

Characterization of Human iPSC-derived Spinal Motor Neurons by Single-cell RNA Sequencing

Louise Thiry,^{a,c} Regan Hamel,^b Stefano Pluchino,^b Thomas Durcan^{a,c} and Stefano Stifani^{a*}

^a Department of Neurology and Neurosurgery, Montreal Neurological Institute-Hospital, McGill University, 3801, rue University, Montreal (Quebec) H3A 2B4, Canada

^b Department of Clinical Neurosciences, Clifford Allbutt Building – Cambridge Biosciences Campus and NIHR Biomedical Research Centre, University of Cambridge, Hills Road, CB2 0HA Cambridge, UK

^c Early Drug Discovery Unit, Montreal Neurological Institute-Hospital, Canada

Abstract—Human induced pluripotent stem cells (iPSCs) offer the opportunity to generate specific cell types from healthy and diseased individuals, allowing the study of mechanisms of early human development, modelling a variety of human diseases, and facilitating the development of new therapeutics. Human iPSC-based applications are often limited by the variability among iPSC lines originating from a single donor, as well as the heterogeneity among specific cell types that can be derived from iPSCs. The ability to deeply phenotype different iPSC-derived cell types is therefore of primary importance to the successful and informative application of this technology. Here we describe a combination of motor neuron (MN) derivation and single-cell RNA sequencing approaches to generate and characterize specific MN subtypes obtained from human iPSCs. Our studies provide evidence for rapid and robust generation of MN progenitor cells that can give rise to a heterogeneous population of MNs. Approximately 58% of human iPSC-derived MNs display molecular characteristics of lateral motor column MNs, with a number of molecularly distinct subpopulations present within this MN group. Roughly 19% of induced MNs resemble hypaxial motor column MNs, while ~6% of induced MNs have features of median motor column MNs. The present study has the potential to improve our understanding of iPSC-derived MN subtype function and dysfunction, possibly leading to improved iPSC-based applications for the study of human MN biology and disease.

This article is part of a Special Issue entitled: Jessell Memorial. © 2020 The Author(s). Published by Elsevier Ltd on behalf of IBRO. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Key words: human iPSCs, spinal motor neurons, single cell RNA sequencing.

INTRODUCTION

Human induced pluripotent stem cells (h-iPSCs) have emerged as a powerful model system to investigate the mechanisms of cell differentiation, function and dysfunction (reviewed in [Haston and Finkbeiner, 2016](#); [Ardhanareeswaran et al., 2017](#); [Amin et al., 2019](#)). Moreover, they offer potential as a physiologically-relevant experimental system for disease modelling and discovery of new therapeutics. A key requirement for fulfilling the full potential of h-iPSCs is the ability to reliably generate specific cell types with defined phenotypic traits. It is often the case that functionally specialized subtypes of cells hold the most promise for fundamental biological studies,

disease modelling or drug discover efforts. Thus, it is of the utmost importance to establish protocols for the differentiation and deep phenotyping of h-iPSC-derived cells that may represent only a small fraction of cells *in vivo*.

Spinal motor neurons (MNs) are a highly specialized group of neurons that reside in the spinal cord and project axons to muscle fibers to control their contraction. The precise coordination of muscle contractions is essential to generate complex motor behaviors such as walking, grasping or breathing. To ensure such refined coordination, spinal MNs must acquire specific identities matching the muscle types they will innervate. During MN development, all spinal MNs emerge from progenitor cells in the MN progenitor (pMN) domain. The pMN domain is located in the medial portion of the ventral neural tube, just ventral to the progenitor domain 2 (p2) and dorsal to the p3 domain, which give rise to V2 and V3 interneurons (INs), respectively ([Alaynick et al., 2011](#); [Stifani, 2014](#)). After the formation of distinct progenitor domains, pMN progenitor cells initially acquire a general MN fate in

*Corresponding author.

E-mail address: stefano.stifani@mcgill.ca (S. Stifani).

Abbreviations: ALS, amyotrophic lateral sclerosis; HMC, hypaxial motor column; iPSCs, induced pluripotent stem cells; LMC, lateral motor column; MMC, median motor column; MN, motor neuron; MNPCs, motor neuron progenitor cells; NPCs, neural progenitor cells; PMC, phrenic motor column; SMA, spinal muscular atrophy.

<https://doi.org/10.1016/j.neuroscience.2020.04.041>

0306-4522/© 2020 The Author(s). Published by Elsevier Ltd on behalf of IBRO.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

response to the expression of specific transcription factors including PAIRED BOX 6 (PAX6) and OLIGODENDROCYTE TRANSCRIPTION FACTOR 2 (OLIG2) (Ericson et al., 1997; Novitsch et al., 2001; Vallstedt et al., 2001). As the spinal cord continues to develop, inductive signals along the rostro-caudal axis further specify developing MNs, leading to the formation of anatomically defined motor columns, each containing phenotypically distinct MNs with specific axonal projections to different muscle targets (Prasad and Hollyday, 1991; Tsuchida et al., 1994; Jessell, 2000; Alaynick et al., 2011; Francius and Clotman, 2014; Stifani, 2014).

Loss of MN function is associated with several devastating neurological diseases including spinal muscular atrophy (SMA) and amyotrophic lateral sclerosis (ALS) (Kanning et al., 2010; Nijssen et al., 2017), and it is also a pathological feature of a number of spinal cord injury cases (Alizadeh et al., 2019). Patient-derived iPSCs have been generated from ALS and SMA patients (e.g., Ebert et al., 2009; Egawa et al., 2012; Chen et al., 2014; Kiskinis et al., 2014) and a number of differentiation protocols have been developed to obtain patient-derived MNs for disease modelling (e.g., Li et al., 2005; Qu et al., 2014; Amoroso et al., 2013; Du et al., 2015). Although some of these protocols facilitate the generation of highly pure MN populations (Du et al., 2015), the molecular characterization of these cells is often limited to the use of pan-MN markers through immunostaining approaches (Amoroso et al., 2013; Chen et al., 2014; Maury et al., 2015; Du et al., 2015). Given the heterogeneity of spinal MNs, our limited understanding of the biological properties of different h-iPSC-derived MNs, and the possibility that not all types of h-MN subtypes may provide disease-relevant experimental models, there is a critical need to improve our understanding of h-iPSC-derived MN biology and function. Here we describe the results of studies applying single-cell RNA sequencing (sc-RNAseq) to characterize the different cell subtypes obtained after subjecting h-iPSCs to a MN differentiation protocol and we describe the implications of these findings for MN disease modeling and drug discovery efforts.

EXPERIMENTAL PROCEDURES

Generation of motor neurons from human iPSCs

Human iPSC line NCRM-1 (male) was obtained from the National Institutes of Health Stem Cell Resource (Bethesda, MD, USA). Cells at low passage number were cultured in mTeSR medium (STEMCELL Technologies; Vancouver, BC, Canada; Cat. No. 85850) in 6-cm culture dishes (Thermo-Fisher Scientific; Waltham, MA, USA; Cat. No. 130181) coated with Matrigel (Thermo-Fisher Scientific; Cat. No. 08-774-552) until they reached 70–80% confluence. To generate neural progenitor cells (NPCs), iPSCs were dissociated with Gentle Cell Dissociation Reagent (STEMCELL Technologies; Cat. No. 07174) and split 1:5 using T25 flasks (Thermo-Fisher Scientific; Cat. No. 12-556-009) coated with poly-L-ornithine (PLO; Sigma-Aldrich; Oakville, ON, Canada; Cat. No. P3655) and laminin

(Sigma-Aldrich; Cat. No. L2020). The mTeSR medium was replaced with a chemically defined neural medium including DMEM/F12 supplemented with GlutaMAX (1/1; Thermo-Fisher Scientific; Cat. No. 35050-061), Neurobasal medium (1/1; Thermo-Fisher Scientific; Cat. No. 21103-049), N2 (0.5X; Thermo-Fisher Scientific; Cat. No. 17504-044), B27 (0.5X; Thermo-Fisher Scientific; Cat. No. 17502-048), ascorbic acid (100 μ M; Sigma-Aldrich; Cat. No. A5960), and antibiotic–antimycotic (1X; Thermo-Fisher Scientific; Cat. No. 15240-062). This medium was supplemented with 3 μ M CHIR99021 (STEMCELL Technologies; Cat. No. 72054), 2 μ M DMH1 (Sigma-Aldrich; Cat. No. D8946) and 2 μ M SB431542 (Tocris Bioscience; Bristol, UK; Cat. No. 1614). The culture medium was changed every other day for 6 days and the resulting NPCs were then differentiated into motor neuron progenitor cells (MNPCs), and subsequently MNs, as described previously (Du et al., 2015), with minor modifications. Briefly, on day 6 NPCs were dissociated with Gentle Cell Dissociation Reagent and split 1:5 with the same medium described above, supplemented with retinoic acid (RA) (0.1 μ M; Sigma-Aldrich; Cat. No. R2625) and purmorphamine (0.5 μ M; Sigma-Aldrich; Cat. No. SML-0868) in combination with 1 μ M CHIR99021, 2 μ M DMH1 and 2 μ M SB431542 reagents. The culture medium was changed every other day for 6 days and the resulting MNPCs were characterized by immunocytochemistry. MNPCs were expanded for 6 days with the same medium containing 3 μ M CHIR99021, 2 μ M DMH1, 2 μ M SB431542, 0.1 μ M RA, 0.5 μ M purmorphamine and 500 μ M valproic acid (VPA; Sigma-Aldrich; Cat. No. P4543). To generate MNs, MNPCs were dissociated and plated at 50,000 cells per well on Matrigel coated coverslips and cultured with the same neural medium supplemented with 0.5 μ M RA and 0.1 μ M purmorphamine. Culture medium was replaced every other day for 6 days and the resulting MNs were characterized by immunocytochemistry. For sc-RNAseq, MNs were dissociated, plated on Matrigel-coated T25 flasks, and cultured with the same neural medium supplemented with 0.5 μ M RA, 0.1 μ M purmorphamine, 0.1 μ M Compound E (Calbiochem; Cat. No. 565790), insulin-like growth factor 1 (10 ng/mL; R&D Systems; Minneapolis, MN; Cat. No. 291-G1-200), brain-derived neurotrophic factor (10 ng/mL; Thermo-Fisher Scientific; Cat. No. PHC7074) and ciliary neurotrophic factor (10 ng/mL; R&D Systems; Cat. No. 257-NT-050) for 6 days before single cell suspension preparation.

