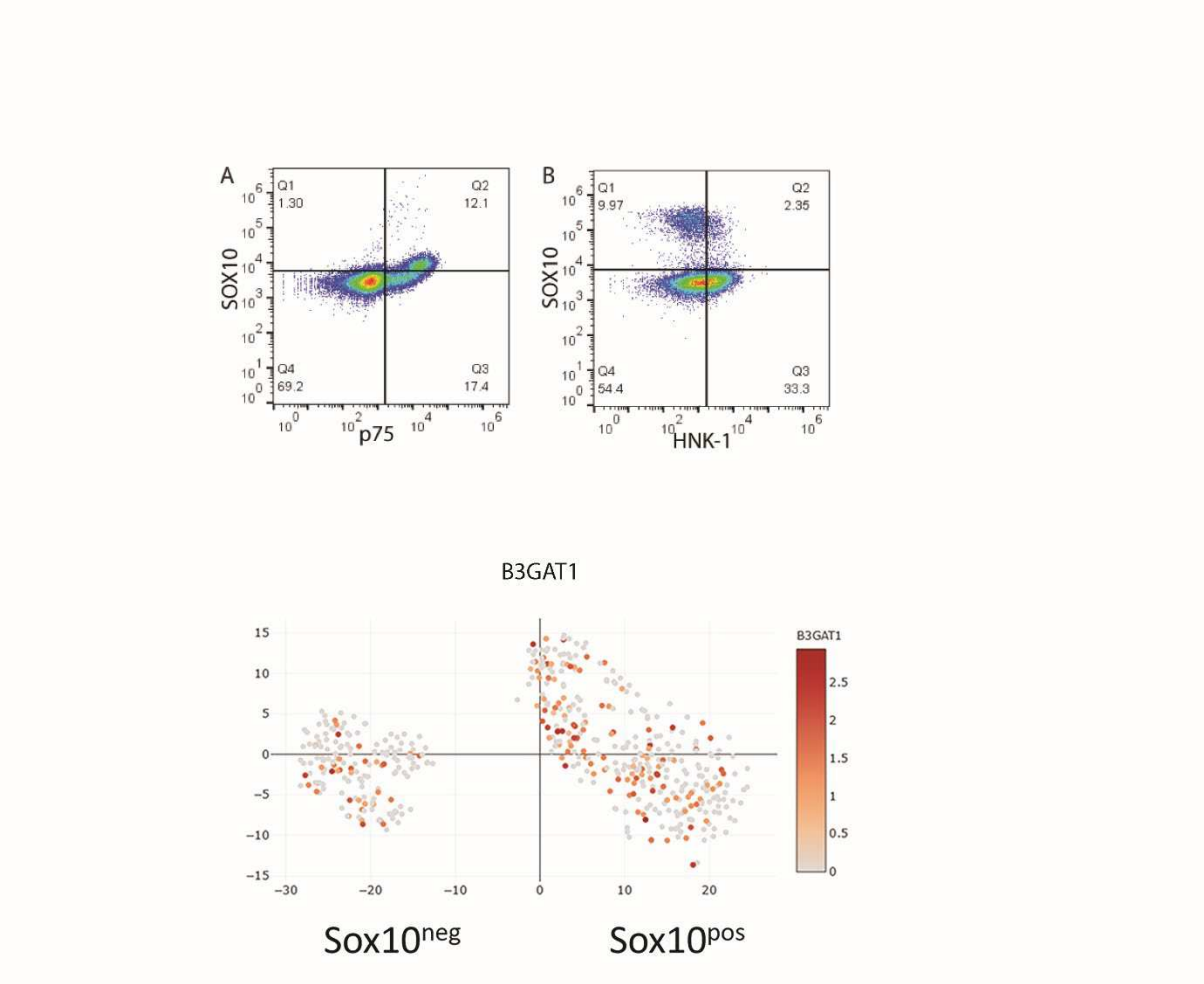


Suppl Fig 4

Multilineage differentiation potential of SOX10pos-SCP-like cells

(A) Brightfield microscopy and (B) Flow cytometric analysis of C11:SOX10 reporter hIPSC line derived SCP-like cells cultured in FCS containing medium for 3 weeks reveals expression of Stro1, CD105, CD90, CD73, CD44 and CD29 expected of mesenchymal stromal cell (MSC). SCP-like cell derived MSC show tri-lineage differentiation potential as indicated by (C) Chrondrogenic (Alcian Blue) (D) Adipogenic (Oil redO), and (E) osteogeneic (Alizarin red) stains. (F) C11:SOX10 reporter hIPSC line derived SCP-like cells readily differentiate into Schwann cells marked by S100beta. Cell nuclei were stained with DAPI. (Scale bars = 100µm) (G) C11:SOX10 reporter hIPSC line derived NCC differentiate into betaIIItub and peripherin double labelled peripheral neurons in contrast to C11:SOX10 reporter hIPSC line derived SCP-like cells. (H) Representative bright-field image of SCP-like cells differentiated into melanocytes at day 16. Right panel, the cell pellet after centrifugation displays a dark tone, indicative of pigmentation.

In support of this < 50% of SOX10^{pos} SCP-like cells were labelled with CD271/p75 antibodies, and < 20% were HNK-1 positive (Suppl Fig 5A & B), two cell surface markers previously reported to be homogeneously expressed on NCCs derived from human pluripotent stem cells [38,39].



Suppl Fig 5

Neural crest cell marker expression in SOX10pos-SCP-like cells

(A) Flow cytometric analysis of non-puromycin selected C11:SOX10 reporter hIPSC line derived SCP-like cell cultures following labelling with canonical NCC-markers SOX10 & HNK1 or SOX10 & p75. (B) scRNAseq UMAP of B3GAT (HNK1) expression in C11:SOX10 reporter hIPSC line derived SOX10pos SCP-like cells and SOX10neg NPB-like cells.

We noticed that the morphology and behaviour of FACS sorted C11:SOX10 hiPSC-derived SOX10^{neg} and SOX10^{pos} cells markedly differed. While sorted SOX10^{neg} cells exhibited a round morphology and self-organised in distinct compact colonies (Fig 2A), sorted SOX10^{pos} cells displayed the bipolar and stellate appearance of SCP-like cells, adopted a spatially uniform distribution across the culture surface, and typically interacted through foot processes (Fig 2A). Unbiased quantification revealed that in comparison to SOX10^{pos} cells, the SOX10^{neg} cells clustered closer together as indicated by the smaller mean distance between neighbours and area occupied by

these cells (Fig 2B). Comparable total nuclei counts between both cell populations indicates these differences were unlikely to be due to intrinsic differences in proliferation rates between SOX10^{pos} and SOX10^{neg} cells. Compared to SOX10^{neg} cells SOX10^{pos} cells also exhibited increased undirected migratory behaviour *in vitro* (Fig 2C), and as indicated by their increased distance travelled and migration speed (Fig 2D-F).

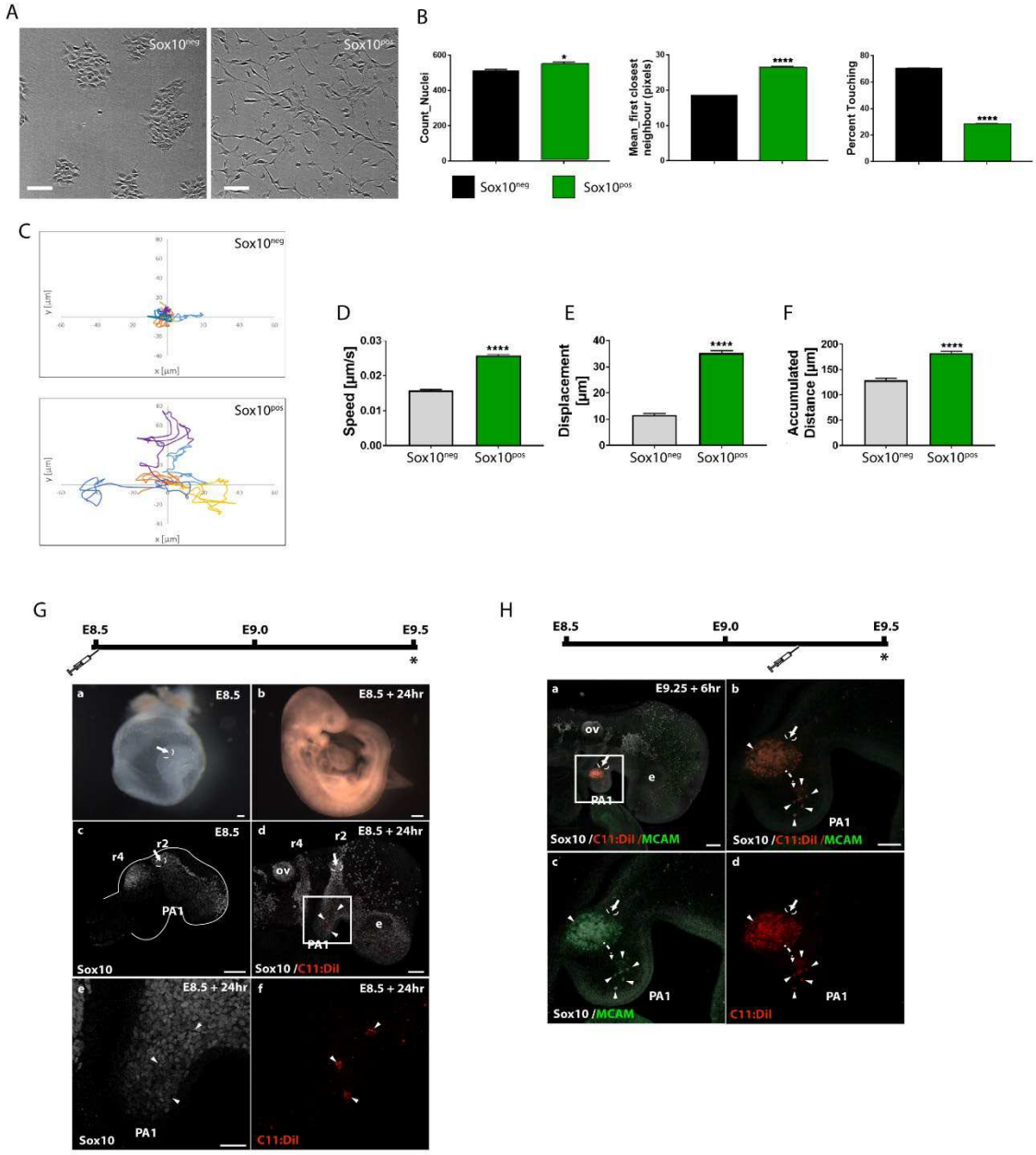


Figure 2: In vitro and in vivo migratory properties of SCP-like cells from C11:SOX10/mMAPLE reporter hiPSC.

(A) Brightfield images of SOX10pos and SOX10neg cells show morphological differences. SOX10neg cells exhibit a round cell morphology and organise themselves in distinct compact colonies while SOX10pos cells have a stellate appearance and are uniformly distributed in the culture dish. Scale bar = 100µm. (B) Graphs showing cell numbers, distance between a cell and its first closest neighbour, and area of the field of view that is occupied by cells (mean±SEM, n=9 fields of view). (C) Representative tracks of 5 randomly chosen SOX10pos and SOX10neg cells over 2 hours from a 16 hour time-lapse experiment involving a minimum of 400 cells. The intersection of the x- and y-axes was taken as the starting point of each cell path. Each tick mark on the axes represents 20µm. (D) Graphs quantify speed, (E) net displacement and (F) total distance travelled by SOX10pos and SOX10neg cells (mean±SEM, n>400 cells). (G) Tracing of human SCP-like cells microinjected into E8.5 embryos. (a) E8.5 embryos within their yolk sacs were microinjected with 20-50 dil labelled human SCP-like cells (white arrow in dashed circle). Scale bar = 100□m. (b) After 24 hrs in culture media cultured embryos develop to approximately E9.5. Scale bar = 100□m. (c) Representative image of an E8.5 embryo labelled for Sox10 depicting the site of SCP-like cell injection (white arrow in dashed circle). Scale bar = 100□m. (d) DiI labelled SCP-like cells (red, arrowheads) incorporate into the host NCC and SCP migratory stream into the first pharyngeal arch (PA1). e, eye; ov, otic vesicle; r2, rhombomere 2, r4, rhombomere 4. Scale bar = 100□m. (e-f) Higher magnification of boxed area in (d). scale bar = 50□□m. (H) Tracing of human SCP-like cells microinjected into E9.25 embryos. (a) DiI labelled SCP-like cells (red) were microinjected into the distal NCC / SCP migratory sream (white arrow in dashed circle) withinPA1. Scale bar = 100□m. (b-d) DiI labelled (red) MCAM+ SCP-like cells (green, arrowheads) incorporate into the host NCC and SCP migratory stream. e, eye; ov, otic vesicle. Scale bar = 50□m.

We concluded that SOX10^{pos} SCP-like cells exhibit different cell-cell adhesion and migration properties from the SOX10^{neg} cells they emerge from and possess the differentiation capacities expected of cranial SCPs.

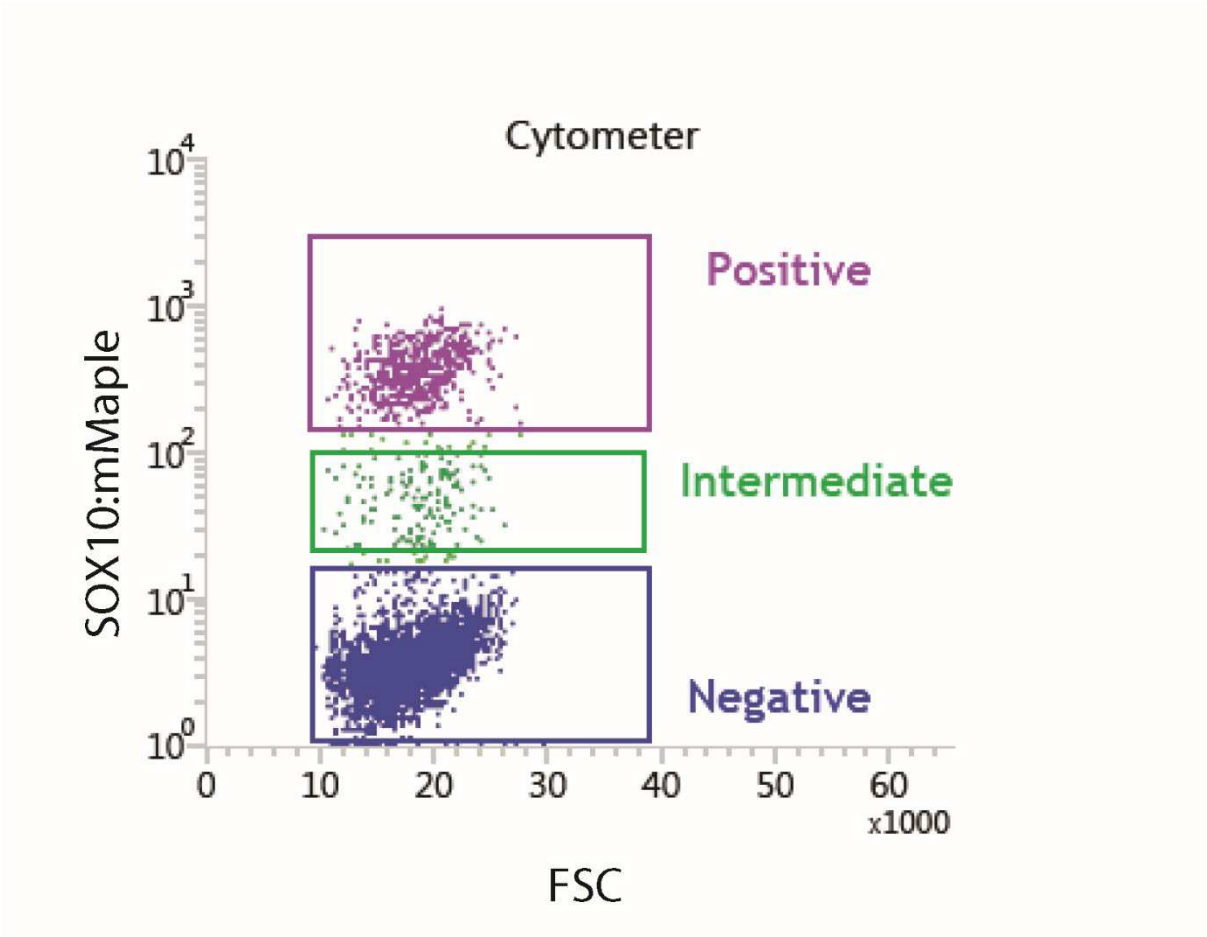
hiPSC derived human cranial Schwann cell precursors-like cells appropriately migrate in mouse embryos.

After cranial SCPs (CSCP) are specified *in vivo* they characteristically migrate along well-characterised pathways into the frontonasal processes, the first and second pharyngeal arches, and the nascent otic vesicle to contribute to the early anlagen of the developing face [13]. To assess whether the SOX10^{pos} human SCP-like cells generated with our protocol possessed this capacity, we microinjected 20-50 dil labelled mMAPLE selected cells into the rhombomere 2 migratory stream of E8.5 mouse embryos. A substantial fraction of injected human SOX10^{pos} SCP-like cells survived for 24hrs and integrated with host SOX10^{pos} expressing cells within the first

pharyngeal arch and frontonasal process (Fig 2G). We concluded that the cranial SCP-like cells (CSCP-like cells) generated with our protocol exhibit the expected *in vivo* migratory behaviour of cranial SCPs. A similar migration behaviour was observed for MCAM sorted SCP-like cells (Fig 2H) (a cell surface marker that permits the prospective isolation of this cell type (see discovery proteomics section below)). Collectively these data led us to conclude that the SOX10 expressing cells generated with our protocol exhibit the *in vivo* migration properties of early cranial CSCP-like cells.

Defining the transcriptome changes that accompany the specification of human SOX10^{pos} cranial Schwann cell precursor-like cells.

Having established that our protocol generates human cranial SCP-like cells we next wished to define the transcriptional make-up of this cell population and the SOX10^{neg} cells they emerge from in greater detail. We therefore deep sequenced total mRNA isolated from FACS sorted SOX10^{pos} and Sox10^{neg} cells (FACS profile in Suppl Fig 6).



Suppl Fig 6

FACS gating of SOX10posand SOX10neg cells destined for multi-omics analyses

Flow cytometry plot outlining the gating strategy used to isolate pure Sox10pos and SOX10neg cells from C11:SOX10 reporter hIPSC subjected to our SCP differentiation medium for 21 days. These cells were next used for bulkRNAseq, metabolomics and discovery proteomics.

This identified 18977 transcripts (Suppl data file 1, GEO accession GSE242011) of which 2794 showed significant differences in expression (|FC|>2, adj.p-value <0.01), with 1767 transcripts that were downregulated and 1027 transcripts that were upregulated in SOX10^{pos} cells (volcano plot Fig 3A, and Suppl data file 2). The canonical SCP gene Cadherin 19 (CDH19) was the most highly upregulated mRNA in SOX10^{pos} SCP-like cells, closely followed by prototypical SCP genes (Soldatov et al. 2019) PMP22, MPZ, GPR17, ERBB3 [9], CXCL12, VEGFB and SOX10 (FC=228, adj.Pval = 3.26x10-24). The human SOX10^{pos} SCP-like cells further showed increased mRNA counts of genes previously reported to be expressed in migratory cranial NCCs such as SOX5, LMO4, ZEB2 (also known as SIP1), TFAP2A, ETS1, MOXD1, PLXNA4, EDNRB, CDH6 [40,41], RXRG [42], CDH2 (N-cadherin), and of regulators of craniofacial mesenchyme differentiation DLX1, DLX2, POSTN, TBX2, NPR3, and LHX1. SOX10^{pos} cells did not express EGR2/KROX20, suggesting the SCP-like cells are not equivalent to boundary cap derived SCPs [43], nor did the cells express OLIG2, suggesting they are dissimilar from motor exit point glia (MEP glia) [44]. SOX10^{neg} cells robustly expressed genes typically expressed in the neural plate border [45–47] such as ZIC1, ZIC2, ZIC4, ZIC5, MSX1, PRDM1, PAX3, GBX2, neural stem cell markers SOX1 and SOX2 [48,49] and genes of pre-migratory neural crest cells such as LMX1b, ID2, ID3 and SOX8, as well as STK11 (LKB1) [50] a gene previously shown to be required for delamination of pre-migratory cranial NCC in the chicken. No significant expression of neuronal markers, such as NeuN and TUBB3, or the pro-neural factors Neurog2 or Ascl1 was detected in either population. Comparison of the RNAseq data of the SOX10^{neg} NPB-like cell population to SOX10^{pos} SCP-like cells highlighted the pronounced shifts in mRNA expression of integrins, cadherins, extracellular matrix remodelling enzymes and transcription factors of the SOX-family (Fig 3B-E).

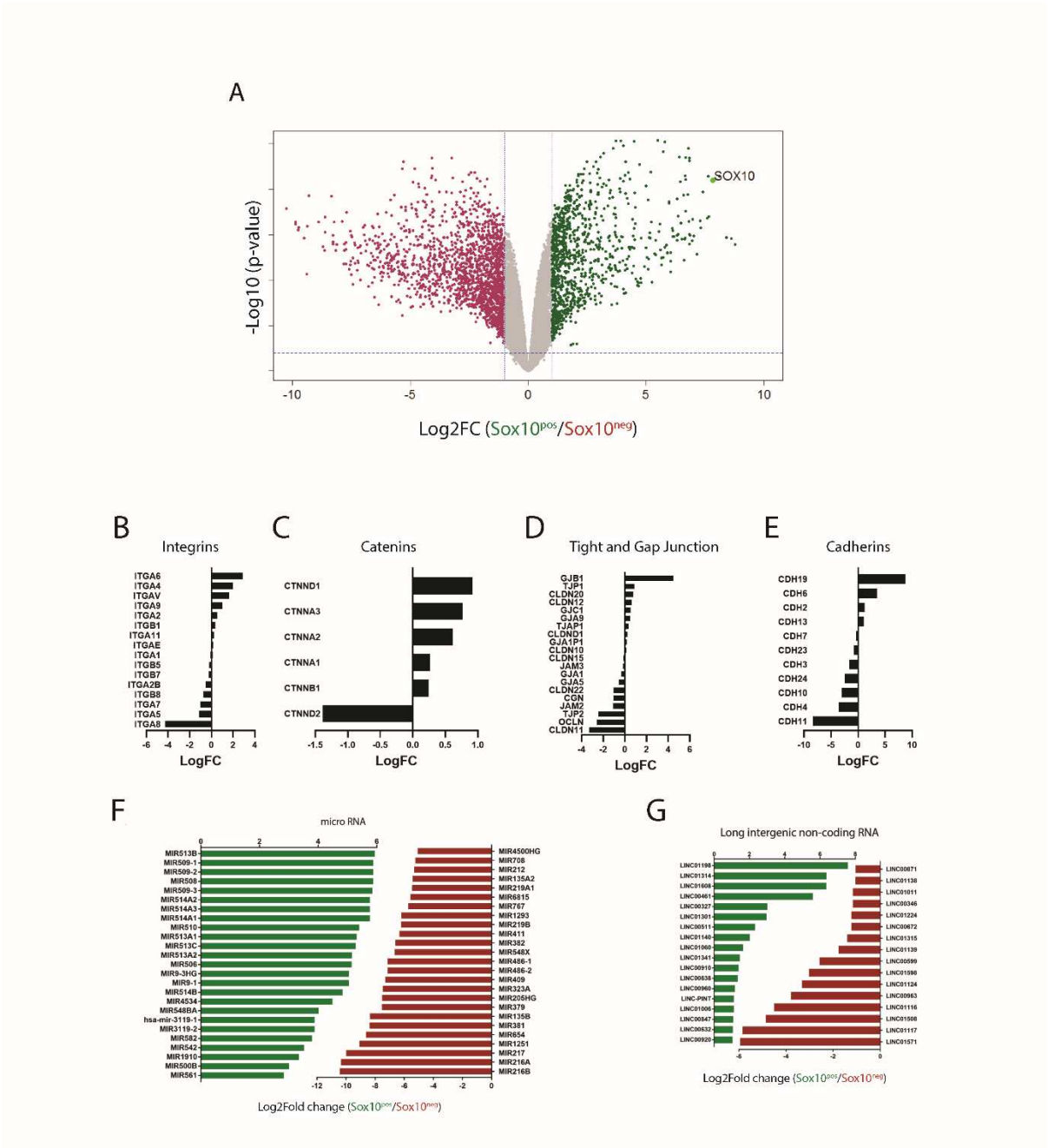


Figure 3: Total RNAseq analysis of SOX10pos and SOX10neg cells derived from C11:SOX10/mMAPLE reporter hiPSC.

(A) Volcano plot of bulk RNA-seq data. Genes are ranked according to their statistical adj.p-value (y-axis) and their relative abundance ratio (log2 fold change SOX10pos /SOX10neg, x-axis). The two vertical dashed blue lines indicate log2 fold change values of 1 and -1. The horizontal blue dashed line indicates an adjusted p-value of 0.01. Genes upregulated in SOX10pos cells with a fold change ≥2 and adj.Pval <0.01 are depicted in green, and those downregulated with a fold change ≤-2 and adj.Pval <0.01 are shown in red. Grey dots represent genes that do not differ significantly in their expression levels between SOX10pos and SOX10neg samples. (B-E) mRNA expression shifts of cell adhesion genes as cells transition from SOX10neg to SOX10pos cells (Log2-fold change). (F) mRNA expression shifts of miRNAs and (G) LncRNAs as cells transition from SOX10neg to SOX10pos cells (Log2-fold change).

