



Universiteit
Leiden
The Netherlands

Confronting ALS: understanding multicellular contribution to neurodegeneration : computational analysis and hiPSCs in vitro modelling as a multidisciplinary approach

Limone, F.

Citation

Limone, F. (2023, September 26). *Confronting ALS: understanding multicellular contribution to neurodegeneration : computational analysis and hiPSCs in vitro modelling as a multidisciplinary approach*. Retrieved from <https://hdl.handle.net/1887/3642489>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

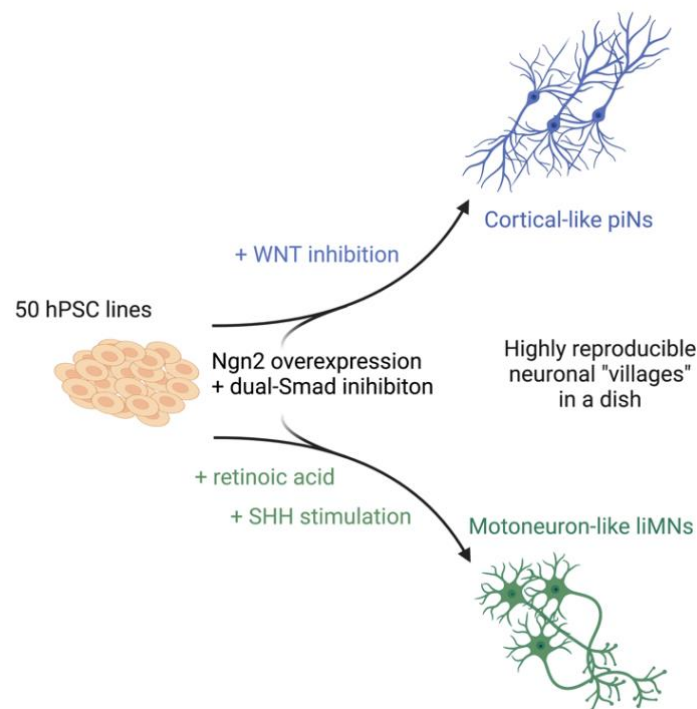
Downloaded from: <https://hdl.handle.net/1887/3642489>

Note: To cite this publication please use the final published version (if applicable).

Chapter 4:

Efficient generation of lower induced Motor Neurons by coupling *Ngn2* expression with developmental cues

In this chapter we describe a new protocol to differentiate motor neurons from human Pluripotent Stem Cells (hPSCs) for a big-scale, high-throughput study of neurodegenerative diseases with the aim to provide a platform for more reproducible modelling of ALS from large cohort of hiPSC lines.



Graphical abstract. Limone et al. induce neuralization of hPSCs into spinal MNs by small molecule patterning and TF overexpression. Multiplexed, pooled single-cell RNAsequencing showcases high reproducibility in dozens of cell lines. These MN villages resemble in vivo spinal MNs and produce disease-relevant MN populations.

This work is published on *Cell Reports*, 2022

<https://doi.org/10.1016/j.celrep.2022.111896>

Earlier versions of this chapter can be found on bioRxiv

<https://doi.org/10.1101/2022.01.12.476020>

Efficient generation of lower induced Motor Neurons by coupling *Ngn2* expression with developmental cues

Francesco Limone^{1,2,3,*}, Jana M. Mitchell^{1,2,9}, Irune Guerra San Juan^{1,2,6,8,9}, Janell L. M. Smith¹, Kavya Raghunathan^{1,2}, Daniel Meyer^{2,4}, Sulagna Dia Ghosh^{1,2,4}, Alexander Couto¹, Joseph R. Klim^{1,2}, Brian J. Joseph^{1,7}, John Gold¹, Curtis J. Mello^{2,4}, James Nemesh^{2,4}, Brittany M. Smith¹, Matthijs Verhage^{6,8}, Steven A. McCarroll^{2,4}, Olli Pietiläinen^{1,2,5}, Ralda Nehme^{1,2} and Kevin Eggan^{1,10,*}

¹Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, MA 02138, USA ²Stanley Centre for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA ³Leiden University medical Center, LUMC, 2333 ZA Leiden, The Netherlands ⁴Department of Genetics, Harvard Medical School, Boston, MA 02115, USA ⁵Neuroscience Center, Helsinki Institute of Life Science, University of Helsinki, Helsinki, Finland ⁶Department of Functional Genomics, Center for Neurogenomics and Cognitive Research, Vrije Universiteit Amsterdam, The Netherlands ⁷Departments of Pathology and Cell Biology, Columbia University Irving Medical Centre, New York, NY 10032, USA ⁸Human genetics, Amsterdam University Medical Center, 1081 HV Amsterdam, The Netherlands ⁹These authors contributed equally ¹⁰Lead Contact. *Correspondence: kevin.eggan@bmrn.com, francesco.limone@fas.harvard.edu

Human pluripotent stem cells (hPSCs) are a powerful tool for disease modelling of hard-to-access tissues (such as the brain). Current protocols either direct neuronal differentiation with small molecules or use transcription-factor-mediated programming. In this study, we couple overexpression of transcription factor Neurogenin2 (*Ngn2*) with small molecule patterning to differentiate hPSCs into lower induced Motor Neurons (liMoNes/liMNs). This approach induces canonical MN markers including motor neuron (MN) specific marker *Hb9/MNX1* activation in >95% of cells. liMNs resemble *bona fide* hPSC-derived MN, exhibit spontaneous electrical activity, express synaptic markers and can contact muscle cells *in vitro*. Pooled, multiplexed single-cell RNA sequencing on 50 hPSC-lines reveals reproducible populations of distinct subtypes of cervical and brachial MNs that resemble their *in vivo*, embryonic counterparts. Combining small molecule patterning with *Ngn2* overexpression facilitates high-yield, reproducible production of disease-relevant MN subtypes, which is fundamental in propelling our knowledge of MN biology and its disruption in disease.

INTRODUCTION

Many groups have recognised the ability of stem cells to differentiate into almost any cell type of the body. This unique capability can facilitate the understanding of basic biology of tissues that are hard to access and that are specifically highly evolved in humans, such as the Central Nervous System (CNS)¹. Most neuronal differentiation schemes mimic developmental embryonic signals by small molecule patterning. The neuralisation of stem cells is achieved by manipulating bone morphogenic protein (BMP) and transforming growth factor β (TGF β), commonly referred to as “dual-Smad inhibition”². This study further showed that different combinations of small molecules used as patterning factors could push neuronal

progenitors towards distinct neuronal fates. From there, many have developed and refined differentiation protocols for specific neuronal subtypes. However, caveats still remain, such as: the incomplete neuralisation of cultures, underlining the need for additional neuralising factors³; the long time needed to generate mature cultures and the heterogeneity in differentiation efficiency amongst cell lines^{4,5}.

To overcome these limitations, others have employed different approaches such as the overexpression of a transcription factors (TFs)⁶. These TFs have been used to generate induced Neurons (iNs) from fibroblasts⁷, and the combination with subtype-specific TFs was able to generate specific types of neurons⁸. These approaches have been transferred to stem cells with one of the more recent reports of *Neurogenin2* (*Ngn2*, *Neurog2*, *Atoh4*) being able to differentiate human Pluripotent Stem Cells (hPSCs) into glutamatergic neurons⁹. These advances allowed reproducible generation of neurons in a shorter time and fewer steps. This approach may, however, skip pivotal developmental steps part of neuronal specification so questions have been raised regarding the identity of the generated populations and the impact of the overexpression of TFs to downstream applications¹⁰.

Previously, we have demonstrated that overexpression of *Ngn2* coupled with small molecule patterning is able to enhance the regional specification of neurons to cortical-like patterned induced Neurons - *piNs*¹¹. Additionally, small molecules have also been reported to enhance efficiency of MN programming^{12,13}. These findings led us to hypothesize that combining *Ngn2* expression with different patterning molecules could generate different neuronal cells.

We wanted to generate spinal Motor Neurons (MNs) for biological modelling of degenerative motoneuron diseases, such as Amyotrophic Lateral Sclerosis (ALS) and Spinal Muscular Atrophy (SMA) that selectively affect these highly specialised neurons¹⁴. MNs reside in the spinal cord and are the only neurons to exit the nervous system and contact skeletal muscles to allow us to breathe and move through a specific synaptic contact, the Neuro-Muscular Junction (NMJ). Protocols to differentiate MNs are based on decades of developmental biology studies^{15,16} and are extensively reviewed elsewhere^{17,18}. Most protocols entail the neuralization inputs described above coupled with ventralising factors like Sonic Hedgehog and/or its agonists (Shh/SAG) and the caudalising effects of retinoids (retinoic acid – RA)^{4,19-21} or, alternatively, the overexpression of a combination of transcription factors: *Ngn2*, *Isl1*, *Lhx3* (i.e. NILs)^{12,22}. Both approaches have proven to be useful for investigating MN biology. However, on one hand directed differentiation produces cultures containing different cell types other than MNs with high line-to-line heterogeneity rendering disease modelling difficult. On the other hand, the overexpression of three TFs produces pure cultures but very specific subtypes of MNs limiting the scalability of these studies since several, specific combinations of TFs are needed to reproduce the diversity of MN subtypes *in vitro*.

Here, we report that the addition of patterning molecules during Ngn2-programming of hPSCs can lead to specification of regionally defined neuronal states. With time in culture, differentially patterned cells developed into morphologically distinct neurons that maintain regionally defined features according to developmental patterning mimicry. A reporter cell line for the MN-specific transcription factor *MX1/Hb9* demonstrated that ~95% of the cells subjected to MN patterning activated this master regulator of MN development. This finding, in combination with the expression of pan-MN markers validated the cellular identity of SAG- and RA-patterned-Ngn2 cells as MN-like cells: the lower induced Motor Neurons (liMoNes/liMNs). liMNs expressed canonical markers and resembled bona fide hiPSC-derived MNs, they were electrophysiologically active and able to form synaptic contact with muscle cells *in vitro*. By leveraging newly developed analysis tools for single-cell RNA-sequencing (scRNAseq) technology that enable analysis of many cell lines cultured in the same dish simultaneously, we demonstrated that our protocol produced several subtypes of disease-relevant diaphragm- and limb-innervating MNs in a robust fashion, that is reproducible across 47 stem cell lines, which resemble primary MNs from the human spinal cord. This combinatorial approach addressed several shortcomings from previously published protocols and will facilitate the understanding of basic spinal MNs biology and its disruption in disease.