Characterization of human iPSC-derived cells by immunocytochemistry

Induced MNPCs and MNs were analyzed by immunocytochemistry, which was performed as described previously (Methot et al., 2018). The following primary antibodies were used: mouse anti-OLIG2 (1/100; Millipore Corp.; Billerica, MA, USA; Cat. No. MABN50); rabbit anti-PAX6 (1/500; Covance; Emeryville, CA, USA; Cat. No. PRB-278P); rabbit anti-HOMEOBOX C4 (HOXC4) (1/150; kindly provided by Dr. Jeremy Dasen, New York University School of Medicine, New York, NY); goat anti-SRY-BOX 1 (SOX1) (1/500; R&D Systems; Cat. No. AF3369); mouse

anti-NK2 HOMEBOX 2 (NKX2.2) (1/100; Developmental Studies Hybridoma Bank – DSHB; Cat. No. 74.5A5-c); rabbit anti-HOMEBOX PROTEIN CHX10 (CHX10) (1/10,000; kindly provided by the late Dr. Thomas Jessell, Columbia University, New York, NY); mouse anti-HOMEBOX PROTEIN HB9 (HB9) (1/30; DSHB; Cat. No. 81.5C10-c); mouse anti-ISLET1 (ISL1) (1/30; DSHB; Cat. No. 39.4D5-c); rabbit anti-LIMB HOMEBOX CONTAINING 3 (LHX3) (1/100; Abcam; Toronto, ON, Canada; Cat. No. ab14555); goat anti-FORKHEAD BOX PROTEIN 1 (FOXP1) (1/100; R&D Systems; Cat. No. AF4534), and guinea pig anti-SCIP (1/16,000; kindly provided by Dr. Jeremy Dasen). Secondary antibodies against primary reagents raised in various species were conjugated to Alexa Fluor 488, Alexa Fluor 555, or Alexa Fluor 647 (1/1000, Invitrogen; Burlington, ON, Canada). For quantification, images were acquired using an Axio Observer Z1 microscope connected to an AxioCam camera, using ZEN software (Zeiss Canada Ltd., Toronto, ON, Canada). For each culture and each time point (e.g., NPCs, MNPCs, MNs), images of > 500 cells in 3 random fields were taken with a 20X objective and analyzed with Image J. The fraction of cells expressing specific marker proteins was calculated as percent of total cells \pm standard error of the mean.

Preparation of single-cell suspensions from induced motor neuron culture

Single-cells in suspension were prepared from MNs cultured for 28 days (Day 0 defined as start of NPC induction), as follows. MNs were first rinsed with 2 mL of dPBS (calcium and magnesium free phosphate buffered saline) containing 0.04% BSA (Sigma-Aldrich; Cat. No. A7906) and dissociated with 2 mL of dissociation reagent containing papain (50U; Sigma-Aldrich; Cat. No. P4762) and Accutase (Thermo-Fischer Scientific; Cat. No. A11105-01). Cells in dissociation reagent were incubated for 15–20 min at 37°C to ensure dissociation of cell clusters before adding 5 mL of DMEM/F12 containing 0.04% BSA and 10% ROCK inhibitor Y-27632 (1/1000; Tocris; Cat. No. 1254), followed by resuspension by gentle pipetting. Cells in suspension were transferred into a 15 mL falcon tube and centrifuged at 1300 rpm for 3 min at room temperature. The cell pellet was resuspended in 2 mL of dPBS containing 0.04% BSA and ROCK-inhibitor (1/1000). Cells were centrifuged again at 1300 rpm for 3 min at room temperature and the cell pellet was resuspended in 500 μ L of the above-mentioned resuspension buffer. A 30- μ m strainer was used to remove cell debris and clumps and the single-cell suspension was transferred into a 2 mL Eppendorf tube placed on ice. Cells were counted to evaluate cell concentration and viability before sc-RNAseq. The targeted cell concentration was 700–1200 cells/ μ L, as recommended in the 10x Genomics guidelines (<https://www.10xgenomics.com/solutions/single-cell/>).

Single cell RNA sequencing and *in silico* analysis

Cells were sequenced at a single-cell level using the microdroplet-based platform, 10x Genomics Chromium

Single Cell 3' Solution (10x Genomics; Pleasanton, Ca, USA), followed by sequencing on a HiSeq4000 system (Illumina; San Diego, CA, USA) at the McGill and Genome Quebec Innovation Center (<https://cesgq.com/en-services>). The 10x cDNA libraries were sequenced at a depth of 50,000 reads per cell. The raw sc-RNAseq data (FASTQ files) were first processed using the Cell Ranger pipeline (10x Genomics) to demultiplex and align the sequences to the human reference genome, GRCh38. FASTQ files were aligned and empty droplet 10x barcodes were filtered out via Cell Ranger. Post-Cell Ranger, gene/cell matrices from 5900 single cells were imported into R. We performed quality control using the Bioconductor-package, scater (McCarthy et al., 2017) to filter out damaged cells based on higher than expected (> 3 median absolute deviations) proportions of mitochondrial to nuclear genes and lower than expected (< 3 median absolute deviations) genes per cell and transcripts per cell, all of which would suggest that the sequenced cells were dead or damaged and thus leaking nuclear transcripts. The normalization was performed by first clustering the cells using the Bioconductor-package, scran (Lun et al., 2016) before computing cluster-based size factors (scran) to finally normalize the dataset. Next, technical noise in the dataset was removed by first modelling the noise (scran) and then performing principal-component analysis (scran). Later principle components account for very little of the variability in the dataset and are thus assumed to represent the random, technical noise while the endogenous variation of co-regulated genes is assumed to be represented in the earlier principle components. Thus, by removing the later principle components, the dataset can be denoised while preserving the biologically-driven variability of gene expression. All software packages are publicly available at the Bioconductor project (<http://bioconductor.org>). The workflow used for the analysis was adapted from a published workflow developed to accommodate droplet-based systems such as 10X Chromium (Lun et al., 2016). This workflow is available on <https://github.com/regan-hamel/h-iPSCs-MNI>.

RESULTS

Generation of an enriched population of human iPSC-derived motor neurons

Spinal MNPCs were generated from the h-iPSC NCRM-1 line by mimicking *in vitro* sequential steps of MN development *in vivo*, including neural induction, followed by caudalization and ventralization of NPCs (Li et al., 2005; Wichterle et al., 2002). Spinal NPCs were induced by inhibiting BONE MORPHOGENETIC PROTEIN/TRANSFORMING GROWTH FACTOR-BETA signalling, using a small-molecule cocktail composed of the CHIR99021, DMH1 and SB431542 reagents, as previously described (Du et al., 2015). After 6 days from the start of the induction protocol, 85 \pm 1% of the induced cells expressed typical NPC markers, including SOX1, and 70 \pm 3% expressed the cervical spinal cord marker HOXC4, indicative of a rostral spinal cell fate (Burke et al., 1995) (Fig. 1A, B). NPCs were ventralized by addi-

tion of the SONIC HEDGEHOG (SHH) agonist, purmorphamine, at a concentration of 0.5 μ M, in the presence of 0.1 μ M RA, resulting in the generation of MNPCs, with $75 \pm 7\%$ of induced cells being PAX6-positive and $77 \pm 5\%$ expressing the MNPC marker OLIG2 (Briscoe et al., 2000; Jessell, 2000; Alaynick et al., 2011) at day 12 (Fig. 1C, D). As expected, a small proportion of spinal ventral IN progenitors was detected in our cultures, with $1.8 \pm 0.3\%$ of the induced cells expressing the p3 domain marker NKX2.2 (Novitsch et al., 2001; Sugimori et al., 2007) and $2.5 \pm 0.8\%$ expressing the p2 domain marker

CHX10 (Debrulle et al., 2019; Alaynick et al., 2011) (Fig. 1E, F).

Continued exposure of MNPCs to RA and purmorphamine led to the generation of an enriched (67 \pm 8%) population of induced cells expressing the typical pan-MN markers ISL1 and HB9 (Amoroso et al., 2013) (Fig. 2A, B), as early as 25 days after the start of the differentiation protocol. Across a number of different MN derivation experiments, the yield of ISL1⁺/HB9⁺ cells varied from ~40% to as high as ~80% (Fig. 2B), possibly the result of different MN survival rates from one culture to

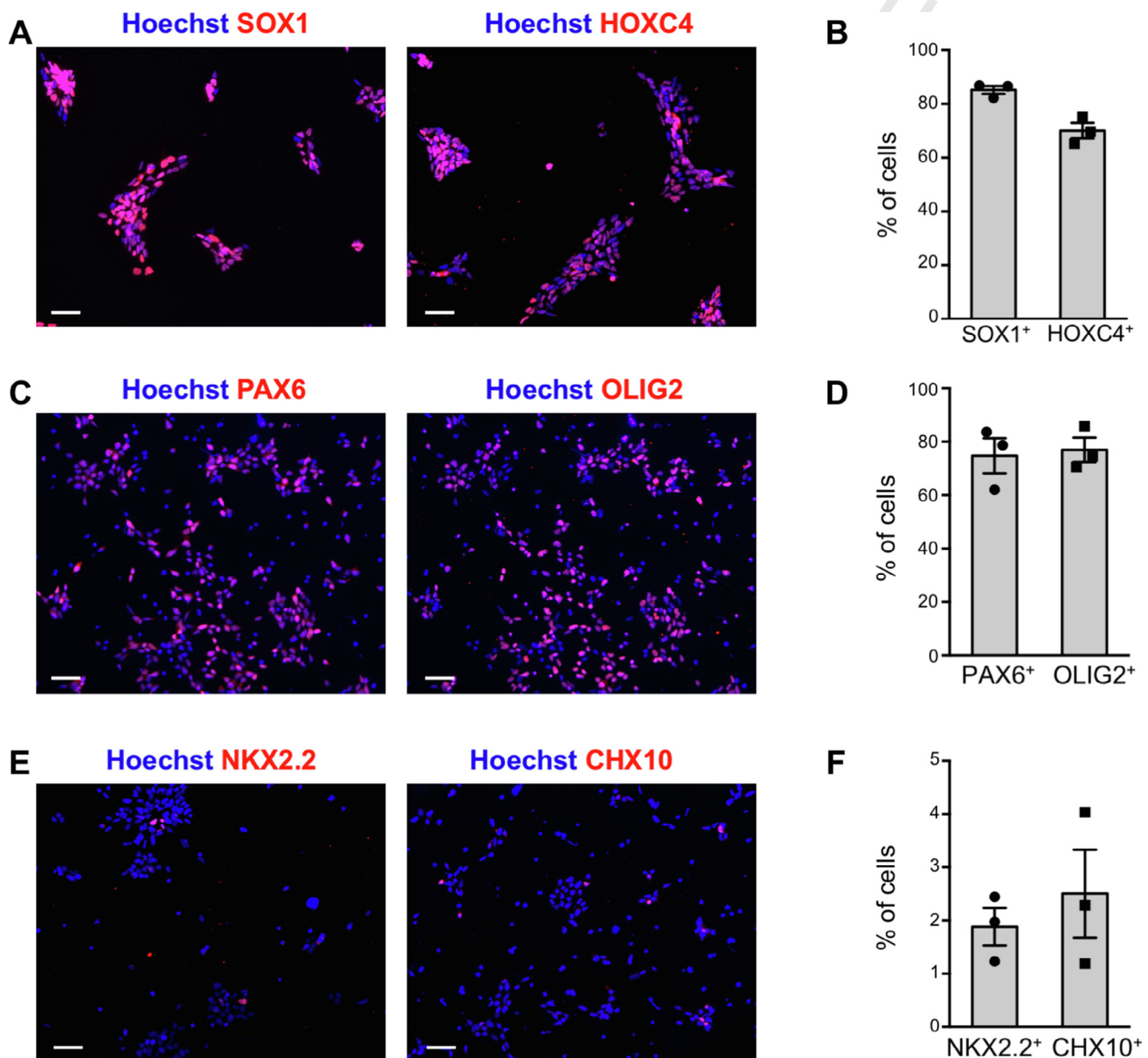


Fig. 1. Generation of an enriched population of human iPSC-derived motor neuron progenitor cells. (A, B) Representative images (A) and quantification (B) of neural progenitor cells (NPCs) subjected to immunocytochemistry with either anti-SOX1 or anti-HOXC4 antibodies 6 days after the start of the differentiation protocol. (C, D) Representative images (C) and quantification (D) of MN progenitor cells (MNPCs) subjected to immunocytochemistry with either anti-OLIG2 or anti-PAX6 antibodies 12 days after the start of the differentiation protocol. (E, F) Representative images (E) and quantification (F) of V3 interneuron progenitors (INPs) subjected to immunocytochemistry with anti-NKX2.2 antibody or V2 INPs visualized with anti-CHX10 antibody, 12 days after the start of the differentiation protocol. For all graphs, $n = 3$ cultures (with > 500 cells in random fields for each culture). Scale bars = 50 μ m.