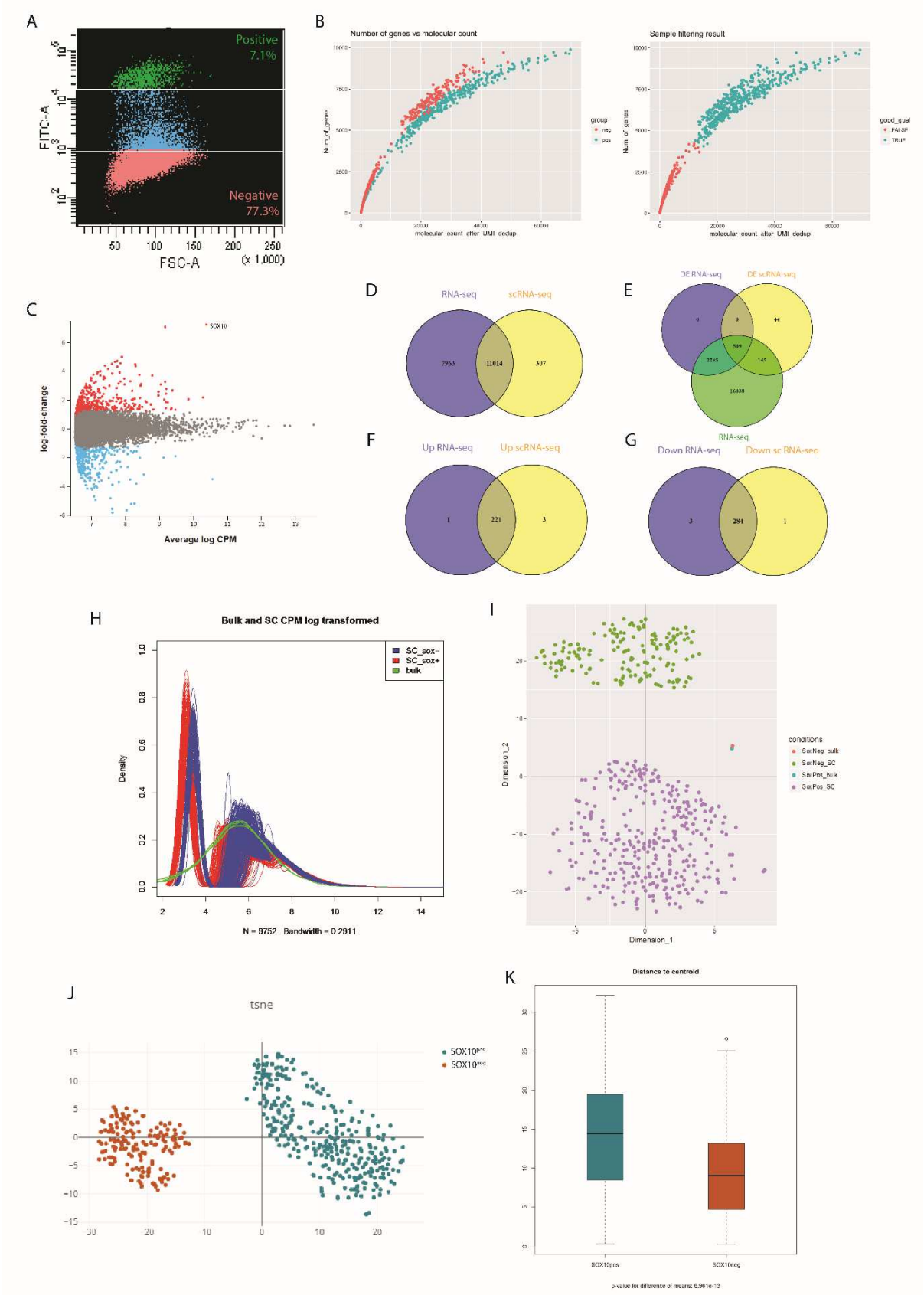
Non coding RNA changes accompanying human cranial SCP-like cell specification.

Having established that the SOX10^{pos} cells exhibit gene expression patterns consistent with cranial SCP-like cells we next isolated and sequenced non-polyadenylated RNAs from the same samples, identifying 570 microRNAs that were significantly up- and down-regulated in SOX10^{pos} and SOX10^{neg} SCPs (|FC|>2, adj.p-value <0.01), top 20 shown in Figure 3C, full list in Suppl data file 3). In SOX10^{pos} cells we noted upregulation of miR-514a, a member of a cluster of miRNAs (miR-506-514) previously implicated in initiating malignant transformation of melanocytes [51]. miR542 3p was previously reported to enhance the proliferation and migration ability of osteosarcoma U2OS cells [52] and was also upregulated in SOX10^{pos} cells. SOX10^{neg} cells were enriched for miR-145 expression, a microRNA known to repress the expression of pluripotency factors such as KLF4, SOX2, and OCT4 [42] and also showed increased expression of miR-217, which inhibits the expression of GRHL2, promotes the expression of key epithelial genes (e.g. E-cadherin, Claudin-3 and Claudin-4), and suppresses EMT [53] by direct inhibition of ZEB1 expression in keratinocytes, suggesting this function may be conserved in human NPB-like cells under our culture conditions. Another miRNA of note in SOX10^{neg} cells is miR-486-5p, which was previously shown to inhibit proliferation, and adipogenic and osteogenic differentiation of MSCs [54]. We further detected differential expression of 109 LincRNAs (top 20 shown in Fig 3D, full list in Suppl data file 3). Linc011908 is the most highly expressed LincRNA in our SCP-like cells in agreement with scRNAseq data of SCPs from human embryos [24]. Another robustly expressed LincRNA in SOX10^{pos} cells is Linc00461, previously shown to act as a sponge for miR216 [55]. Interestingly, in bladder cancer downregulation of miR216 results in increased Linc01198 and increased Sema3d expression [56], similar to what we observe here as NPB-like cells transition into SCP-like cells, suggesting a similar functional connection between these two classes of non-coding RNAs may exist in hIPSC-derived SCP-like cells.

ScRNAseq analysis of human cranial SOX10^{pos} SCP-like cells and SOX10^{neg} NPB-like cells.

To obtain insight into the inter-cellular gene expression heterogeneity of SOX10^{pos} CSCP-like cells and SOX10^{neg} NPB-like cell populations we next performed Cell-Seq2

[57,58]. 337 SOX10^{pos} cells and 165 SOX10^{neg} FACS sorted (Suppl Fig 7A) deposited cells passed quality control (see Suppl Fig 7B, material & methods, and Suppl materials & methods) and the requirement of detection of at least 12000 molecules after UMI deduplication and expression of at least 4000 genes per cell (scRNAseq data in Suppl data file 4).


A
Positive
7.1%
Negative
77.3%
FITC-A
FSC-A
(x 1,000)
B
Number of genes vs molecular count
Num_of_genes
molecular_count_after_UMI_dedup
group
neg
pos
Sample filtering result
good_qual
FALSE
TRUE
C
SOX10
log-fold-change
Average log CPM
D
RNA-seq
scRNA-seq
7963
11014
307
E
DE RNA-seq
DE scRNA-seq
44
145
F
Up RNA-seq
Up scRNA-seq
1
221
3
G
Down RNA-seq
Down sc RNA-seq
3
284
1
H
Bulk and SC CPM log transformed
SC_sox−
SC_sox+
bulk
Density
N = 9752 Bandwidth = 0.2911
I
conditions
SorNeg_bulk
SorNeg_SC
SorPos_bulk
SorPos_SC
Dimension_1
Dimension_2
J
tsne
SOX10pos
SOX10neg
K
Distance to centroid
SOX10pos
SOX10neg
p-value for difference of means: 6.961e-13

Suppl Fig 7

ScRNAseq quality control

(A) FACS profile indicating the gating strategy used for isolating cells destined for Cellseq. (B) Cell samples were filtered by requiring at least 12000 molecules counted after UMI deduplication and at least 4000 genes detected. Post quality control, 337 SOX10pos cells and 165 SOX10neg- cells were deemed suitable for further analysis. (C) Genes that were either not expressed with at least 1 count in 10% of samples or were not annotated with a description from the Homo sapiens Bioconductor database were filtered out. Post filtering, of 21540 genes, 11321 were retained for further analysis. (C) Volcano plot of average log CPM values (x-axis) and their relative abundance ratio (log2 fold change) between SOX10pos and SOX10neg (y-axis). Genes upregulated in SOX10pos cells with a fold change ≥2 and adj.p-value <0.05 are depicted in red, and those downregulated with a fold change ≤-2 and adj.p-value<0.05 are shown in blue. Grey dots represent proteins in the arrays that do not differ significantly in expression between SOX10 positive and negative cells. (D) Venn diagram showing the number of transcripts in bulk only (left) and single cell only (right) and the number of genes common to both datasets (centre). (E) Venn diagram showing the number of differentially expressed genes between SOX10pos and SOX10neg cells in bulk (left) only and single cell only (right) in comparison to the total number of transcripts in the bulk dataset. (F) Venn diagram showing the number of upregulated transcripts in SOX10pos cells in bulk only (left), single cell only (right) and the number of transcripts common to both (centre). (G) Venn diagram showing the number of upregulated transcripts in SOX10neg cells in bulk only (left), single cell only (right) and the number of transcripts common to both (centre). (H) The density of log CPM values for post-filtered bulk RNA-seq data (green), SOX10pos (red) and SOX10neg (blue) data are shown for each replicate/cell. (I) t-SNE map of quality-filtered, bulk and scRNA-seq data. (J) t-SNE map of quality-filtered single-cell transcriptomes from SOX10 positive and negative groups. (K) Boxplot of Euclidean distances between each point and the centroid of the cluster.

To compare the scRNAseq data with the bulk RNAseq dataset we collapsed the expression levels of each gene into a single average value across all cells within each group (Suppl data file 5), revealing 698 transcripts that showed a significant difference in abundance (t-test (|FC|>2, BH adj.p-value <0.05) between $SOX10^{pos}$ and $SOX10^{neg}$ cells (Suppl Fig 7C) of which 342 were upregulated in $SOX10^{pos}$ cells (red dots) and 356 (blue dots) were enriched in $SOX10^{neg}$ cells. Of transcripts that passed the scRNA-seq QC requirements 97% were detected in the filtered and normalised bulk RNA-seq dataset (Suppl Fig 7D) and 73% of the differentially expressed transcripts in the scRNA-seq dataset were also differentially expressed in the bulk RNA-seq dataset (Suppl Fig 7E). Of these, 505 transcripts (99.2%), were similarly up- or down-regulated in $SOX10^{pos}$ cells in both the scRNA-seq and bulk RNA-seq datasets (Suppl Fig

7F&G), underscoring the high level of concordance between these datasets. We next performed unsupervised hierarchical clustering of the top 500 most variable transcripts between the SOX10^{pos} and SOX10^{neg} cells (ranked by their BCV), revealing six transcript clusters (Fig 4A).

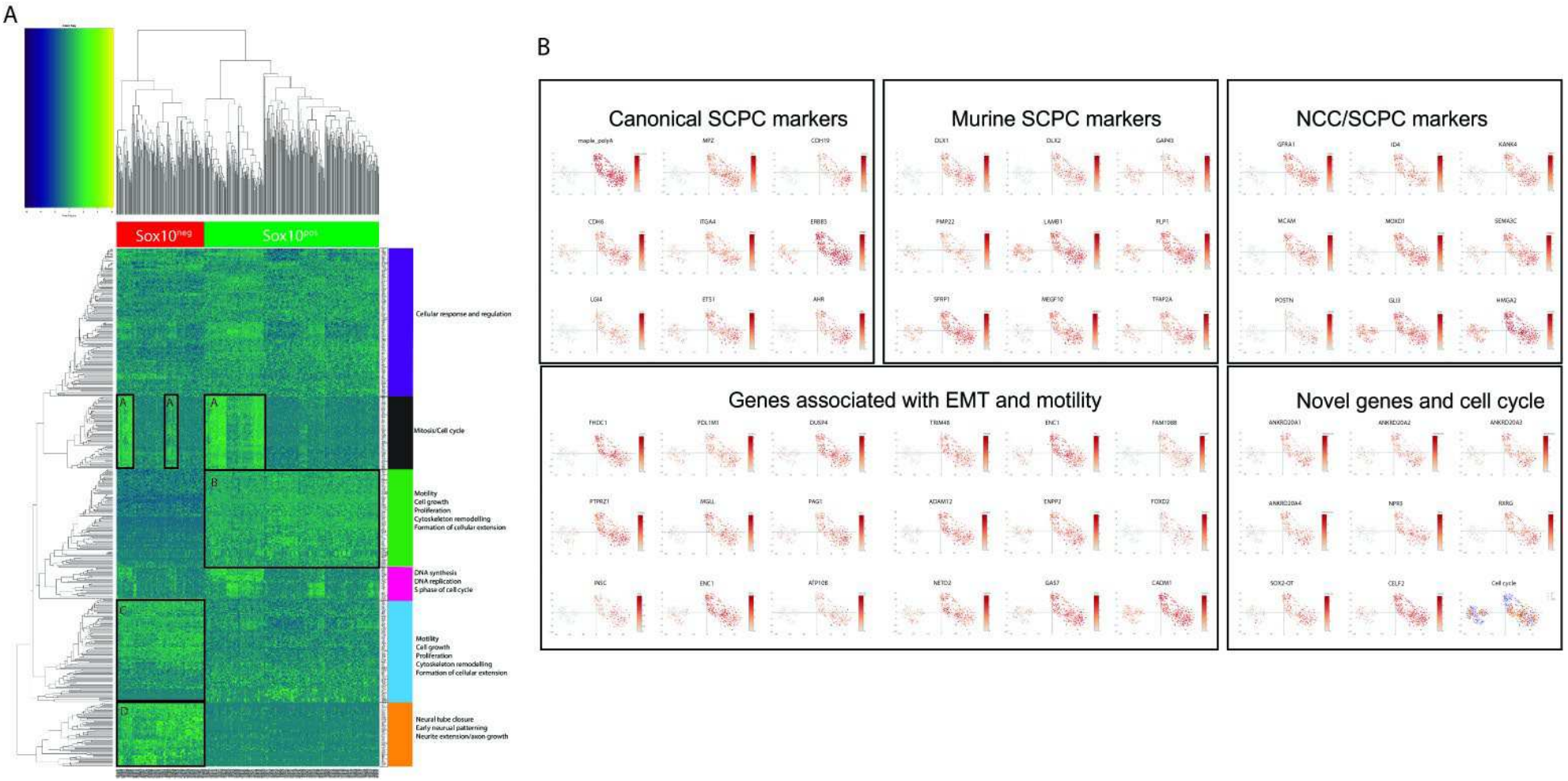

A
Sox10neg
Sox10pos
A
A
A
B
C
D
Cellular response and regulation
Mitosis/Cell cycle
Motility
Cell growth
Proliferation
Cytoskeleton remodelling
Formation of cellular extension
DNA synthesis
DNA replication
S phase of cell cycle
Motility
Cell growth
Proliferation
Cytoskeleton remodelling
Formation of cellular extension
Neural tube closure
Early neurual patterning
Neurite extension/axon growth
B
Canonical SCPC markers
Murine SCPC markers
NCC/SCPC markers
Genes associated with EMT and motility
Novel genes and cell cycle

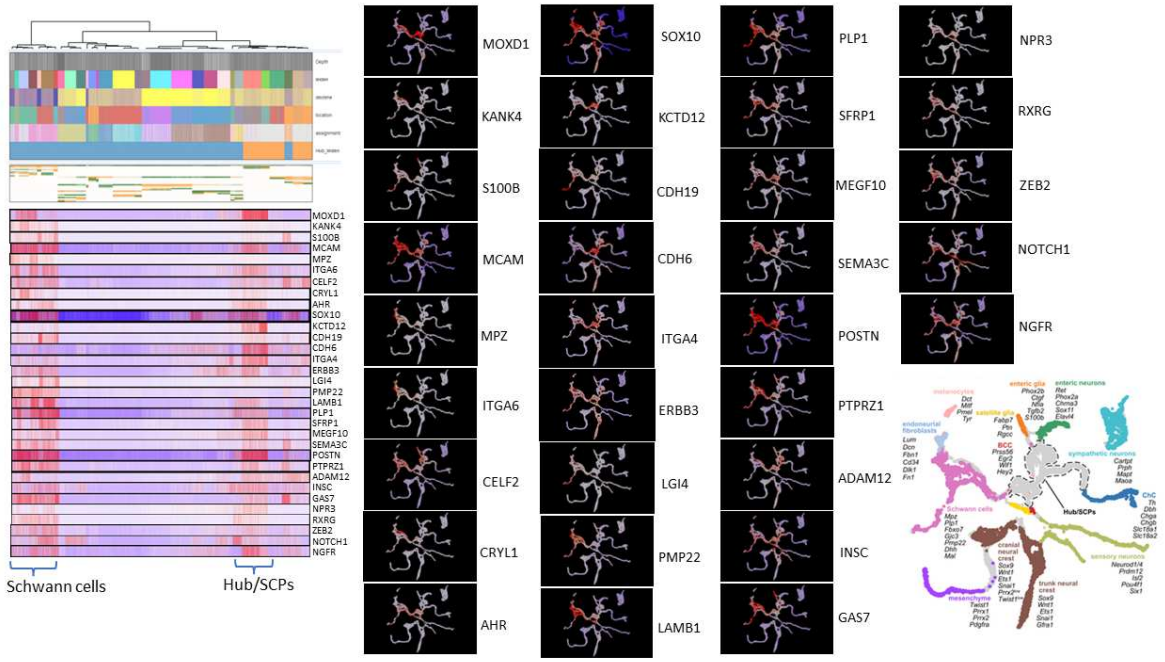
Figure 4: Single cell RNAseq analysis of SOX10pos and SOX10neg cells derived from C11:SOX10/mMAPLE reporter hiPSC.

(A) Hierarchical clustering by SOX10 expression group of all single cells (n = 402 single cells, top 500 variable genes) analysed by scRNA-seq. Left Y-axis axis shows the relationships between the genes according to similarity in expression pattern. Right Y-axis contains gene names and enriched biological processes in each of the subclusters marked by coloured boxes. (B) Gene expression projection of canonical SCP markers, murine SCP markers, NCC/SCP markers, genes associated with EMT and cell motility, and novel or cell cycle genes on tSNE clustered scRNAseq data from SOX10neg to SOX10pos cells.

EnrichR analysis [59] revealed Cluster 1 (purple) is enriched for mRNAs associated with cellular response and regulation; cluster 2 (black) for mRNAs associated with mitosis and cell cycle that were upregulated in subsets of cells from SOX10^{neg} and SOX10^{pos} cells (Box A), and cluster 3 (green) enriched for mRNAs associated with motility and cell growth, including multiple mRNAs of genes associated with SCPs (ITGA4, CDH19, MPZ, SOX10, MCAM) and NCCs (LMO4, NR2F2, NGFR, SEMA3C, RHOB, ITGA6, GFRA1, ETS1, CDH6, TFAP2A, DLX1, DLX2, NPR3, LRP2) that were all upregulated in SOX10^{pos} cells (Box B). Cluster 4 (purple) is enriched for mRNA associated with S-phase cell cycle progression. Transcripts related to cell proliferation (CCND2, LEF1, TIMP1, FGFR1) were enriched in cluster 5 (blue) and showed slightly higher expression in SOX10^{neg} cells compared to SOX10^{pos} cells (Box C). SOX10^{neg} cells (Box D) exhibited increased expression of transcripts in cluster 6 (orange) that are related to neural tube identity (ZIC1, ZIC2, PAX3, GBX2) and negative regulation of migration (DACH1, TPM1, STC1, CYP1B1, SLIT2, COL3A1). To quantify the degree of heterogeneity within and between the SOX10^{pos} and SOX10^{neg} single cell populations we conducted dimensionality reduction of the post filtered single cell transcriptomics dataset with t-Distributed Stochastic Neighbour Embedding (t-SNE), revealing two well separated populations with little evidence of subclusters within the SOX10^{pos} or SOX10^{neg} populations (Suppl Fig 7I&J). To visualize the gene expression heterogeneity between cells we next projected gene expression data on the tSNE plots revealing the expected homogeneous expression of SOX10-mMAPLE, as well as robust and homogeneous expression of the canonical SCP genes MPZ, CDH19 [34], CDH6 [60,61], ITGA4, neuregulin-1 receptor ERBB3 [62–65], LGI4 [18], ETS1 [66], and AHR [67,68] in the SOX10 expressing cell cluster (Fig 4B). Homogeneous upregulation of DLX1 and DLX2 [69] in the SOX10^{pos} SCP-like cells further exemplifies their uniform cranial

identity. While expression of SCP genes GAP43 and PMP22 was enriched in the SOX10^{pos} cluster their expression was less homogenous than perhaps expected, but not clustered (Fig 4B). We further noted strong and homogeneous upregulation of MCAM, LAMB1, PLP1, SFRP1 [70–73], MEGF10 [74], TFAP2A [75], POSTN [76], ID4 [77], LMO4 [78,79], SFRP1A [70–73] and GLI3 in SOX10^{pos} cells, genes that were previously found to be enriched in E12.5 mouse SCPs [18]. In addition to these well-established SCP genes we detected homogenous upregulation of the GDNF co-receptor GFRA1a [80], as well as NPR3 [81], ZEB2 [82], and FOXP2 [83], genes that were previously found to be expressed in developing and immature Schwann cells in rodent and avian models (Fig 4B). The SCP-like cells also expressed the NCC genes KANK4, RXRG [42], SEMA3C [84], and MOXD1, as well as genes not previously reported in SCPs such as FHDC1 [85,86] and PDLIM1 [87], that are involved in actin re-organization [85,86,88]. DUSP4, a gene previously reported in mouse melanoblasts [89], HMGA2, a gene linked to EMT, three human specific ANKRD20A family members [90], and FSTL5, a gene previously reported as a pericyte marker [91] were also homogeneously enriched in the SCP-like cells. Expression of the splicing factor CELF2 and SOX2-OT [92] a LincRNA expressed in the developing face in mouse embryos, were also robustly upregulated in the human SCP-like cells generated with our protocol (Fig 4B). In agreement with our Q-PCR and bulk RNAseq data sets the SOX10 negative NPB-like cell population exhibited robust and homogeneous expression of ZIC1, ZIC2, ZIC3, ZIC4, ZIC5, ID2 [93], ID3, CDH11, SNAI2, SLIT2, FOXB1 and EYA4 [94], as well as SEMA3A [95,96], EPHB1 [84], EPHA4, COL3A1, OLFM3 [97] all genes previously reported to be expressed in the NPB or in pre-migratory NCC populations [98–101]. STC1, CDON, DACH1 and CYP1B1 showed very similar expression patterns to these genes, but were previously not reported in these cells (Suppl Fig 9A). A weblink for interactive interrogation of gene expression data is provided (GLIMMA plot Suppl data file 16). Cells within the SOX10^{pos} group exhibited greater intra-group heterogeneity than the SOX10^{neg} population, as indicated by the Euclidian distance of individual cells from the cluster centroids (p-value 6.961x10-13), suggesting that SOX10^{pos} cranial SCP-like cells are transcriptionally more diverse than SOX10^{neg} neural plate border-like cells (Suppl Fig7K). To investigate to what extent this was a function of cell cycle differences we next used the cyclone method from scran to predict cell cycle phases (Scialdone et al. 2015), revealing roughly similar distributions of cells in the G2/M and S phases of the cell cycle, and a conspicuous absence of a G1 phase in both the SOX10^{pos} and SOX10^{neg}

cell populations (Fig 4B). Importantly, we find substantial concordance between differentially expressed genes in our scRNAseq data set (adjP<0.05) and SCP marker genes obtained from primary human SCP and cycling SCPs from the developing adrenal gland from 7-8 week old human embryos subjected to single cell and single nuclei RNAseq [24,102,103] (Suppl datafile 6). Furthermore, a large proportion of the top upregulated genes in our human CSCPs marked the hub-SCP population and Schwann cell clusters identified in the scRNAseq datasets of murine SCP development [5] when analysed with the transcriptional profiling resource provided by the Adameyko laboratory (Suppl fig 8).