RESULTS

Ngn2-driven neuralization can be directed to different neuronal fates by small molecules patterning

Given that the combination of patterning molecules with Ngn2 expression could generate cortical excitatory neurons¹¹, we wondered whether the protocol could be repurposed with alternative patterning factors to generate other types of neurons. To test this hypothesis, we used an overexpression system in which a doxycycline inducible tetO-Ngn2-T2A-Puro/rtTA lentiviral system is used to infect hPSCs for strong overexpression of the neuralising factor Neurogenin2⁹. We started by substituting WNT inhibition, used to generate cortical cells (piNs)¹¹, with ventralising SAG and caudalising RA to induce a ventral-posterior fate and ultimately produce lower-induced Motor Neurons (liMNs) (Figures 1A-B).

To test if the patterning induced regionally specified neuronal states, we selected markers pivotal for early neuronal development that are divergent between cortex and spinal cord (Figures 1C). To this end we collected RNA and performed RT-qPCR at day 4, a stage described as Neuronal Progenitor Cell (NPC)-like¹¹, to assess the expression of these markers. While rostro-dorsalising WNT inhibition induced the expression of master regulators of cortical development *EMX1*, *FOXG1*, *OTX1* and *OTX2* (Figure 1D), the caudal-ventral patterning induced the expression of posterior markers *HOXB4* and *HOXC6*, of cholinergic

master regulator *ISL1*²³ and of *MNX1* (Hb9), expressed by spinal motor neurons in the nervous system²⁴ (Figure 1E). Importantly, caudal-ventral patterning reduced the expression of *OTX1* and *OTX2*, transcription factors that regulate the schism between the cortex and posterior regions of the CNS²⁵. In line with previous studies, dual-Smad inhibition in combination with Ngn2 resulted in loss of pluripotency markers, *OCT4* and *SOX2*, and acquisition of pan-neuronal markers, *PAX6* and *TUBB3* (Figures S1A-B)¹¹.

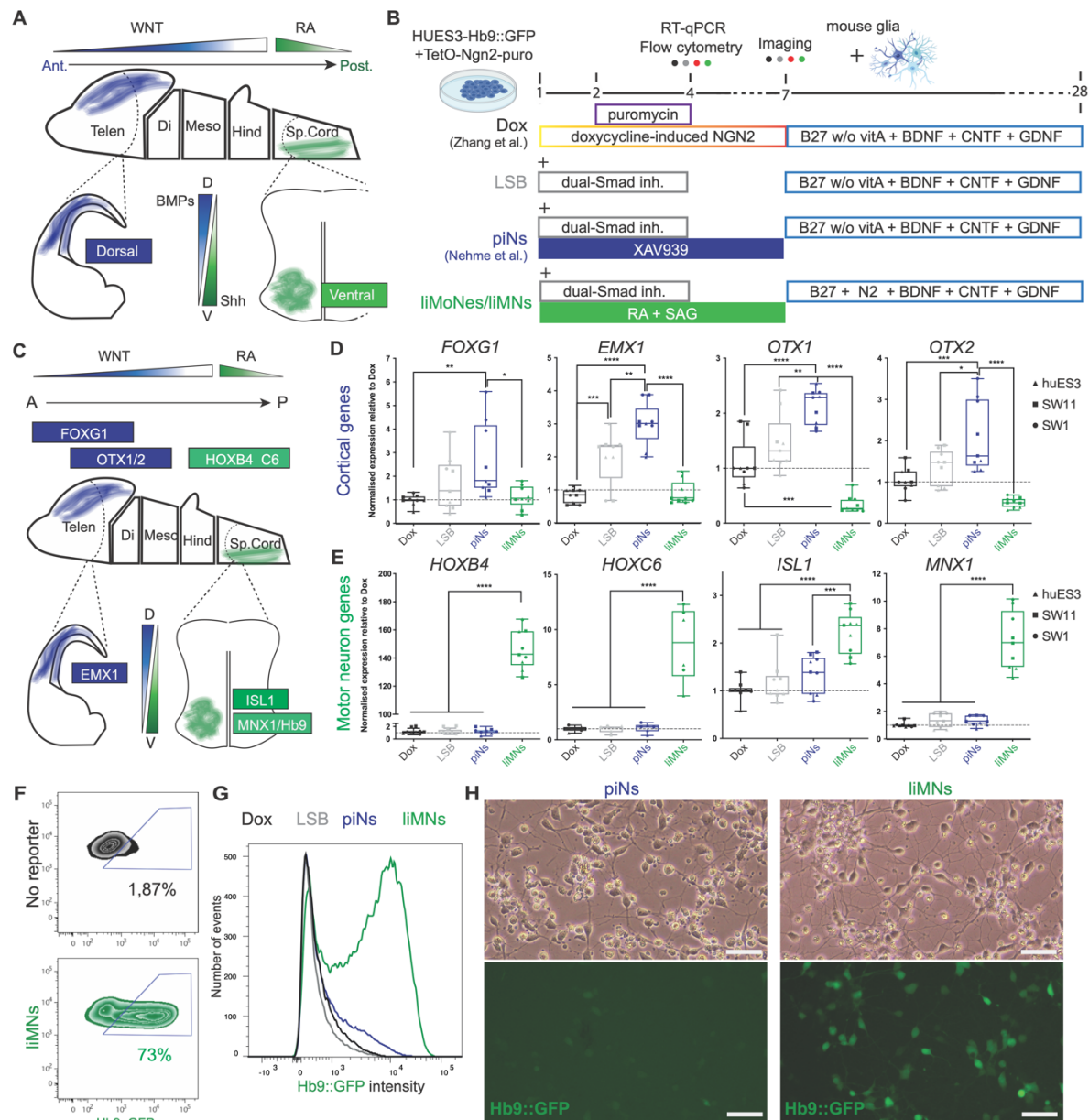


Figure 1 Ngn2-driven neuralization can be directed to different neuronal fates by small molecules patterning. **(A)** Diagram of known developmental cues used to design patterning strategy. **(B)** Differentiation schemes used for comparison of divergent Ngn2-driven trajectories: Dox – original Ngn2 overexpression from Zhang et al. 2013; LSB – Ngn2 overexpression coupled with neuralising dual-Smad inhibition (LDN193189, SB431542); piNs – cortical-like patterned induced Neurons (Nehme et al. 2018); liMoNes/liMNs – lower induced Motor Neurons generated by Ngn2-overexpression and ventro-caudal patterning (Retinoic Acid and Smoothed Agonist). **(C)** Genes selected as master regulators of anterior-dorsal, cortical development and ventro-caudal, spinal cord development. **(D)** RT-qPCR quantification for induction of cortical genes after rostro-dorsalising WNT inhibition at day 4 (three cell lines in $n=3$ technical replicates each, p-values from one-way ANOVA). **(E)** RT-qPCR quantification demonstrating induction of spinal genes after caudal-ventralising induction of SAG and RA at day 4 (three cell lines in $n=3$ technical replicates each, one-way ANOVA). **(F)** Flow cytometry quantification of Hb9::GFP positive cells at day 4. **(G)** Hb9::GFP intensity at day 4 of differentiation demonstrating higher total intensity of the Hb9::GFP signal in liMNs **(H)** Hb9::GFP expression day 7 post-induction in piNs and liMNs, the majority of liMNs express the reporter (scale bar 50 μ m).

To further confirm the regional specification of NPCs, we took advantage of a reporter line that expresses GFP under the murine, MN-specific, Hb9 promoter^{26,27} inserted into human embryonic stem cell line used to validate differentiation protocols^{19,28-30}. Flow cytometry analysis confirmed that by day 4 after induction, more than 70% of cells treated with RA and SAG were GFP positive (Figures 1F). Strikingly, not only was the percentage of GFP⁺ cells higher, but the intensity of GFP signal also increased (Figure 1G and S1C), in agreement with higher levels of *MNX1/Hb9* RNA. By day 7, cells subjected to RA and SAG showed strong *Hb9::GFP* expression whereas only a fewer, dimmer GFP positive cells were visible in the other conditions (Figures 1H and S1D). Taken together, these data suggests that differential patterning coupled with Ngn2-overexpression leads to the specification of different neuronal fates, including MN.

Neuronal fates induced by patterned Ngn2 expression maintained throughout differentiation

We then proceeded to confirm that regional specification was maintained long-term after neurogenesis. For this purpose, we extended *in vitro* culturing by replating cells in neuronally supportive conditions (Figure 2A). First, we analysed cell morphology by microscopy. Patterning produced neurons with strikingly different morphology; with piNs showing small, polarised cell bodies and MN-patterned cells showing a wider soma with a multipolar shape with one extended axon-like structure (Figure 2B and S2A-B), strikingly reminiscent of the morphology of cortical pyramidal neurons and spinal, ventral-horn motor neurons *in vivo*, respectively³¹.

To confirm that the regional identity specified by patterning was maintained, we collected RNA at day 30 of differentiation and investigated the expression of genes known to be specifically expressed in either glutamatergic neurons of the cortex or cholinergic MNs of the spinal cord (Figure 2C). We confirmed that caudalisation repressed cortical genes *SATB2* and *TBR1* (Figure 2D). Expression of posterior markers *HOXB4* and *HOXC6* was sustained in caudalised cells and suppressed in piNs (Figure 2E). Moreover, mature ventralised cells expressed the MN-specific TF, *MNX1/Hb9* and higher transcript levels of the main component of the cholinergic machinery, Choline Acetyltransferase (*CHAT*) (Figure 2E), while maintaining expression of pan-neuronal markers (Figure S2C). According to this polarised gene expression, expression of the *Hb9::GFP* reporter was also maintained through-out differentiation only in RA- and SAG-patterned cells, reaching a peak of ~95% at day 7 (Figure 2F-G and S2E), and was then slightly downregulated as seen in early development of MNs of the spinal cord *in vivo*³². To further ensure their MN identity and overcome some of the

limitations of the reporter, we combined the Hb9::GFP reporter with staining for Islet1 and SMI32, the triad recognised as the human pan-Motor Neuron staining¹⁹ and confirmed that 80% of the cells co-expressed at least two of these markers (Figure 2 H-I and S2D). The data so far confirmed that coupling of Ngn2 overexpression with patterning factors can produce regionally specified neurons and we define the ventralised and caudalised cultures as lower-induced Motor Neurons: *liMoNes/liMNs*.