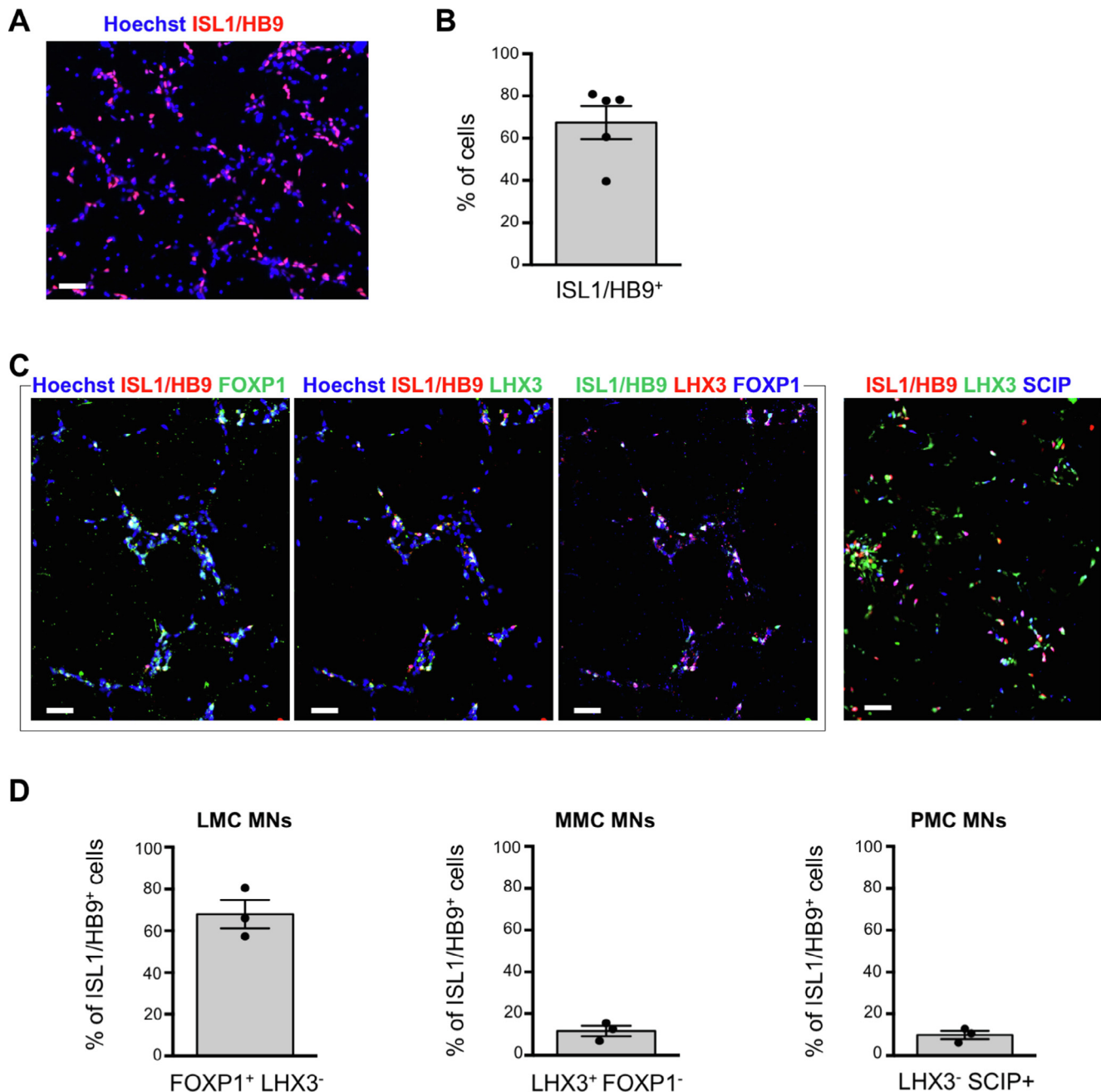


Fig. 2. Different subpopulations of human iPSC-derived motor neurons identified by immunocytochemistry. **(A, B)** Representative image **(A)** and quantification **(B)** of h-IPSC-derived MNs visualized by immunocytochemistry with the pan-MN markers HB9 and ISL1 combined (ISL1/HB9) 25 days after the start of the differentiation protocol ($n = 5$ cultures; > 500 cells in random fields for each culture). **(C, D)** Representative images **(C)** and quantification **(D)** of MN subpopulations identified through expression of different combinations of anti-ISL1/HB9, anti-LHX3, anti-FOXP1, and anti-SCIP antibodies, as revealed by immunocytochemistry. ISL1⁺/HB9⁺/FOXP1⁺/LHX3⁻ cells are defined as LMC MNs. ISL1⁺/HB9⁺/LHX3⁺/FOXP1⁻ cells are defined as MMC MNs. ISL1⁺/HB9⁺/SCIP⁺/LHX3⁻ cells are defined as PMC MNs. For all graphs in **(D)**, $n = 3$ cultures (with > 500 cells in random fields for each culture). LMC, lateral motor column, MMC, median motor column, PMC, phrenic motor column. Scale bars = 50 μ m.

another, and/or variability of cell culture reagents from one lot to another. The majority ($68 \pm 7\%$) of iPSC-derived ISL1⁺/HB9⁺ MNs did not express the LHX3 protein (Thaler et al., 2002; Agalliu et al., 2009) and instead expressed FOXP1, commonly considered as a marker of limb innervating MNs of the lateral motor column (LMC) (Dasen et al., 2003, 2008; Amoroso et al., 2013)

(Fig. 2C, D). Among the ISL1⁺/HB9⁺ cells that did not express FOXP1, $15 \pm 3\%$ did not express LHX3 either, suggesting that this fraction of MNs had a molecular profile (FOXP1⁻/LHX3⁻) similar to that of MNs of the hypaxial motor column (HMC), containing MNs innervating respiratory muscles *in vivo* (data not shown). Consistent with this possibility, we observed that $10 \pm 2\%$ of

ISL1⁺/HB9⁺ MNs expressed the transcription factor POU CLASS 3 HOMEODOMAIN 1/OCTAMER BINDING PROTEIN 6/SCIP (SCIP), known to be enriched *in vivo* in MNs of the phrenic motor column (PMC), a component of the HMC (Bermingham et al., 1996; Philippidou et al., 2012; Machado et al., 2014) (Fig. 2C, D). Lastly, only 12 ± 3% of the h-iPSC-derived MNs were FOXP1⁺ and expressed LHX3, thus exhibiting a profile (FOXP1⁺/LHX3⁺) resembling that of MNs of the median motor column (MMC) (Fig. 2C, D). Together, these observations suggest that h-iPSC-derived MNPCs can be differentiated into a population of heterogeneous spinal MNs comprising cells exhibiting molecular profiles suggestive of LMC (68%), HMC (15%), and MMC (11%) MN fates, thus underscoring the ability of h-iPSCs to give rise to heterogeneous MN cultures.

Characterization of cellular heterogeneity of human iPSC-derived MN culture using single-cell RNA sequencing

To further characterize the cellular composition of h-iPSC-derived MN cultures, we performed microdroplet-based sc-RNAseq, which enables a precise characterization of the genome-wide expression profile of individual cells (Stegle et al., 2015). We chose to analyze day-28 MN cultures, because at this relatively early stage of *in vitro* differentiation, induced MNs have not yet begun to coalesce into large cell clusters, as they do after continued culture (Calder et al., 2015; data not shown), which would make the isolation of single cells technically challenging. After sc-RNAseq, FASTQ files were acquired from 5900 single cells. We used the Bioconductor software (Lun et al., 2016) to perform quality control steps to filter out damaged cells, followed by principal-component analysis and t-distributed stochastic neighbor embedding (t-SNE) analysis of the remaining cells. A total of 4584 cells passed quality control, with a median of 5421 mapped reads per cell and a median of 2215 genes expressed per cell. The percentage of mitochondrial genes present in most cells was less than 2%.

The quality-controlled 4584 cells were divided into 14 shared nearest neighbour graph-based clusters (scran) that exhibited distinct gene expression patterns, consistent with cellular heterogeneity in the h-iPSC-derived MN culture (Fig. 3A). As shown in Fig. 3B, the majority of induced cells were mainly contained in clusters #4 (28% of the cells), #6 (15%), #10 (10%) and #11 (9%). In order to characterize the different cell types in the identified clusters, we performed integrative analysis to compare the cell proportions and gene expression differences between each cluster. Based on differential genes enriched in each cluster, 6 major types of cells were identified (Fig. 3C). Clusters #1, #3, and #11 contained a large proportion of cells identified as “MNPCs and/or MNs”, based on their expression of specific markers such as *NEUROGENIN2* (*NEUROG2*), *OLIG2*, *HB9*, *ISL1*, *ISL2*, and *CHOLINE ACETYLTRANSFERASE* (*CHAT*). Cells of cluster #13 were annotated as “undifferentiated stem-like cells” based on high expression of the pluripotency markers, *OCTAMER BINDING PROTEIN 4* (*OCT4*) and *NANOG* (Nichols

and Smith, 2012) (Fig. 3D). Clusters #10 and #12 were annotated as “NPCs” based on the expression of *SOX1*, *SOX2* and *MARKER OF PROLIFERATION KI-67* (*MKI67*) (Kan et al., 2007; Graham et al., 2003; Scholzen and Gerdes, 2000). Clusters #5 and #6 were annotated as “INs”, with high expression of specific marker genes for IN progenitors such as *SOX14*, *NKX2.2*, *VISUAL SYSTEM HOMEODOMAIN 1* (*VSX1*), *CHX10*, and *SINGLE-MINDED HOMOLOG 1* (*SIM1*) (Ericson et al., 1997; Novitsch et al., 2001; Sugimori et al., 2007; Alaynick et al., 2011; Debrulle et al., 2019), as well as the mature IN marker *CALBINDIN 1* (*CALB1*) (Alvarez et al., 2005). Cluster #2 was annotated as “astrocytic glial cells” based on high expression of *S100 CALCIUM BINDING PROTEIN B* (*S100β*) and *SOX9* (Rosengren et al., 1986; Sun et al., 2017). Cluster #4 represented a mix of MNPCs, astrocytic glial cells and oligodendrocytes, with high expression of *NEUROG2*, *OLIG2*, *HB9*, *ISL1*, *ISL2*, *VESICULAR ACETYLCHOLINE TRANSPORTER* (*VACHT*), *S100β* and *SOX9*, as well as *PLATELET DERIVED GROWTH FACTOR RECEPTOR ALPHA* (*PDGFRα*) and *GALACTOSYL CERAMIDASE* (*GALC*) (Hall et al., 1996; Miller, 2002). Clusters #8, #9, and #14 represented “apoptotic cells”, with high expression of *BCL2 ASSOCIATED X PROTEIN* (*BAX*) (Wei et al., 2001; D’Orsi et al., 2015) and *NERVE GROWTH FACTOR* (*NGF*) (Frade and Barde, 1999).