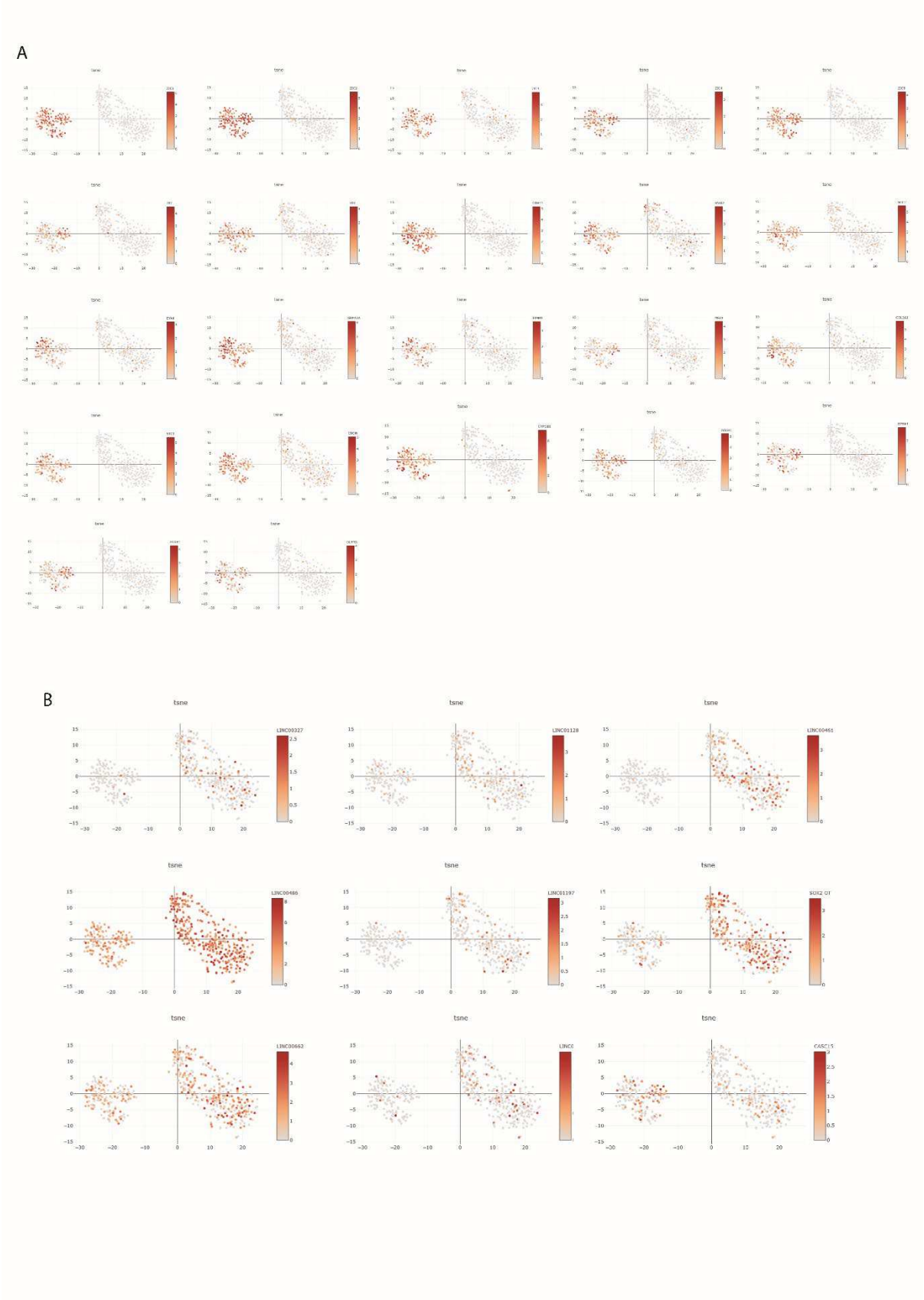
Suppl Fig 8

Gene expression concordance between top human SCP-like cells and mouse hub-SCPs and Schwann cell lineage.

Gene-based embedding, and SCENIC-based UMAP embedding of the top differentially expressed genes in human SCP-like cells (from bulk and scRNAseq) in lineage traced murine NCC and Schwann cell lineages through the online visualisation tool at https://protect-au.mimecast.com/s/C_kdCJyBq7FYQA02cGTDAu?domain=adameykolab.hifo.meduniwien.ac.at/, as outlined in (Kastriti et al. 2022) that leverages data from (Soldatov et al. 2019).

A subset of LncRNAs detected in the bulk RNAs seq data were also found in the scRNAseq data set (Suppl Fig 9B), both independently validating the bulk RNA seq

data, and revealing that some LincRNAs were robustly upregulated in almost all SCP-like cells whereas others showed more restricted upregulation.



Suppl Fig 9

Single cell analysis of neural plate border and LncRNA expression.

(A) Gene expression projection of selected neural plate border marker genes on scRNAseq t-SNE plots (B) gene expression projection of selected LncRNAs on scRNAseq t-SNE plots.

Currently little is known about the function of the miRNAs and LincRNAs that are enriched in the SOX10^{pos} and SOX10^{neg} cell populations, but functional validation of their roles and functions is now possible with the experimental platform described here and, assessing their expression in very early human embryo sections via *in situ* hybridization, RNAscope or spatial transcriptomics should prove illuminating.

Alternative splicing events accompanying the specification of human SOX10^{pos} cranial SCP-like cells.

It is well established that craniofacial development is particularly sensitive to mutations in spliceosome proteins and associated factors [104–117]. For example, in mouse embryos deletion of splicing factors *RBFOX2* [104] or *ESRP1* [105] results in craniofacial defects that are linked to aberrant splicing, yet little is known about alternative splicing (AS) in human SCP development. To address this, we first examined the expression of 1,636 RNA binding proteins (RBPs), compiled from lists of RBPs [106] splicing factors [107], and spliceosomal proteins [108] (Suppl data file 7) in our bulk RNAseq dataset. This identified 16 RBPs that show a greater than 4-fold change in expression between SOX10^{pos} and SOX10^{neg} cells (adjusted p-value cut-off of 0.01) (Fig 5A). To investigate the prevalence and nature of alternative splicing events in SOX10^{pos} SCP-like cells and SOX10^{neg} NPB-like cells we mined the bulk RNA-seq data with a recently developed pipeline that identifies both classical two-dimensional (2D) AS patterns as well as more complex patterns [109]. This identified 23,997 splicing events, including a large number (n = 6,944, 29%) of complex events. To identify significant differential splicing events, isoform proportions (Ψ) were fit using a Dirichlet-multinomial regression [110], and a likelihood ratio test (Mh: $\Psi \sim \{ SOX10^{pos}, SOX10^{neg} \}$ vs. Ma: $\Psi \sim 1$) was conducted, identifying 4,983 events, including 2,153 (43%) complex events (i.e. detection of at least one isoform with $\Psi \in [10\%, 90\%]$ in any condition (Suppl data file 8). The table in Figure 5B depicts the top five most frequently detected complex splicing patterns (denoted as C1…C5). We detected 200 differentially spliced events

with at least one isoform undergoing a |ΔΨ| >= 10%, between the SOX10^{pos} SCP-like cells and the SOX10^{neg} NPB-like cells at an adjusted p-value cut-off of 0.05 (Suppl data file 8). Strikingly, the largest fraction of these events (47%) localized to annotated protein domains within their host genes (Fig 5C). GO analysis of the host genes revealed an enrichment of processes related to EMT and cell migration (cytoskeletal organization, actin-filament based process, and cell junction organization), neuronal development, and splicing (FDR < 0.05; Fig 5D, Suppl data file 9). Very few differentially spliced genes were also differentially expressed, and most events involved coding switches (Fig 5E), suggesting that differential splicing promotes proteomic diversity during the transition into SOX10^{pos} SCP-like cells.

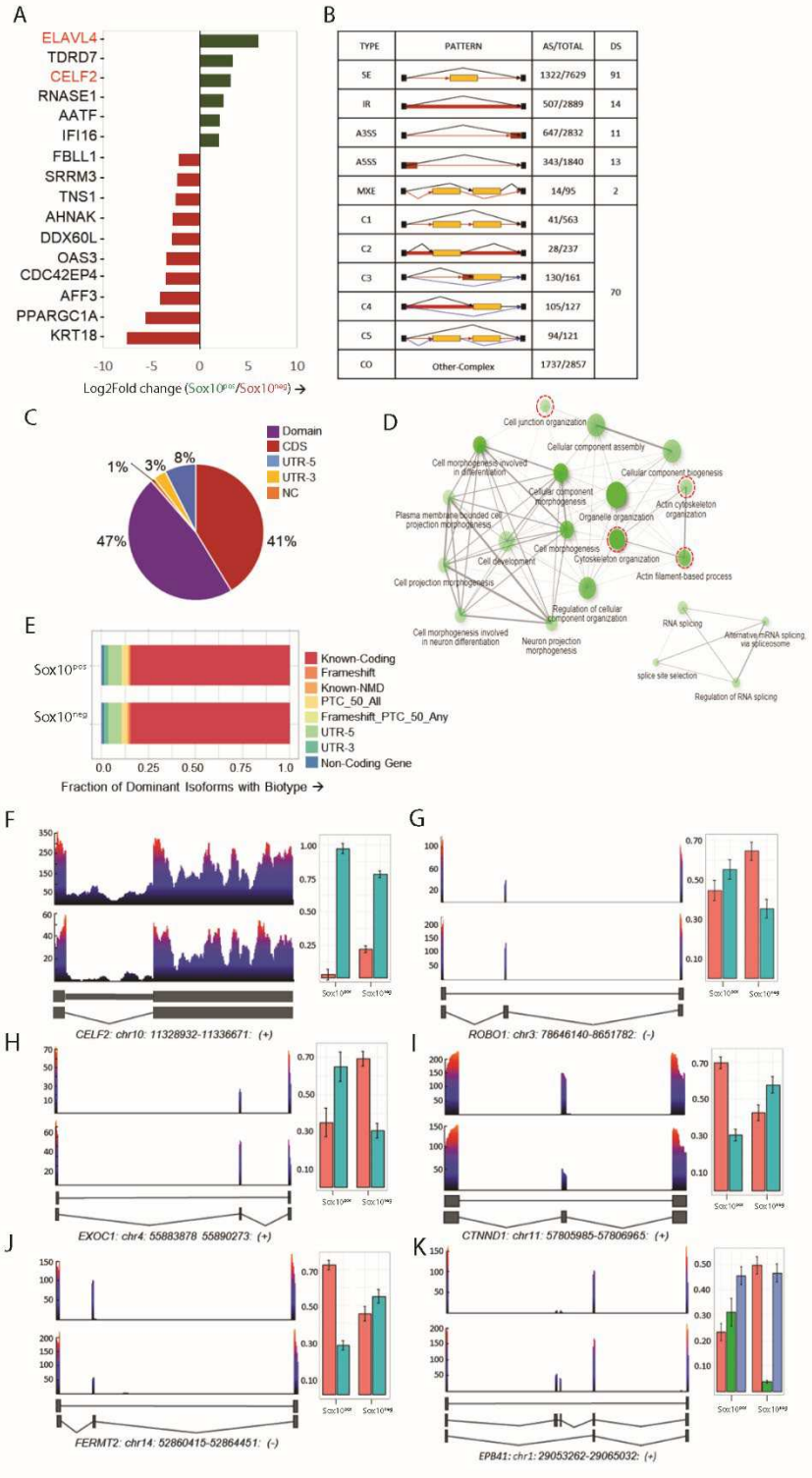
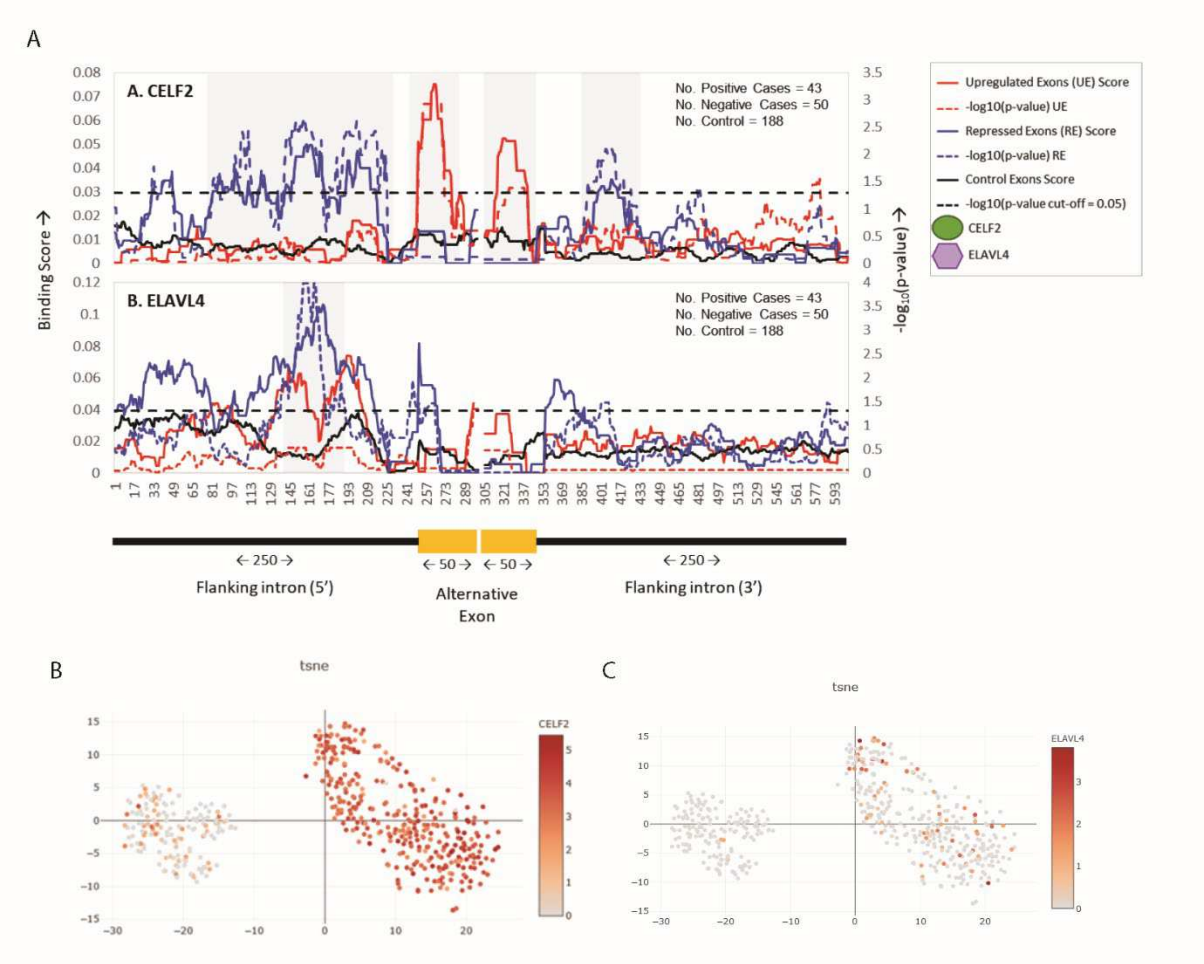

A
ELAVL4
TDRD7
CELF2
RNASE1
AATF
IFI16
FBLL1
SRRM3
TNS1
AHNAK
DDX60L
OAS3
CDC42EP4
AFF3
PPARGC1A
KRT18
-10 -5 0 5 10
Log2Fold change (Sox10pos/Sox10neg) →
B
TYPE PATTERN AS/TOTAL DS
SE 1322/7629 91
IR 507/2889 14
A3SS 647/2832 11
A5SS 343/1840 13
MXE 14/95 2
C1 41/563
C2 28/237
C3 130/161
C4 105/127
C5 94/121
70
CO Other-Complex 1737/2857
C
Domain
CDS
UTR-5
UTR-3
NC
1%
3%
8%
47%
41%
D
Cell junction organization
Cellular component assembly
Cell morphogenesis involved in differentiation
Cellular component biogenesis
Cellular component morphogenesis
Actin cytoskeleton organization
Plasma membrane bounded cell projection morphogenesis
Organelle organization
Cell morphogenesis
Cytoskeleton organization
Cell development
Actin filament-based process
Cell projection morphogenesis
Regulation of cellular component organization
Cell morphogenesis involved in neuron differentiation
Neuron projection morphogenesis
RNA splicing
Alternative mRNA splicing, via spliceosome
splice site selection
Regulation of RNA splicing
E
Sox10pos
Sox10neg
Known-Coding
Frameshift
Known-NMD
PTC_50_All
Frameshift_PTC_50_Any
UTR-5
UTR-3
Non-Coding Gene
0.0 0.25 0.50 0.75 1.0
Fraction of Dominant Isoforms with Biotype →
F
CELF2: chr10: 11328932-11336671: (+)
G
ROBO1: chr3: 78646140-8651782: (-)
H
EXOC1: chr4: 55883878 55890273: (+)
I
CTNND1: chr11: 57805985-57806965: (+)
J
FERMT2: chr14: 52860415-52864451: (-)
K
EPB41: chr1: 29053262-29065032: (+)

Figure 5: Splicing differences between SOX10^{neg} and SOX10^{pos} cells derived from C11:SOX10/mMAPLE reporter hiPSC.

(A) Significantly differentially expressed (log_2Fold-Change >= 2; p-adj < 0.01) RNA-binding proteins detected in the data. (B) The top most frequently detected splicing patterns based on the total number of events representative of these patterns. Of the total, the number of AS and DS events are indicated. (C) The distribution of DS events across different functional regions in their host genes. (D) Network representation of the top 20 most significantly enriched GO terms in DS protein-coding genes. (E) Potential Biotypes of the spliced isoforms generated for each event. ● Know-Coding: At-least one full-length transcript supporting this isoform is an annotated protein-coding transcript. ● UTR-5: Event lies in the 5'-UTR ● UTR-3: Event lies in the 3'-UTR ● Frameshift: Isoform introduces a frameshift as compared to the best matching annotated protein-coding transcript. ● Frameshift_PTC_50_Any: Isoform introduces a frameshift and a pre-termination codon (PTC) that agrees with 50nt NMD rule is introduced in at least one translation frame. ● PTC_50_All: PTC(s) agreeing with the 50nt NMD rule are introduced in all translation frames. ● Non-Coding-Gene: Event is hosted by a non-coding gene. Coverage diagrams for differential splicing events in (F) *CELF2*, (G) *ROBO1*, (H) *EXOC1,* (I) *CTNND1*, (J) *FERMT2*, and (K) *EPB41* genes. In each sub-diagram, the lower and upper coverage panel represent SOX10^{neg} and SOX10^{pos} samples respectively. The Y-axis indicates the number of reads. The splicing diagrams (dark-grey) represent event isoforms in order (i.e. isoform 1 to n). The Bar-plots indicate the fitted proportions for the event isoforms in order. The Y-axis indicates the proportion.

Two tissue-specific splicing factors (*CELF2* (Guo et al. 2020) and *ELAVL4* were particularly strongly upregulated in the SOX10^{pos} population (Suppl data files 2 & 4). *ELAVL4* (Fold-Change = 64; p-adj = 1.19e-11) is a neuronal factor, best known for regulating the stability of neuronal mRNAs by binding A-rich elements in 3'-UTRs [111,112], but it can also regulate the splicing of a subset of pre-mRNA targets [112–114]. *CELF2* (Fold-Change = 8; p-adj = 1.09e-24) is a well-known splicing regulator previously implicated in heart and muscle development [115,116] and T-cell activation [117,118]. Interestingly, C*ELF2* was also identified as an abundantly expressed putative regulatory splicing factor in a high-throughput transcriptomic analysis of mouse craniofacial development [119]. Since RBPs generally exert a position-dependent effect on splicing outcomes we constructed maps of *CELF2* and *ELAVL4* binding in the regions around regulated exons. To extract the top preferred binding 6-mers for *ELAVL4* and *CELF2* we processed CLIP-seq [118] and RBNS datasets [120] according to

published protocols (Suppl data file 10 and Suppl methods). In the *ELAVL4* map, we observed a significant peak in the upstream flanking intronic region of repressed exons, consistent with its previously reported ability to mediate exon repression [114]. For *CELF2*, we also observed strong peaks in the flanking intronic regions of repressed exons, consistent with its reported repressive mechanisms [118], that became significant when we generated evolutionary conservation-weighted binding maps [121] (Suppl Fig 10), indicating that *CELF2* likely represses exons by binding flanking conserved intronic sites.



Suppl Fig 10

Conservation Score weighted splicing maps of exon skipping events

Conservation Score weighted splicing maps of exon skipping event regulation by (A.) CELF2 and (B.) ELAVL4. The fraction of 6-mers matching preferred binding 6-mers for each protein were counted in 30bp windows within the regulated alternative exons and surrounding intronic regions and compared using the Wilcoxon Rank-Sum test against a control set of alternative exons. The 6-mer counts were scaled by the mean PhastCons100 conservation scores of the sequence. (C) Projection of CELF2 and ELAVL4 expression at a single cell level.

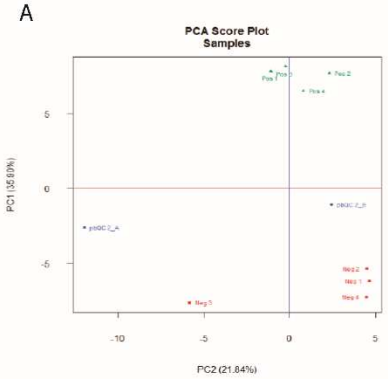
We further detected a strong and significant *CELF2* peak within the first (5') 50nts of upregulated exons. This CELF2-mediated regulation by exonic binding has not been investigated before, although its exonic binding has been reported in CLIP-seq studies [118]. Collectively these observations suggest that during the specification of SCP-like cells *ELAVL4* primarily functions to repress exons, possibly by interfering with 3' splice site recognition, while *CELF2* can enhance both inclusion and repression, depending on its binding location and/or recruitment of spliceosome components. Our data also revealed a known auto-regulatory intron retention (IR) event in the 3'-UTR of the *CELF2* gene itself (Fig 5F). The resulting IR mRNA has higher stability, and thus enhances translation and CELF2 protein levels [117]. In agreement with this we observed a strong (FC = 4.1) upregulation of CELF2 protein in our proteomics analysis (Suppl data file 13) in keeping with the observation that the proportion of this IR CELF2 isoform increases in the SOX10^{pos} population (ΔPSI = +17.5%). The greatest change was detected for a highly conserved 45nt micro-exon of the *EXOC1* gene, which was increasingly skipped in the SOX10^{pos} population (ΔΨ = -34.7%, Fig 5F). The encoded protein, EXOC1/SEC3 forms an essential subunit of the evolutionarily conserved exocyst complex involved in vesicle exocytosis, but has also been implicated in regulating cell migration and EMT in other cellular contexts [122–125]. Figure 5F depicts another micro-exon-skipping event in the *CTNND1* gene. This conserved 18nt micro-exon was increasingly included in the SOX10^{pos} population (Δψ = +27.3%). *CTNND1* encodes the guanine-nucleotide-dissociation inhibitor, p120, which specifically binds to and inhibits the small GTPase RhoA involved in actin organization, myosin contractility, and cellular morphological polarization [126,127]. Isoforms of p120 are particularly well studied, and the inclusion of this exon introduces a 6 amino acid insertion in its second Rho-binding domain that is necessary for its inhibitory

interaction with RhoA [128,129]. Interestingly, we detected the co-dominant expression of both isoforms in the SOX10^{pos} cells, suggesting this may constitute a mechanism for balancing p120's inhibitory activity, since spatially segregated activation of RhoA signalling is required during EMT and migration. Other genes with established involvements in regulating EMT and migration that exhibit differential and alternative splicing events (ROBO1, *EPB41, FERMT2,* and *RAI14* [130–133], were also observed (Fig 5F). Collectively, these data reveal that modulation of the splicing program is prevalent in human cranial SCP-like cells and likely plays a significant, albeit still largely unexplored, role in controlling their migratory phenotype and cell-cell interactions.

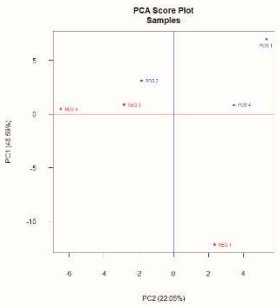
Metabolomics identifies a decreased reliance on glycolysis in human SOX10^{pos} cranial SCP-like cells.

It is increasingly recognized that elevated glycolytic activity in oxygen rich environments (the Warburg effect) is an important regulator of signalling in early embryo development, and this was previously linked to the onset of NCC migration in the chicken [134]. To investigate the nature of the metabolic changes that accompany the transition from human SOX10^{neg} NPB-like cells into migratory SOX10^{pos} SCP-like cells we cultured equal quantities of pure FACS sorted SOX10^{pos} and SOX10^{neg} populations and carried out metabolomic analysis on metabolites extracted from the cells and from culture medium that had been exposed to the cells for 24 hrs. LC-MS detected 124 metabolites in SOX10^{pos} and SOX10^{neg} cell lysates of which 26 showed a significant difference in abundance (adj.p-value <0.05, Suppl data file 11). Principal component analysis (PCA) of these metabolites (Suppl Fig 11A) indicated that replicates of SOX10^{pos} cells (green) clustered and were clearly distinct from replicates of SOX10^{neg} cells (red).