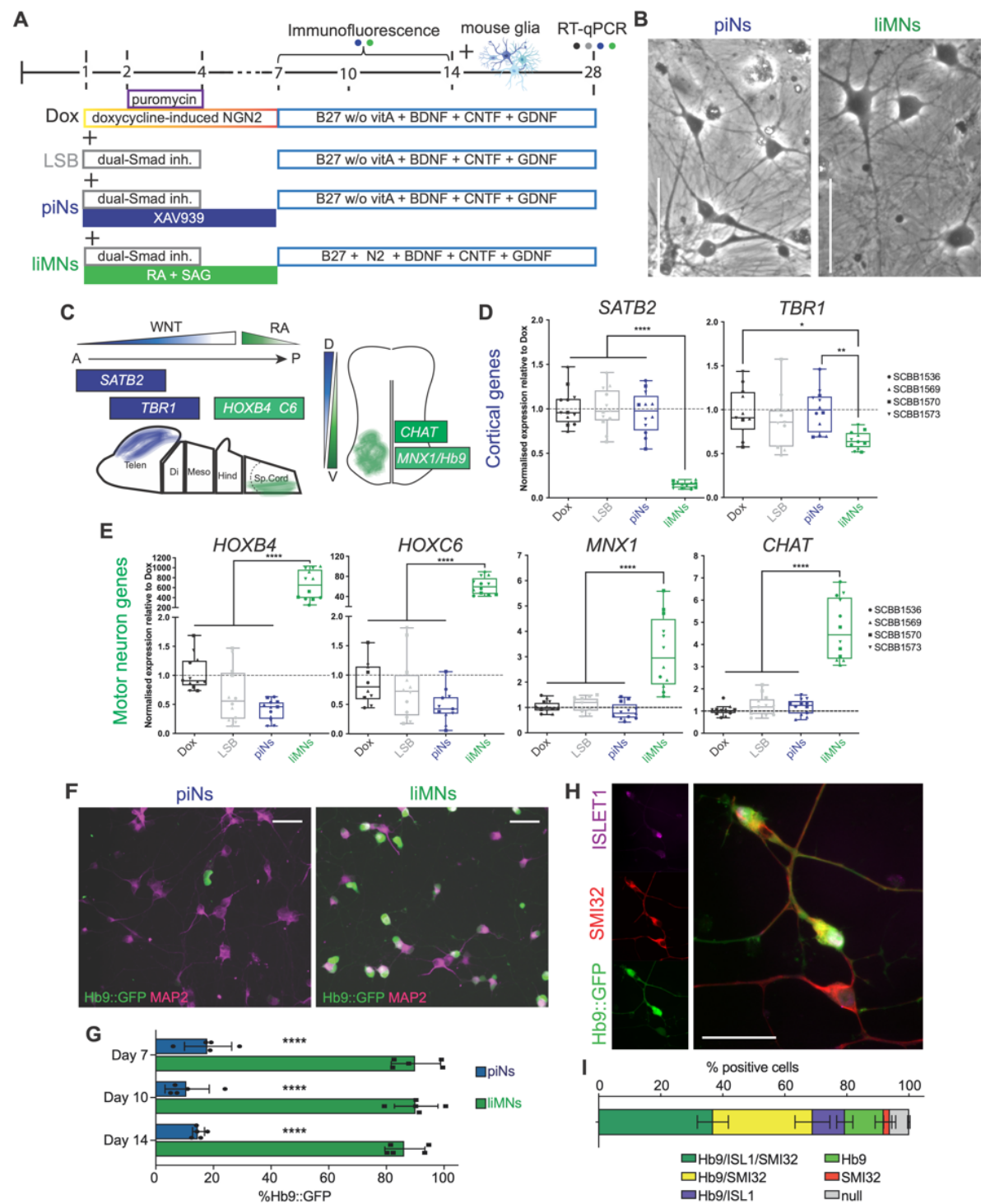
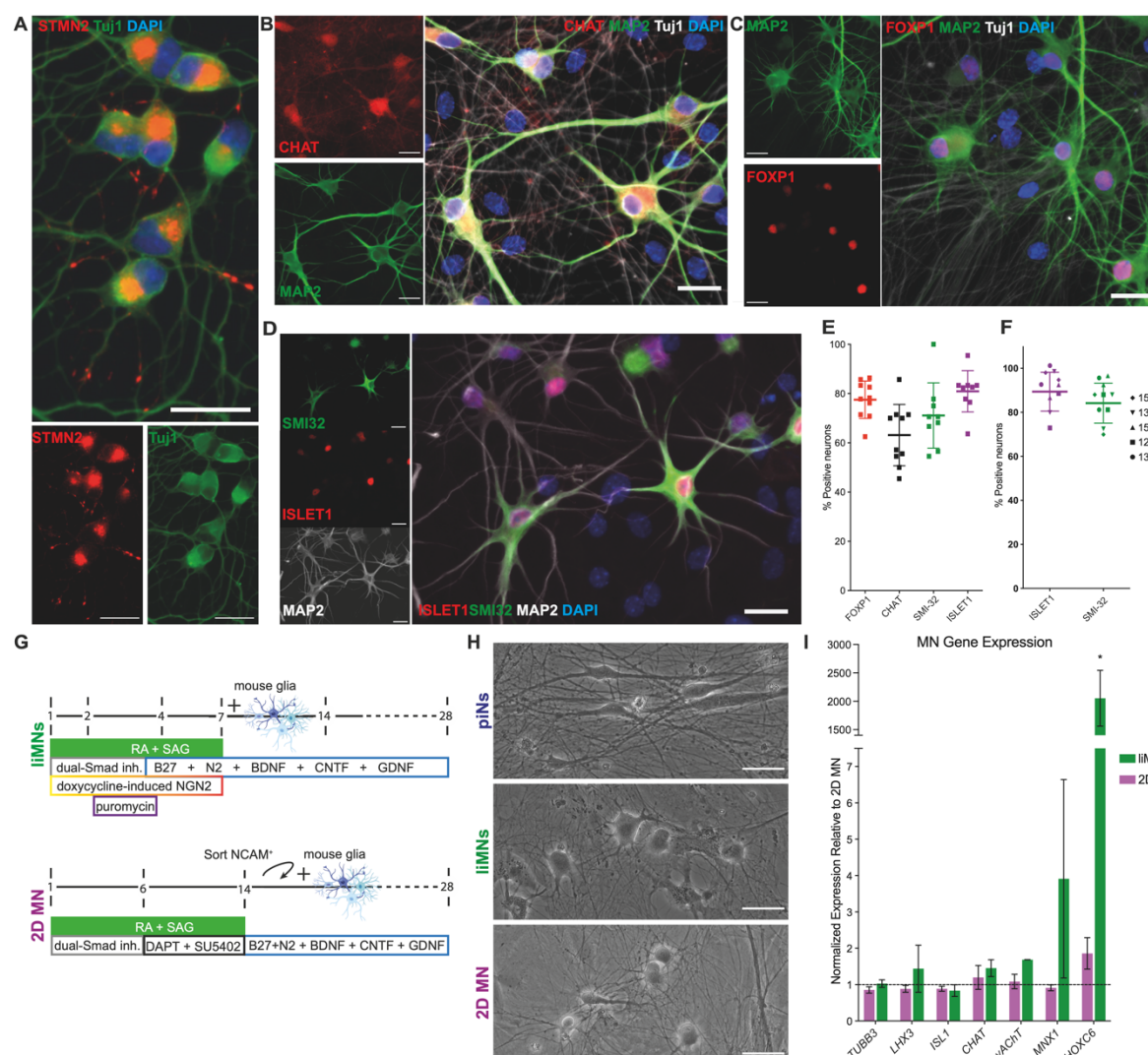


Figure 2 Patterned Ngn2-induced neuronal fate is maintained throughout the differentiation (A) Differentiation schemes for neuronal maturation after one-week of patterning: Dox – original Zhang et al. 2013; LSB – Ngn2 with dual-Smad inhibition; piNs – cortical-like piNs (Nehme et al. 2018); liMNs – lower induced Motor Neurons. (B) Brightfield image at day 30 of piNs and liMNs (scale bar 100 μ m). (C) Diagram of genes specifically expressed in either anterior-dorsal cortical neurons or ventro-caudal, spinal cord motor neurons. (D) RT-qPCR quantification for induction of cortical genes at day 30 (four cell lines in $n=3$ technical replicates, one-way ANOVA). (E) RT-qPCR quantification for spinal cord genes at day 28 (four cell lines in $n=3$ technical replicates, one-way ANOVA). (F) Hb9::GFP reporter expression at day 14 post-induction in piNs and liMNs (scale bar 50 μ m). (G) Quantification of Hb9::GFP reporter expression at day 7, 10 and 14 post-induction in piNs (blue) and liMNs (green) by immunofluorescence ($n=5$, p-values from t-test at each time point). (H) IF analysis for pan-MN SMI-32, Islet1 and Hb9::GFP reporter expression at day 7 post-induction (scale bar 50 μ m). (I) Quantification H ($n=3$ replicates).

liMNs reproducibly express canonical pan-Motor Neuron markers and resemble bona fide hPSC-MN

Given that neuralisation by Ngn2 overexpression can be directed to different neuronal fates and maintained during *in vitro* culture, we wanted to confirm the expression of key motor neuron markers at the protein level. As early as day 14, liMNs expressed the ventral horn motor neuron specific marker Stathmin2 (STMN2) (Figure 3E)^{33,34}. By day 30, liMNs expressed Cholinergic Acetyltransferase (ChAT) (Figure 3B) and limb-innervating marker Foxp1 (Figure 3C)³⁵. Moreover, liMNs showed reactivity for antibodies against the transcription factor Islet1 along with SMI-32, that recognises spinal MN-enriched neurofilament heavy chain (Figure 3D). Indeed, 60-90% of cells express at least one of these markers (figure 3E), while MN markers were robustly and reproducibly expressed by 80-90% of cells by different cell lines (Figure 3F).