To confirm this cluster annotation, we correlated genes specifically expressed in different clusters with defined gene ontology (GO) terms. Genes that were expressed in clusters associated with NPCs were significantly enriched for neurogenesis ($p = 3.366e^{-07}$) and nervous system development ($p = 8.511e^{-05}$) categories. Genes that were specifically expressed in clusters associated with MNs were significantly enriched for MN axon guidance pathways ($p = 3.084e^{-05}$). Genes specifically expressed in the IN clusters were significantly enriched for ventral spinal cord IN specification ($p = 8.826e^{-05}$) and neuronal differentiation ($p = 3.004e^{-06}$). Genes specifically expressed in the astrocytic glial cell clusters were significantly enriched for immune system processes ($p = 6.116e^{-06}$) and immune responses ($p = 8.909e^{-05}$). Genes that were enriched in clusters associated with apoptotic cells were significantly enriched for regulation of programmed cell death ($p = 1.039e^{-10}$) and neuronal apoptotic processes ($p = 7.76e^{-05}$). Overall, cells that went through sc-RNAseq analysis unscathed were mainly enriched in MNs (35%), INs (28%) and glial cells (22%) (Fig. 3D). As shown in Fig. 3E, the vast majority (96.2%) of INs were identified as V2 INs expressing *SOX14*, *CHX10* and *GATA BINDING PROTEIN 3* (*GATA3*). Ventral V3 INs, expressing *NKX2.2* and *SIM1*, represented only 0.7% of the IN population, and the more dorsal IN populations were mainly absent: specifically, V1 INs [*ENGRAILED1* (*EN1*)⁺/*EVEN-SKIPPED HOMEODOMAIN 1* (*EVX1*)⁺] accounted for 1.9% of cells; V0 INs (*EN1*[−]/*EVX1*⁺) accounted for 0.5% of cells; and dorsal INs [*LADYBIRD HOMEODOMAIN 1* (*LBX1*)⁺/*T CELL LEUKEMIA HOMEODOMAIN 3* (*TLX3*)⁺, or *LBX1*⁺/*PAX2*⁺]

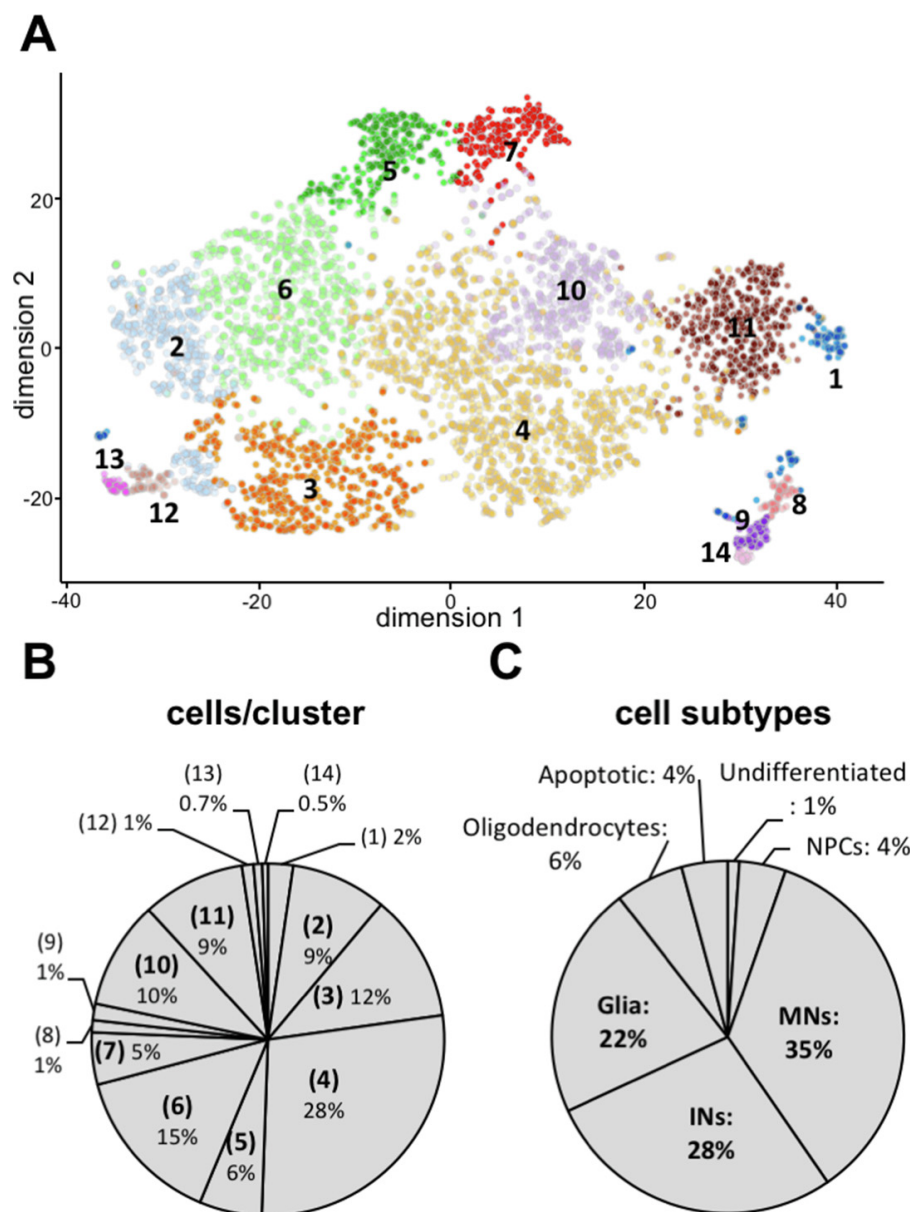


Fig. 3. Heterogeneous composition of human iPSC-derived motor neuron culture defined by single cell RNA sequencing. **(A)** t-SNE plot of h-iPSC derived MN culture after 28 days of differentiation, showing 14 cell shared nearest neighbour graph-based cell clusters. **(B)** Percentage of total cells present in each cluster. **(C)** Relative proportion of iPSCs, NPCs, INs, MNs and glial cells, identified based on the specific expression of the following genes: *NANOG* and *OCT4* for iPSCs; *SOX1*, *SOX2*, and *MKI67* for NPCs; *PAX2*, *PAX3*, *LBX1*, *EVX1*, *EN1*, *CHX10*, *GATA3*, *SOX14*, *SIM1*, *TLX3* for INs; *NEUROG2*, *OLIG2*, *HB9*, *ISL1*, *ISL2*, *CHAT* for MNs; *S100B*, *SOX9* for glial cells. **(D)** t-SNE plot colored for MNPCs and MNs (orange dots), identified based on the specific expression of the following genes: *NEUROG2*, *HB9*, *ISL1*, *ISL2*, *CHAT*. Blue dots represent the rest of the cells. **(E)** Proportion of each of the spinal interneuron (IN) subtypes identified based on combinatorial gene expression patterns: dorsal INs (dINs), *LBX1*⁺/*PAX2*⁺/*TLX3*⁺; V0 INs, *EVX1*⁺/*EN1*⁺; V1 INs, *EVX1*⁺/*EN1*⁺; V2 INs, *CHX10*⁺/*SOX14*⁺/*GATA3*⁺; V3 INs, *SIM1*⁺/*NKX2.2*⁺.

2015), these findings suggest that MN viability may be affected by the single-cell isolation procedure more than other cell types, leading to a more selective loss of MNs during sc-RNAseq compared to other cell types.

Different subtypes of human iPSC-derived motor neurons identified by single-cell RNA sequencing

We next sought to identify the different MN subtypes derived from h-iPSCs. To this end, we selected the cells collectively characterized as MNPCs and/or MNs; namely, with high expression of *NEUROG2*, *OLIG2*, *HB9*, *ISL1*, *ISL2*, and *CHAT*, as described above. A total of 1629 such cells passed quality control, with a median of 6322 mapped reads per cell and a median of 2483 genes expressed per cell. The percentage of mitochondrial genes present in most cells was less than 2%. The 1629 cells were divided into 8 shared nearest neighbour graph-based clusters exhibiting distinct gene expression patterns, suggesting heterogeneity in the MNPC/MN population (Fig. 4A). As shown in Fig. 4B, the majority of cells were mainly found in clusters #2 (26% of the cells), #1 (23%), #5 (20%) and #6 (16%).

To identify the different MN subtypes within these clusters, we performed integrative analysis to compare the cell proportions and gene expression differences between each cluster. Based on differential genes enriched in each cluster, 6 major types of cells were identified (Fig. 4C). Clusters #1, #2, #5, and #7 harboured a large proportion of cells identified as LMC MNs, based on their expression of *FOXP1* and lack of expression of *LHX3* (Fig. 4D, top t-SNE plot). Conversely, MMC MNs, defined as expressing *LHX3*

and lacking *FOXP1* expression, were mainly contained in clusters #6 and #8 (Fig. 4D, middle t-SNE plot). Furthermore, spinal HMC MNs, expressing *HOMEBOX A5* (*HOXA5*) and lacking both *FOXP1* and *LHX3*, were found in clusters #5 and #6 (Fig. 4D, bottom t-SNE plot). In addition, cluster #4 represented MNPCs with high expression of *NKX6.1*, *OLIG2* and

for 0.5% of cells. Together, these results show that h-iPSC-derived cultures obtained under MN differentiation experimental conditions and subjected to sc-RNAseq contain mainly MNs and V2 INs. Furthermore, when compared to immunocytochemistry results shown in Figs. 1 and 2, as well as previous similar immunocytochemistry studies by other groups (e.g., Amoroso et al., 2013; Maury et al., 2015; Du et al.,