A



B



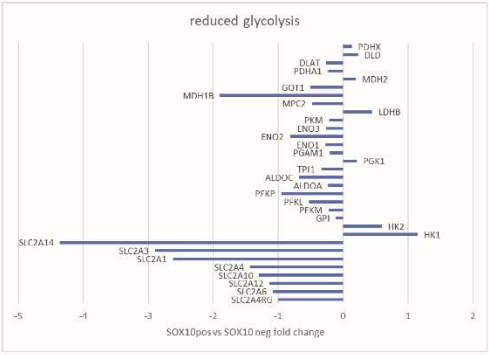
C

Metabolite	p-value	BH Adjusted p-value	fold change
D-Fructose 1,6-bisphosphate	**7.08E-07**	**8.64E-05**	**-2.26262**
alpha-D-Glucose 1,6-bisphosphate	**1.35E-06**	**8.64E-05**	**-2.27762**
D-Glucose 6-phosphate	**2.78E-06**	**0.000119**	**-1.71757**
D-Fructose-6-phosphate	**6.89E-06**	**0.00022**	**-1.72108**
Guanosine monophosphate	**1.16E-05**	**0.000297**	**1.298**
D-Ribulose-5-phosphate	**0.00019**	**0.004044**	**-2.98341**
Thiamin diphosphate	**0.000529**	**0.008721**	**1.626454**
D-Ribose-5-phosphate	**0.000545**	**0.008721**	**-3.28884**
Glutathione	**0.000614**	**0.008739**	**0.914249**
Guanosine diphosphate-mannose	**0.000743**	**0.009512**	**-0.30655**
Arginine	**0.001761**	**0.020497**	**-0.59628**
Phosphoenolpyruvic acid	**0.003305**	**0.022612**	**1.87015**
Glycerophosphoethanolamine (GPE)	**0.003382**	**0.022612**	**1.249327**
deoxy Guanosine monophosphate	**0.002995**	**0.022612**	**1.044191**
Cystathionine	**0.003221**	**0.022612**	**0.282626**
Pantothenic acid	**0.003236**	**0.022612**	**-0.48957**
Guanosine triphosphate	**0.003533**	**0.022612**	**-0.78666**
Biotin	**0.002652**	**0.022612**	**-1.59207**
D-Sorbitol	**0.003392**	**0.022612**	**-2.21473**
Phenylacetylglycine	**0.00273**	**0.022612**	**-2.45504**
Gluconic acid	**0.004036**	**0.0246**	**-1.52481**
Mannitol	**0.006024**	**0.034233**	**-2.72366**
Adenosine diphosphate	**0.00803**	**0.039624**	**0.817693**
Uridine diphosphate	**0.008049**	**0.039624**	**0.815709**
Adenosine-5'-triphosphate	**0.008031**	**0.039624**	**-0.29678**
Glycosylphosphatidylinositol (GPI)	**0.008914**	**0.04226**	**0.506705**

D

Metabolite	p-value	BH Adjusted p-value	fold change
Glucose pk1	**2.18E-05**	**0.001941**	**5.813877**
Aspartic acid	0.004332	0.19279	3.772813
L-glutamic acid	0.015333	0.454872	5.087146
Unidentified 1	0.025434	0.565903	-1.86251
Fructose	0.032844	0.58462	-0.4979
Lysine	0.047885	0.598095	-0.33009

E



Suppl Fig 11

Metabolomics

(A) PCA score plot of median normalised cell matrix of 124 features identified during metabolomic analysis of cell lysates from SOX10pos (green) and SOX10neg (red) groups. Pooled biological quality control samples (PBQCs) fell in between the 2 groups as expected. (B) PCA score plot of median normalised cell matrix of 89 features identified during metabolomic analysis of spent media from SOX10pos (green) and SOX10neg (red) groups. Pooled biological quality control samples (PBQCs) fell in between the 2 groups as expected. (C) List of differentially expressed (based on p-value) metabolites in SOX10pos and SOX10neg cell lysates in decreasing order of significance. Blue rows show metabolites that were upregulated in SOX10neg samples and orange rows show metabolites that were upregulated in SOX10pos samples. Metabolites in bold are significant based on BH-adjusted p-value<0.05. (D) List of differentially expressed (based on p-value) metabolites in SOX10pos and SOX10neg cell lysates in decreasing order of significance. Blue rows show metabolites that were upregulated in SOX10neg samples and orange rows show metabolites that were upregulated in SOX10pos samples. Metabolites in bold are significant based on BH-adjusted p-value<0.05. (E) Differential expression of mRNA species coding for glycolytic enzymes between SOX10pos and SOX10neg cells (Bulk RNAseq data, log-fold change shown).

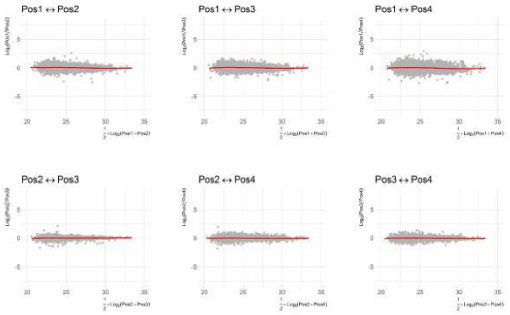
Of the 26 significantly differentially abundant metabolites (Suppl Fig S10C) 10 were upregulated in the SOX10^{pos} population and 16 were upregulated in the SOX10^{neg} population. Noticeably, metabolites of the glycolysis/gluconeogenesis pathways (D-Fructose 1,6-bisphosphate, D-Glucose 6-phosphate, D-Fructose 6-phosphate) showed significantly higher abundance in the SOX10^{neg} population. Metabolomics analysis of medium collected from SOX10^{pos} and SOX10^{neg} cells (PCA plot in Suppl Fig S10B), identified 89 metabolites, of which only glucose exhibited a significant 5.8-fold difference (BH-adjusted p-value < 0.05) (Suppl Fig S11D). Collectively these data indicated that, in contrast to NCC, *in vitro* generated human SCP-like cells display a decreased dependence on glycolysis as compared to SOX10^{neg} NPB-like cells. In agreement with this data, analysis of the bulk RNA seq dataset showed that almost all genes of the glycolytic pathway were downregulated in SOX10^{pos} SCP-like cells as compared to SOX10^{neg} cells (Suppl Fig S11E, full metabolomics data in Suppl data file 12)

Discovery proteomics reveals MCAM/CD146 as a specific marker for human SCP-like cells and identifies novel putative regulators of cell motility and EMT.

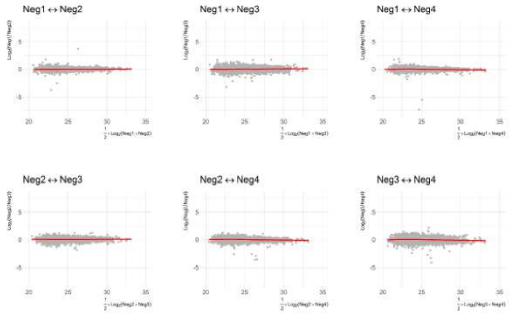
We next applied quantitative proteomics to define the protein make-up of SOX10^{neg} progenitors and migratory SOX10^{pos} SCP-like cells, and to identify cell surface proteins that specifically mark these cells. Pairwise comparison indicated good biological reproducibility of independent differentiation experiments (Suppl Fig S12A i and ii) and Multi-Dimensional-Scaling demonstrated that SOX10^{neg} and SOX10^{pos} cells of biological replicates are similar but readily discriminated in the first dimension (Suppl Fig S12B).

A

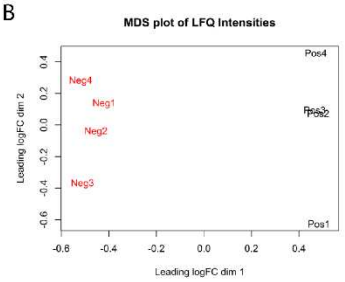
i



ii

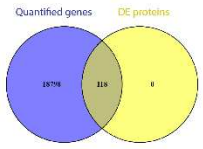


B

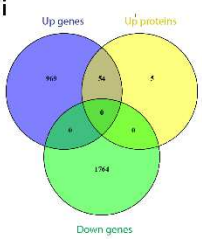


C

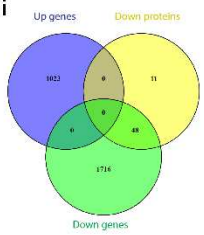
i



ii



iii



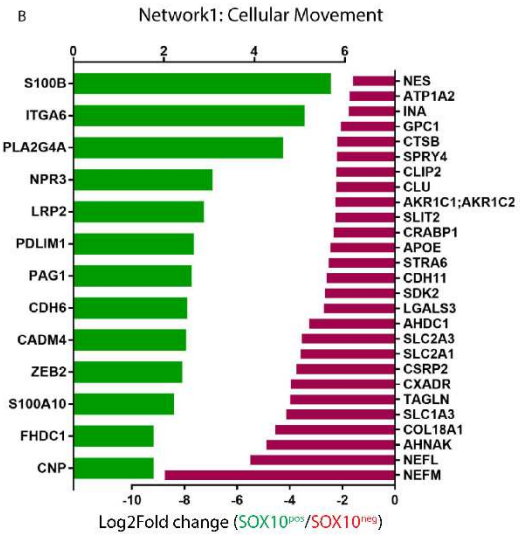
D

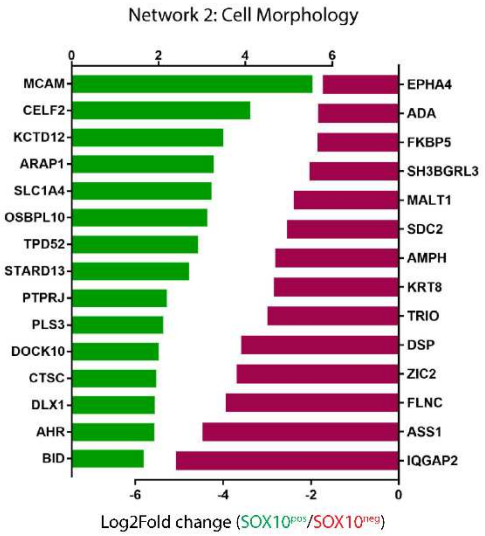
A

Top Networks

ID	Associated Network Functions	Score
1	Cellular Movement, Metabolic Disease, Neurological Disease	80
2	Cell Morphology, Cellular Assembly and Organization, Cellular Function and Maintenance	52
3	Cellular Development, Cellular Growth and Proliferation, Reproductive System Development and Function	45
4	Neurological Disease, Organismal Injury and Abnormalities, Skeletal and Muscular Disorders	29
5	Cellular Movement, Lipid Metabolism, Small Molecule Biochemistry	9

B





Suppl Fig 12

Proteomics

(A) Pairwise M (log2 ratio, y-axis) A (normalized mean average, x-axis) scatterplots (MAplots) of (i) SOX10 positive (ii) SOX10 negative replicates. Red line is the locally weighted scatterplot smoothing (loess) curve fitted to the MAplot. Each data point represents the protein intensity. All replicates within experimental groups show tightly distributed intensity values about y = 0, indicating good biological reproducibility. (B) Multi-dimensional scaling (MDS) plots of values over dimension 1 and 2 with samples colored and labeled by experimental groups (SOX10pos, black; SOX10neg, red). Four biological replicates of SOX10posand SOX10neg samples cluster close together within their experimental group and the first dimension separates SOX10posand SOX10neg cells. (C) i) Venn diagram showing the overlap of all identified transcripts and differentially expressed proteins between SOX10pos and SOX10neg cells. ii) Venn diagrams showing the overlap between significantly upregulated and downregulated transcripts between SOX10pos and SOX10neg cells with up regulated proteins and (iii) with down regulated proteins in SOX10pos cells.(D) The five most highly enriched associated networks (IPA analysis) in the cohort of differentially expressed proteins between SOX10posand SOX10neg. (E) The identity of the 40 and 29 significantly upregulated proteins in SOX10pos cells (green) and SOX10neg cells (red) are associated with cellular movement and cell morphology respectively.

Quantitative proteomics yielded a total of 3804 quantified protein groups of which 117 showed a statistically significant difference in abundance (FDR <0.01) between the two sample cohorts (Fig 6A, full list of proteins provided in Suppl data file 13). Within this cohort 59 and 58 protein groups were differentially more abundant in SOX10^{pos} SCP-like cells and SOX10^{neg} NPB-like cells, respectively (Fig 6 A, Suppl data file 14). As expected, SOX10 was within the top 10 differentially expressed protein groups upregulated in SOX10^{pos} cells, alongside several well-known SCP marker proteins such as CDH6, S100B and MPZ [47,135,136] and cranial NCC markers such as DLX1, ZEB2, MOXD1. Conversely, the neural plate border specifier ZIC2 [137], and the tight junction protein ZO2 [138] upregulated in SOX10^{neg} cells. To evaluate the concordance between differentially expressed proteins and mRNAs we compared the proteomics data with transcripts from the bulk RNA-seq dataset, revealing that 54 proteins (representing 91.5% of significantly upregulated proteins) were also upregulated at the transcript levels (Suppl Fig S12C i) and 48 proteins (representing 81.3% of significantly down-regulated proteins) were down-regulated at the transcript level (Suppl Fig S12Cii). Many of these differentially expressed proteins also marked the SOX10^{pos} SCP-like cells and SOX10^{neg} NPB-like cells in our scRNAseq dataset. Ingenuity Pathway Analysis (IPA) revealed that the 117 differentially expressed proteins were enriched for proteins involved in cellular movement (40) and cell morphology (29)

(Suppl Fig S12D). There were no overlapping proteins between these two networks, and collectively its members accounted for 58.5% of all 117 differentially expressed proteins. This enrichment for proteins associated with cellular movement and cell morphology fits well with the changes in morphology and motility of SOX10^{pos} SCP-like cells we observed and provides an opportunity to identify known and novel proteins that may regulate motility of human SCP-like cells. Of note, we found several significantly upregulated proteins in SOX10^{pos} cells that are associated with GTPases, including STARD13, ARAP1 and DOCK10, proteins that are known to play a role in modulating cytoskeletal dynamics and cell movement in other cell types [139,140]. In addition to these new putative regulators of motility human SOX10^{pos} SCP-like also express higher levels of CDH2 and CDH6, proteins known to regulate NCC migration in mouse and Xenopus [61,141,142], whereas SOX10^{neg} NPB-like cells express higher levels of CDH11, a protein that inhibits NCC migration in Xenopus when overexpressed [143].

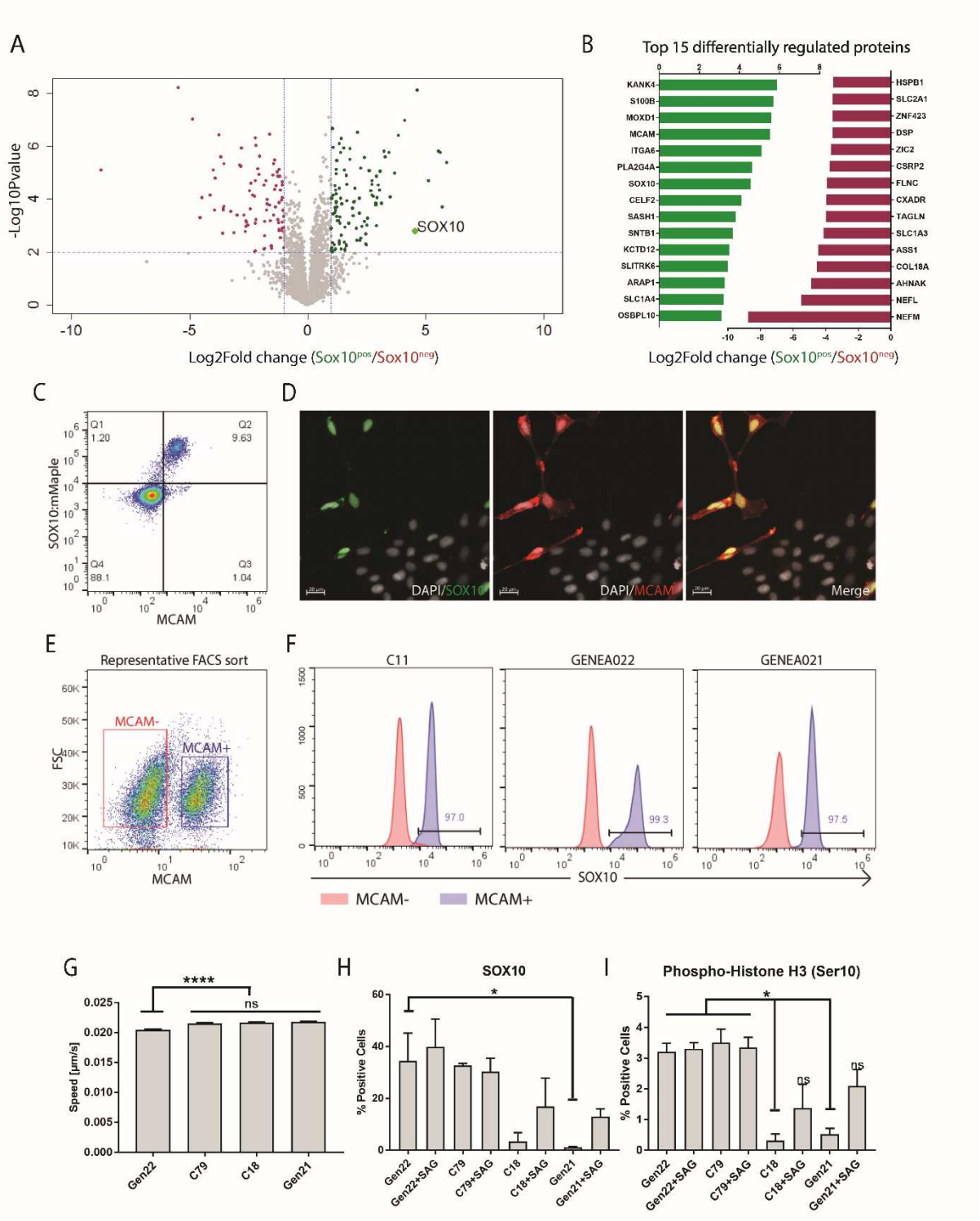
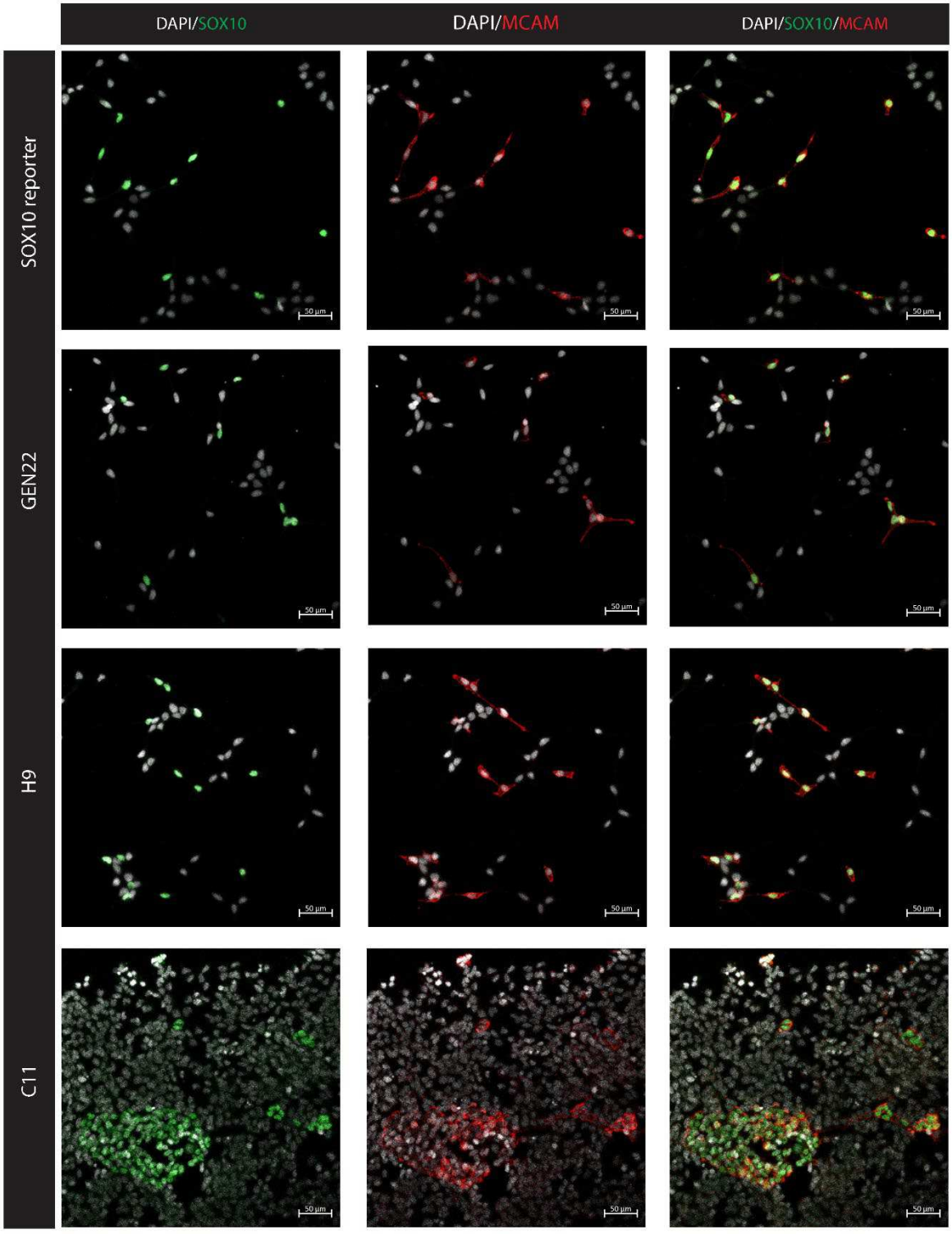

A
-Log10Pvalue
SOX10
Log2Fold change (Sox10pos/Sox10neg)
B
Top 15 differentially regulated proteins
KANK4
S100B
MOXD1
MCAM
ITGA6
PLA2G4A
SOX10
CELF2
SASH1
SNTB1
KCTD12
SLITRK6
ARAP1
SLC1A4
OSBPL10
HSPB1
SLC2A1
ZNF423
DSP
ZIC2
CSRP2
FLNC
CXADR
TAGLN
SLC1A3
ASS1
COL18A
AHNAK
NEFL
NEFM
Log2Fold change (Sox10pos/Sox10neg)
C
SOX10:mMaple
MCAM
Q1 1.20
Q2 9.63
Q4 88.1
Q3 1.04
D
DAPI/SOX10
DAPI/MCAM
Merge
E
Representative FACS sort
MCAM-
MCAM+
FSC
MCAM
F
C11
GENEA022
GENEA021
97.0
99.3
97.5
SOX10
MCAM-
MCAM+
G
Speed [µm/s]
****
ns
Gen22
C79
C18
Gen21
H
SOX10
% Positive Cells
*
Gen22
Gen22+SAG
C79
C79+SAG
C18
C18+SAG
Gen21
Gen21+SAG
I
Phospho-Histone H3 (Ser10)
% Positive Cells
*
ns
ns
Gen22
Gen22+SAG
C79
C79+SAG
C18
C18+SAG
Gen21
Gen21+SAG

Figure 6: Proteomic analysis identifies MCAM as a novel cell surface marker of SOX10pos SCP-like cells that enables isolation and disease modelling.

(A) A volcano plot of differentially abundant proteins based on fold change versus t-test probability. The vertical axis (y-axis) indicates –log10(FDR-adjusted P-value). The horizontal axis (x-axis) indicates relative abundance ratio (log2 fold change) between SOX10pos and SOX10neg. Cutoff values are illustrated by dotted vertical (at ±1.0 log2 fold) and horizontal lines (FDR-adjusted P-value = 0.01). Data points in red are proteins that satisfied both -fold and FDR-adjusted P-value cutoff criteria. (B) Significantly upregulated (green) or downregulated (red) plasma membrane associated proteins identified by IPA in SOX10pos cells compared with SOX10neg cells. (C) Flow cytometry and (D) Immunocytochemical analyses confirm co-expression of MCAM and SOX10 in C11:SOX10 hiPSC derived Day 21 unsorted SCP-like cell cultures. (E) Representative MCAM antibody FACS sort of untagged C11 hiPSC-derived SCP-like cells. (F) FACS analyses of SOX10 expression of MCAM-sorted SCP-like cells derived from hiPSC (C11), euploid hESC (GENEA022) and Down syndrome hESC (GENEA021). (G) Migration speeds (µm/s) of MCAMpos SCP-like cells cells derived from control and DS-hESC and hiPSC lines (mean±SEM, n>400 cells). (H) Quantification of SOX10 expression in MCAM-sorted SCP-like cells derived from control and DS-hESC and hiPSC lines in the absence or presence of 0.5 µM SAG. (I) phospho-histone H3 expression in MCAM-sorted SCP-like cells derived from control and DS-hESC and hiPSC lines in the absence or presence of 0.5 µM SAG (25 random fields counted per cell line per condition with samples de-identified to the experimenter). Data represented as the mean number (from duplicate experiments) of mitotic figures ± SEM. Statistical significance was determined by a Student's t-test.

To identify cell surface markers specific to SOX10^{pos} SCP-like cells we employed IPA assisted classification of differentially upregulated proteins associated with the plasma membrane and cell surface (Fig 6B), and identified CDH6, ITGA6, and MCAM. Amongst these MCAM (CD146) was the most highly differentially expressed cell surface protein in SOX10^{pos} SCP-like cells, and this correlated well with our scRNAseq data set, which also revealed uniform and strong expression of MCAM transcripts in SOX10^{pos} SCP-like cells (see Fig 4B). To test the utility of MCAM in isolating human SOX10^{pos} SCP-like cells we labelled Day 21 single cell dissociated C11:SOX10 hiPSC-derived SCP cultures with MCAM/CD146 antibodies. Flow cytometric analysis demonstrated that MCAM antibodies almost exclusively marked SOX10:mMaple expressing cells (Fig 6C), and only 1% of SOX10^{pos} MCAM^{neg} and 1% of MCAM^{neg} SOX10^{pos} cells. In agreement with this data immunofluorescent analysis revealed strong MCAM immunoreactivity on the plasma-membrane of all SOX10^{pos} cells whereas MCAM antibodies did not label SOX10^{neg} cells (Fig 6D). MCAM immunoreactivity was also detected in migrating human SOX10^{pos} SCP-like cells that were microinjected into E9.25 mouse embryos and cultured for 6 hrs (Fig 2G), and

MCAM mRNA is also expressed in murine (Suppl Fig 8)[5] and human embryo-derived SCPs [24]. MCAM antibody-mediated sorting of SCP-like cells derived from the unmodified C11 hiPSC cell line (Fig 6E) and the hESC line Genea022 (Fig 6F) exposed to our SCP differentiation medium for 21 days labelled >99 % of SOX10 expressing cells (IHC data shown in Suppl Fig 13).


DAPI/SOX10
DAPI/MCAM
DAPI/SOX10/MCAM
SOX10 reporter
GEN22
H9
C11
50 µm

Suppl Fig 13

SOX10 expression in MCAM sorted SCP-like cells derived from human pluripotent stem cell lines.

SCP-like cells (day 21) derived from the C11-SOX10 reporter line, its parental control line C11, and the healthy control hESC lines GEN22 and H9, double labelled with MCAM and SOX10 antibodies. Nuclei stained with DAPI. (Scale bar 50 □M).