We next wanted to confirm that liMNs resembled cells defined by the scientific community as bona fide hiPSC-derived motor neurons. We thus differentiated MNs following a conventional, widely used method using just small molecule patterning factors (2D MN)³⁶. Briefly, stem cells were subjected to neuralising dual-Smad inhibition followed by DAPT and SU5402 while caudalised and ventralised with RA and SAG. Differentiated neurons were separated from the mixed cultures by sorting for cell surface marker N-CAM 14 days post-neuronal induction³⁶, and then cultured in neuronal differentiation media, under similar conditions to liMNs for 14 more days (Figure 3G, Figure S3A-B). We then compared the morphologies of the conventional 2D MNs and liMNs by imaging. We found that liMNs were morphologically similar to 2D MN, with large multipolar cell bodies, and very distinct from cortical cells (Figure 3H). Moreover, liMNs and 2D MN expressed similar patterns of pan-MN staining (Figure S3C-D). Remarkably, RT-qPCR analysis revealed that liMNs expressed comparable levels of other motor-neuron markers and even higher transcript levels of limb-innervating motor neurons marker *HOXC6* (Figure 3I). These results confirmed that liMNs resemble one kind of bona fide hiPSC-derived motor neurons defined by the broader scientific community.



liMNs form active synaptic networks and contact muscle cells *in vitro*

We next set out to assess liMNs functional properties and ability to form synapses. liMNs expressed both pre- and post-synaptic molecules Synaptophysin and PSD-95 (Figure 4A, Figure S4A-B) and displayed abundant staining for Synapsin and axonal AnkyrinG, similarly to piNs (Figure S4C). Multielectrode arrays (MEAs) analyses showed that cultures have a steady increase in spiking rates over time (Figure S4D-E). Treating cells with potassium-gated channel opener Retigabine, a potential therapeutic agent for ALS^{37,38}, silenced cultures underlining the usefulness of liMNs as model for therapeutic strategies in neurodegenerative diseases (Figure 4B).

MNs are the only neurons to connect with muscles through a highly specific synapse: the NMJ. To test the ability of liMNs to form NMJ-like structures we established co-cultures with murine muscle cells in compartmentalised microfluidic devices where neurons grown in one chamber can extend axons through groves that connect to muscle cells (Figure 4C). Staining showed that liMNs extended neurites to the second chamber, contact muscle cells and form structures expressing pre-synaptic protein Synapsin (Figure 4Dⁱ-Dⁱⁱ and S5A-E), a sign of an early development of contact.

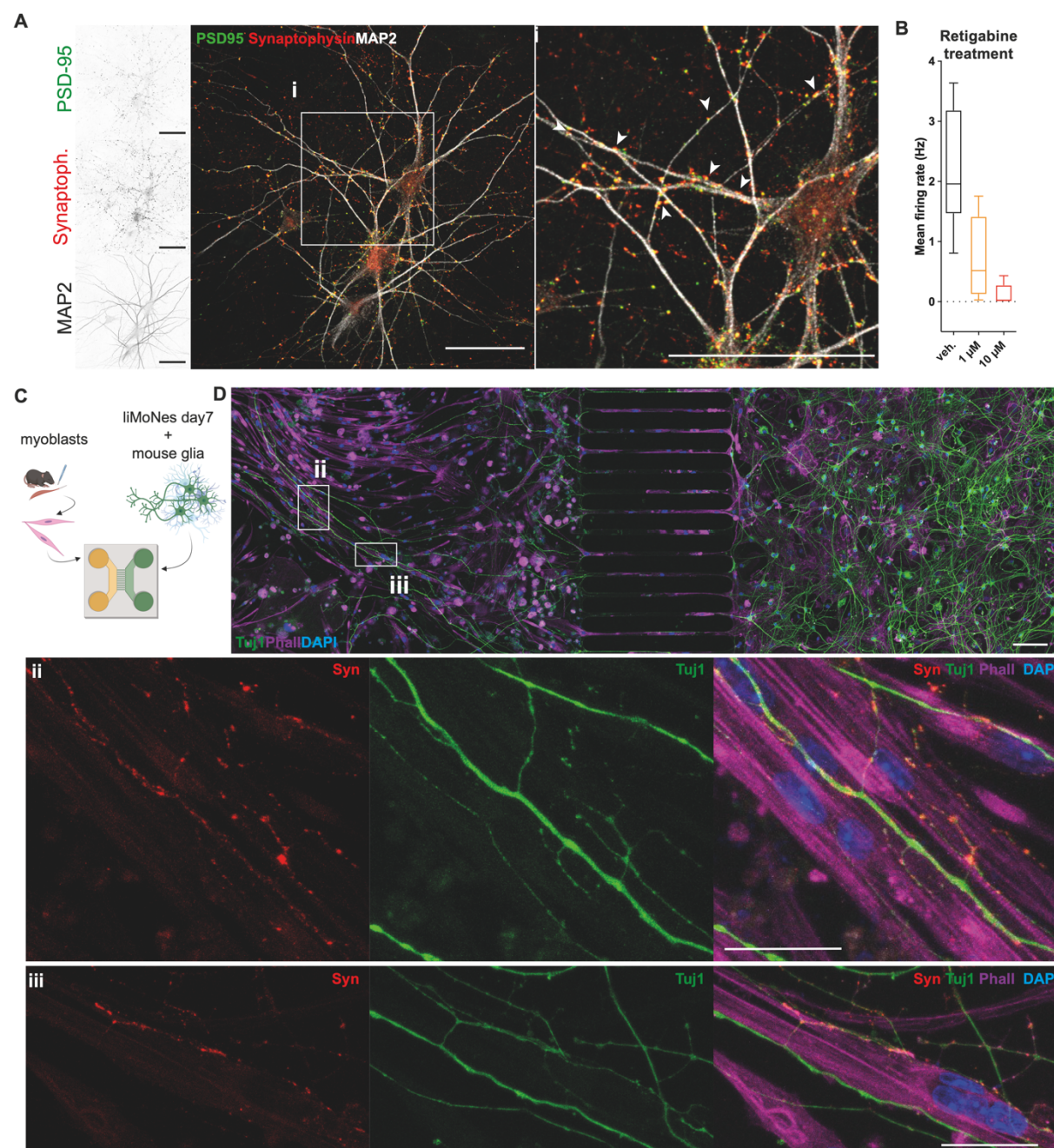


Figure 4 liMoNes can form active synaptic structures *in vitro* (A) Day50 liMNs express pre- and post-synaptic density proteins (scale bar 50 μ m). (B) Mean number of spikes in day50 cultures treated with raising concentrations of Retigabine ($n=6$). (C) Diagram of co-culture experiments of liMNs and primary murine myoblasts in microfluidic devices. (D) Immunofluorescence of co-culture of liMNs and primary murine myoblasts showing glia-liMNs co-cultures (right), where neurons extend axons through the channels (middle), contacting primary muscle cells (left). (Dⁱ-Dⁱⁱ) Insets of (D) showing liMNs forming synaptic-like contacts with muscles cells (scale bar 50 μ m).

scRNA-seq confirms expression of MN-specific genes and reproducibility of the protocol

After confirming the MN-like properties of liMNs, we set out to further characterise their molecular identity and reproducibility by single cell RNA sequencing. We coupled sequencing with two newly developed technologies: *Census-seq* and *Dropulation*^{39,40} to enable the characterization of lines from many different donors in a single experiment. These methods utilise the intrinsic variability of single nucleotide polymorphisms (SNPs) within a population as a barcode to assign identities in a mixed culture - a “village” - of multiple donors, similarly to pooled CRISPR-Cas9 barcoded screens⁴¹⁻⁴³. More precisely, *Census-seq* allows population-scale, quantitative identity assignment from a mixed group of donors³⁹, *Dropulation* can assign identities at a single cell level in a “village” for scRNAseq studies⁴⁰. With this aim in mind, we produced liMNs “villages”: 50 embryonic stem cell lines, previously subjected to whole-genome sequencing, were separately differentiated into liMNs. At day 7 post-induction, postmitotic cells were pooled in equal numbers to make up “villages” containing all donors in one dish. Using genotypes from WGS we were able to reassign the donor identities in a mixed village (Figure 5A).

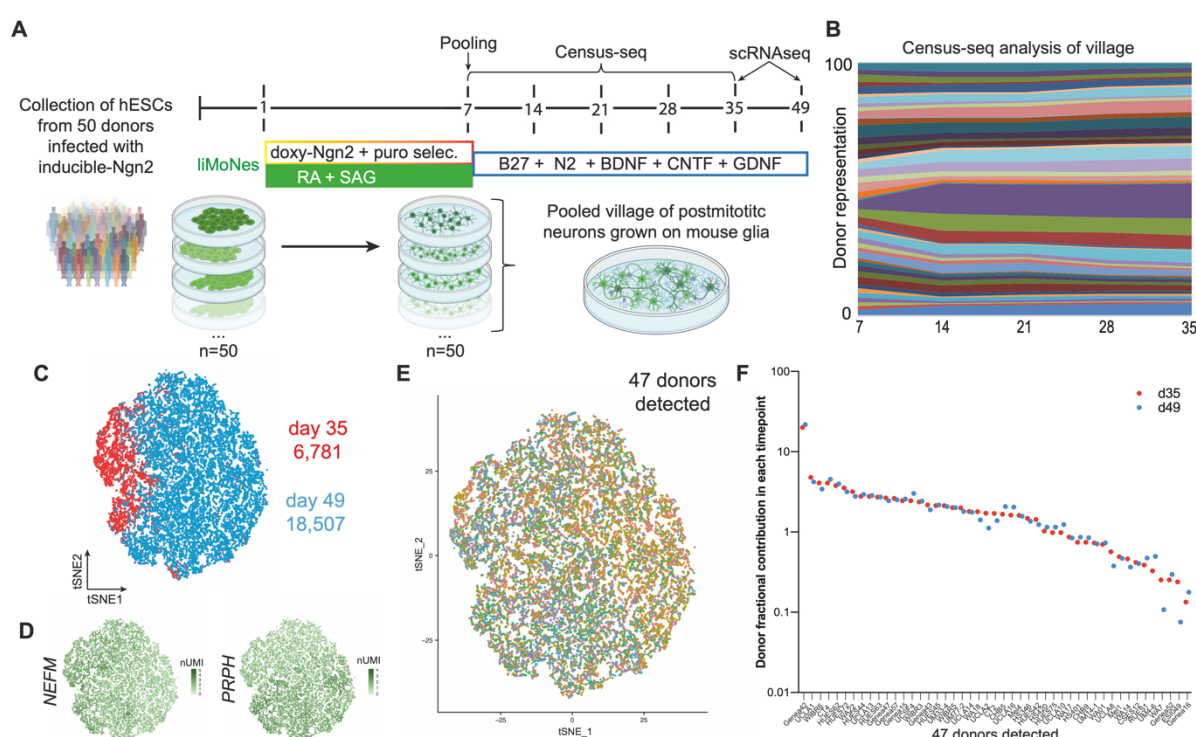


Figure 5 scRNA-seq confirms expression of MN-specific genes and reproducibility of the protocol **(A)** Pooling strategy and village construction for *Census-seq* and *Dropulation* analysis. **(B)** Sandplot of *Census-seq* analysis showing balanced representation of 47 detected donors throughout several days post-induction. **(C)** t-SNE projection of scRNAseq analysis of 25,288 cells of two timepoints of mature liMNs differentiation. **(D)** t-SNE projection with expression of markers for neurons of the peripheral nervous system. **(E)** t-SNE projection of 25,288 cells depicting donor's identity of each cell from 47 donors detected by *Dropulation* analysis. **(F)** Fraction representation of 47 donors in the two timepoints of mature liMNs differentiation.