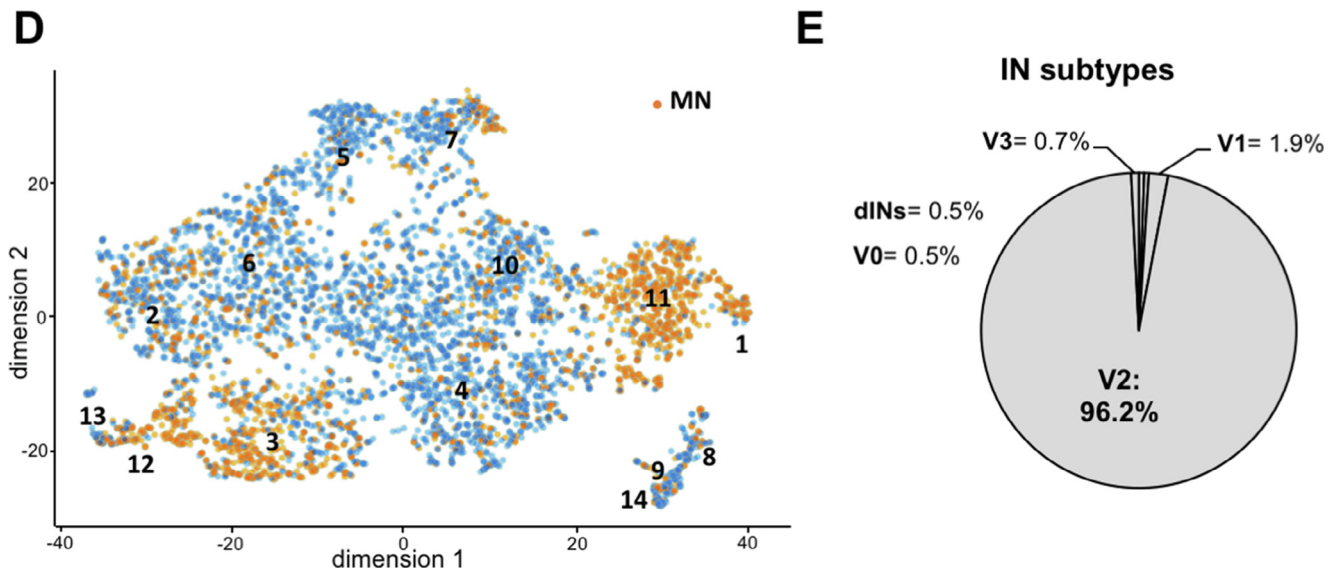


Fig. 3 (continued)

LHX3 (Briscoe et al., 2000), while cluster #3 was annotated as “apoptotic cells”, with high expression of *BAX*, *CASPASE 3* (*CASP3*), *NGF* and *FAS CELL SURFACE DEATH RECEPTOR* (*FAS*) (Raoul et al., 1999).

To confirm this cluster annotation, we searched for genes specifically expressed in each cell type that were enriched for the expected appropriate GO terms (Fig. 4E). For example, genes that were specifically expressed in the MNPC cluster were significantly associated with mitotic cell cycle ($p = 2.596e^{-49}$), neuronal migration ($p = 3.193e^{-07}$), and pattern specification processes ($p = 7.509e^{-05}$). Genes that were expressed in the apoptotic cell cluster were significantly enriched for neuronal apoptotic processes ($p = 6.596e^{-05}$) and regulation of neuronal death ($p = 6.111e^{-05}$). Genes expressed in the LMC clusters were significantly enriched for locomotor behavior ($p = 9.601e^{-09}$), trans-synaptic signaling ($p = 1.341e^{-07}$) and spinal cord development ($p = 7.078e^{-05}$). Genes specifically expressed in the MMC clusters were significantly enriched for MN axon guidance ($p = 3.192e^{-07}$), G-protein coupled acetylcholine receptor signaling pathways ($p = 3.192e^{-07}$) and spinal cord development ($p = 4.941e^{-05}$). Genes that were specifically expressed in the HMC clusters were significantly enriched for sensory system development ($p = 3.835e^{-06}$), MN axon guidance ($p = 8.383e^{-05}$) and respiratory system development ($p = 7.284e^{-05}$). Interestingly, MNs expressing the specific brainstem marker genes, *MAF BZIP TRANSCRIPTION FACTOR B* (*MAFB*) and *HOXB1*, were also identified and were mainly contained in cluster #1. However, cells in this cluster did not express the typical hindbrain markers *ORTHODENTICLE HOMEODOMAIN 2* (*OTX2*), *T-BOX TRANSCRIPTION FACTOR 20* (*TBX20*) or *PAIRED-LIKE HOMEODOMAIN 2A/B* (*PHOX2A/B*), suggesting the presence of incompletely caudalized neurons, rather than more mature brainstem MNs. Overall, LMC MNs

were the most predominant (58%), followed by HMC MNs (19%), and incompletely caudalized MNs (14%). The majority of HMC MNs (63%) expressed genes such as *SCIP* or *TEASHIRT ZINC FINGER HOMEODOMAIN 1* (*TSHZ1*), known to be enriched in phrenic MNs, and they were thus annotated as PMC MNs (Fig. 4C). MNPCs were also detected (6%), whereas we did not detect MNs of the spinal accessory column (SAC) (Fig. 4C). Together, these results provide evidence for different subtypes of h-iPSC-derived MNs, as well as other neural cell types that normally arise during the transition from undifferentiated progenitors to committed precursors to differentiated neurons in the spinal cord.

Different subtypes of human iPSC-derived LMC motor neurons identified by single-cell RNA sequencing

LMC MNs were observed in four separate clusters (#1, #2, #5, and #7) (Fig. 4D). We next asked whether this observation could be due to the presence of distinct subsets of LMC MNs. Indeed, LMC MNs acquire specialized “pool” identities that determine their ability to form selective connections with limb muscles *in vivo*. Dasen and colleagues showed that HOX protein-regulated mechanisms operate within MNs and assign rostro-caudal motor pool positions, as well as pool identities at a single segmental level (Dasen et al., 2005). Consequently, subsets of LMC MNs can be distinguished by their expression of specific HOX proteins (Carpenter, 2002; Lance-Jones et al., 2001; Dasen et al., 2005). In addition, LIM homeodomain proteins define the medial and lateral subdivisions of the LMC (Sockanathan and Jessell, 1998; Kania and Jessell, 2003; Roussio et al., 2008). Early born LMC MNs co-express ISL1 and ISL2, and at later stages ISL1 specifies medial LMC (LMCm) MNs. Later born LMC MNs down-regulate ISL1 and express LHX1, whose expression

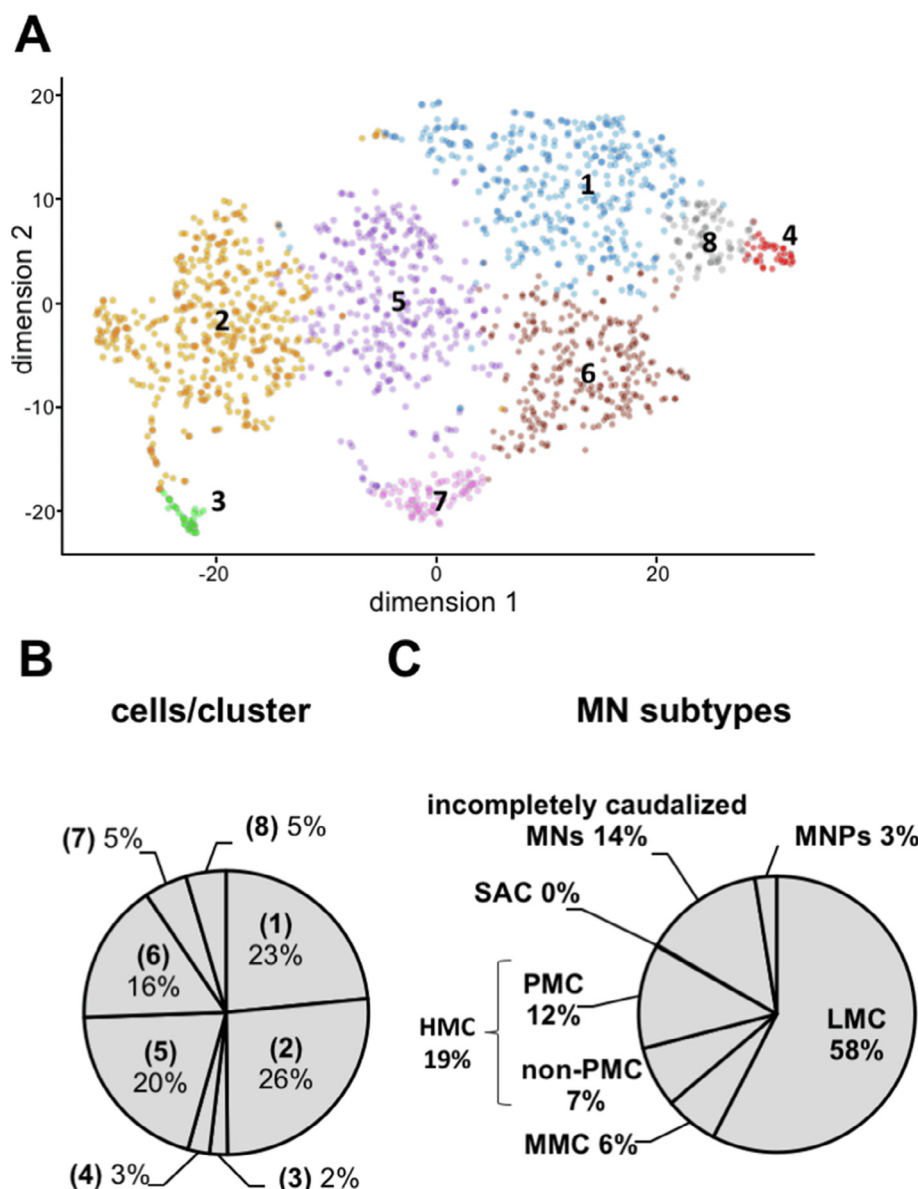


Fig. 4. Identification of motor neuron subtypes with single-cell RNA sequencing. **(A)** t-SNE plot of all cells identified as MNs in h-iPSC derived MN culture after 28 days of differentiation, showing 8 shared nearest neighbour graph-based cell clusters. **(B)** Percentage of total cells present in each cluster. **(C)** Proportion of each MN subtype, identified by their expression of specific genes: MNPs, *OLIG2*⁺; LMC MNs, *FOXP1*⁺/*LHX3*⁺; MMC MNs, *FOXP1*⁺/*LHX3*⁺; cervical HMC MNs, *FOXP1*⁺/*LHX3*⁺/*HOXA5*⁺; PMC MNs, cervical HMC expressing *SCIP* and/or *TSHZ1*; SAC MNs, *PHOX2B*⁺; incompletely caudalized MNs, *MAFB*⁺, *HOXB1*⁺. **(D)** t-SNE plot colored for the 3 major subtypes of spinal MNs (orange dots): LMC, MMC, and cervical HMC, as defined above. **(E)** GO analysis of the most differentially expressed genes between each cluster. Three «Biological Process» GO categories with a *p* value < e-05 are shown for each cluster. The number of differentially expressed genes (DEGs) associated with each category is reported, and each category is colored in grey if *p* = e-05 or red if *p* < e-05. dev., development; syst., system; Ach-R = acetylcholine receptor.