We concluded that MCAM antibodies allow the isolation of SOX10 expressing SCP-like cells from human pluripotent stem cell lines subjected to our SCP differentiation protocol.

MCAM sorted human pluripotent stem cell derived SCP-like cells can be used for disease modelling.

To exemplify the utility of MCAM sorted SCP-like cells we next generated SCP-like cells from the Down syndrome hESC line Genea021 [31] and FACS sorted the cells with MCAM antibodies. MCAM sorting again efficiently separated $SOX10^{pos}$ cells from $SOX10^{neg}$ cells with >97% accuracy (Fig 6C, E and F), similar to what we observed for the C11:SOX10-reporter line. Interestingly, SCP-like cell generation was less efficient than in the sibling euploid cells (Genea22) (Fig 6H), and the MCAM-sorted Down syndrome SCP-like cells displayed profoundly decreased proliferation (as indicated by reduced fraction of cells labelled with phospho-Histone 3 antibodies) as compared to euploid cells (Fig 6I). Interestingly, this defective proliferation of Down syndrome SCP-like cells could be rescued by stimulation of the Hedgehog pathway with the small molecule SAG (Fig 6H and I), reaching levels comparable to euploid cells. We concluded that MCAM antibodies can be used to isolate pure SOX10 expressing SCP-like cells from control and disease affected human pluripotent stem cell cultures subjected to SCP differentiation conditions. We further show that such disease-specific SCP-like cells can be used to investigate neurocristopathies such as Down syndrome, and perhaps other genetic diseases that involve SCP, such as neuroblastoma or neurofibromatosis.

Discussion

Animal models have provided pivotal insights into the molecular mechanisms that govern the specification, migration, differentiation and cell fates of NCC and SCP during early embryogenesis [5,11,18]. While the origin of the different types of SCPs is increasingly better understood [5,144], very little is known about the earliest steps of human SCP specification. Since the neural plate border, from which NCCs and SCPs are derived, is a heterogeneous population of cells that *in vivo* is patterned by highly dynamic local morphogen gradients and mechanical forces, it is challenging to accurately mimic this process *in vitro* with human pluripotent stem cell cultures. In this study we therefore opted to expose hiPSC cultures to a constant growth factor and small molecule regime free of retinoic acid (to prevent posterization) rather than attempting to optimise sequential differentiation steps that foster cell state transitions with varying efficiencies. We reasoned that this strategy would allow us to better model the cell intrinsic driven processes of SCP generation, while recognizing that this approach would likely not permit us to capture the full gamut of cell types that are specified from the NPB *in vivo and* would compromise SCP generation efficiency. Our one step differentiation regimen fostered the emergence of a relatively homogeneous SOX10^{pos} SCP-like cell population over a 3-week period, as exemplified by our scRNAseq analysis. Interestingly, this developmental time frame is remarkably similar to that observed for SCP specification in early human embryos (Carnegie stage 9)[145]. The most highly expressed genes in our SCP-like cells are also abundant in primary SCPs derived from the developing adrenal gland of week 7-8 human embryos (Carnegie stages 18-23), 5-week human embryo-derived SCPs [24], and highly enriched in murine SCP-hub cells and early Schwann cells [5]. Since SCP specification during early human embryogenesis has not been studied to date we anticipate that mapping of the human cranial SCP specific proteins, genes, miRNA, LncRNAs and splicing events reported in this study will prove useful in future validation studies in human embryos or primates, and will provide a resource for identifying coding and non-coding genes linked to diseases that involve SCPs (such as neurofibromatosis). Our experimental approach of sorting SOX10^{pos} and SOX10^{neg} cells was not designed to determine whether the SCP were derived from an intermediate NCC population or were directly specified from the NPB-like cells. It is highly likely that inclusion of the SCP specifier molecule NRG1 and the SCP survival factor FGF2 into the medium from

the start of differentiation is a strong contributor to this differentiation trajectory, as previously shown by others [22]. Future lineage tracing experiments and pseudo-temporal modelling of the entire culture (instead of performing CELLseq2 on FACS sorted populations) over temporally closely spaced time points should help resolve this outstanding question. In this respect it was notable that neither the SOX10^{neg} nor the SOX10^{pos} populations contained cells in the G1 cell cycle phase, challenging the notion that this cell cycle phase is a prerequisite for EMT. The data presented in this study suggest that EMT and motility of CSCPs may be related to alternative splicing changes and expression of non-coding RNAs. Our data indicate that the SCP-like cells generated with our one-step protocol are highly similar to the SCP generated by Kim et al., who employed pre-neuralized cultures that were exposed to NRG1, both in terms of the expression of key SCP genes such as CDH19, DHH, MPZ, ITGA4, absence of KROX20 expression, and the absence of NCC genes such as SLUG, SNAI2, MSX2, as well as in terms of cell morphology and expandability. We further find that acquisition of SOX10 expression in our system is accompanied by decreased *SNAIL2* expression and upregulation of CDH6, an observation that is consistent with the repression of chicken Cadherin6B expression by Snail2 observed in chicken NCCs [40]. Furthermore, our SOX10^{pos} cells gain *TFAP2a* [146] and lose CDH11 (transcript and protein) expression, two events that were previously shown to control SOX10 expression during Xenopus NCC development [147]. The substantial overlap between our bulk and scRNAseq data and the single cell datasets of murine SCPs generated by Soldatov et al. and Kastriti et al. [4,5] further highlights the substantial degree of transcriptional similarity between mouse SCPs and the human SCP-like cells generated with our protocol. The transcriptomic, metabolomic and proteomic characterisation of both the starting SOX10^{neg} NPB-like population and the SOX10^{pos} SCP-like cell population described in this study provide several new insights into human SCP biology. Firstly, our data show that SOX10 expressing SCP-like cells exhibit dramatic changes in alternative splicing of cell adhesion molecules likely involved in the motility of SCPs [148], including events that impact the inclusion or exclusion of functional protein domains (e.g. DOCK6). Our data also revealed the increased expression of non-coding RNAs such as members of the miR-506-514 cluster, previously implicated in metastatic melanoma, elevated miR-9 [149,150] which is known to mediate downregulation of E-cadherin, and increased Linc00461 which has been associated with increased cell migration in multiple cancers [151–153]. We further

noted a strong downregulation of LINC01116 in SOX10^{pos} SCP-like cells, a LincRNA that inhibits the migration and invasion of squamous cell carcinoma and nasopharyngeal carcinomas [154] by blocking the epithelial-mesenchymal transition process and that represses MOXD1 [155], a gene that is strongly upregulated in our SOX10^{pos} SCP-like cells, suggesting a similar mechanism may operate in cranial SCP-like cells. Our data further reveal that cranial SCP-like cells are less reliant on glycolysis as compared to the NPB-like cells they originate from, in contrast to what is observed for NCC. Because the expression of HOX genes, which is virtually absent in our cranial SCP-like cells, can be readily induced by retinoic acid it is likely that these cells may also prove useful for studying cell types derived from posteriorly specified SCPs, such as enteric neurons and boundary cap cells. In this study we identify MCAM as a specific cell surface marker for human cranial SCP-like cells. MCAM is a protein that can control the expression of integrins and cell adhesion, and was previously detected in migrating avian trunk NCC [156], in NCCs differentiated from a mouse *Sox10-IRES-Venus* ESC line [157], in a cranial neural crest cell line established from E8.5 mouse embryos [158], and in FACS sorted GFP^{pos} cranial NCCs isolated from E9.5 Wnt1Cre; Z/EG mouse embryos [159]. In agreement with this data the murine single cell expression data sets of Soldatov et al. and and Kastriti et al. show that MCAM expression marks a large proportion of murine CNCC/SCPs at E8.5 [4,5]. Here we show that MCAM antibodies can be used to purify SOX10^{pos} human cranial SCP-like cells to near homogeneity via FACS or MACS, irrespective of the hiPSC/hESC line used. In this paper we use this approach to sort Down syndrome CSCP-like cells, identifying defects in specification, migration and proliferation of cranial DS SPC-like cells as compared to a sibling euploid control, that can at least in part be overcome by Hedgehog pathway stimulation with SAG. These data are consistent with previous observations made for NCC in the Ts65Dn mouse model of DS and reinforces the notion that defective Hedgehog signalling [160] is a likely contributor to the craniofacial defects associated with Down syndrome, as well as perhaps other features of DS such as Hirschsprung's disease, congenital heart defects, teeth and tongue abnormalities [20], and altered skin pigmentation that likely involve SCPs. Collectively, the insights and datasets provided in this study should prove to be valuable resources for progressing investigations into the role of SCPs in normal and abnormal early human development, neurocristopathies, peripheral neuropathies [102,161], and the pathogenesis of neuroblastoma [162], and neurofibromatosis [23,26].

Acknowledgements

We would like to acknowledge the contribution of the Stem Cells consortium (https://data.bioplatforms.com/organization/pages/bpa-stemcells/consortium) in the generation of data used in this publication. The Initiative is supported by funding from Bioplatforms Australia through the Australian Government National Collaborative Research Infrastructure Strategy (NCRIS). Suzy Butchers is greatly acknowledged for enabling data accessibility. We would also like to acknowledge the Monash Biomedical Proteomics facility and funding from the Australian Research Council's Centre for Research Excellence "Stem Cells Australia". This work was in part performed at the Australian Institute of Bioengineering and Nanotechnology node of the Australian National Fabrication Facility, a company established under the National Collaborative Research Infrastructure Strategy to provide nano and micro-fabrication facilities for Australia's researchers. This research was supported by grants NHMRC 2020 MRFF Childhood Cancer Research Application APP2007404, NHMRC 2020 Stem Cell Therapies Mission APP2009101 (EW &QS), Swiss National Science Foundation, Early Postdoc Mobility Grant no. 165165 (CV), Professor Lars Keld Nielsen is supported by the Novo Nordisk Foundation (Grant numbers: NNF20CC0035580 and NNF14OC0009473).

Materials and Methods

Maintenance of Human Pluripotent Stem Cells

Human pluripotent stem cells were cultured as colonies under feeder free conditions on ECM (Sigma Aldrich) and fed daily with mouse embryonic fibroblast conditioned medium supplemented with 90ng/ml bFGF (Miltenyi Biotec) and β-mercaptoethanol. Cells were manually passaged every 7 days. For feeder-free cultures used for differentiation experiments, cells were grown on dishes coated with ECM in mTeSR1 medium (StemCell Technologies) with daily media changes.

Reporter line generation

sgRNAs were designed using a then-available CRISPR design web resource (http://crispr.mit.edu) and cloned into the px459.2 vector as described (Ran et al. 2013). To construct the reporter cassette, 5' and 3' homology arms were amplified by PCR from C11 hiPSC genomic DNA and cloned into a selectable plasmid containing IRES-mMaple-P2A-Puro-FRT-poly(A)-PGK1-HSV.TK-2A-NeoR-bGHpoly(A)-FRT cassette. hiPSC line C11 cells were electroporated using Amaxa 4D (program CA-137, P3 kit) with sgRNA/SpCas9-encoding and reporter targeting vectors, and correctly targeted clones were identified by targeted allele-specific PCR after geneticin selection (400ug/mL, 5 days) and selected as previously described (Ran et al. 2013).

SCP-like cell Derivation from Human Pluripotent Stem Cells

For SCP differentiation, human pluripotent stem cell colonies maintained in mTESR were re-plated as single cells at a density of 90,000 cells/cm^2 onto ECM coated culture dishes. The next day, culture medium was switched to SCP differentiation medium containing 2xB27 supplement without retinoic acid, 8ng/ml β-FGF, 1% GlutaMAX, 1% mNEAA, 10ng/ml Heregulin β-1, 3μM CHIR99021, 10μM SB431542, 1xPenstrep, 0.11mM β-mercaptoethanol, in DMEM/F12. Cells were passaged every 4-5 days and maintained in differentiation medium for 21 days.

In vitro migration of SCP-like cells

Cells were plated on ECM coated surfaces at 1500 cells/cm^2. The next day, time-lapse analysis was performed using the Operetta High Content Imaging System (PerkinElmer). Images were captured every 3 minutes with a 20x air objective both in

bright field and digital phase contrast over 16 hours. Morphology and kinetic properties of migration were calculated from at least 400 cells per condition using the Harmony high content imaging and analysis software (http://www.perkinelmer.com). Cell masks were created using digital phase contrast channel, method M (Diameter: 35µm; Splitting coefficient: 0.60; Common threshold: 0.35).

Cell sorting

Cells were dissociated and resuspended in FACS buffer (PBS+/+, 1% FBS, 1mM EDTA). Magnetic activated cell sorting was performed using the Neural Crest Stem Cell Microbeads kit (Miltenyi Biotec) according to the manufacturer's instructions. For FACS, cells were incubated with validated and concentration optimised primary antibodies diluted in FACS buffer for 45min (or as per manufacturer's instructions) on ice. Following three washes with FACS buffer, cells were incubated with appropriate Alexa Fluor secondary antibody (Invitrogen) at a 1:1000 dilution for 1 hour on ice. Cells were rinsed three times in FACS buffer prior to sorting.

Flow cytometric analysis

Cells were collected and spun down at 300x*g* for three minutes. Pellets were washed once with PBS+/+ to remove excess medium. Cells were fixed in 4% PFA for 10 minutes at RT and blocked with 5% goat serum or fetal bovine serum. For intracellular and nuclear markers, cells were permeabilised with 0.1% Triton X-100 for 15 minutes at RT minutes before blocking. Blocked cells were incubated with validated and concentration-optimised primary antibodies diluted in 1% goat serum or fetal bovine serum in PBS+/+for 45 minutes on ice. Following three washes with PBS+/+, cells were incubated with appropriate Alexa Fluor secondary antibody (Invitrogen) at a 1:1000 dilution for 1 hour on ice. An Accuri C6 Flow Cytometer (BD) and CFlow Sampler software (https://www.bdbiosciences.com) were used for analysis.

Differentiation of SCP-like cells

For differentiation into mesenchymal stem cells SCP-like cells (or NCCs generated according to [38] were cultured in FBS medium containing 10% FBS (Hyclone, GE Healthcare), 1% non-essential amino acids, 1% GutaMAX in DMEM/F12. Medium was changed on alternate days and cells were enzymatically passaged once every 7 days. On day 21, cells were harvested for flow cytometry analysis. Differentiation into

osteocytes, adipocytes and chondrocytes was performed using the commercially available differentiation kits (GIBCO) following the manufacturer's instructions.

Differentiation into peripheral neuronal cells was carried out by plating cells at a density of $1x10^5$ cells/cm^2 on surfaces coated with poly-L-ornithine (100µg/ml, Sigma Aldrich) and mouse laminin (1ug/ml, Thermofisher Scientific) NC differentiation medium. The next day, medium was changed to peripheral neuron differentiation medium containing 1% N2 supplement, 1ml penicillin-streptomycin, 10ng/ml BDNF (Peprotech), 10ng/ml NGF (Peprotech), 10ng/ml GDNF (Peprotech), 10ng/ml NT3 (Peprotech), 50ug/ml sodium L-ascorbate and 0.5 mM cAMP (Sigma Aldrich) in DMEM/F12, as previously described (Menendez et al., 2013). Cells were maintained in this medium for 21 days with media changes every alternate day and cells were fixed for immunocytochemical analysis to determine cell identity.

Differentiation into Schwann cells was carried out by plating cells at a density of $9x10^4$ cells/cm^2 on surfaces coated with ECM in NC differentiation medium. The next day, medium was changed to Schwann cell differentiation medium containing 1xN2 supplement, 10ng/ml CNTF (Peprotech), 20ng/ml neuregulin (Peprotech), 10ng/ml bFGF (Miltenyi Biotech) and 0.5 mM cAMP in DMEM/F12 as previously described (Lee et al., 2010). Cells were maintained in this medium for 21 days with media change every alternate day after which cells were fixed for immunocytochemical analysis to determine cell identity.

Differentiation into melanocytes was performed as previously described (Mica et al. 2013) with appropriate modifications. Briefly, SCPs were cultured on Matrigel-coated culture dishes. The next day, the culture medium was replaced with melanocyte derivation medium containing 1x N2, 1x B27, 0.005% BSA, 2 mM Glutamax, 0.11 mM b-mercaptoethanol, 3 µM CT 99021, 20 ng/ml FGF2 (Peprotech), 0.5 mM dbcAMP (Tocris Biosciences), 25 ng/ml BMP4, and 100 nM EDN3 (Tocris Biosciences) in DMEM/F12 and Neurobasal medium (1:1 mix). The medium was replaced every other day. The confluent cells were dissociated using Accutase (Millipore) treatment and passaged at 1:6 ratio.

Immunocytochemistry

Cells were fixed in 4% paraformaldehyde for 15 minutes at room temperature (RT) and blocked with 5% goat serum or fetal bovine serum (GIBCO) in 1X PBS+/+ (Hyclone, GE Healthcare). For intracellular or nuclear proteins, 0.1% Triton X-100 (Sigma Aldrich) was first used to permeabilise the cells at RT for 15 minutes before blocking. The blocked cells were incubated overnight at 4^{o}C with validated and concentration optimised primary antibody diluted in 1% goat serum or fetal bovine serum in PBS+/+. This was followed by three washes with PBS+/+ and incubation with the appropriate Alexa Fluor secondary antibody (Invitrogen) at a 1:1000 dilution. Cells were washed three times with PBS+/+ and nuclei were labelled using DAPI (1:5000 in 1X PBS+/+) and embedded in Vectashield Antifade Mounting Medium (Vector Laboratories). An Olympus IX51 fluorescence microscope or ZEISS 710 Confocal microscope and ImageJ (https://imagej.net) or ZEN Microscope and Imagining software (https://www.zeiss.com) was used for image capture and analysis. Mouse embryos were fixed in 4% paraformaldehyde and immunostained whole or following cryosectioning at 12µm. Embryos and cryosections were blocked in 10% DAKO serum-free blocking solution in PBST (0.1% Triton X-100). Primary antibodies used are listed in Table S1 and the following secondary antibodies were used: anti-rabbit AlexaFluor 488 (Life Technologies #A21206) 1:200; anti-rabbit AlexaFluor 555 (Life Technologies #A31572); anti-mouse AlexaFluor 647 (Life Technologies #A31571). Immunolabelled samples were mounted in Prolong Diamond Anti-fade with DAPI (Life Technologies #P36962). Images were acquired using a Zeiss LSM 800 confocal microscope.

Mouse embryo culture and microinjection

To obtain embryos of defined gestational ages, mice were mated in the evening, and the morning of vaginal plug formation was counted as embryonic day (E) 0.5. Pregnant dams were humanely euthanized at relevant days post vaginal plug detection by CO_2 inhalation and cervical dislocation. Embryos were removed from the uterus in fresh Tyrode's solution. Reichert's membrane was dissected from the embryos being careful to keep the yolk sac intact. SOX10:mMaple positive cells grown in selection media were harvested, labelled with 50uM DiI at 37^{o}C for 10min and washed twice with PBS. Labelled cells were resuspended at a concentration of 1X10^{5} cells/ml in growth media supplemented with 1%BSA and 0.02% phenol red. Pulled microinjection needled were

pre-treated with Sigmacote (Sigma) for 1hr prior to flushing with 2% BSA and loading with labelled cells. Microinjection was performed with an Eppendorf FemtoJet (pi, 250hPa; ti, 0.1s) with labelled cells implanted into the tissue adjacent the neural tube or within the pharyngeal arch territory. Injected embryos were immediately placed into preheated and equilibrated DR50 culture media (50% rat serum (Envigo), 50% DMEM/F12/Glutamax with antibiotics). Embryos were grown in glass roller culture bottles (BTC Engineering) for 6-24 hrs at 37^{o}C in a gas atmosphere of 5% O_2, 5% CO_2, and 90% N_2.

Multi-omics analyses

Six replicates of a single clone of the SOX10 reporter line were maintained as undifferentiated colonies for over five passages prior to initiation of differentiation and represent six biological replicates. These six replicates were subjected to our SCP differentiation protocol. On day 21 of differentiation, cells were enzymatically passaged and separated into two fractions per replicate. One fraction was replated and subjected to 4 µg/ml Puromycin selection for 3 days to obtain pure populations of SOX10^{pos} cells. Cells that did not express mMaple were considered to be SOX10^{neg} and these were isolated from the second fraction by FACS sorting using a BD influx cell sorter. These pure cell preparations were next plated separately at identical densities (1x10^5 cells/cm^2) and allowed to recover and expand for another 48 hours. Cells were harvested and processed for each omic analysis as described below.

Bulk RNAseq

Culture medium was aspirated from SOX10 positive and negative cells. Cultures were washed three times with PBS-/- and lysed in Trizol reagent (Thermo Fisher Scientific). Total RNA was extracted using PureLink RNA Mini Kit (Thermo Fisher Scientific) involving a DNAse treatment step according to manufacturer's instructions. RNA quantity was measured using a NanoDrop 1000 (Thermo Fisher Scientific) and RNA quality was analysed using a 2100 Bioanalyser (Agilent Genomics). All samples had RIN>8. RNA samples were submitted to the Australian Genome Research Facility, Victoria (http://www.agrf.org.au) for cDNA synthesis, DNA library preparation and sequencing. Paired end FASTQ libraries were quality checked using FASTX toolkit (http://hannonlab.cshl.edu/fastx_toolkit/index.html by Hannon Lab) and MultiQC [163]. Reads were aligned to the hg38 genome using Subread package v1.20.6 [164] under R

v3.2.2. Annotated (Ensembl 86) raw gene counts were considered for differential expression analysis if all replicates in a comparison group ($SOX10^{neg}$ or $SOX10^{pos}$) expressed at least 1 CPM or greater. The filtered gene counts were TMM scaled [165] and Limma's Voom function was used [166] to model the mean-variance relationship of the data. Resulting log2 CPM counts were fitted to a linear model with Voom's precision weights and further empirical Bayes moderation [167]. Pairwise DEG top tables from moderated t-tests were produced and genes were considered to be differentially expressed for false discovery rate adjusted [168] p-values (referred to as adj.Pval < 0.01).

Single cell RNAseq (Cell-Seq2)

$SOX10^{pos}$ and $SOX10^{neg}$ single cells were deposited into individual wells of a 384 well plate using a BD Aria Flow Cytometer. Plates containing sorted cells were immediately sealed with AlumaSeal® II film (Sigma Aldrich), centrifuged at 1000*g* and frozen at -80 ^{o}C. For DNA library preparation we essentially followed the CEL-Seq2 protocol by Hashimshony and colleagues [57] except that the CEL-Seq2 primer was modified with a 6 bp unique molecular identifier (UMI) consisting of 6 random nucleotides directly upstream of a cell specific index sequence, and minor changes in reagents use (see Suppl data file 15). The final library was eluted in 20 μl of 10 mM Trizma hydrochloride solution (Sigma-Aldrich) and submitted to the Ramaciotti Center for Genomics (https://www.ramaciotti.unsw.edu.au/) for sequencing. Raw sequences were aligned using the *Subread* (v1.5.1) software with the ENSEMBL v86 Homo Sapiens genome as reference. Summarising into counts with deduplication of unique molecular identifiers was performed using scPipe (v0.0.9000) [169]. Adjusted library size normalisation was performed with library size adjustment factors from the *computeSumFactors* function of the *scater* (v1.4.0) package for R [170]. Expression was quantified by adjusted counts per million (CPM) values in the log2 scale. The biological coefficient of variation (BCV) was estimated using the *estimateDisp* function of the *edgeR* (v3.18.1) package for R. Comparisons of variability between genes was based on the ratio between point and trended estimates of BCV for the genes [171]. t-SNE plots were computed using Rt-sne (v0.13) package for R-statistical-software [172].

Label-free Quantitative Proteomics

Sample preparation - Medium was aspirated from adherent SOX10:mMaple positive and negative cells. Cells were washed three times with ice-cold PBS containing protease inhibitor cocktail (1:100, Sigma #P8340) and phosphatase inhibitor cocktail (1:500 Sigma #P5726). Cells were lysed in RIPA buffer containing protease and phosphatase inhibitors. The lysate was vigorously vortexed to disrupt gelatinous DNA and then centrifuged for 15 minutes @ 14000x*g* at 4°C. The supernatant was collected for proteomic analysis.

Sample processing and LC-MS –The experimental procedures used in this study have been performed as reported in detail in the 'Protein Fractionation-free Workflow' section previously [173,174]. Briefly, sample lysates were enriched for proteins and a protein assay was used to estimate protein concentrations [173,174]. The protein pellets after protein enrichment were resuspended in 50 mM Tris-HCl, 7 M urea, 2 M thiourea, and 2% (w/v) CHAPS, pH 8.8. Equal amounts of proteins from each replicate were subjected to disulfide bond reduction with 10 mM DTT at 4°C overnight (~16 to 18 h) and for 2 h at ~22°C. Reduced samples were alkylated with 50 mM iodoacetamide at ~22°C for 2 h in the dark. Samples were precipitated with methanol, digested by LysC and trypsin combinatorial digestion strategy and analysed by nUHPLC-LTQ-OrbitrapElite using LCMS conditions as described previously [173].