To ensure that the donor composition remained balanced, cells were harvested once a week to collect genomic DNA for low-coverage sequencing. *Census-seq* analyses showed that we could detect 47 of the 50 donors originally pooled and confirmed that donor distribution remained consistent for four weeks (Figure 5B). Neurons were harvested at day 35 and day 49 for scRNA-seq and *Dropulation* analysis. Libraries generated from 25,288 cells demonstrated strong expression of neuronal markers, especially of the peripheral nervous system (PNS), *NEFM* and *PRPH* (Figure 5C-D and S6A). liMNs did not express cycling cells markers (Figure S6B), nor markers of ventral, spinal interneuronal pools V1, V2a, V2b, V3 nor mid-dorsal spinal interneurons V0 (Figure S6C-E). liMNs expressed MN-enriched *STMN2*, *NEFH*, *ISL1* and *MNX1* (Figure S6F)^{19,33,34,44} and low but detectable expression of cholinergic genes *ACHE*, *SLC5A7* (Cht1), *SLC18A3* (vAChT) (Figure S6G). Finally, we detected expression of *AGRN* and *NRG1*, expressed by MNs to form NMJs (Figure S6H).

Using the newly devised *Dropulation* analytical pipeline, we assigned donor identity to barcoded droplets. Initial *t*-SNE clustering showed an even distribution of each donor (Figure 5E) and we confirmed that the contribution of each donor remained constant at both timepoints (Figure 5F) underlying the robustness and reproducibility of the protocol. We therefore confirmed that our protocol can reproducibly generate MN-like cells from many cell lines.

Cell villages confirm polarization generated by differential patterning of *Ngn2* differentiation

To unbiasedly confirm that differential patterning strategies could generate different neuronal fates we then compared single cell libraries from liMNs to libraries similarly generated from piNs (Figure S7A). *t*-SNE clustering showed a clear separation of piNs and liMNs (Figure 6A). All cells expressed neuronal markers (Figure 6B and S7B) but piNs expressed higher levels of genes of dorsal, cortical and glutamatergic cells (Figure 6C and S7C), whereas liMNs expressed higher levels of genes of ventral, spinal and cholinergic cells (Figure 6D and S7D), confirming that the two different patterning strategies preferentially upregulate genes connected to these distinct cellular identities in a strongly polarised manner (Figure 6E). Interestingly, *HOX* genes, mostly expressed in the midbrain and in the spinal cord and known markers of caudalisation, were highly expressed in liMNs and barely detected in piNs (Figure 6F-G). We assigned donor identity to barcoded droplets with *Dropulation* and showed an even distribution of each donor across the different clusters (Figure S7E-F) underlying the robustness and reproducibility of these protocols.

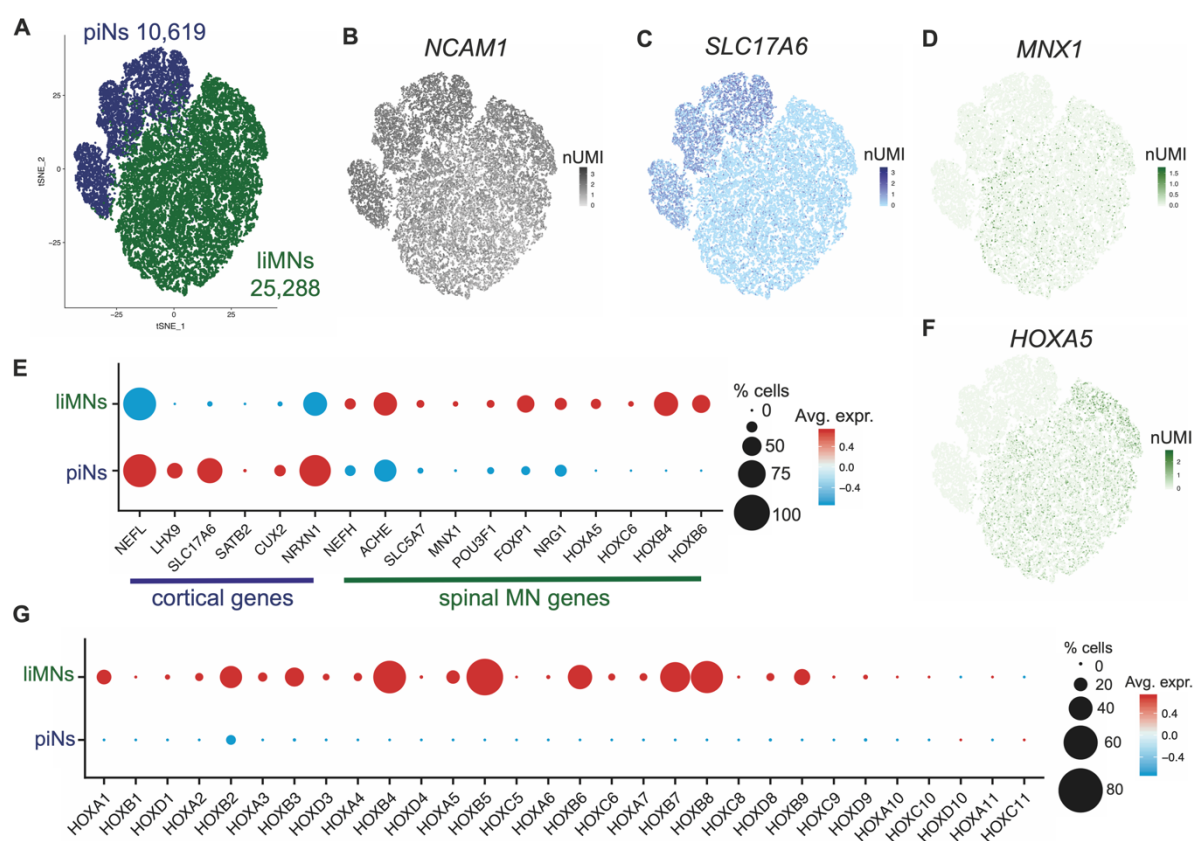


Figure 6 Confirmed divergent neuronal fate of piNs and liMNs (A) *t*-SNE projection of scRNAseq analysis of 25,288 cells of two timepoints of piNs and liMNs differentiation. (B) *t*-SNE projection with expression of neuronal marker. (C) *t*-SNE projection with expression of cortical-enriched marker. (D) *t*-SNE projection with expression of MN-specific marker. (E) Dotplot for differential gene expression of markers specific to either cortical excitatory neurons or spinal MNs. (F) *t*-SNE projection with expression of brachial MN-specific HOX gene expression. (G) Dotplot for gene expression of all retinoid-dependent HOX genes in piNs and liMNs.

Ventro-caudal patterning of *Ngn2* can produce different MN subtypes

In vivo motor neurons are classified in subtypes (a.k.a. pools or columns) according to their position along the cord and the anatomical part of the body they innervate. Four groups lie in spinal cord areas developmentally regulated by retinoids: 1. Medial Motor Column (MMC), along the entire spine, connects to axial musculature to maintain posture, 2. cervical Spinal Accessory Column (SAC) innervates head and neck, 3. Phrenic Motor Column (PMC), also cervical, innervates the diaphragm, 4. Lateral Motor Column (LMC), at brachial level on the cervico-thoracic boundary, connects to forelimbs and is divided in ventral-innervating, medial or dorsal-innervating, lateral LMC (Figure 7A)¹⁶. Remarkably, we were able to find markers specific to these pools in our dataset: a small group of *PHOX2B*-expressing SAC-like cells, wide expression of PMC-enriched *ALCAM* and *POU3F1* (*SCIP*) and markers of both lateral- and medial-LMCs: *FOXP1* and *LHX1* (Figure S8A-D).

We wondered if the discrete expression of these markers shaped subgroups with different transcriptomic profiles. We decided to unbiasedly identify subclusters and found four groups: liMNs 0,1,2,3 (Figure 7B). Intriguingly, markers of MN pools segregated within the

groups demarcating an *ALCAM*⁺ group, an *LHX1*⁺ and a *FOXP1*⁺ groups, and a small *PHOX2B*⁺ group (Figure 7C and S8E-G). No expression of MMC markers was found (Figure S8H) consistent with reports identifying this population as less responsive to certain patterning factors^{19,45}. Differential genes expression analysis for genes specifically expressed in each subgroup unbiasedly confirmed regional specification consistent with the markers described above (Figure S8I, Table S1). We observed two additional features: expression of markers of anterior digit-innervating MNs *FIGN* and *CPNE4* in a small percentage of cells (Figure S8J)^{46,47}; and expression of *HOX* genes activated in response to retinoids⁴⁸ and specifically in cervical/brachial MNs⁴⁹ (Figure S8K-L). Taking advantage of the *Dropulation* technology, we investigated the distribution of donors within each subcluster and surprisingly found that each of the 47 donors distributed evenly within clusters highlighting the robustness and reproducibility of the protocol (Figure 7D and S9A-B).

To ensure that liMNs resembled cervico-brachial MNs, we integrated our data with a recently published scRNA-seq dataset generated from human embryonic spinal cord⁴⁴, and visualised the resulting dataset using UMAP (Uniform Manifold Approximation and Projection). First, we confirmed we could identify neurons and progenitors of different spinal lineages matching the cell types identified in Rayon et al. (Figure S10A-D). In the integrated analysis, liMoNes clustered closely to embryonic post-mitotic MNs (MNs) (Figure 7E), while they clustered separately from both sensory neurons and dorsal interneurons (Figure 7F) and partially closer to ventral interneurons (Figure 7G and S10E), further validating the MN-like fate of liMNs. We then isolated MN-like cells from the integrated dataset and analysed them separately from the rest of the spinal cord, liMNs and primary MN clustered separately from progenitor cells (pMNs) (Figure 7H). Consistent with *HOX* expression, liMNs clustered more closely to MN of brachial origin, consistent with the more caudal position of samples in the primary human dataset and therefore low expression of more hindbrain markers (Figure S10F-G), and from mid-to-late stages of development (Figure 7I). Taken together, our integrative analyses with human embryonic spinal cord cells not only confirms the MN identity of liMoNes, but also demonstrates that are composed of a plethora of motor neuron subtypes that intrinsically recapitulates pools and columns identified in the cervical and brachial spinal cord and that these subtypes can be robustly generated in a myriad of cells lines.