‘cervical LMCm’ (33%, orange), ‘immature cervical LMC’, which were not identified as more developmentally mature LMCI or LMCm (34%, purple), and ‘lumbar LMC’ (5%, pink) (Fig. 5A). The majority of the LMC MNs (91%) were identified as cervical LMC MNs, co-expressing *HOX6* genes and *ALDH1A2*, and were mainly contained in clusters #1, #2, and #5 (Fig. 5B). In contrast, the LMC MNs contained in cluster #7 were identified as lumbar LMC MNs, lacking *HOX6* genes and expressing *HOXD8* and *HOX9* genes: this group represented only 5% of the LMC MNs and 1.1% of the total number of MNs (Fig. 5C). Among the cervical LMC MNs, the rostral-most cells, expressing higher levels of *HOX5* genes and lower levels of *HOXC8*, were mainly contained in clusters #1 and #2 (Fig. 5D), while the caudal-most cervical LMC MNs (*HOXC8*^{high}/*HOX5*^{low}) were limited to cluster #1 (Fig. 5E). In addition, clusters #2 and #7 contained cells expressing the highest levels of *ISL1*, defining LMCI MNs (Fig. 5F). In contrast, cells of clusters #1 and #5 expressed higher levels of *LHX1*, specifying LMCm neurons (Fig. 5G). As depicted in Fig. 5H, cluster #1 contained a majority of cervical LMCI MNs (57%, blue) while cluster #2 contained mainly cervical LMCm MNs (64%, orange). Cluster #5 contained a majority of cervical LMC MNs that were not identified as either LMCI or LMCm (53%, purple). The lumbar LMC MNs accounted for 41% of the LMC found in cluster #7 (pink).

To extend this analysis, we considered the previous demonstration that the specific *HOX* profiles of LMC MNs direct the differential expression of the transcription factors *RUNX1*, *SCIP*, and *PEA3*, which in turn defines four groups of MNs with distinct rostro-caudal and intrasegmental positions within the cervical LMC (Dasen et al., 2005). However, none of the LMC MNs expressed either *RUNX1* or *PEA3* and only a few expressed *SCIP*, suggesting a developmentally immature state of LMC MNs after 28 days *in vitro*. In summary, cervical LMC MNs were the most predominant (91%), while

defines lateral LMC (LMCI) MNs (Kania and Jessell, 2003).

To characterize different human-iPSC-derived LMC MN subtypes, we performed integrative analysis to compare *ISL1*, *LHX1* and *HOX* genes expression differences between each cluster. Based on differential genes enriched in each cluster, 4 major types of LMC MNs were identified: ‘cervical LMCI’ (24%, blue colour),

defines four groups of MNs with distinct rostro-caudal and intrasegmental positions within the cervical LMC (Dasen et al., 2005). However, none of the LMC MNs expressed either *RUNX1* or *PEA3* and only a few expressed *SCIP*, suggesting a developmentally immature state of LMC MNs after 28 days *in vitro*. In summary, cervical LMC MNs were the most predominant (91%), while

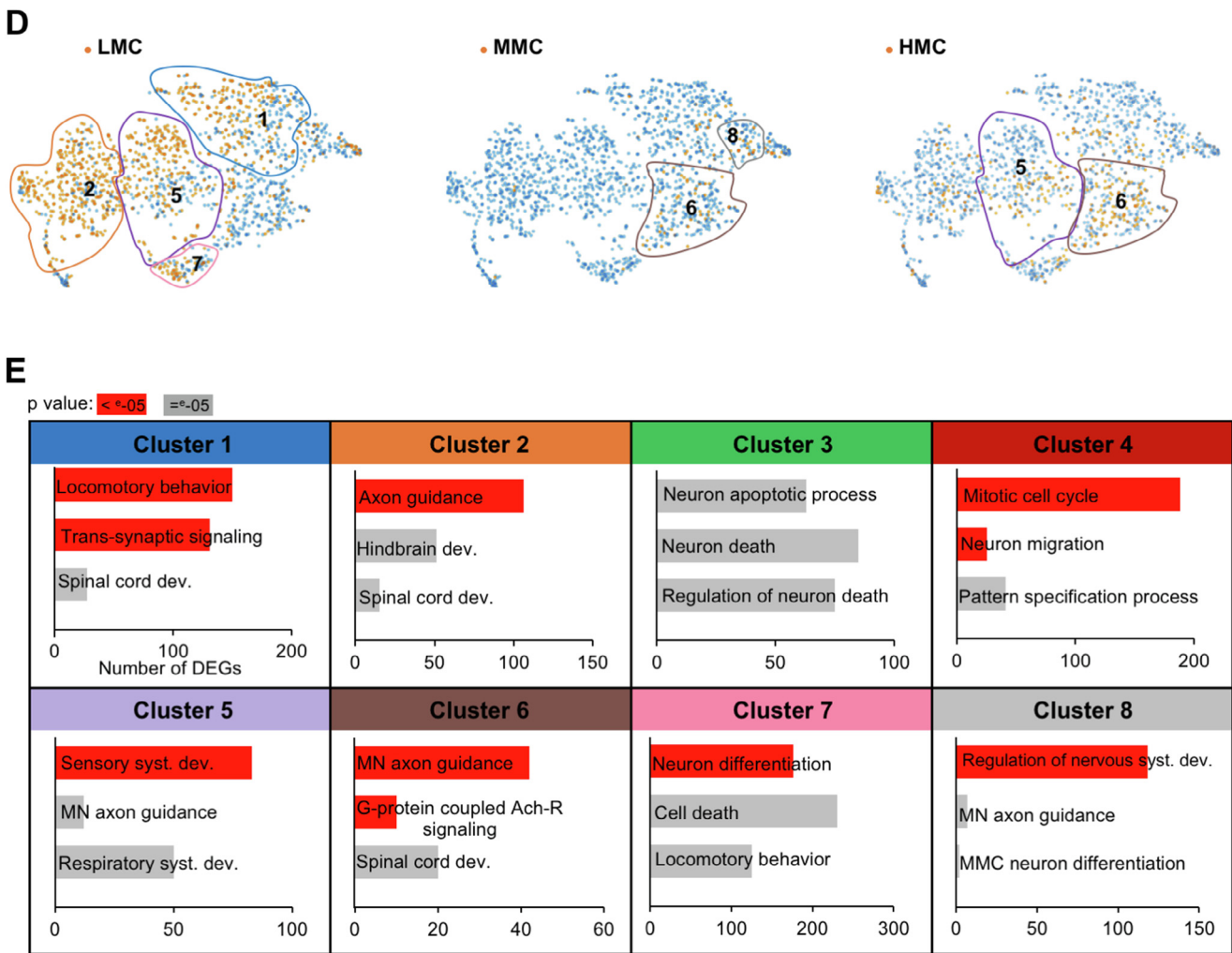


Fig. 4 (continued)

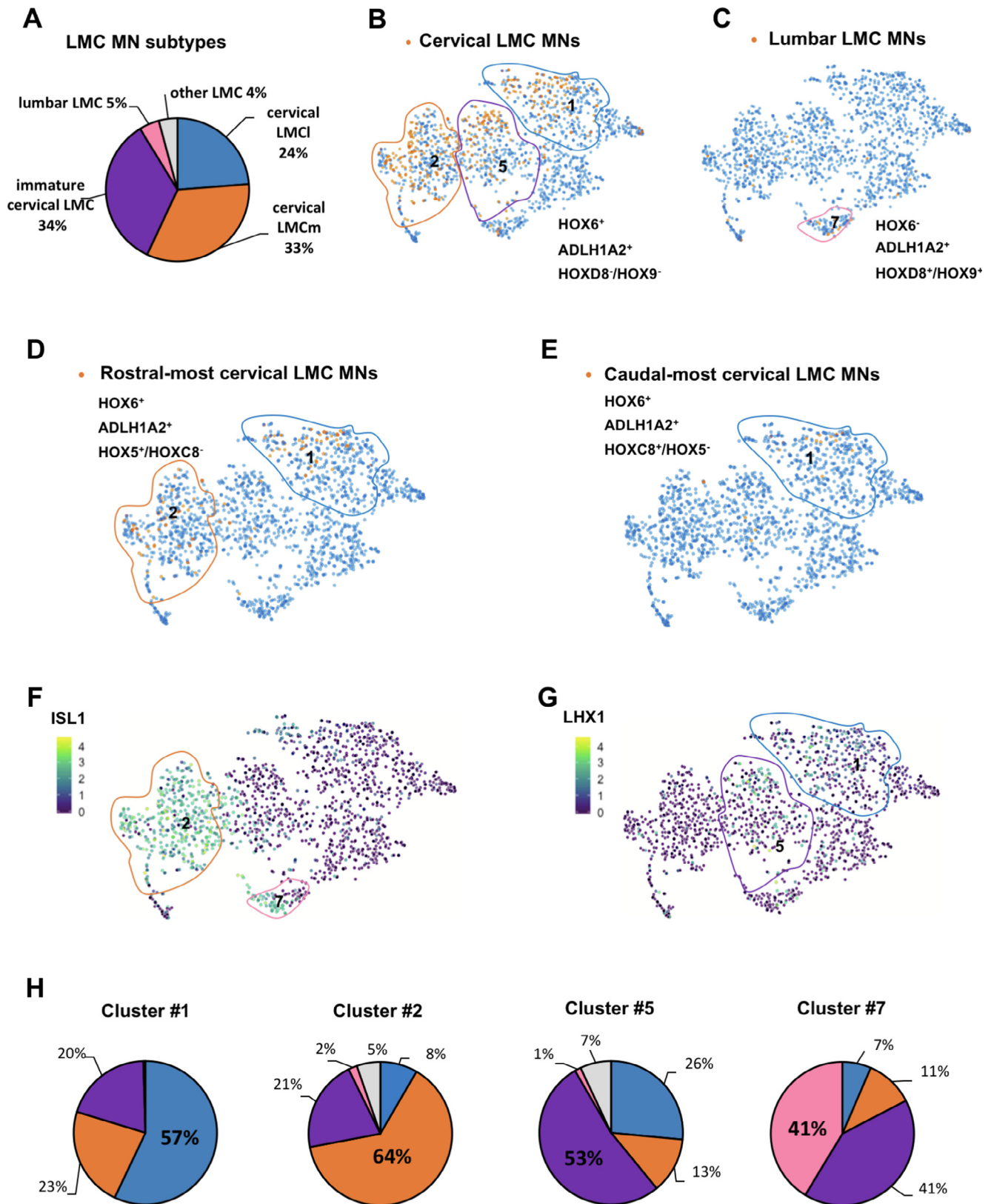
only 5 % were identified as lumbar LMC MNs under the examined experimental conditions.

DISCUSSION

In most studies to date, the characterization of h-iPSC-derived MNs has relied mainly on the expression of specific molecular markers by either immunocytochemistry or real-time PCR approaches (Li et al., 2005; Patani et al., 2011; Amoroso et al., 2013; Chen et al., 2014; Maury et al., 2015; Du et al., 2015). Nascent spinal MNs are characterized by transient co-expression of LIM homeodomain (LIM-HD) transcription factors like HB9, ISL1/2, and LHX3 (Sharma et al., 1998; Arber et al., 1999; Thaler et al., 1999). Thus, most MN derivation studies rely on a combination of HB9 and ISL1/2 immunostaining (operationally referred to as pan-MN staining – Amoroso et al., 2013) to characterize the MN phenotype *in vitro*. However, LIM-HD proteins are gradually down-regulated during motor neuron development *in vivo*, suggesting that these proteins may not be ideal markers to monitor the gradual differentiation of induced MNs *in vitro*. The *CHAT* and *VACHT* genes are

expressed by all cholinergic neurons including spinal MNs and they have been used as markers for more developmentally mature MNs (Habecker and Landis, 1994). These generic MN markers are frequently used in combination with more specific motor column MN markers, such as FOXP1 or SCIP, to identify particular MN subtypes. Even these combined approaches, however, fall short of providing tools for precisely defining the molecular profiles of all h-iPSC-derived MNs, especially when considering that common differentiation protocols give rise to heterogenous neural cell cultures (Li et al., 2005; Patani et al., 2011; Amoroso et al., 2013; Chen et al., 2014; Maury et al., 2015; Du et al., 2015). Based on these observations, the present study aimed at developing protocols to perform untargeted characterization of gene expression in all cell subtypes present in cultures of h-iPSC derived spinal MNs.