Equal quantities of proteins from four independently prepared matched replicate sets of $SOX10^{pos}$ and $SOX10^{neg}$ cell lysates (~3 million cells) were individually digested and subjected to nanoUltraHPLC-LTQ-OrbitrapElite analysis as previously described [173]. Protein identification and label-free quantification were performed using MaxQuant v1.5.2.8 [175], with default parameters. Briefly, carbamidomethylcysteine was set as fixed and protein N-terminal acetylation, methionine oxidation and asparagine\glutamine deamidation were set as variable modifications. Enzyme specificity was set to trypsin, allowing for cleavage N-terminal to proline. Second peptide and match between runs were enabled. The false discovery rate (FDR) at the peptide and protein level were set to 1% and both unique and razor peptides were used for label-free quantification. Perseus version 1.5.3.0 [176] was used for statistical analysis of LFQ peak intensities. The normalized LFQ data values were retained for all protein groups that were identified in at least three $SOX10^{pos}$ and\or $SOX10^{neg}$ replicates. The LFQ intensities were logarithmized and missing values were imputed by drawing random numbers from a normal distribution to simulate signals from low

abundant proteins. The width parameter for the normal distribution was chosen as 0.3 of the standard deviation (σ) of all measured values and the center was shifted towards low abundance by 1.8 times this σ. LFQ intensities were grouped by conditions and a two-sided *t*-test was applied with a permutation-based FDR cutoff of 1% (q-value = 0.01) and the S0 parameter was set to 2.0.

Metabolomics

Culture medium was collected 24 hours after the previous media change from $SOX10^{pos}$ and $SOX10^{neg}$ cells and cultures. Collected medium was snap frozen in liquid nitrogen. Cells were washed once in MilliQ water. Sufficient liquid nitrogen was added to cover the base of the culture surface. Next, cells were incubated with 600 µl (per 10 cm^2 surface area) of ice cold 9:1 MeOH:$CHCl_3$ containing internal standards for 10 minutes on ice. The lysate was transferred to a clean tube, incubated on ice for another 5 minutes, and centrifuged at 16,100 ×g for 5 minutes at 4°C following which the supernatant was collected. The supernatant and spent media samples were submitted to Metabolomics Australia (http://www.metabolomics.net.au/) for metabolite extraction and analysed by Liquid Chromatography-Mass Spectrometry (LC-MS) or Gas Chromatography-Mass Spectrometry (GC-MS) followed by identification and statistical analysis [177,178]. Standard mixes were run within this analysis to help in identification of metabolites. Raw data files were analysed using Agilent MassHunter Quantitative software to check retention times of compounds and to verify peak integrations were correct. Compounds were identified based on retention time of the specific compounds in the standard mixes and ion masses and quantified based on areas under the curve. Statistical analysis of the output matrix in R via log transformation and median normalisation was carried out before multivariate analysis and student's t-test was performed. All features were identified based on retention time of the standards and masses of ions. Features with p-values or BH adjusted p-values of less than 0.05 were considered as significantly different between groups.

Quantitative RT-PCR analysis

Cells were lysed in TRIzol reagent (Ambion) followed by purification using the PureLink RNA Mini Kit (Ambion) according to the manufacturer's instructions. One microgram of RNA was used for reverse transcription with the iScript cDNA synthesis kit (Bio-Rad) according to the manufacturer's specifications. Quantitative RT-PCR

(qPCR) analysis was performed in biological triplicate and technical duplicate using the CFX96 real-time PCR detection system (Bio-Rad) with SsoFast™ EvaGreen® Supermix mix (Bio-Rad). Gene expression was quantified relative to *GAPDH* expression. Sequence of primers used are listed in Table S2.

Additional Resources used for data analysis and deposition

Differentially expressed genes identified by bulk RNA-seq were uploaded into GOrilla (http://cbl-gorilla.cs.technion.ac.il/) for Gene Ontology analysis using default settings. Differentially expressed proteins identified from proteomic analysis were uploaded into Ingenuity Pathway Analysis software (IPA, www.ingenuity.com) along with the gene identifiers and corresponding fold change and p-value. Each identifier was mapped to its corresponding gene object in the Ingenuity knowledge base. Networks were then algorithmically generated using default settings based on their connectivity and a score was assigned. The score is a numerical value used to rank networks according to how relevant they are to the genes in the input dataset. Venny 2.1 (http://bioinfogp.cnb.csic.es/tools/venny/) was used to generate Venn diagrams showing overlap between omics datasets. Cell profiler (http://cellprofiler.org/) was used to quantify morphological differences observed between $SOX10^{pos}$ and $SOX10^{neg}$ cells. Graphs and statistics were generated using either Prism 5 software (GraphPad) or Rstudio. The data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus [179] and are accessible through GEO Series accession number GSE242011 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE242011).

Author contributions

E.J.W. conceived the project, designed experiments, performed data analysis and interpretation, and wrote the manuscript; A.B. performed the majority of wet-lab experiments, wrote the manuscript, and performed data analysis; C.M., I.P. and Q.S. performed and analysed all mouse microinjection experiments; S.S., D.A-Z. and S.H. N. supervised, and conducted all Cell-Seq2 analyses; K.A.D, M.H. and J.G. designed, conducted and analysed discovery proteomics experiments; K.T. and L.N. conducted and supervised the splicing analyses; H.L., G.P., N.G., O.H., I.S., C.V. and C.GI. performed supporting wet-lab experiments and C.GI in manuscript preparation.; D.O. contributed to reporter cell line design and generation., JCW supervised MSC differentiation, C.W., D.T., R.S., I.P., D.D.S., and M.M. coordinated and analysed RNAseq and metabolomics data. All authors edited and approved the manuscript.

Figure Legends

Figure 1: ***One-step generation of SCP-like cells from C11:SOX10/mMAPLE reporter hiPSC.***

(A) Flow cytometric analysis of mMAPLE expression in C11:SOX10 reporter hIPSC line exposed to our SCP differentiation medium for 0,10,16 and 21 days. (B) mMAPLE fluorescence emerges in small clustered cell populations and is photo-switchable to red fluorescence following 30 sec exposure to 405nm light. (C) Flow cytometric validation of the concordance between mMAPLE expression and SOX10 expression in C11:SOX10 hiPSC subjected to 21 days of SCP differentiation medium. (D) Flow cytometric analysis of mMAPLE expression before and after Puromycin selection. (E) Flow cytometric analysis of mMAPLE expression of sorted SOX10/mMAPLE expressing cells cultured in differentiation medium for 10 passages. (F) Phase contrast (G) mMAPLE fluorescence and (H) Validation of nuclear SOX10 expression via IHC in mMAPLE positive cells counterstained with DAPI (nuclei) Scalebar= 50 μm (I) mMAPLE expressing cells co-express SCP markers CDH19 and Gap43 (J) Q-PCR analysis of the temporal kinetics of SCP and NPB mRNA species in C11:SOX10 hIPSC subjected to SCP-differentiation medium for 21 days (One-way ANOVA. *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001).

Figure 2: ***In vitro and in vivo migratory properties of SCP-like cells from C11:SOX10/mMAPLE reporter hiPSC.***

(A) Brightfield images of $SOX10^{pos}$ and $SOX10^{neg}$ cells show morphological differences. $SOX10^{neg}$ cells exhibit a round cell morphology and organise themselves in distinct compact colonies while $SOX10^{pos}$ cells have a stellate appearance and are uniformly distributed in the culture dish. Scale bar = 100μm. (B) Graphs showing cell numbers, distance between a cell and its first closest neighbour, and area of the field of view that is occupied by cells (mean±SEM, n=9 fields of view). (C) Representative tracks of 5 randomly chosen $SOX10^{pos}$ and $SOX10^{neg}$ cells over 2 hours from a 16 hour time-lapse experiment involving a minimum of 400 cells. The intersection of the x- and y-axes was taken as the starting point of each cell path. Each tick mark on the axes represents 20μm. (D) Graphs quantify speed, (E) net displacement and (F) total distance travelled by $SOX10^{pos}$ and $SOX10^{neg}$ cells (mean±SEM, n>400 cells). (G)

Tracing of human SCP-like cells microinjected into E8.5 embryos. (a) E8.5 embryos within their yolk sacs were microinjected with 20-50 dil labelled human SCP-like cells (white arrow in dashed circle). Scale bar = 100μm. (b) After 24 hrs in culture media cultured embryos develop to approximately E9.5. Scale bar = 100μm. (c) Representative image of an E8.5 embryo labelled for Sox10 depicting the site of SCP-like cell injection (white arrow in dashed circle). Scale bar = 100μm. (d) DiI labelled SCP-like cells (red, arrowheads) incorporate into the host NCC and SCP migratory stream into the first pharyngeal arch (PA1). e, eye; ov, otic vesicle; r2, rhombomere 2, r4, rhombomere 4. Scale bar = 100μm. (e-f) Higher magnification of boxed area in (d). scale bar = 50 μm. (H) Tracing of human SCP-like cells microinjected into E9.25 embryos. (a) DiI labelled SCP-like cells (red) were microinjected into the distal NCC / SCP migratory sream (white arrow in dashed circle) withinPA1. Scale bar = 100μm. (b-d) DiI labelled (red) MCAM+ SCP-like cells (green, arrowheads) incorporate into the host NCC and SCP migratory stream. e, eye; ov, otic vesicle. Scale bar = 50μm.

Figure 3: ***Total RNAseq analysis of SOX10^{pos} and SOX10^{neg} cells derived from C11:SOX10/mMAPLE reporter hiPSC.***

(A) Volcano plot of bulk RNA-seq data. Genes are ranked according to their statistical adj.p-value (y-axis) and their relative abundance ratio (log2 fold change SOX10^{pos} /SOX10^{neg}, x-axis). The two vertical dashed blue lines indicate log2 fold change values of 1 and -1. The horizontal blue dashed line indicates an adjusted p-value of 0.01. Genes upregulated in SOX10^{pos} cells with a fold change ≥2 and adj.Pval <0.01 are depicted in green, and those downregulated with a fold change ≤-2 and adj.Pval <0.01 are shown in red. Grey dots represent genes that do not differ significantly in their expression levels between SOX10^{pos} and SOX10^{neg} samples. (B-E) mRNA expression shifts of cell adhesion genes as cells transition from SOX10^{neg} to SOX10^{pos} cells (Log2-fold change). (F) mRNA expression shifts of miRNAs and (G) LncRNAs as cells transition from SOX10^{neg} to SOX10^{pos} cells (Log2-fold change).

Figure 4: ***Single cell RNAseq analysis of SOX10^{pos} and SOX10^{neg} cells derived from C11:SOX10/mMAPLE reporter hiPSC.***

(A) Hierarchical clustering by SOX10 expression group of all single cells (n = 402 single cells, top 500 variable genes) analysed by scRNA-seq. Left Y-axis axis shows the relationships between the genes according to similarity in expression pattern. Right

Y-axis contains gene names and enriched biological processes in each of the subclusters marked by coloured boxes. (B) Gene expression projection of canonical SCP markers, murine SCP markers, NCC/SCP markers, genes associated with EMT and cell motility, and novel or cell cycle genes on tSNE clustered scRNAseq data from $SOX10^{neg}$ to $SOX10^{pos}$ cells.

Figure 5: ***Splicing differences between $SOX10^{neg}$ and $SOX10^{pos}$ cells derived from C11:SOX10/mMAPLE reporter hiPSC.***

(A) Significantly differentially expressed (log_2Fold-Change >= 2; p-adj < 0.01) RNA-binding proteins detected in the data. (B) The top most frequently detected splicing patterns based on the total number of events representative of these patterns. Of the total, the number of AS and DS events are indicated. (C) The distribution of DS events across different functional regions in their host genes. (D) Network representation of the top 20 most significantly enriched GO terms in DS protein-coding genes. (E) Potential Biotypes of the spliced isoforms generated for each event.
● Know-Coding: At-least one full-length transcript supporting this isoform is an annotated protein-coding transcript. ● UTR-5: Event lies in the 5'-UTR ● UTR-3: Event lies in the 3'-UTR ● Frameshift: Isoform introduces a frameshift as compared to the best matching annotated protein-coding transcript. ● Frameshift_PTC_50_Any: Isoform introduces a frameshift and a pre-termination codon (PTC) that agrees with 50nt NMD rule is introduced in at least one translation frame. ● PTC_50_All: PTC(s) agreeing with the 50nt NMD rule are introduced in all translation frames. ● Non-Coding-Gene: Event is hosted by a non-coding gene. Coverage diagrams for differential splicing events in (F) *CELF2*, (G) *ROBO1*, (H) *EXOC1,* (I) *CTNND1*, (J) *FERMT2*, and (K) *EPB41* genes. In each sub-diagram, the lower and upper coverage panel represent $SOX10^{neg}$ and $SOX10^{pos}$ samples respectively. The Y-axis indicates the number of reads. The splicing diagrams (dark-grey) represent event isoforms in order (i.e. isoform 1 to n). The Bar-plots indicate the fitted proportions for the event isoforms in order. The Y-axis indicates the proportion.

Figure 6: ***Proteomic analysis identifies MCAM as a novel cell surface marker of SOX10^{pos} SCP-like cells that enables isolation and disease modelling.***

(A) A volcano plot of differentially abundant proteins based on fold change versus t-test probability. The vertical axis (y-axis) indicates –log10(FDR-adjusted P-value). The horizontal axis (x-axis) indicates relative abundance ratio (log2 fold change) between SOX10^{pos} and SOX10^{neg}. Cutoff values are illustrated by dotted vertical (at ±1.0 log2 fold) and horizontal lines (FDR-adjusted P-value = 0.01). Data points in red are proteins that satisfied both -fold and FDR-adjusted P-value cutoff criteria. (B) Significantly upregulated (green) or downregulated (red) plasma membrane associated proteins identified by IPA in SOX10^{pos} cells compared with SOX10^{neg} cells. (C) Flow cytometry and (D) Immunocytochemical analyses confirm co-expression of MCAM and SOX10 in C11:SOX10 hiPSC derived Day 21 unsorted SCP-like cell cultures. (E) Representative MCAM antibody FACS sort of untagged C11 hiPSC-derived SCP-like cells. (F) FACS analyses of SOX10 expression of MCAM-sorted SCP-like cells derived from hiPSC (C11), euploid hESC (GENEA022) and Down syndrome hESC (GENEA021). (G) Migration speeds (μm/s) of MCAM^{pos} SCP-like cells cells derived from control and DS-hESC and hiPSC lines (mean±SEM, n>400 cells). (H) Quantification of SOX10 expression in MCAM-sorted SCP-like cells derived from control and DS-hESC and hiPSC lines in the absence or presence of 0.5 μM SAG. (I) phospho-histone H3 expression in MCAM-sorted SCP-like cells derived from control and DS-hESC and hiPSC lines in the absence or presence of 0.5 μM SAG (25 random fields counted per cell line per condition with samples de-identified to the experimenter). Data represented as the mean number (from duplicate experiments) of mitotic figures ± SEM. Statistical significance was determined by a Student's t-test.

Supplementary figure legends:

Suppl Fig 1

Generation of the C11:SOX10 hiPSC line.

(A) Schematic of the knock-in targeting construct and location of the primers used for identification of correctly targeted clones (red arrows). (B) The targeted C11:SOX10 hiPSC line expresses the pluripotency markers NANOG, OCT4 and TRA1-60 (Scale bars = 100µm). (C) Flanking regions of the SOX10 stop codon/upstream 3' UTR were analysed for potential sgRNA target regions. The sgRNA used in this study along with its cutting site and associated PAM sequence is marked. (D) PCR analysis of genomic DNA of the targeted C11:SOX10 hiPSC line with primers anchored within the exogenous cassette sequence within the targeting construct and flanking SOX10 locus sequences exemplifies the correct targeting of the reporter cassette. (E) The correctly-targeted C11:SOX10 hiPSC line displays a normal karyotype (example of 1 of the 15 analyzed metaphase spreads). (F) SURVEYOR nuclease assay demonstrates cutting efficiency of the chosen sgRNA on *SOX10* target site in HEK293FT cells. (G) Sanger sequencing cannot detect any mutations or indels in the top four predicted off-target binding sites for the selected gRNA in genomic DNA isolated from the established C11:SOX10 hiPSC line.

Suppl Fig 2

SCP differentiation protocol applied to other human pluripotent stem cell lines

Flowcytometric analysis of SOX10 expression (using SOX10 antibody labelling) of Gen22 and H9 human ES cells exposed to our one-step SCP differentiation medium for 21 days.

Suppl Fig 3

Hox gene expression in SOX10^{pos}-SCP-like cells generated in the absence or presence of retinoic acid.

qPCR analysis of HOX gene expression in reporter line derived Puromycin selected SOX10^{pos}-SCP-like cells maintained in differentiation medium either with or without RA. Data are shown as mean of relative expression normalized to GAPDH ± SEM of three independent experiments (N=3). Unpaired t-test with Welch's correlation not

assuming equal standard deviation (ns not significant, * p<0.05, ** p<0.01, *** p<0.001).

Suppl Fig 4

Multilineage differentiation potential of SOX10^{pos}-SCP-like cells

(A) Brightfield microscopy and (B) Flow cytometric analysis of C11:SOX10 reporter hIPSC line derived SCP-like cells cultured in FCS containing medium for 3 weeks reveals expression of Stro1, CD105, CD90, CD73, CD44 and CD29 expected of mesenchymal stromal cell (MSC). SCP-like cell derived MSC show tri-lineage differentiation potential as indicated by (C) Chrondrogenic (Alcian Blue) (D) Adipogenic (Oil redO), and (E) osteogeneic (Alizarin red) stains. (F) C11:SOX10 reporter hIPSC line derived SCP-like cells readily differentiate into Schwann cells marked by S100beta. Cell nuclei were stained with DAPI. (Scale bars = 100µm) (G) C11:SOX10 reporter hIPSC line derived NCC differentiate into betaIIItub and peripherin double labelled peripheral neurons in contrast to C11:SOX10 reporter hIPSC line derived SCP-like cells. (H) Representative bright-field image of SCP-like cells differentiated into melanocytes at day 16. Right panel, the cell pellet after centrifugation displays a dark tone, indicative of pigmentation.

Suppl Fig 5

Neural crest cell marker expression in SOX10^{pos}-SCP-like cells

(A) Flow cytometric analysis of non-puromycin selected C11:SOX10 reporter hIPSC line derived SCP-like cell cultures following labelling with canonical NCC-markers SOX10 & HNK1 or SOX10 & p75. (B) scRNAseq UMAP of B3GAT (HNK1) expression in C11:SOX10 reporter hIPSC line derived SOX10^{pos} SCP-like cells and SOX10^{neg} NPB-like cells.

Suppl Fig 6

FACS gating of SOX10^{pos}and SOX10^{neg} cells destined for multi-omics analyses

Flow cytometry plot outlining the gating strategy used to isolate pure Sox10^{pos} and SOX10^{neg} cells from C11:SOX10 reporter hIPSC subjected to our SCP differentiation

medium for 21 days. These cells were next used for bulkRNAseq, metabolomics and discovery proteomics.

Suppl Fig 7

ScRNAseq quality control

(A) FACS profile indicating the gating strategy used for isolating cells destined for Cellseq. (B) Cell samples were filtered by requiring at least 12000 molecules counted after UMI deduplication and at least 4000 genes detected. Post quality control, 337 SOX10pos cells and 165 SOX10neg- cells were deemed suitable for further analysis. (C) Genes that were either not expressed with at least 1 count in 10% of samples or were not annotated with a description from the Homo sapiens Bioconductor database were filtered out. Post filtering, of 21540 genes, 11321 were retained for further analysis. (C) Volcano plot of average log CPM values (x-axis) and their relative abundance ratio (log2 fold change) between $SOX10^{pos}$ and $SOX10^{neg}$ (y-axis). Genes upregulated in $SOX10^{pos}$ cells with a fold change ≥2 and adj.p-value <0.05 are depicted in red, and those downregulated with a fold change ≤-2 and adj.p-value<0.05 are shown in blue. Grey dots represent proteins in the arrays that do not differ significantly in expression between SOX10 positive and negative cells. (D) Venn diagram showing the number of transcripts in bulk only (left) and single cell only (right) and the number of genes common to both datasets (centre). (E) Venn diagram showing the number of differentially expressed genes between $SOX10^{pos}$ and $SOX10^{neg}$ cells in bulk (left) only and single cell only (right) in comparison to the total number of transcripts in the bulk dataset. (F) Venn diagram showing the number of upregulated transcripts in $SOX10^{pos}$ cells in bulk only (left), single cell only (right) and the number of transcripts common to both (centre). (G) Venn diagram showing the number of upregulated transcripts in $SOX10^{neg}$ cells in bulk only (left), single cell only (right) and the number of transcripts common to both (centre). (H) The density of log CPM values for post-filtered bulk RNA-seq data (green), $SOX10^{pos}$ (red) and $SOX10^{neg}$ (blue) data are shown for each replicate/cell. (I) t-SNE map of quality-filtered, bulk and scRNA-seq data. (J) t-SNE map of quality-filtered single-cell transcriptomes from SOX10 positive and negative groups. (K) Boxplot of Euclidean distances between each point and the centroid of the cluster.

Suppl Fig 8

Gene expression concordance between top human SCP-like cells and mouse hub-SCPs and Schwann cell lineage.

Gene-based embedding, and SCENIC-based UMAP embedding of the top differentially expressed genes in human SCP-like cells (from bulk and scRNAseq) in lineage traced murine NCC and Schwann cell lineages through the online visualisation tool at https://protect-au.mimecast.com/s/C_kdCJyBq7FYQA02cGTDAu?domain=adameykolab.hifo.meduniwien.ac.at/, as outlined in (Kastriti et al. 2022) that leverages data from (Soldatov et al. 2019)

Suppl Fig 9

Single cell analysis of neural plate border and LncRNA expression.

(A) Gene expression projection of selected neural plate border marker genes on scRNAseq t-SNE plots (B) gene expression projection of selected LncRNAs on scRNAseq t-SNE plots.

Suppl Fig 10

Conservation Score weighted splicing maps of exon skipping events

Conservation Score weighted splicing maps of exon skipping event regulation by (A.) CELF2 and (B.) ELAVL4. The fraction of 6-mers matching preferred binding 6-mers for each protein were counted in 30bp windows within the regulated alternative exons and surrounding intronic regions and compared using the Wilcoxon Rank-Sum test against a control set of alternative exons. The 6-mer counts were scaled by the mean PhastCons100 conservation scores of the sequence. (C) Projection of CELF2 and ELAVL4 expression at a single cell level.

Suppl Fig 11

Metabolomics

(A) PCA score plot of median normalised cell matrix of 124 features identified during metabolomic analysis of cell lysates from SOX10^{pos} (green) and SOX10^{neg} (red) groups. Pooled biological quality control samples (PBQCs) fell in between the 2 groups as expected. (B) PCA score plot of median normalised cell matrix of 89 features identified during metabolomic analysis of spent media from SOX10^{pos} (green) and SOX10^{neg} (red) groups. Pooled biological quality control samples (PBQCs) fell in between the 2 groups as expected. (C) List of differentially expressed (based on p-value) metabolites in SOX10^{pos} and SOX10^{neg} cell lysates in decreasing order of significance. Blue rows show metabolites that were upregulated in SOX10^{neg} samples and orange rows show metabolites that were upregulated in SOX10^{pos} samples. Metabolites in bold are significant based on BH-adjusted p-value<0.05. (D) List of differentially expressed (based on p-value) metabolites in SOX10^{pos} and SOX10^{neg} cell lysates in decreasing order of significance. Blue rows show metabolites that were upregulated in SOX10^{neg} samples and orange rows show metabolites that were upregulated in SOX10^{pos} samples. Metabolites in bold are significant based on BH-adjusted p-value<0.05. (E) Differential expression of mRNA species coding for glycolytic enzymes between SOX10^{pos} and SOX10^{neg} cells (Bulk RNAseq data, log-fold change shown).

Suppl Fig 12

Proteomics

(A) Pairwise M (log2 ratio, y-axis) A (normalized mean average, x-axis) scatterplots (MAplots) of (i) SOX10 positive (ii) SOX10 negative replicates. Red line is the locally weighted scatterplot smoothing (loess) curve fitted to the MAplot. Each data point represents the protein intensity. All replicates within experimental groups show tightly distributed intensity values about y = 0, indicating good biological reproducibility. (B) Multi-dimensional scaling (MDS) plots of values over dimension 1 and 2 with samples colored and labeled by experimental groups (SOX10^{pos}, black; SOX10^{neg}, red). Four biological replicates of SOX10^{pos}and SOX10^{neg} samples cluster close together within their experimental group and the first dimension separates SOX10^{pos}and SOX10^{neg} cells. (C) i) Venn diagram showing the overlap of all identified transcripts and differentially expressed proteins between SOX10^{pos} and SOX10^{neg} cells. ii) Venn

diagrams showing the overlap between significantly upregulated and downregulated transcripts between SOX10^{pos} and SOX10^{neg} cells with up regulated proteins and (iii) with down regulated proteins in SOX10^{pos} cells.(D) The five most highly enriched associated networks (IPA analysis) in the cohort of differentially expressed proteins between SOX10^{pos}and SOX10^{neg}. (E) The identity of the 40 and 29 significantly upregulated proteins in SOX10^{pos} cells (green) and SOX10^{neg} cells (red) are associated with cellular movement and cell morphology respectively.

Suppl Fig 13

SOX10 expression in MCAM sorted SCP-like cells derived from human pluripotent stem cell lines.

SCP-like cells (day 21) derived from the C11-SOX10 reporter line, its parental control line C11, and the healthy control hESC lines GEN22 and H9, double labelled with MCAM and SOX10 antibodies. Nuclei stained with DAPI. (Scale bar 50 μM).

Supplementary data file legends:

Supplementary data file 1: Bulk RNAseq comparison of SOX10^{pos} and SOX10^{neg} cells. Sheet 1: Ad.P Value ranked Sheet 2: Log fold change ranked.

Supplementary data file 2: Significantly up- and downregulated genes in SOX10^{pos} and SOX10^{neg} cells.

Supplementary data file 3: Significantly up- and downregulated non-coding RNAs in SOX10^{pos} and SOX10^{neg} cells.