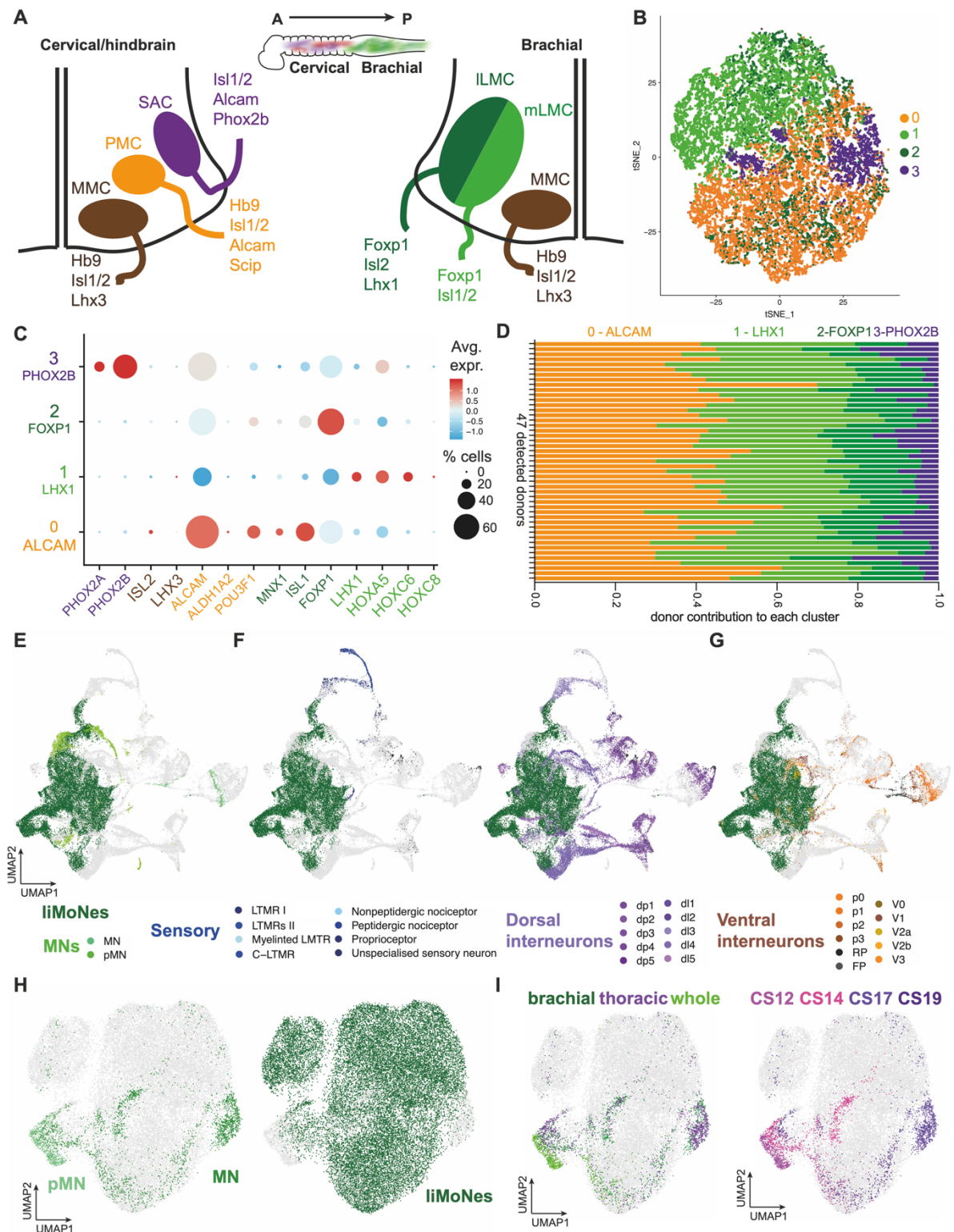


Figure 7 Ventro-caudal patterning of NGN2 can produce different MN subtypes. (A) Diagram of known pools of MN subtypes along mammalian spinal cord. (B) t-SNE projection of four, unbiasedly identified subclusters in the 25,288 cells analysed. (C) Dotplot for differential gene expression of MN subtype-specific markers in the four cervico-brachial MN groups. (D) Fraction of each donor's share between the identified subclusters as calculated by *Dropulation*. (E) t-SNE projection of integrated datasets: liMoNes and MNs and pMNs (progenitors) from human embryonic spinal cord Rayon et al. 2021. (F) t-SNE projection of integrated datasets: liMoNes, sensory neurons and dorsal interneurons from Rayon et al. 2021. (G) t-SNE projection of integrated datasets: liMoNes and ventral interneurons from Rayon et al. 2021. (H) t-SNE projection of integrated datasets with MNs only. (I) t-SNE projection of integrated datasets with MNs only with regionality and timepoints (Carnegie Stage) from Rayon et al. highlighted.

DISCUSSION

In this study we describe a rapid and efficient protocol to generate human MN-like cells from hPSCs by combining the overexpression of neuralising factor Ngn2⁹ and ventralising and caudalising small molecules patterning in human ESC/iPSCs^{28,50}. We demonstrated that different patterning molecules can direct Ngn2-driven neuralisation into the specification of distinct neuronal fates that are maintained during *in vitro* culture. In particular, we show that ventral-caudal patterning induces expression of the MN-specific TF *MX1*/Hb9 in >90% of differentiated cells bypassing the previously used sorting methods to isolate MN from mixed cultures and in a shorter period of time. The ventro-caudalised cells expressed pan-MN markers as identified *in vivo* and resembled bona fide hPSC-derived MN giving them a lower motor neuron identity - hence lower induced Motor Neuron (liMoNes/liMNs). liMNs generated electrophysiologically active cultures capable to form early contact points with muscle cells *in vitro*. By leveraging newly developed single-cell RNA-sequencing analyses tools, we demonstrated that this protocol could successfully generate a previously reported hard-to-produce neuronal cell type by a straightforward one-step programming. Additionally, we showed that the differentiation scheme is highly scalable and reproducible across 47 cell lines, and that the generated cultures contain a diverse population of disease-relevant MN subtypes that closely in part resemble their human, embryonic, *in vivo* counterpart.

The protocol described here enabled us to overcome some of the main issues reported in previously published differentiation schemes based on small molecules patterning. Specifically, we showed how with a single step induction, we were able to generate in only seven days, a pure population of post-mitotic neurons in which virtually all cells expressed the MN-specific marker *MX1*/Hb9, whereas most protocols reported at least two weeks of differentiation to achieve partial expression of this reporter¹⁷. Moreover, we demonstrated how the enforced expression of one transcription factor can achieve complete neuralisation of cells to avoid the heterogeneous generation of other cell types on a cell-line-to-cell-line dependent manner⁴ and how this method could be replicated in dozens of pluripotent lines. This single-step, 7-day induction protocol would allow the generation of defined motor neuron cultures for *in vitro* modelling studies and avoid time-consuming and expensive cell-sorting step to select relevant cell types from mixed ventral-caudal populations³⁶. Intriguingly, very few reports showed NMJ-like structures *in vitro* from human MNs⁵¹⁻⁵⁶ and so far only one has established a system that allows it in culture conditions that resemble human physiology⁵⁷. The combination of our highly pure, accelerated protocol and this report could allow further understanding of NMJs in a physiologically relevant, human context.

Our study is among the first reports to highlight the malleability of Ngn2-based reprogramming and its ability to be directed to differential states by small molecules patterning mimicking embryonic development. We have thoroughly demonstrated in a previous report

that patterning can direct Ngn2 towards a cortical-like state¹¹, but this is the first side-by-side, systematic comparison of the ability of this programming method to diverge into different neuronal fates. Others have reported that overexpression of Ngn2 alone is able to produce an admixture of different neuronal subtypes of both the central and peripheral nervous system¹⁰, confirming that Ngn2-driven neuralisation yields several neuronal subtypes. Here, we expand on this biology showing that small molecule patterning can direct the multipotent neuralising ability of Nuerogenin2 to populations of regionally specified neurons in a robust, reproducible manner.

Many molecular studies have shown how retinoids can specifically act as epigenetic modulators, open chromatin domains in neural progenitor cells consistent with spinal cord identity and aid posteriorisation in MN differentiation systems¹³. Moreover, transcriptomic and epigenomic studies along NIL-based MN differentiation have shown that Ngn2 acts independently of the Isl1-Lhx3 heterodimers, upregulating neuralising factors that in turn open sites of chromatin that allow further specification into MN-fates⁵⁸. Intriguingly, others have reported that overexpression of Ngn2 in fibroblasts coupled with patterning factors could generate small populations of cholinergic neurons, hinting at the malleability of this system⁵⁹. We speculate that the addition of patterning molecules to Ngn2-programming permits the opening of chromatin at sites of MN-specific genes usually achieved by the overexpression of other TFs forming a permissive epigenetic landscape that allows specification into motor neuron identity. Other groups have reported that in other TF-based differentiation systems, addition of RA can upregulate sets of genes that the TFs alone could not achieve⁶⁰, confirming that combinatorial approaches might aid specification into desired cell types.

The use of only one transcription factor combined with small quantities of inexpensive patterning molecules renders this protocol amenable to large-scale, high-throughput studies compared to previous studies¹². The combinatorial use of multiple TFs often induces the generation of extremely specified subtypes of MNs^{22,30} that, even though pure and well-defined, limit the ability of hPSC to differentiate into the intrinsic admixture of MNs generated by retinoids/Shh and only elicits the transcriptomic programs of restricted pools³⁰. Moreover, others have demonstrated how combinations of multiple transcription factors might take longer time to develop hPSC into neurons when compared to Ngn2 alone and that the timing of overexpression could interfere with the subtypes of neurons generated⁶⁰. Here we propose that a short pulse of Ngn2 overexpression coupled with patterning molecules not only reduces the number of TF needed to direct the specification of neuralisation but also allows intrinsic developmental processes to take place and generate myriad MN subtypes seen in spinal cord development, as shown by the similarities with our cells and primary samples. Given the differential susceptibility of subtypes of MNs to degenerate in certain diseases like ALS⁶¹,

having both resistant and susceptible populations of MNs reproducibly generated in one dish could help to further understand the dynamic process of neurodegeneration.