We chose to analyze day-28 h-iPSC-derived MN cultures because, at this relatively early stage of *in vitro* differentiation, induced MNs already express the typical pan-MN markers ISL1 and HB9. Furthermore, at this stage of differentiation, h-iPSC-derived MNs have not yet coalesced into large cell clusters, as they typically



do at more developmentally mature stages under standard *in vitro* derivation protocols. The presence of these large MN clusters would make it technically challenging to isolate single MNs without causing cell damage due to the need of harsher dissociation techniques. The sc-RNAseq results presented in this study show that h-iPSCs can give rise, in 28 days of differentiation, to heterogeneous neural cell cultures containing a majority of spinal MNPCs and MNs, as well as other spinal neural cells such as INs and glia. We observed that immunocytochemical analysis of h-iPSC-derived cultures after 25 days of differentiation showed a relative proportion of induced MNs (67%) larger than the fraction of MNs detected using sc-RNAseq (35%). This situation could be the result of a specific loss of MNs over other cell types during single-cell suspension preparation prior to sequencing, as MNs may be more sensitive to mechanical and enzymatic cell dissociation.

Strategies to generate and characterize different h-iPSC-derived MNs are important to be able to study the function and dysfunction of distinct human MNs, including their pathologies in MN diseases such as ALS (Frey et al., 2000; Stifani, 2014). However, little is known about the diversity of induced human MNs *in vitro*. The present immunocytochemistry and sc-RNAseq data strongly suggest that h-iPSCs can give rise to spinal MNs of the lateral, median, hypaxial and phrenic motor columns *in vitro*. Both approaches revealed that the majority (58%) of MNs exhibit a molecular profile similar to that of LMC MNs. This finding is consistent with the enrichment in LMC MN identity previously shown by Amoroso and coworkers through the combined use of RA, the SHH pathway activator Smoothened Agonist, and purmorphamine (Amoroso et al., 2013). It is worth mentioning that other studies (e.g., Davis-Dusenbery et al., 2014; Qu et al., 2014; Maury et al., 2015) reported generation of mixed populations of LMC and MMC subtypes in human stem cell-derived MNs *in vitro*. At the other end of the spectrum, enhanced specification of the MMC MN fate was achieved with the omission of retinoids from the differentiation medium (Patani et al., 2011).

In the present study, we have identified four separate pools of LMC MNs by combining the expression of previously identified LMC MN markers, such as *FOXP1* and *ALDH1A2*, with the differential expression of *HOX* family genes, *ISL1* (a marker of LMCm MNs) and *LHX1* (a marker of LMCI MNs). The scRNA-seq data used to obtain a catalog of distinct *in vitro* generated human LMC MNs will undoubtedly provide a publicly available

resource to the community that will enhance efforts aimed at characterizing MN subtype-specific markers. Similarly, our present observation that MNs exhibiting a molecular profile corresponding to HMC MNs represent ~19% of the h-iPSC-derived MNs, and that ~12% of these cells express either *SCIP* or *TSHZ1*, two PMC MN markers, provides a foundation for future characterization of the identity, and possibly diversity, of human HMC MNs. This information is currently lacking. In summary, the results of the present studies show that untargeted assessment of gene expression by RNA profiling on a single cell level can provide a global measure of spinal MN diversity in h-iPSC-derived MN cultures.

Improved means to precisely phenotype the different MN subtypes generated from human iPSCs are expected to promote advancements in various fields of research. They will facilitate the study of the biology of different types of human MNs, enabling a better characterization of their physiological properties. Moreover, they may lead to the development of strategies to enrich for defined MN subtypes that may offer more attractive options for modeling specific MN diseases and/or injury conditions. In turn, deeply phenotyped disease-relevant MN subtypes may provide ideal experimental model systems to develop cellular assays with increased potential to facilitate early-stage drug discovery efforts in ALS and other MN diseases.

DATA AVAILABILITY

All data underlying the results as well as the R workflow used for the analysis are available on GitHub: <https://github.com/regan-hamel/h-iPSCs-MNI>.

AUTHOR CONTRIBUTIONS

LT performed all experiments, data analysis, figure preparation and wrote the first draft of the manuscript. LT and RH analyzed the sc-RNAseq data. SS, TD and SP conceived the study plan. SS designed the experiments and supervised data analysis and manuscript writing.

COMPETING INTERESTS

The authors declare no competing interests.

Fig. 5. Identification of LMC motor neuron subtypes by single-cell RNA sequencing. **(A)** Proportion of different LMC MN subtypes identified by their expression of specific genes: cervical LMCI MNs (blue), *HOX6*⁺/*ALDH1A2*⁺/*HOX5*⁺/*LHX1*⁺, or *HOX6*⁺/*ALDH1A2*⁺/*HOXC8*⁺/*LHX1*⁺; cervical LMCm MNs (orange), *HOX6*⁺/*ALDH1A2*⁺/*HOXC8*⁺/*ISL1*⁺; lumbar LMC MNs (pink), *HOX6*⁺/*ALDH1A2*⁺/*HOXD8*⁺/*HOX9*⁺. Cervical LMC neurons that were not identified as mature LMCm or LMCI MNs were annotated as 'immature cervical LMC' (purple). LMC neurons that did not belong to any of these categories were annotated as "other LMC" (grey). **(B, C)** t-SNE plot of MNs, colored for cervical **(B)** or lumbar **(C)** LMC MNs (orange dots), identified by their expression of specific marker genes: cervical LMC MNs, *FOXP1*⁺/*LHX3*⁺/*HOX6*⁺/*ALDH1A2*⁺/*HOXD8*⁺/*HOX9*⁺; lumbar LMC MNs, *FOXP1*⁺/*LHX3*⁺/*HOX6*⁺/*ALDH1A2*⁺/*HOXD8*⁺/*HOX9*⁺. **(D, E)** t-SNE plot of MNs colored for the rostral-most **(D)** or the caudal-most **(E)** cervical LMC MNs (orange dots), identified by their expression of the following specific marker genes: rostral-most cervical LMC MNs, *HOXA5*⁺/*HOXC5*⁺/*HOXC8*⁺; caudal-most cervical LMC MNs, *HOXA5*⁺/*HOXC5*⁺/*HOXC8*⁺. **(F, G)** t-SNE plot of cells identified as MNs colored according to their level of expression of the transcription factor-encoding genes *ISL1* **(F)** or *LHX1* **(G)**. **(H)** Proportion of each LMC MN subtypes identified as described above, in each of the four clusters previously annotated as LMC MN clusters (#1, #2, #5, and #7). LMC, lateral motor column; LMCI, lateral LMC; LMCm, medial LMC.

FUNDING INFORMATION

These studies were funded in part by Canadian Institutes for Health Research and Fonds de la recherche en Santé-Québec under the frame of E-Rare-3, the ERA-Net for Research on Rare Diseases (SS), the Douglas Avrieth MNI-Cambridge Neuroscience Collaboration Initiative (SS, SP, RH), and the Montreal Neurological Institute Tony Proudfoot Postdoctoral Fellowship Award (LT).

UNCITED REFERENCE

Dasen and Jessell (2009).

ACKNOWLEDGMENTS

We thank Dr. Ragoussis' team at the McGill University and Genome Quebec Innovation Centre for scRNA library preparation, as well as Yeman Tang, Carol Chen, and Rita Lo for invaluable advice and assistance.

REFERENCES

Agalliu D, Takada S, Agalliu I, McMahon AP, Jessell TM (2009) Motor neurons with axial muscle projections specified by Wnt4/5 signaling. *Neuron* 61(5):708–720.

Alaynick WA, Jessell TM, Pfaff SL (2011) SnapShot: spinal cord development. *Cell* 146(1):178–178.e1.

Alizadeh A, Dyck SM, Karimi-Abdolrezaee S (2019) Traumatic spinal cord injury: an overview of pathophysiology, models and acute injury mechanisms. *Front Neurol* 10:282.

Alvarez FJ, Jonas PC, Sapir T, Hartley R, Berrocal MC, Geiman EJ, Goulding M (2005) Postnatal phenotype and localization of spinal cord V1 derived interneurons. *J Comp Neurol* 493(2):177–192.

Amin N, Tan X, Ren Q, Zhu N, Botchway BOA, Hu Z, Fang M (2019) Recent advances of induced pluripotent stem cells application in neurodegenerative diseases. *Prog Neuropsychopharmacol Biol Psychiatry* 95 109674.

Amoroso MW, Croft GF, Williams DJ, O'Keefe S, Carrasco MA, Davis AR, Wichterle H (2013) Accelerated high-yield generation of limb-innervating motor neurons from human stem cells. *J Neurosci* 33(2):574–586.

Arber S, Han B, Mendelsohn M, Smith M, Jessell TM, Sockanathan S (1999) Requirement for the homeobox gene Hb9 in the consolidation of motor neuron identity. *Neuron* 23(4):659–674.

Ardhanareeswaran K, Mariani J, Coppola G, Abyzov A, Vaccarino FM (2017) Human induced pluripotent stem cells for modelling neurodevelopmental disorders. *Nat Rev Neurol* 13(5):265–278.

Bermingham Jr JR, Scherer SS, O'Connell S, Arroyo E, Kalla KA, Powell FL, Rosenfeld MG (1996) Tst-1/Oct-6/SCIP regulates a unique step in peripheral myelination and is required for normal respiration. *Genes Dev* 10(14):1751–1762.

Briscoe J, Pierani A, Jessell TM, Ericson J (2000) A homeodomain protein code specifies progenitor cell identity and neuronal fate in the ventral neural tube. *Cell* 101(4):435–445.

Burke AC, Nelson CE, Morgan BA, Tabin C (1995) Hox genes and the evolution of vertebrate axial morphology. *Development* 121(2):333–346.

Calder EL, Tchieu J, Steinbeck JA, Tu E, Keros S, Ying SW, Jaiswal MK, Cornacchia D, Goldstein PA, Tabar V, Studer L (2015) Retinoic acid-mediated regulation of GLI3 enables efficient motoneuron derivation from human ESCs in the absence of extrinsic SHH activation. *J Neurosci* 35(33):11462–11481.

Carpenter EM (2002) Hox genes and spinal cord development. *Dev Neurosci* 24:24–34.

Chen H, Qian K, Du Z, Cao J, Petersen A, Liu H, Zhang SC (2014) Modeling ALS with iPSCs reveals that mutant SOD1 misregulates

neurofilament balance in motor neurons. *Cell Stem Cell* 14(6):796–809.

Dasen JS, De Camilli A, Wang B, Tucker PW, Jessell TM (2008) Hox repertoires for motor neuron diversity and connectivity gated by a single accessory factor, FoxP1. *Cell* 134(2):304–316.

Dasen JS, Jessell TM (2009) Hox networks and the origins of motor neuron diversity. *Curr Top Dev Biol* 88:169–200.