Supplementary data file 4: scRNAseq data (Cellseq2) of FACS sorted SOX10pos and SOX10neg cells.

Supplementary data file 5: Collapsed scRNAseq dataset of FACS sorted SOX10pos and SOX10neg cells.

Supplementary data file 6: Comparison of human SCP-like scRNAseq data with published scRNAseq and snRNAseq datasets.

Supplementary data file 7: List of RNA binding proteins, splicing factors and spliceosomal proteins used for RNAseq analysis of SOX10^{pos} and SOX10^{neg} cells.

Supplementary data file 8: Significant differential splicing events between SOX10^{pos} and SOX10^{neg} cells.

Supplementary data file 9: GO analysis of splicing events in SOX10^{pos} cells.

Supplementary data file 10: CLIP-seq and RBNS datasets used for assessing position dependent splicing events.

Supplementary data file 11: Metabolomics analysis of SOX10^{pos} and SOX10^{neg} cells.

Supplementary data file 12: GC-MS and LC-MS metabolomics data of SOX10^{pos} and SOX10^{neg} cells.

Supplementary data file 13: Proteomics data of SOX10^{pos} and SOX10^{neg} cells.

Supplementary data file 14: Statistically significantly differentially expressed proteins in SOX10^{pos} and SOX10^{neg} cells.

Supplementary data file 15: 384 Cellseq primers reference

Supplementary data file 16: Glimma plot ZIP file for gene expression interrogation of Cell seq data.

Supplementary Table legends:

Supplementary Table 1: Antibodies and concentrations used in various applications.

Supplementary Table 2: Primer sequences used.

Supplementary materials and methods:

Reporter line generation.

Because SOX10 is haplo-insufficient [180,181], we inserted an IRES-linked puromycin-selectable mMaple [182] reporter (structure of the targeting construct including 800 bp homology arms outlined in Suppl Fig 1A) 129 bps downstream from the stop codon in exon 4 of the SOX10 locus using co-delivery by transient transfection of CRISPR-Cas9 and locus- specific gRNA (sequence and target location in Suppl Fig 1C) with demonstrated cutting efficiency (Suppl Fig 1F) into the previously generated C11-hIPSC line [32]. The clonally derived C11:SOX10-IRES-mMaple-Puro hiPSC line (here referred to as C11:SOX10) was expanded, and shown to be karyotypically normal (Suppl Fig 1E). PCR analysis of genomic DNA with primers anchored within the targeting construct and exogenous flanking sequences (location of primers shown in Suppl Fig 1A) revealed the expected amplicons that exemplified the correct targeting of the reporter cassette (Suppl Figure 1D). The C11:SOX10 reporter line expressed the pluripotency markers NANOG, OCT4 and TRA-1-60 (Suppl Fig 1B) and Sanger sequencing of PCR-amplicons of the top four predicted off-target human genome cutting sites for the selected gRNA revealed the absence of indels or mutations at these sites (Suppl Fig 1G). Previously Horikiri et al [183] reported the generation of a similar SOX10-Nano-Lantern Reporter iPSC line that was, however, not offering purification of NCC/SCP populations via antibiotic selection.

Detection, Quantification and Differential Analysis of AS events

Detection and quantification of the alternative splicing (AS) events was performed according to the ASTA-P pipeline described in GitHub - uqktiwar/ASTAP: A pipeline for the detection, quantification and statistical analysis of complex alternative splicing events. For filtering the splicing event isoforms at least five reads were required to uniquely map to a unique feature of an isoform in at least five samples (N_{th} = 5, N_s = 5). Isoforms without any unique features were retained if all their junctions were expressed with at least five reads in at least five samples.

For differential testing, we filtered the events to retain those expressed with a minimum total read count of 20 in both conditions. The isoform proportions were modelled using

DRIMSeq [110]. We identified the differentially spliced events by comparing the full model (~ Group ϵ [$SOX10^{neg}$, $SOX10^{pos}$]) to the null model (intercept only) using a likelihood ratio test. To further identify the differentially regulated isoforms of these events, we used the stageR package [184] to assign FDR values at the isoform level. Significantly differentially spliced isoforms were selected as those with an event level FDR < 0.05, and expressing at least one isoform with |Δψ| >= 10% and isoform level FDR < 0.05.

Functional Analyses of Differentially Spliced Events

For functional interpretation, apart from a GO enrichment analysis (and GO network construction) of their host genes using the ShinyGO R package [185], we also mapped the events to transcript regions, annotated protein domains/ features, and assessed their coding potential. In order to assign events to gene regions, we compared event coordinates to the coordinates of different functional regions (5'-UTR, CDS, and 3'-UTR) within the best coding transcripts for each gene, indicated in the APPRIS database [186]. For events lying in CDS regions, we further determined their overlap with annotated Pfam [187] and Uniprot [188] domain records available through the UCSC table browser [189] to assign them to "Annotated Domain" regions.

To predict coding potentials of the event isoforms, we first compared each isoform to the structures of the annotated protein-coding transcripts expressed by their host genes. If at least one protein-coding transcript was structurally consistent with an isoform, we annotated it as "Known-Coding". For the remainder of isoforms, if the isoform introduced a frameshift as compared to the best matching protein-coding transcript, we annotated it as "Frameshift". Further, for predicting nonsense mediated decay (NMD), if an isoform was structurally consistent with an annotated NMD transcript and contained the annotated NMD stop codon, we annotated it as "Known-NMD". Otherwise, if the isoform included a pre-termination codon (PTC), which agrees with the 50nt rule [190], in at least one translation frame, we annotate it as "NMD_50_Any". If PTCs are included in all frames, we annotate it as "NMD_50_All". If the isoform did not conform to any of these rules (no NMD, no PTC, no frameshift), we annotated it as "Predicted-Coding". Only events lying in the CDS regions were analysed for their coding potentials, the rest were annotated as "UTR-5", "UTR-3" or "Non-Coding-Gene" depending on their gene region annotations.

Splicing Factor Binding Models

We determined binding models for the splicing factors that underwent significant changes in expression (|Log_2Fold-Change| >= 2, FDR < 0.05) between the NPB and CNCC populations. For different factors, we used different source datasets and their respective published protocols to extract their preferred binding k-mers.

Splicing Map Generation

We constructed splicing maps for only the 2D exon skipping events, since these dominated (91/201) the differentially spliced events set. In order to construct the splicing maps, we compiled three sets of exon-skipping events – positively regulated, negatively regulated, and a control set. The regulated events were identified as those with at least one isoform having a |Δψ| >= 10% at an FDR < 0.05. The control set comprised of events where all isoforms had a ψ ϵ [10%, 90%] and a |Δψ| < 5% at an event level FDR > 0.5, and no genomic overlap with the regulated set. For each skipped exon, we extracted the 50 and 250 base pairs (bps) of sequence from the 5p and 3p ends of the exon and its flanking introns respectively.

Next, we counted the occurrences of a splicing factor's top k-mers in a 30bp window, sliding it along the sequences 1bp at a time. We assigned the fraction of nucleotides (out of 30) covered by the k-mers as the score for each position. We compared the scores for the regulated sets with that for the control set at each position using the Wilcoxon rank sum test, followed by FDR adjustment, to find the significantly scoring positions. For the conservation weighted splicing map, we first retrieved the phastCons100way [121] scores for the sequence positions using the GenomicScores R package [191]. Following this, the position specific scores were weighted by the average conservation score of the counting window.

CEL-Seq2 protocol and modifications: The CEL-Seq2 6 bp index sequences were replaced by 384 × 8 bp indices (Suppl data file 15) designed to have a Hamming distance of >2, to avoid wrong allocation of sequencing reads (with up to 2 sequencing errors in the index sequence). We further used the original CEL-Seq2 primer sequence containing an extended T7 promotor binding site as well as longer Illumina 5'-adapter sequence:

CGATTGAGGCCGGTAATACGACTCACTATAGGGGTTCAGAGTTCTACAGTCCG ACGATCNNNNNNGTAGCTCATTTTTTTTTTTTTTTTTTTTTTTTV. The 8 bp index is underlined. All other regions are constant. Additionally, AMPure XP beads (Beckman Coulter - A63880) were replaced by NucleoMag NGS Clean-up and Size select magnetic beads (Macherey-Nagel - 7449970.5) and the Phusion® High-Fidelity PCR Master Mix was replaced with HF Buffer (NEB M0531) by KAPA HiFi HotStart ReadyMix (KapaBiosystems KK2602). Plates containing cells were thawed on ice and briefly centrifuged (1 min at 1500 rpm). To lyse the cells and anneal the mRNA capture primer, each cell was reconstituted in 1.2 μl of primer/lysis mix containing a well specific cell index consisting of 20 nM indexed polydT primer, 1:6,000,000 dilution of ERCC RNA spike-in mix (Ambion - 4456740), 1 mM dNTPs (NEB), 1.2 units SUPERaseIN Rnase Inhibitor (Thermo Fisher - AM2696), DEPC water (Thermo Fisher - AM9920)], mixed and incubated 5 minutes on ice. The plate was then incubated at 65 °C for 5 minutes and immediately chilled on ice for at least 2 minutes before adding 0.8 ml reverse transcription reaction mix (2 ml RT reaction: 1× First Strand buffer (Invitrogen), 20 mM DTT (Invitrogen), 4 units RNaseOUT (Invitrogen), 10 units SuperScript II (Invitrogen)). The plate was incubated at 42 °C for 1 hour, 70 °C for 10 minutes and chilled to 4 °C to generate first strand cDNA. For second strand cDNA synthesis 10 mL of second strand reaction mix were added 2.31 μl Second strand buffer (Invitrogen 10812-014), 0.3 μL DNA Polymerase I from E.coli (Invitrogen 18010-025), 0.08 μL E. coli DNA ligase (Invitrogen 18052-019), 0.08 μL RnaseH (E.coli) (Invitrogen 18021-071), 7 μL DEPC water (Thermo Fisher – AM9920)]. The plate was incubated at 16 °C for 2 hours to generate double stranded cDNA. All samples were pooled and cleaned using a 1.2× NucleoMag NGS Clean-up and Size select magnetic beads (Macherey-Nagel) according to manufacturer's instructions. To reduce the amount of beads for each 100 µL pooled sample 20 µL beads and 100 µL Bead binding buffer (20% PEG8000, 2.5 M NaCl, pH5.5) was added. The cDNA was eluted in 25.6 ml DEPC water and kept with beads for the following in-vitro-transcription (IVT) reaction with 38.4 μL of IVT reaction mix 6.4 μL of each of the following: A,G,C,U, 10× T7 buffer, T7 enzyme (MEGAscript T7 Transcription Kit – Ambion AM1334)] and incubated at 37 °C for 13 hours and then chilled and kept at 4 °C. To remove leftover primers 24 μL ExoSAP-IT PCR Product Clean-Up (Affymetrix -78200) was added and the sample was incubated at 37 °C for 15 minutes and then chilled and kept at 4 °C.

Chemical heat fragmentation was performed by adding 22 µl of 10× Fragmentation buffer (RNA fragmentation reagents - AM8740) to the sample and incubation in pre-heated thermal cycler at 94 °C for 2.5 minutes followed by immediate chilling on ice and addition of 11 µL of Fragmentation Stop buffer (RNA fragmentation reagents - AM8740). The fragmented amplified RNA was purified using 1.8× RNAClean XP beads (Beckman coulter - A63987) according to manufacturer's instruction and eluted in 6 ml DEPC water of which 5 µL (no beads) were transferred to a fresh tube for library preparation. The fragmented RNA was transcribed into cDNA using 5'-tagged random hexamer primers 9 (GCCTTGGCACCCGAGAATTCCANNNNNN) introducing a partial Illumina adapter as also described in CEL-Seq2 [57]. To remove RNA secondary structure and anneal the mRNA capture primer 1 µL of tagged random hexamer (100 mM) and 0.5 µL of 10 mM dNTPs (dNTP solution set NEB - N0446S) were added to the sample and incubated at 65 °C for 5 minutes and immediately chilled on ice for at least 2 minutes before adding 4 µL reverse transcription reaction mix (10 ml RT reaction: 1× First Strand buffer (Invitrogen - 18064-014), 20 mM DTT (Invitrogen - 18064-014), 4 units RNaseOUT (Invitrogen 10777-019), 10 units SuperScript II (Invitrogen 18064-014)). The PCR primers introduce the full-length adaptor sequence required for Illumina sequencing. PCR was performed in 12.5 µL using half of the ranhexRT sample as a template (1× KAPA HiFi HotStart ReadyMix (KapaBiosystems), 400 nM each primer). The final PCR amplified Library was subjected to two consecutive 1× NucleoMag NGS Clean-up and Size select magnetic beads (Macherey-Nagel) according to manufacturer's instructions.

References

1. Knecht, A. K. & Bronner-Fraser, M. Induction of the neural crest: a multigene process. *Nat. Rev. Genet.* **3**, 453–461 (2002).
2. Dong, Z. *et al.* Neu differentiation factor is a neuron-glia signal and regulates survival, proliferation, and maturation of rat schwann cell precursors. *Neuron* **15**, 585–596 (1995).
3. Nagoshi, N. *et al.* Ontogeny and Multipotency of Neural Crest-Derived Stem Cells in Mouse Bone Marrow, Dorsal Root Ganglia, and Whisker Pad. *Cell Stem Cell* **2**, 392–403 (2008).
4. Soldatov, R. *et al.* Spatiotemporal structure of cell fate decisions in murine neural crest. *Science* **364**, eaas9536 (2019).
5. Kastriti, M. E. *et al.* Schwann cell precursors represent a neural crest-like state with biased multipotency. *EMBO J.* **41**, e108780 (2022).
6. Jacob, C. Transcriptional control of neural crest specification into peripheral glia. *Glia* **63**, 1883–1896 (2015).
7. Jessen, K. R. & Mirsky, R. The origin and development of glial cells in peripheral nerves. *Nat. Rev. Neurosci.* **6**, 671–682 (2005).
8. Dyachuk, V. *et al.* Parasympathetic neurons originate from nerve-associated peripheral glial progenitors. *Science* **345**, 82–87 (2014).
9. Espinosa-Medina, I. *et al.* Parasympathetic ganglia derive from Schwann cell precursors. *Science*.
10. Joseph, N. M. *et al.* Neural crest stem cells undergo multilineage differentiation in developing peripheral nerves to generate endoneurial fibroblasts in addition to Schwann cells. *Development* **131**, 5599–5612 (2004).
11. Adameyko, I. *et al.* Schwann cell precursors from nerve innervation are a cellular origin of melanocytes in skin. *Cell* **139**, 366–379 (2009).

12. Nitzan, E., Pfaltzgraff, E. R., Labosky, P. A. & Kalcheim, C. Neural crest and Schwann cell progenitor-derived melanocytes are two spatially segregated populations similarly regulated by Foxd3. *Proc. Natl. Acad. Sci.* **110**, 12709–12714 (2013).

13. Lumb, R., Buckberry, S., Secker, G., Lawrence, D. & Schwarz, Q. Transcriptome profiling reveals expression signatures of cranial neural crest cells arising from different axial levels. *BMC Dev. Biol.* **17**, 5 (2017).

14. Freter, S., Fleenor, S. J., Freter, R., Liu, K. J. & Begbie, J. Cranial neural crest cells form corridors prefiguring sensory neuroblast migration. *Dev. Camb. Engl.* **140**, 3595–3600 (2013).

15. Li, W. *et al.* Peripheral nerve-derived CXCL12 and VEGF-A regulate the patterning of arterial vessel branching in developing limb skin. *Dev. Cell* **24**, 359–371 (2013).

16. Mukouyama, Y., Shin, D., Britsch, S., Taniguchi, M. & Anderson, D. J. Sensory nerves determine the pattern of arterial differentiation and blood vessel branching in the skin. *Cell* **109**, 693–705 (2002).

17. Sudiwala, S. & Knox, S. M. The emerging role of cranial nerves in shaping craniofacial development. *Genes. N. Y. N 2000* **57**, e23282 (2019).

18. Furlan, A. *et al.* Multipotent peripheral glial cells generate neuroendocrine cells of the adrenal medulla. *Science* **357**, eaal3753 (2017).

19. Furlan, A. & Adameyko, I. Schwann cell precursor: a neural crest cell in disguise? *Dev. Biol.* **444 Suppl 1**, S25–S35 (2018).

20. Kaukua, N. *et al.* Glial origin of mesenchymal stem cells in a tooth model system. *Nature* **513**, 551–554 (2014).

21. Mayor, R. & Theveneau, E. The neural crest. *Dev. Camb. Engl.* **140**, 2247–2251 (2013).

22. Kim, H.-S. *et al.* Schwann Cell Precursors from Human Pluripotent Stem Cells as a Potential Therapeutic Target for Myelin Repair. *Stem Cell Rep.* **8**, 1714–1726 (2017).

23. Carrió, M. *et al.* Reprogramming Captures the Genetic and Tumorigenic Properties of Neurofibromatosis Type 1 Plexiform Neurofibromas. *Stem Cell Rep.* **12**, 411–426 (2019).

24. Dong, R. *et al.* Single-Cell Characterization of Malignant Phenotypes and Developmental Trajectories of Adrenal Neuroblastoma. *Cancer Cell* **38**, 716-733.e6 (2020).

25. Hanemaaijer, E. S. *et al.* Single-cell atlas of developing murine adrenal gland reveals relation of Schwann cell precursor signature to neuroblastoma phenotype. *Proc. Natl. Acad. Sci. U. S. A.* **118**, e2022350118 (2021).

26. Mo, J. *et al.* Humanized neurofibroma model from induced pluripotent stem cells delineates tumor pathogenesis and developmental origins. *J. Clin. Invest.* **131**, e139807,-139807 (2021).

27. Tsubota, S. & Kadomatsu, K. Origin and initiation mechanisms of neuroblastoma. *Cell Tissue Res.* **372**, 211–221 (2018).

28. Finzsch, M. *et al.* Sox10 is required for Schwann cell identity and progression beyond the immature Schwann cell stage. *J. Cell Biol.* **189**, 701–712 (2010).

29. Dong, Z. *et al.* Schwann cell development in embryonic mouse nerves. *J. Neurosci. Res.* **56**, 334–348 (1999).

30. Thomson, J. A. *et al.* Embryonic Stem Cell Lines Derived from Human Blastocysts. *Science* **282**, 1145–1147 (1998).

31. Dumevska, B., Bosman, A., McKernan, R., Schmidt, U. & Peura, T. Derivation of human embryonic stem cell line Genea022. *Stem Cell Res.* **16**, 472–475 (2016).

32. Briggs, J. A. *et al.* Integration-Free Induced Pluripotent Stem Cells Model Genetic and Neural Developmental Features of Down Syndrome Etiology. *STEM CELLS* **31**, 467–478 (2013).

33. Ferronha, T. *et al.* LMO4 is an Essential Cofactor in the Snail2-Mediated Epithelial-to-Mesenchymal Transition of Neuroblastoma and Neural Crest Cells. *J. Neurosci.* **33**, 2773–2783 (2013).

34. Takahashi, M. & Osumi, N. Identification of a novel type II classical cadherin: rat cadherin19 is expressed in the cranial ganglia and Schwann cell precursors during development. *Dev. Dyn. Off. Publ. Am. Assoc. Anat.* **232**, 200–208 (2005).

35. Bitgood, M. J. & McMahon, A. P. Hedgehog and Bmp genes are coexpressed at many diverse sites of cell-cell interaction in the mouse embryo. *Dev. Biol.* **172**, 126–138 (1995).

36. Bonnamour, G., Soret, R. & Pilon, N. Dhh-expressing Schwann cell precursors contribute to skin and cochlear melanocytes, but not to vestibular melanocytes. *Pigment Cell Melanoma Res.* **34**, 648–654 (2021).

37. Parmantier, E. *et al.* Schwann cell-derived Desert hedgehog controls the development of peripheral nerve sheaths. *Neuron* **23**, 713–724 (1999).

38. Menendez, L. *et al.* Directed differentiation of human pluripotent cells to neural crest stem cells. *Nat. Protoc.* **8**, 203–212 (2013).

39. Menendez, L., Yatskievych, T. A., Antin, P. B. & Dalton, S. Wnt signaling and a Smad pathway blockade direct the differentiation of human pluripotent stem cells to multipotent neural crest cells. *Proc. Natl. Acad. Sci.* **108**, 19240–19245 (2011).

40. Taneyhill, L. A., Coles, E. G. & Bronner-Fraser, M. Snail2 directly represses cadherin6B during epithelial-to-mesenchymal transitions of the neural crest. *Development* **134**, 1481–1490 (2007).

41. Coles, E. G., Taneyhill, L. A. & Bronner-Fraser, M. A critical role for Cadherin6B in regulating avian neural crest emigration. *Dev. Biol.* **312**, 533–544 (2007).

42. Simões-Costa, M., Tan-Cabugao, J., Antoshechkin, I., Sauka-Spengler, T. & Bronner, M. E. Transcriptome analysis reveals novel players in the cranial neural crest gene regulatory network. *Genome Res.* **24**, 281–290 (2014).

43. Maro, G. S. *et al.* Neural crest boundary cap cells constitute a source of neuronal and glial cells of the PNS. *Nat. Neurosci.* **7**, 930–938 (2004).

44. Fontenas, L. & Kucenas, S. Motor Exit Point (MEP) Glia: Novel Myelinating Glia That Bridge CNS and PNS Myelin. *Front. Cell. Neurosci.* **12**, (2018).

45. Sen, R. *et al.* Kat2a and Kat2b Acetyltransferase Activity Regulates Craniofacial Cartilage and Bone Differentiation in Zebrafish and Mice. *J. Dev. Biol.* **6**, 27 (2018).

46. Shull, L. C. *et al.* The conserved and divergent roles of Prdm3 and Prdm16 in zebrafish and mouse craniofacial development. *Dev. Biol.* **461**, 132–144 (2020).

47. Kim, Y. J. *et al.* Generation of multipotent induced neural crest by direct reprogramming of human postnatal fibroblasts with a single transcription factor. *Cell Stem Cell* **15**, 497–506 (2014).

48. Zhang, J. & Jiao, J. Molecular Biomarkers for Embryonic and Adult Neural Stem Cell and Neurogenesis. *BioMed Res. Int.* **2015**, 727542 (2015).

49. Graham, V., Khudyakov, J., Ellis, P. & Pevny, L. SOX2 functions to maintain neural progenitor identity. *Neuron* **39**, 749–765 (2003).

50. Creuzet, S. E. *et al.* LKB1 signaling in cephalic neural crest cells is essential for vertebrate head development. *Dev. Biol.* **418**, 283–296 (2016).

51. Sun, Z., Zhang, Y., Zhang, R., Qi, X. & Su, B. Functional divergence of the rapidly evolving miR-513 subfamily in primates. *BMC Evol. Biol.* **13**, 255 (2013).

52. LI, H., LIU, H., PEI, J., WANG, H. & LV, H. miR-542-3p overexpression is associated with enhanced osteosarcoma cell proliferation and migration ability by targeting Van Gogh-like 2. *Mol. Med. Rep.* **11**, 851–856 (2015).

53. Willems, B. *et al.* The Wnt Co-Receptor Lrp5 Is Required for Cranial Neural Crest Cell Migration in Zebrafish. *PloS One* **10**, e0131768 (2015).

54. Kim, Y. J. *et al.* miR-486-5p induces replicative senescence of human adipose tissue-derived mesenchymal stem cells and its expression is controlled by high glucose. *Stem Cells Dev.* **21**, 1749–1760 (2012).

55. Peng, Y., Wu, W., Shang, Z., Li, W. & Chen, S. Inhibition of lncRNA LINC00461/miR-216a/aquaporin 4 pathway suppresses cell proliferation, migration, invasion, and chemoresistance in glioma. *Open Life Sci.* **15**, 532–543 (2020).

56. Sun, Y., Zhu, D., Xing, H., Hou, Y. & Liu, Y. Screening of characteristic biomolecules related to bladder cancer based on construction of ceRNA regulation network. *World J. Urol.* **38**, 2835–2847 (2020).

57. Hashimshony, T., Wagner, F., Sher, N. & Yanai, I. CEL-Seq: Single-Cell RNA-Seq by Multiplexed Linear Amplification. *Cell Rep.* **2**, 666–673 (2012).

58. Picelli, S. *et al.* Full-length RNA-seq from single cells using Smart-seq2. *Nat. Protoc.* **9**, 171–181 (2014).

59. Kuleshov, M. V. *et al.* Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* **44**, W90–W97 (2016).

60. Inoue, T. *et al.* Analysis of mouse Cdh6 gene regulation by transgenesis of modified bacterial artificial chromosomes. *Dev. Biol.* **315**, 506–520 (2008).

61. Inoue, T., Chisaka, O., Matsunami, H. & Takeichi, M. Cadherin-6 Expression Transiently Delineates Specific Rhombomeres, Other Neural Tube Subdivisions, and Neural Crest Subpopulations in Mouse Embryos. *Dev. Biol.* **183**, 183–194 (1997).

62. Riethmacher, D. *et al.* Severe neuropathies in mice with targeted mutations in the ErbB3 receptor. *Nature* **389**, 725–730 (1997).

63. Britsch, S. *et al.* The transcription factor Sox10 is a key regulator of peripheral glial development. *Genes Dev.* **15**, 66–78 (2001).

64. Monje, P. V., Athauda, G. & Wood, P. M. Protein Kinase A-mediated Gating of Neuregulin-dependent ErbB2-ErbB3 Activation Underlies the Synergistic Action of cAMP on Schwann Cell Proliferation*. *J. Biol. Chem.* **283**, 34087–34100 (2008).

65. Torii, T. *et al.* In vivo knockdown of ErbB3 in mice inhibits Schwann cell precursor migration. *Biochem. Biophys. Res. Commun.* **452**, 782–788 (2014).