Limitations of this study

One of the strengths of this protocol is its high reproducibility and accelerated nature. However, the method still lacks pivotal positional, geographical, sequentially timed signals that generate the milieu of motoneurons in the spinal cord. This method thus produces MNs that do not exactly reproduce transcriptomic profiles of columns *in vivo*, for example the incomplete co-expression of HB9 and Islet1 in young neurons that is only achieved later into the differentiation, the discrepancy between liMNs identities at protein and RNA-levels, or the incomplete overlap of certain markers between liMNs and their *in vivo* counterparts. Considering different concentrations and timing of patterning molecules and also exploring RA-independent ways of generating MNs^{45,62}, would be an important follow-up to this study. Furthermore, our microfluidic system did not show clustering of postsynaptic acetylcholine receptors (AChR) on muscle cells. Formation of mature NMJ contacts has been a primary limitation of *in vitro* hPSC-derived MNs and muscle co-cultures as observed *in vivo*⁶³. Towards optimization of this model, a recent study has shown that supplementation of agrin and laminin increased clustering of AChR in *in vitro* human co-cultures system⁵⁷, adapting this system to our protocol might provide essential steps for the further maturation of these synaptic structures for liMNs. These discrepancies are shortcomings of accelerated systems like ours, nonetheless, our approach provides a platform for the study of the biology of different MN subtypes and their functionality in health and disease in a scalable, highly reproducible manner never achieved before.

REFERENCES

1. Penney, J., Ralvenius, W.T., and Tsai, L.H. (2020). Modeling Alzheimer's disease with iPSC-derived brain cells. *Mol Psychiatry* 25, 148-167. 10.1038/s41380-019-0468-3.
2. Chambers, S.M., Fasano, C.A., Papapetrou, E.P., Tomishima, M., Sadelain, M., and Studer, L. (2009). Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. *Nat Biotechnol* 27, 275-280. 10.1038/nbt.1529.
3. Chambers, S.M., Qi, Y., Mica, Y., Lee, G., Zhang, X.J., Niu, L., Bilsland, J., Cao, L., Stevens, E., Whiting, P., et al. (2012). Combined small-molecule inhibition accelerates developmental timing and converts human pluripotent stem cells into nociceptors. *Nat Biotechnol* 30, 715-720. 10.1038/nbt.2249.
4. Boulting, G.L., Kiskinis, E., Croft, G.F., Amoroso, M.W., Oakley, D.H., Wainger, B.J., Williams, D.J., Kahler, D.J., Yamaki, M., Davidow, L., et al. (2011). A functionally characterized test set of human induced pluripotent stem cells. *Nat Biotechnol* 29, 279-286. 10.1038/nbt.1783.
5. Kampmann, M. (2020). CRISPR-based functional genomics for neurological disease. *Nat Rev Neurol* 16, 465-480. 10.1038/s41582-020-0373-z.
6. Mertens, J., Marchetto, M.C., Bardy, C., and Gage, F.H. (2016). Evaluating cell reprogramming, differentiation and conversion technologies in neuroscience. *Nat Rev Neurosci* 17, 424-437. 10.1038/nrn.2016.46.

7. Vierbuchen, T., Ostermeier, A., Pang, Z.P., Kokubu, Y., Sudhof, T.C., and Wernig, M. (2010). Direct conversion of fibroblasts to functional neurons by defined factors. *Nature* 463, 1035-1041. 10.1038/nature08797.
8. Son, E.Y., Ichida, J.K., Wainger, B.J., Toma, J.S., Rafuse, V.F., Woolf, C.J., and Eggan, K. (2011). Conversion of mouse and human fibroblasts into functional spinal motor neurons. *Cell Stem Cell* 9, 205-218. 10.1016/j.stem.2011.07.014.
9. Zhang, Y., Pak, C., Han, Y., Ahlenius, H., Zhang, Z., Chanda, S., Marro, S., Patzke, C., Acuna, C., Covy, J., et al. (2013). Rapid single-step induction of functional neurons from human pluripotent stem cells. *Neuron* 78, 785-798. 10.1016/j.neuron.2013.05.029.
10. Lin, H.C., He, Z., Ebert, S., Schornig, M., Santel, M., Nikolova, M.T., Weigert, A., Hevers, W., Kasri, N.N., Taverna, E., et al. (2021). NGN2 induces diverse neuron types from human pluripotency. *Stem Cell Reports*. 10.1016/j.stemcr.2021.07.006.
11. Nehme, R., Zuccaro, E., Ghosh, S.D., Li, C., Sherwood, J.L., Pietilainen, O., Barrett, L.E., Limone, F., Worringer, K.A., Kommineni, S., et al. (2018). Combining NGN2 Programming with Developmental Patterning Generates Human Excitatory Neurons with NMDAR-Mediated Synaptic Transmission. *Cell Rep* 23, 2509-2523. 10.1016/j.celrep.2018.04.066.
12. Hester, M.E., Murtha, M.J., Song, S., Rao, M., Miranda, C.J., Meyer, K., Tian, J., Boulting, G., Schaffer, D.V., Zhu, M.X., et al. (2011). Rapid and efficient generation of functional motor neurons from human pluripotent stem cells using gene delivered transcription factor codes. *Mol Ther* 19, 1905-1912. 10.1038/mt.2011.135.
13. Mazzoni, E.O., Mahony, S., Peljto, M., Patel, T., Thornton, S.R., McCuine, S., Reeder, C., Boyer, L.A., Young, R.A., Gifford, D.K., and Wichterle, H. (2013). Saltatory remodeling of Hox chromatin in response to rostrocaudal patterning signals. *Nat Neurosci* 16, 1191-1198. 10.1038/nn.3490.
14. Han, S.S., Williams, L.A., and Eggan, K.C. (2011). Constructing and deconstructing stem cell models of neurological disease. *Neuron* 70, 626-644. 10.1016/j.neuron.2011.05.003.
15. Jessell, T.M. (2000). Neuronal specification in the spinal cord: inductive signals and transcriptional codes. *Nat Rev Genet* 1, 20-29. 10.1038/35049541.
16. Stifani, N. (2014). Motor neurons and the generation of spinal motor neuron diversity. *Front Cell Neurosci* 8, 293. 10.3389/fncel.2014.00293.
17. Sances, S., Bruijn, L.I., Chandran, S., Eggan, K., Ho, R., Klim, J.R., Livesey, M.R., Lowry, E., Macklis, J.D., Rushton, D., et al. (2016). Modeling ALS with motor neurons derived from human induced pluripotent stem cells. *Nat Neurosci* 19, 542-553. 10.1038/nn.4273.
18. Davis-Dusenbery, B.N., Williams, L.A., Klim, J.R., and Eggan, K. (2014). How to make spinal motor neurons. *Development* 141, 491-501. 10.1242/dev.097410.
19. Amoroso, M.W., Croft, G.F., Williams, D.J., O'Keeffe, S., Carrasco, M.A., Davis, A.R., Roybon, L., Oakley, D.H., Maniatis, T., Henderson, C.E., and Wichterle, H. (2013). Accelerated high-yield generation of limb-innervating motor neurons from human stem cells. *J Neurosci* 33, 574-586. 10.1523/JNEUROSCI.0906-12.2013.
20. Dimos, J.T., Rodolfa, K.T., Niakan, K.K., Weisenthal, L.M., Mitsumoto, H., Chung, W., Croft, G.F., Saphier, G., Leibel, R., Golland, R., et al. (2008). Induced pluripotent stem cells generated from patients with ALS can be differentiated into motor neurons. *Science* 321, 1218-1221. 10.1126/science.1158799.
21. Kiskinis, E., Sandoe, J., Williams, L.A., Boulting, G.L., Moccia, R., Wainger, B.J., Han, S., Peng, T., Thams, S., Mikkilineni, S., et al. (2014). Pathways disrupted in human ALS motor neurons identified through genetic correction of mutant SOD1. *Cell Stem Cell* 14, 781-795. 10.1016/j.stem.2014.03.004.
22. De Santis, R., Garone, M.G., Pagani, F., de Turris, V., Di Angelantonio, S., and Rosa, A. (2018). Direct conversion of human pluripotent stem cells into cranial motor neurons using a piggyBac vector. *Stem Cell Res* 29, 189-196. 10.1016/j.scr.2018.04.012.
23. Pfaff, S.L., Mendelsohn, M., Stewart, C.L., Edlund, T., and Jessell, T.M. (1996). Requirement for LIM homeobox gene *Isl1* in motor neuron generation reveals a motor neuron-dependent step in interneuron differentiation. *Cell* 84, 309-320. 10.1016/s0092-8674(00)80985-x.
24. Arber, S., Han, B., Mendelsohn, M., Smith, M., Jessell, T.M., and Sockanathan, S. (1999). Requirement for the homeobox gene *Hb9* in the consolidation of motor neuron identity. *Neuron* 23, 659-674. 10.1016/s0896-6273(01)80026-x.
25. Muhr, J., Graziano, E., Wilson, S., Jessell, T.M., and Edlund, T. (1999). Convergent inductive signals specify midbrain, hindbrain, and spinal cord identity in gastrula stage chick embryos. *Neuron* 23, 689-702. 10.1016/s0896-6273(01)80028-3.
26. Di Giorgio, F.P., Carrasco, M.A., Siao, M.C., Maniatis, T., and Eggan, K. (2007). Non-cell autonomous effect of glia on motor neurons in an embryonic stem cell-based ALS model. *Nat Neurosci* 10, 608-614. 10.1038/nn1885.
27. Wichterle, H., Lieberam, I., Porter, J.A., and Jessell, T.M. (2002). Directed differentiation of embryonic stem cells into motor neurons. *Cell* 110, 385-397. 10.1016/s0092-8674(02)00835-8.