Dasen JS, Liu JP, Jessell TM (2003) Motor neuron columnar fate imposed by sequential phases of hox-c activity. *Nature* 425(6961):926–933.

Dasen JS, Tice BC, Brenner-Morton S, Jessell TM (2005) A Hox regulatory network establishes motor neuron pool identity and target-muscle connectivity. *Cell* 123:477–491.

Davis-Dusenbery BN, Williams LA, Klim JR, Eggan K (2014) How to make spinal motor neurons. *Development* 141(3):491–501.

Debrulle S, Baudouin C, Hidalgo-Figueroa M, Pelosi B, Francius C, Rucchin V, Clotman F (2019) Vsx1 and Chx10 paralogs sequentially secure V2 interneuron identity during spinal cord development. *Cell Mol Life Sci*.

D'Orsi B, Kilbride SM, Chen G, Perez Alvarez S, Bonner HP, Pfeiffer S, Prehn JH (2015) Bax regulates neuronal Ca²⁺ homeostasis. *J Neurosci* 35(4):1706–1722.

Du ZW, Chen H, Liu H, Lu J, Qian K, Huang CL, Zhang SC (2015) Generation and expansion of highly pure motor neuron progenitors from human pluripotent stem cells. *Nat Commun* 6:6626.

Ebert AD, Yu J, Rose Jr FF, Mattis VB, Lorson CL, Thomson JA, Svendsen CN (2009) Induced pluripotent stem cells from a spinal muscular atrophy patient. *Nature* 457(7227):277–280.

Egawa N, Kitaoka S, Tsukita K, Naitoh M, Takahashi K, Yamamoto T, Adachi F, Kondo T, Okita K, Asaka I, Aoi T, Watanabe A, Yamada Y, Morizane A, Takahashi J, Ayaki T, Ito H, Yoshikawa K, Yamawaki S, Suzuki S, Watanabe D, Hioki H, Kaneko T, Makioka K, Okamoto K, Takuma H, Tamaoka A, Hasegawa K, Nonaka T, Hasegawa M, Kawata A, Yoshida M, Nakahata T, Takahashi R, Marchetto MC, Gage FH, Yamanaka S, Inoue H (2012) Drug screening for ALS using patient-specific induced pluripotent stem cells. *Sci Transl Med* 4(145):145ra104.

Ericson J, Rashbass P, Schedl A, Brenner-Morton S, Kawakami A, van Heyningen V, Briscoe J (1997) Pax6 controls progenitor cell identity and neuronal fate in response to graded shh signaling. *Cell* 90(1):169–180.

Frade JM, Barde YA (1999) Genetic evidence for cell death mediated by nerve growth factor and the neurotrophin receptor p75 in the developing mouse retina and spinal cord. *Development* 126(4):683–690.

Francius C, Clotman F (2014) Generating spinal motor neuron diversity: a long quest for neuronal identity. *Cell Mol Life Sci* 71(5):813–829.

Frey D, Schneider C, Xu L, Borg J, Spooren W, Caroni P (2000) Early and selective loss of neuromuscular synapse subtypes with low sprouting competence in motoneuron diseases. *J Neurosci* 20(7):2534–2542.

Graham V, Khudyakov J, Ellis P, Pevny L (2003) SOX2 functions to maintain neural progenitor identity. *Neuron* 39(5):749–765.

Habecker BA, Landis SC (1994) Noradrenergic regulation of cholinergic differentiation. *Science (New York, N.Y.)* 264(5165):1602–1604.

Hall A, Giese NA, Richardson WD (1996) Spinal cord oligodendrocytes develop from ventrally derived progenitor cells that express PDGF alpha-receptors. *Development* 122(12):4085–4094.

Haston KM, Finkbeiner S (2016) Clinical trials in a dish: the potential of pluripotent stem cells to develop therapies for neurodegenerative diseases. *Ann Rev Pharmacol Toxicol* 56:489–510.

Jessell TM (2000) Neuronal specification in the spinal cord: inductive signals and transcriptional codes. *Nat Rev Genet* 1(1):20–29.

Kan L, Jalali A, Zhao LR, Zhou X, McGuire T, Kazanis I, Kessler JA (2007) Dual function of Sox1 in telencephalic progenitor cells. *Dev Biol* 310(1):85–98.

- Kania A, Jessell TM (2003) Topographic motor projections in the limb imposed by LIM homeodomain protein regulation of ephrin-A: EphA interactions. *Neuron* 38(1):581–596.
- Kanning KC, Kaplan A, Henderson CE (2010) Motor neuron diversity in development and disease. *Ann Rev Neurosci* 33:409–440.
- Kiskinis E, Sandoe J, Williams LA, Boulting GL, Moccia R, Wainger BJ, Eggan K (2014) Pathways disrupted in human ALS motor neurons identified through genetic correction of mutant SOD1. *Cell Stem Cell* 14(6):781–795.
- Lance-Jones C, Omelchenko N, Bailis A, Lynch S, Sharma K (2001) Hoxd10 induction and regionalization in the developing lumbosacral spinal cord. *Development* 128:2255–2268.
- Li XJ, Du ZW, Zarnowska ED, Pankratz M, Hansen LO, Pearce RA, Zhang SC (2005) Specification of motoneurons from human embryonic stem cells. *Nat Biotechnol* 23(2):215–221.
- Lun AT, McCarthy DJ, Marioni JC (2016) A step-by-step workflow for low-level analysis of single-cell RNA-seq data with bioconductor. *F1000research* 5:2122.
- Machado CB, Kanning KC, Kreis P, Stevenson D, Crossley M, Nowak M, Lieberam I (2014) Reconstruction of phrenic neuron identity in embryonic stem cell-derived motor neurons. *Development* 141(4):784–794.
- Maury Y, Come J, Piskowski RA, Salah-Mohellibi N, Chevalere V, Peschanski M, Nedelec S (2015) Combinatorial analysis of developmental cues efficiently converts human pluripotent stem cells into multiple neuronal subtypes. *Nat Biotechnol* 33(1):89–96.
- McCarthy DJ, Campbell KR, Lun ATL, Willis QF (2017) Scater: pre-processing, quality control, normalisation and visualisation of single-cell RNA-seq data in R. *Bioinformatics* 33:1179–1186.
- Methot L, Soubannier V, Hermann R, Campos E, Li S, Stifani S (2018) Nuclear factor-kappaB regulates multiple steps of gliogenesis in the developing murine cerebral cortex. *Glia* 66(12):2659–2672.
- Miller RH (2002) Regulation of oligodendrocyte development in the vertebrate CNS. *Prog Neurobiol* 67(6):451–467.
- Nichols J, Smith A (2012) Pluripotency in the embryo and in culture. *Cold Spring Harb Perspect Biol* 4(8):a008128.
- Nijssen J, Comley LH, Hedlund E (2017) Motor neuron vulnerability and resistance in amyotrophic lateral sclerosis. *Acta Neuropathol* 133(6):863–885.
- Novitsch BG, Chen AI, Jessell TM (2001) Coordinate regulation of motor neuron subtype identity and pan-neuronal properties by the bHLH repressor Olig2. *Neuron* 31(5):773–789.
- Patani R, Hollins AJ, Wishart TM, Puddifoot CA, Alvarez S, de Lera AR, Chandran S (2011) Retinoid-independent motor neurogenesis from human embryonic stem cells reveals a medial columnar ground state. *Nat Commun* 2:214.
- Philippidou P, Walsh CM, Aubin J, Jeannotte L, Dasen JS (2012) Sustained Hox5 gene activity is required for respiratory motor neuron development. *Nat Neurosci* 15(12):1636–1644.
- Prasad A, Hollyday M (1991) Development and migration of avian sympathetic preganglionic neurons. *J Comp Neurol* 307(2):237–258.
- Qu Q, Li D, Louis KR, Li X, Yang H, Sun Q, Wang F (2014) High-efficiency motor neuron differentiation from human pluripotent stem cells and the function of islet-1. *Nat Commun* 5:3449.
- Raoul C, Henderson CE, Pettmann B (1999) Programmed cell death of embryonic motoneurons triggered through the fas death receptor. *J Cell Biol* 147(5):1049–1062.
- Rosengren LE, Kjellstrand P, Aurell A, Haglid KG (1986) Irreversible effects of dichloromethane on the brain after long term exposure: a quantitative study of DNA and the glial cell marker proteins S-100 and GFA. *Br J Ind Med* 43(5):291–299.
- Rousso DL, Gaber ZB, Wellik D, Morrissey EE, Novitsch BG (2008) Coordinated actions of the forkhead protein Foxp1 and Hox proteins in the columnar organization of spinal motor neurons. *Neuron* 59:226–240.
- Scholzen T, Gerdes J (2000) The ki-67 protein: from the known and the unknown. *J Cell Physiol* 182(3):311–322.
- Sockanathan S, Jessell TM (1998) Motor neuron-derived retinoid signaling specifies the subtype identity of spinal motor neurons. *Cell* 94:503–514.
- Stegle O, Teichmann SA, Marioni JC (2015) Computational and analytical challenges in single-cell transcriptomics. *Nat Rev Genet* 16(3):133–145.
- Stifani N (2014) Motor neurons and the generation of spinal motor neuron diversity. *Front Cell Neurosci* 8:293.
- Sugimori M, Nagao M, Bertrand N, Parras CM, Guillemot F, Nakafuku M (2007) Combinatorial actions of patterning and HLH transcription factors in the spatiotemporal control of neurogenesis and gliogenesis in the developing spinal cord. *Development* 134(8):1617–1629.
- Sun W, Cornwell A, Li J, Peng S, Osorio MJ, Aalling N, Nedergaard M (2017) SOX9 is an astrocyte-specific nuclear marker in the adult brain outside the neurogenic regions. *J Neurosci* 37(17):4493–4507.
- Thaler J, Harrison K, Sharma K, Lettieri K, Kehrl J, Pfaff SL (1999) Active suppression of interneuron programs within developing motor neurons revealed by analysis of homeodomain factor HB9. *Neuron* 23(4):675–687.
- Thaler JP, Lee SK, Jurata LW, Gill GN, Pfaff SL (2002) LIM factor Lhx3 contributes to the specification of motor neuron and interneuron identity through cell-type-specific protein-protein interactions. *Cell* 110(2):237–249.
- Tsuchida T, Ensini M, Morton SB, Baldassare M, Edlund T, Jessell TM, Pfaff SL (1994) Topographic organization of embryonic motor neurons defined by expression of LIM homeobox genes. *Cell* 79(6):957–970.
- Vallstedt A, Muhr J, Pattyn A, Pierani A, Mendelsohn M, Sander M, Ericson J (2001) Different levels of repressor activity assign redundant and specific roles to Nkx6 genes in motor neuron and interneuron specification. *Neuron* 31(5):743–755.
- Wei MC, Zong WX, Cheng EH, Lindsten T, Panoutsakopoulou V, Ross AJ, Korsmeyer SJ (2001) Proapoptotic BAX and BAK: a requisite gateway to mitochondrial dysfunction and death. *Science* 292(5517):727–730.
- Wichterle H, Lieberam I, Porter JA, Jessell TM (2002) Directed differentiation of embryonic stem cells into motor neurons. *Cell* 110(3):385–397.

(Received 16 January 2020, Accepted 24 April 2020)
(Available online xxxx)