66. Parkinson, D. B., Langner, K., Namini, S. S., Jessen, K. R. & Mirsky, R. β-Neuregulin and Autocrine Mediated Survival of Schwann Cells Requires Activity of Ets Family Transcription Factors. *Mol. Cell. Neurosci.* **20**, 154–167 (2002).

67. Wu, P.-Y. *et al.* Novel Endogenous Ligands of Aryl Hydrocarbon Receptor Mediate Neural Development and Differentiation of Neuroblastoma. *ACS Chem. Neurosci.* **10**, 4031–4042 (2019).

68. Frericks, M., Meissner, M. & Esser, C. Microarray analysis of the AHR system: Tissue-specific flexibility in signal and target genes. *Toxicol. Appl. Pharmacol.* **220**, 320–332 (2007).

69. Merlo, G. R. *et al.* The role of Dlx homeogenes in early development of the olfactory pathway. *J. Mol. Histol.* **38**, 347–358 (2007).

70. Leimeister, C., Bach, A. & Gessler, M. Developmental expression patterns of mouse sFRP genes encoding members of the secreted frizzled related protein family. *Mech. Dev.* **75**, 29–42 (1998).

71. Xu, Q., D'Amore, P. A. & Sokol, S. Y. Functional and biochemical interactions of Wnts with FrzA, a secreted Wnt antagonist. *Dev. Camb. Engl.* **125**, 4767–4776 (1998).

72. Esteve, P., Morcillo, J. & Bovolenta, P. Early and dynamic expression of cSfrp1 during chick embryo development. *Mech. Dev.* **97**, 217–221 (2000).

73. Terry, K., Magan, H., Baranski, M. & Burrus, L. W. Sfrp-1 and sfrp-2 are expressed in overlapping and distinct domains during chick development. *Mech. Dev.* **97**, 177–182 (2000).

74. Wu, H.-H. *et al.* Glial precursors clear sensory neuron corpses during development via Jedi-1, an engulfment receptor. *Nat. Neurosci.* **12**, 1534–1541 (2009).

75. Stewart, H. J. S. *et al.* Developmental regulation and overexpression of the transcription factor AP-2, a potential regulator of the timing of Schwann cell generation. *Eur. J. Neurosci.* **14**, 363–372 (2001).

76. Sonnenberg-Riethmacher, E., Miehe, M. & Riethmacher, D. Promotion of periostin expression contributes to the migration of Schwann cells. *J. Cell Sci.* **128**, 3345–3355 (2015).

77. Stewart, H. j. s *et al.* Helix-loop-helix proteins in Schwann cells: A study of regulation and subcellular localization of Ids, REB, and E12/47 during embryonic and postnatal development. *J. Neurosci. Res.* **50**, 684–701 (1997).

78. Kenny, D. A., Jurata, L. W., Saga, Y. & Gill, G. N. Identification and characterization of LMO4, an LMO gene with a novel pattern of expression during embryogenesis. *Proc. Natl. Acad. Sci.* **95**, 11257–11262 (1998).

79. Lee, S.-K. *et al.* The LIM domain-only protein LMO4 is required for neural tube closure. *Mol. Cell. Neurosci.* **28**, 205–214 (2005).

80. Paratcha, G., Ledda, F. & Ibáñez, C. F. The Neural Cell Adhesion Molecule NCAM Is an Alternative Signaling Receptor for GDNF Family Ligands. *Cell* **113**, 867–879 (2003).

81. Kishimoto, I. *et al.* C-type natriuretic peptide is a Schwann cell-derived factor for development and function of sensory neurones. *J. Neuroendocrinol.* **20**, 1213–1223 (2008).

82. Quintes, S. *et al.* Zeb2 is essential for Schwann cell differentiation, myelination and nerve repair. *Nat. Neurosci.* **19**, 1050–1059 (2016).

83. Pohl, B. S. & Knöchel, W. Temporal and spatial expression patterns of FoxD2 during the early development of Xenopus laevis. *Mech. Dev.* **111**, 181–184 (2002).

84. De Winter, F. *et al.* The expression of the chemorepellent Semaphorin 3A is selectively induced in terminal Schwann cells of a subset of neuromuscular synapses that display limited anatomical plasticity and enhanced vulnerability in motor neuron disease. *Mol. Cell. Neurosci.* **32**, 102–117 (2006).

85. Young, K. G., Thurston, S. F., Copeland, S., Smallwood, C. & Copeland, J. W. INF1 Is a Novel Microtubule-associated Formin. *Mol. Biol. Cell* **19**, 5168–5180 (2008).

86. Thurston, S. F., Kulacz, W. A., Shaikh, S., Lee, J. M. & Copeland, J. W. The ability to induce microtubule acetylation is a general feature of formin proteins. *PloS One* **7**, e48041 (2012).

87. Ahn, B. Y. *et al.* Glioma invasion mediated by the p75 neurotrophin receptor (p75NTR/CD271) requires regulated interaction with PDLIM1. *Oncogene* **35**, 1411–1422 (2016).

88. Liu, Z. *et al.* PDZ and LIM domain protein 1(PDLIM1)/CLP36 promotes breast cancer cell migration, invasion and metastasis through interaction with α-actinin. *Oncogene* **34**, 1300–1311 (2015).

89. Colombo, S., Champeval, D., Rambow, F. & Larue, L. Transcriptomic Analysis of Mouse Embryonic Skin Cells Reveals Previously Unreported Genes Expressed in Melanoblasts. *J. Invest. Dermatol.* **132**, 170–178 (2012).

90. Wolf, D. *et al.* Ankyrin repeat-containing N-Ank proteins shape cellular membranes. *Nat. Cell Biol.* **21**, 1191–1205 (2019).

91. Kumar, A. *et al.* Specification and Diversification of Pericytes and Smooth Muscle Cells from Mesenchymoangioblasts. *Cell Rep.* **19**, 1902–1916 (2017).

92. Messemaker, T. C. *et al.* Allele-specific repression of Sox2 through the long non-coding RNA Sox2ot. *Sci. Rep.* **8**, 386 (2018).

93. Martinsen, B. J. & Bronner-Fraser, M. Neural Crest Specification Regulated by the Helix-Loop-Helix Repressor Id2. *Science* **281**, 988–991 (1998).

94. Streit, A. The preplacodal region: an ectodermal domain with multipotential progenitors that contribute to sense organs and cranial sensory ganglia. *Int. J. Dev. Biol.* **51**, 447–461 (2007).

95. Osborne, N. J., Begbie, J., Chilton, J. K., Schmidt, H. & Eickholt, B. J. Semaphorin/neuropilin signaling influences the positioning of migratory neural crest cells within the hindbrain region of the chick. *Dev. Dyn. Off. Publ. Am. Assoc. Anat.* **232**, 939–949 (2005).

96. Koestner, U., Shnitsar, I., Linnemannstöns, K., Hufton, A. L. & Borchers, A. Semaphorin and neuropilin expression during early morphogenesis of Xenopus laevis. *Dev. Dyn.* **237**, 3853–3863 (2008).

97. Lee, H.-S. & Tomarev, S. I. Optimedin induces expression of N-cadherin and stimulates aggregation of NGF-stimulated PC12 cells. *Exp. Cell Res.* **313**, 98–108 (2007).

98. Schwarz, Q. *et al.* Plexin A3 and plexin A4 convey semaphorin signals during facial nerve development. *Dev. Biol.* **324**, 1–9 (2008).

99. Cheng, H.-J. *et al.* Plexin-A3 Mediates Semaphorin Signaling and Regulates the Development of Hippocampal Axonal Projections. *Neuron* **32**, 249–263 (2001).

100. Suto, F. *et al.* Plexin-A4 Mediates Axon-Repulsive Activities of Both Secreted and Transmembrane Semaphorins and Plays Roles in Nerve Fiber Guidance. *J. Neurosci.* **25**, 3628–3637 (2005).

101. Yaron, A., Huang, P.-H., Cheng, H.-J. & Tessier-Lavigne, M. Differential Requirement for Plexin-A3 and -A4 in Mediating Responses of Sensory and Sympathetic Neurons to Distinct Class 3 Semaphorins. *Neuron* **45**, 513–523 (2005).

102. Majd, H. *et al.* Deriving Schwann cells from hPSCs enables disease modeling and drug discovery for diabetic peripheral neuropathy. *Cell Stem Cell* **30**, 632-647.e10 (2023).

103. Jansky, S. *et al.* Single-cell transcriptomic analyses provide insights into the developmental origins of neuroblastoma. *Nat. Genet.* **53**, 683–693 (2021).

104. Cibi, D. M. *et al.* Neural crest-specific deletion of Rbfox2 in mice leads to craniofacial abnormalities including cleft palate. *eLife* **8**, e45418 (2019).

105. Lee, S. *et al.* Cleft lip and cleft palate in Esrp1 knockout mice is associated with alterations in epithelial-mesenchymal crosstalk. *Development* **147**, dev187369 (2020).

106. Gerstberger, S., Hafner, M. & Tuschl, T. A census of human RNA-binding proteins. *Nat. Rev. Genet.* **15**, 829–845 (2014).

107. Bebee, T. W. *et al.* The splicing regulators Esrp1 and Esrp2 direct an epithelial splicing program essential for mammalian development. *eLife* **4**, e08954 (2015).

108. Cvitkovic, I. & Jurica, M. S. Spliceosome Database: a tool for tracking components of the spliceosome. *Nucleic Acids Res.* **41**, D132–D141 (2013).

109. Tiwari, K. & Nielsen, L. K. ASTA-P: a pipeline for the detection, quantification and statistical analysis of complex alternative splicing events. 2023.08.28.555224 Preprint at https://doi.org/10.1101/2023.08.28.555224 (2023).

110. Nowicka, M. & Robinson, M. D. DRIMSeq: a Dirichlet-multinomial framework for multivariate count outcomes in genomics. *F1000Research* **5**, 1356 (2016).

111. Bronicki, L. M. & Jasmin, B. J. Emerging complexity of the HuD/ELAVl4 gene; implications for neuronal development, function, and dysfunction. *RNA* **19**, 1019–1037 (2013).

112. Fragkouli, A. *et al.* Neuronal ELAVL proteins utilize AUF-1 as a co-partner to induce neuron-specific alternative splicing of APP. *Sci. Rep.* **7**, 44507 (2017).

113. Zhu, H., Hasman, R. A., Barron, V. A., Luo, G. & Lou, H. A Nuclear Function of Hu Proteins as Neuron-specific Alternative RNA Processing Regulators. *Mol. Biol. Cell* **17**, 5105–5114 (2006).

114. Ince-Dunn, G. *et al.* Neuronal Elav-like (Hu) Proteins Regulate RNA Splicing and Abundance to Control Glutamate Levels and Neuronal Excitability. *Neuron* **75**, 1067–1080 (2012).

115. Dasgupta, T. & Ladd, A. N. The importance of CELF control: molecular and biological roles of the CUG-BP, Elav-like family of RNA-binding proteins. *WIREs RNA* **3**, 104–121 (2012).

116. Blech-Hermoni, Y., Stillwagon, S. j. & Ladd, A. n. Diversity and conservation of CELF1 and CELF2 RNA and protein expression patterns during embryonic development. *Dev. Dyn.* **242**, 767–777 (2013).

117. Mallory, M. J. *et al.* Induced transcription and stability of CELF2 mRNA drives widespread alternative splicing during T-cell signaling. *Proc. Natl. Acad. Sci.* **112**, E2139–E2148 (2015).

118. Ajith, S. *et al.* Position-dependent activity of CELF2 in the regulation of splicing and implications for signal-responsive regulation in T cells. *RNA Biol.* **13**, 569–581 (2016).

119. Hooper, J. E., Jones, K. L., Smith, F. J., Williams, T. & Li, H. An Alternative Splicing Program for Mouse Craniofacial Development. *Front. Physiol.* **11**, 1099 (2020).

120. Dominguez, D. *et al.* Sequence, Structure, and Context Preferences of Human RNA Binding Proteins. *Elsevier* (2018).

121. Siepel, A. *et al.* Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res.* **15**, 1034–1050 (2005).

122. Zago, G., Biondini, M., Camonis, J. & Parrini, M. C. A family affair: A Ral-exocyst-centered network links Ras, Rac, Rho signaling to control cell migration. *Small GTPases* **10**, 323–330 (2019).

123. Sakurai-Yageta, M. *et al.* The interaction of IQGAP1 with the exocyst complex is required for tumor cell invasion downstream of Cdc42 and RhoA. *J. Cell Biol.* **181**, 985–998 (2008).

124. Qian, X. *et al.* Sec3 knockdown inhibits TGF-β induced epithelial-mesenchymal transition through the down-regulation of Akt phosphorylation in A549 cells. *Biochem. Biophys. Res. Commun.* **519**, 253–260 (2019).

125. Rosse, C. *et al.* An aPKC-Exocyst Complex Controls Paxillin Phosphorylation and Migration through Localised JNK1 Activation. *PLOS Biol.* **7**, e1000235 (2009).

126. Lawson, C. D. & Ridley, A. J. Rho GTPase signaling complexes in cell migration and invasion. *J. Cell Biol.* **217**, 447–457 (2017).

127. Machacek, M. *et al.* Coordination of Rho GTPase activities during cell protrusion. *Nature* **461**, 99–103 (2009).

128. Aho, S. *et al.* Specific sequences in p120ctn determine subcellular distribution of its multiple isoforms involved in cellular adhesion of normal and malignant epithelial cells. *J. Cell Sci.* **115**, 1391–1402 (2002).

129. Pieters, T., Roy, F. van & Hengel, J. van. Functions of p120ctn isoforms in cell-cell adhesion and intracellular signaling. *Front. Biosci.-Landmark* **17**, 1669–1694 (2012).

130. Wieczorek, K. *et al.* Filamin A upregulation correlates with Snail-induced epithelial to mesenchymal transition (EMT) and cell adhesion but its inhibition increases the migration of colon adenocarcinoma HT29 cells. *Exp. Cell Res.* **359**, 163–170 (2017).

131. Chen, X. *et al.* Supervillin promotes epithelial-mesenchymal transition and metastasis of hepatocellular carcinoma in hypoxia via activation of the RhoA/ROCK-ERK/p38 pathway. *J. Exp. Clin. Cancer Res.* **37**, 128 (2018).

132. Kinslechner, K. *et al.* Malignant Phenotypes in Metastatic Melanoma are Governed by SR-BI and its Association with Glycosylation and STAT5 Activation. *Mol. Cancer Res.* **16**, 135–146 (2018).

133. Gu, M. *et al.* Downregulation of RAI14 inhibits the proliferation and invasion of breast cancer cells. *J. Cancer* **10**, 6341–6348 (2019).

134. Bhattacharya, D., Azambuja, A. P. & Simoes-Costa, M. Metabolic Reprogramming Promotes Neural Crest Migration via Yap/Tead Signaling. *Dev. Cell* **53**, 199-211.e6 (2020).

135. Clay, M. R. & Halloran, M. C. Cadherin 6 promotes neural crest cell detachment via F-actin regulation and influences active Rho distribution during epithelial-to-mesenchymal transition. *Dev. Camb. Engl.* **141**, 2506–2515 (2014).

136. Knecht, A. K. & Bronner-Fraser, M. DBHR, a gene with homology to dopamine beta-hydroxylase, is expressed in the neural crest throughout early development. *Dev. Biol.* **234**, 365–375 (2001).

137. Elms, P., Siggers, P., Napper, D., Greenfield, A. & Arkell, R. Zic2 is required for neural crest formation and hindbrain patterning during mouse development. *Dev. Biol.* **264**, 391–406 (2003).

138. Grego-Bessa, J., Hildebrand, J. & Anderson, K. V. Morphogenesis of the mouse neural plate depends on distinct roles of cofilin 1 in apical and basal epithelial domains. *Development* dev.115493 (2015) doi:10.1242/dev.115493.

139. Miura, K. *et al.* ARAP1: a point of convergence for Arf and Rho signaling. *Mol. Cell* **9**, 109–119 (2002).

140. Sadok, A. & Marshall, C. J. Rho GTPases: masters of cell migration. *Small GTPases* **5**, e29710 (2014).

141. Scarpa, E. *et al.* Cadherin Switch during EMT in Neural Crest Cells Leads to Contact Inhibition of Locomotion via Repolarization of Forces. *Dev. Cell* **34**, 421–434 (2015).

142. Xu, X. *et al.* Modulation of mouse neural crest cell motility by N-cadherin and connexin 43 gap junctions. *J. Cell Biol.* **154**, 217–230 (2001).

143. Borchers, A., David, R. & Wedlich, D. Xenopus cadherin-11 restrains cranial neural crest migration and influences neural crest specification. *Dev. Camb. Engl.* **128**, 3049–3060 (2001).

144. Solovieva, T. & Bronner, M. Schwann cell precursors: Where they come from and where they go. *Cells Dev.* **166**, 203686 (2021).

145. O'Rahilly, R. & Müller, F. Developmental stages in human embryos: revised and new measurements. *Cells Tissues Organs* **192**, 73–84 (2010).

146. Manohar, S., Camacho-Magallanes, A., Echeverria, C. & Rogers, C. D. Cadherin-11 Is Required for Neural Crest Specification and Survival. *Front. Physiol.* **11**, 563372 (2020).

147. Koehler, A. *et al.* Loss of Xenopus cadherin-11 leads to increased Wnt/β-catenin signaling and up-regulation of target genes c-myc and cyclin D1 in neural crest. *Dev. Biol.* **383**, 132–145 (2013).

148. McKeown, S. J., Wallace, A. S. & Anderson, R. B. Expression and function of cell adhesion molecules during neural crest migration. *Dev. Biol.* **373**, 244–257 (2013).

149. Ma, L. *et al.* miR-9, a MYC/MYCN-activated microRNA, regulates E-cadherin and cancer metastasis. *Nat. Cell Biol.* **12**, 247–256 (2010).

150. Streicher, K. L. *et al.* A novel oncogenic role for the miRNA-506-514 cluster in initiating melanocyte transformation and promoting melanoma growth. *Oncogene* **31**, 1558–1570 (2012).

151. Ji, D., Wang, Y., Li, H., Sun, B. & Luo, X. Long non-coding RNA LINC00461/miR-149-5p/LRIG2 axis regulates hepatocellular carcinoma progression. *Biochem. Biophys. Res. Commun.* **512**, 176–181 (2019).

152. Dong, L., Qian, J., Chen, F., Fan, Y. & Long, J. LINC00461 promotes cell migration and invasion in breast cancer through miR-30a-5p/integrin β3 axis. *J. Cell. Biochem.* **120**, 4851–4862 (2019).

153. Yang, Y. *et al.* LINC00461, a long non-coding RNA, is important for the proliferation and migration of glioma cells. *Oncotarget* **8**, 84123–84139 (2017).

154. Wu, J. *et al.* Knockdown of LINC01116 inhibits cell migration and invasion in head and neck squamous cell carcinoma through epithelial-mesenchymal transition pathway. *J. Cell. Biochem.* **121**, 867–875 (2020).

155. Douka, K. *et al.* Cytoplasmic long noncoding RNAs are differentially regulated and translated during human neuronal differentiation. *RNA* **27**, 1082–1101 (2021).

156. Alais, S. *et al.* HEMCAM/CD146 downregulates cell surface expression of β1 integrins. *J. Cell Sci.* **114**, 1847–1859 (2001).

157. Motohashi, T. *et al.* Gene array analysis of neural crest cells identifies transcription factors necessary for direct conversion of embryonic fibroblasts into neural crest cells. *Biol. Open* **5**, 311–322 (2016).

158. Ishii, M. *et al.* A Stable Cranial Neural Crest Cell Line from Mouse. *Stem Cells Dev.* **21**, 3069–3080 (2012).

159. Lumb, R. & Schwarz, Q. Sympathoadrenal neural crest cells: The known, unknown and forgotten? *Dev. Growth Differ.* **57**, 146–157 (2015).

160. Wu, S. C.-Y., Grindley, J., Winnier, G. E., Hargett, L. & Hogan, B. L. M. Mouse Mesenchyme forkhead 2 (Mf2): expression, DNA binding and induction by sonic hedgehog during somitogenesis. *Mech. Dev.* **70**, 3–13 (1998).

161. Mukherjee-Clavin, B. *et al.* Comparison of three congruent patient-specific cell types for the modelling of a human genetic Schwann-cell disorder. *Nat. Biomed. Eng.* **3**, 571–582 (2019).

162. Kerosuo, L. *et al.* Enhanced expression of MycN/CIP2A drives neural crest toward a neural stem cell-like fate: Implications for priming of neuroblastoma. *Proc. Natl. Acad. Sci.* **115**, E7351–E7360 (2018).

163. Ewels, P., Magnusson, M., Lundin, S. & Käller, M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinforma. Oxf. Engl.* **32**, 3047–3048 (2016).

164. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinforma. Oxf. Engl.* **30**, 923–930 (2014).

165. Robinson, M. D. & Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* **11**, R25 (2010).

166. Law, C. W., Chen, Y., Shi, W. & Smyth, G. K. voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* **15**, R29 (2014).

167. Smyth, G. K. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.* **3**, Article3 (2004).

168. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B Methodol.* **57**, 289–300 (1995).

169. Tian, L. *et al.* scPipe: A flexible R/Bioconductor preprocessing pipeline for single-cell RNA-sequencing data. *PLOS Comput. Biol.* **14**, e1006361 (2018).

170. McCarthy, D. J., Campbell, K. R., Lun, A. T. L. & Wills, Q. F. Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. *Bioinforma. Oxf. Engl.* **33**, 1179–1186 (2017).

171. McCarthy, D. J., Chen, Y. & Smyth, G. K. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.* **40**, 4288–4297 (2012).

172. Maaten, L. van der & Hinton, G. Visualizing Data using t-SNE. *J. Mach. Learn. Res.* **9**, 2579–2605 (2008).

173. Dave, K. A. *et al.* A Comprehensive Proteomic View of Responses of A549 Type II Alveolar Epithelial Cells to Human Respiratory Syncytial Virus Infection. *Mol. Cell. Proteomics MCP* **13**, 3250–3269 (2014).

174. Hastie, M. L. *et al.* The human respiratory syncytial virus nonstructural protein 1 regulates type I and type II interferon pathways. *Mol. Cell. Proteomics MCP* **11**, 108–127 (2012).

175. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372 (2008).

176. Tyanova, S. *et al.* The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* **13**, 731–740 (2016).

177. O'Callaghan, S. *et al.* PyMS: a Python toolkit for processing of gas chromatography-mass spectrometry (GC-MS) data. Application and comparative study of selected tools. *BMC Bioinformatics* **13**, 115 (2012).

178. Mu, Y. *et al.* Simultaneous determination of nine phenolic compounds in imitation wild Dendrobium officinale samples using ultrahigh-performance liquid chromatography–tandem mass spectrometry. *Front. Nutr.* **10**, (2023).

179. Edgar, R., Domrachev, M. & Lash, A. E. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* **30**, 207–210 (2002).

180. Kuhlbrodt, K., Herbarth, B., Sock, E., Hermans-Borgmeyer, I. & Wegner, M. Sox10, a novel transcriptional modulator in glial cells. *J. Neurosci. Off. J. Soc. Neurosci.* **18**, 237–250 (1998).

181. Pingault, V. *et al.* SOX10 mutations in patients with Waardenburg-Hirschsprung disease. *Nat. Genet.* **18**, 171–173 (1998).

182. McEvoy, A. L. *et al.* mMaple: A Photoconvertible Fluorescent Protein for Use in Multiple Imaging Modalities. *PLOS ONE* **7**, e51314 (2012).

183. Horikiri, T. *et al.* SOX10-Nano-Lantern Reporter Human iPS Cells; A Versatile Tool for Neural Crest Research. *PLOS ONE* **12**, e0170342 (2017).

184. Van den Berge, K. *et al.* Observation weights unlock bulk RNA-seq tools for zero inflation and single-cell applications. *Genome Biol.* **19**, 24 (2018).

185. Ge, S. X., Jung, D. & Yao, R. ShinyGO: a graphical gene-set enrichment tool for animals and plants. *Bioinformatics* **36**, 2628–2629 (2020).

186. Rodriguez, J. M. *et al.* APPRIS 2017: principal isoforms for multiple gene sets. *Nucleic Acids Res.* **46**, D213–D217 (2018).

187. Finn, R. D. *et al.* Pfam: the protein families database. *Nucleic Acids Res.* **42**, D222–D230 (2014).

188. The UniProt Consortium. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res.* **47**, D506–D515 (2019).

189. Karolchik, D. *et al.* The UCSC Table Browser data retrieval tool. *Nucleic Acids Res.* **32**, D493–D496 (2004).

190. Hug, N., Longman, D. & Cáceres, J. F. Mechanism and regulation of the nonsense-mediated decay pathway. *Nucleic Acids Res.* **44**, 1483–1495 (2016).

191. Puigdevall, P. & Castelo, R. GenomicScores: seamless access to genomewide position-specific scores from R and Bioconductor. *Bioinformatics* **34**, 3208–3210 (2018).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- TableS1.xlsx
- TableS2.xlsx
- supplementarydatafile1.xlsx
- supplementarydatafile2.xlsx
- supplementarydatafile3.xlsx
- supplementarydatafile4.csv
- supplementarydatafile5.xlsx
- supplementarydatafile6.xlsx
- supplementarydatafile7.xlsx
- supplementarydatafile8.xlsx
- supplementarydatafile9.xlsx
- supplementarydatafile10.xlsx
- supplementarydatafile11.xlsx
- supplementarydatafile12.xlsx
- supplementarydatafile13.xlsx
- supplementarydatafile14.xlsx
- supplementarydatafile15384CelSeqprimersreference.xlsx