28. Di Giorgio, F.P., Boulting, G.L., Bobrowicz, S., and Eggan, K.C. (2008). Human embryonic stem cell-derived motor neurons are sensitive to the toxic effect of glial cells carrying an ALS-causing mutation. *Cell Stem Cell* 3, 637-648. 10.1016/j.stem.2008.09.017.
29. Machado, C.B., Kanning, K.C., Kreis, P., Stevenson, D., Crossley, M., Nowak, M., Iacovino, M., Kyba, M., Chambers, D., Blanc, E., and Lieberam, I. (2014). Reconstruction of phrenic neuron identity in embryonic stem cell-derived motor neurons. *Development* 141, 784-794. 10.1242/dev.097188.
30. Mazzoni, E.O., Mahony, S., Closser, M., Morrison, C.A., Nedelec, S., Williams, D.J., An, D., Gifford, D.K., and Wichterle, H. (2013). Synergistic binding of transcription factors to cell-specific enhancers programs motor neuron identity. *Nat Neurosci* 16, 1219-1227. 10.1038/nn.3467.
31. Abdel-Maguid, T.E., and Bowsher, D. (1984). Classification of neurons by dendritic branching pattern. A categorisation based on Golgi impregnation of spinal and cranial somatic and visceral afferent and efferent cells in the adult human. *J Anat* 138 (Pt 4), 689-702.
32. Thaler, J.P., Koo, S.J., Kania, A., Lettieri, K., Andrews, S., Cox, C., Jessell, T.M., and Pfaff, S.L. (2004). A postmitotic role for Isl-class LIM homeodomain proteins in the assignment of visceral spinal motor neuron identity. *Neuron* 41, 337-350. 10.1016/s0896-6273(04)00011-x.
33. Guerra San Juan, I., Nash, L.A., Smith, K.S., Leyton-Jaimes, M.F., Qian, M., Klim, J.R., Limone, F., Dorr, A.B., Couto, A., Pintacuda, G., et al. (2022). Loss of mouse *Stmn2* function causes motor neuropathy. *Neuron*. 10.1016/j.neuron.2022.02.011.
34. Yadav, A., Matson, K.J.E., Li, L., Hua, I., Gaur, P., Alkaslasi, M., Hasan, S., Galuta, A., A., D., Ameri, S., et al. (2022). The Human Motoneuron Expression Signature is Defined by ALS-Related Genes. *bioRxiv [PREPRINT]*. <https://doi.org/10.1101/2022.03.25.485808>.
35. Dasen, J.S., De Camilli, A., Wang, B., Tucker, P.W., and Jessell, T.M. (2008). Hox repertoires for motor neuron diversity and connectivity gated by a single accessory factor, FoxP1. *Cell* 134, 304-316. 10.1016/j.cell.2008.06.019.
36. Klim, J.R., Williams, L.A., Limone, F., Guerra San Juan, I., Davis-Dusenbery, B.N., Mordes, D.A., Burberry, A., Steinbaugh, M.J., Gamage, K.K., Kirchner, R., et al. (2019). ALS-implicated protein TDP-43 sustains levels of STMN2, a mediator of motor neuron growth and repair. *Nat Neurosci* 22, 167-179. 10.1038/s41593-018-0300-4.
37. Wainger, B.J., Kiskinis, E., Mellin, C., Wiskow, O., Han, S.S., Sandoe, J., Perez, N.P., Williams, L.A., Lee, S., Boulting, G., et al. (2014). Intrinsic membrane hyperexcitability of amyotrophic lateral sclerosis patient-derived motor neurons. *Cell Rep* 7, 1-11. 10.1016/j.celrep.2014.03.019.
38. Wainger, B.J., Macklin, E.A., Vucic, S., McIllduff, C.E., Paganoni, S., Maragakis, N.J., Bedlack, R., Goyal, N.A., Rutkove, S.B., Lange, D.J., et al. (2021). Effect of Ezogabine on Cortical and Spinal Motor Neuron Excitability in Amyotrophic Lateral Sclerosis: A Randomized Clinical Trial. *JAMA Neurol* 78, 186-196. 10.1001/jamaneurol.2020.4300.
39. Mitchell, J.M., Nemesh, J., Ghosh, S., Handsaker, R.E., Mello, C.J., Meyer, D., Raghunathan, K., de Rivera, H., Tegtmeyer, M., Hawes, D., et al. (2020). Mapping genetic effects on cellular phenotypes with "cell villages". *bioRxiv [PREPRINT]*. <https://doi.org/10.1101/2020.06.29.174383>.
40. Wells, M.F., Nemesh, J., Ghosh, S., Mitchell, J.M., Mello, C.J., Meyer, D., Raghunathan, K., Tegtmeyer, M., Hawes, D., Neumann, A., et al. (2021). Natural variation in gene expression and Zika virus susceptibility revealed by villages of neural progenitor cells. *bioRxiv [PREPRINT]*. <https://doi.org/10.1101/2021.11.08.467815>
41. Shalem, O., Sanjana, N.E., Hartenian, E., Shi, X., Scott, D.A., Mikkelsen, T., Heckl, D., Ebert, B.L., Root, D.E., Doench, J.G., and Zhang, F. (2014). Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* 343, 84-87. 10.1126/science.1247005.
42. Wang, T., Wei, J.J., Sabatini, D.M., and Lander, E.S. (2014). Genetic screens in human cells using the CRISPR-Cas9 system. *Science* 343, 80-84. 10.1126/science.1246981.
43. Ashoti, A., Limone, F., van Kranenburg, M., Alemany, A., Baak, M., Vivie, J., Piccioni, F., Dijkers, P.F., Creighton, M., Eggan, K., and Geijsen, N. (2022). Considerations and practical implications of performing a phenotypic CRISPR/Cas survival screen. *PLoS One* 17, e0263262. 10.1371/journal.pone.0263262.
44. Rayon, T., Maizels, R.J., Barrington, C., and Briscoe, J. (2021). Single-cell transcriptome profiling of the human developing spinal cord reveals a conserved genetic programme with human-specific features. *Development* 148. 10.1242/dev.199711.
45. Patani, R., Hollins, A.J., Wishart, T.M., Puddifoot, C.A., Alvarez, S., de Lera, A.R., Wyllie, D.J., Compston, D.A., Pedersen, R.A., Gillingwater, T.H., et al. (2011). Retinoid-independent motor neurogenesis from human embryonic stem cells reveals a medial columnar ground state. *Nat Commun* 2, 214. 10.1038/ncomms1216.
46. Mendelsohn, A.I., Dasen, J.S., and Jessell, T.M. (2017). Divergent Hox Coding and Evasion of Retinoid Signaling Specifies Motor Neurons Innervating Digit Muscles. *Neuron* 93, 792-805 e794. 10.1016/j.neuron.2017.01.017.

47. Blum, J.A., Klemm, S., Shadrach, J.L., Gattenplan, K.A., Nakayama, L., Kathiria, A., Hoang, P.T., Gautier, O., Kaltschmidt, J.A., Greenleaf, W.J., and Gitler, A.D. (2021). Single-cell transcriptomic analysis of the adult mouse spinal cord reveals molecular diversity of autonomic and skeletal motor neurons. *Nat Neurosci* 24, 572-583. 10.1038/s41593-020-00795-0.
48. Philippidou, P., and Dasen, J.S. (2013). Hox genes: choreographers in neural development, architects of circuit organization. *Neuron* 80, 12-34. 10.1016/j.neuron.2013.09.020.
49. Dasen, J.S., Tice, B.C., Brenner-Morton, S., and Jessell, T.M. (2005). A Hox regulatory network establishes motor neuron pool identity and target-muscle connectivity. *Cell* 123, 477-491. 10.1016/j.cell.2005.09.009.
50. Li, X.J., Du, Z.W., Zarnowska, E.D., Pankratz, M., Hansen, L.O., Pearce, R.A., and Zhang, S.C. (2005). Specification of motoneurons from human embryonic stem cells. *Nat Biotechnol* 23, 215-221. 10.1038/nbt1063.
51. Rigamonti, A., Repetti, G.G., Sun, C., Price, F.D., Reny, D.C., Rapino, F., Weisinger, K., Benkler, C., Peterson, Q.P., Davidow, L.S., et al. (2016). Large-Scale Production of Mature Neurons from Human Pluripotent Stem Cells in a Three-Dimensional Suspension Culture System. *Stem Cell Reports* 6, 993-1008. 10.1016/j.stemcr.2016.05.010.
52. Demestre, M., Orth, M., Fohr, K.J., Achberger, K., Ludolph, A.C., Liebau, S., and Boeckers, T.M. (2015). Formation and characterisation of neuromuscular junctions between hiPSC derived motoneurons and myotubes. *Stem Cell Res* 15, 328-336. 10.1016/j.scr.2015.07.005.
53. Guo, X., Gonzalez, M., Stancescu, M., Vandenberg, H.H., and Hickman, J.J. (2011). Neuromuscular junction formation between human stem cell-derived motoneurons and human skeletal muscle in a defined system. *Biomaterials* 32, 9602-9611. 10.1016/j.biomaterials.2011.09.014.
54. Lin, C.Y., Yoshida, M., Li, L.T., Ikenaka, A., Oshima, S., Nakagawa, K., Sakurai, H., Matsui, E., Nakahata, T., and Saito, M.K. (2019). iPSC-derived functional human neuromuscular junctions model the pathophysiology of neuromuscular diseases. *JCI Insight* 4. 10.1172/jci.insight.124299.
55. Puttonen, K.A., Ruponen, M., Naumenko, N., Hovatta, O.H., Tavi, P., and Koistinaho, J. (2015). Generation of Functional Neuromuscular Junctions from Human Pluripotent Stem Cell Lines. *Front Cell Neurosci* 9, 473. 10.3389/fncel.2015.00473.
56. Umbach, J.A., Adams, K.L., Gundersen, C.B., and Novitch, B.G. (2012). Functional neuromuscular junctions formed by embryonic stem cell-derived motor neurons. *PLoS One* 7, e36049. 10.1371/journal.pone.0036049.
57. Stoklund Dittlau, K., Krasnow, E.N., Fumagalli, L., Vandoorne, T., Baatsen, P., Kerstens, A., Giacomazzi, G., Pavie, B., Rossaert, E., Beckers, J., et al. (2021). Human motor units in microfluidic devices are impaired by FUS mutations and improved by HDAC6 inhibition. *Stem Cell Reports* 16, 2213-2227. 10.1016/j.stemcr.2021.03.029.
58. Velasco, S., Ibrahim, M.M., Kakumanu, A., Garipier, G., Aydin, B., Al-Sayegh, M.A., Hirsekorn, A., Abdul-Rahman, F., Satija, R., Ohler, U., et al. (2017). A Multi-step Transcriptional and Chromatin State Cascade Underlies Motor Neuron Programming from Embryonic Stem Cells. *Cell Stem Cell* 20, 205-217 e208. 10.1016/j.stem.2016.11.006.
59. Liu, M.L., Zang, T., Zou, Y., Chang, J.C., Gibson, J.R., Huber, K.M., and Zhang, C.L. (2013). Small molecules enable neurogenin 2 to efficiently convert human fibroblasts into cholinergic neurons. *Nat Commun* 4, 2183. 10.1038/ncomms3183.
60. Nickolls, A.R., Lee, M.M., Espinoza, D.F., Szczot, M., Lam, R.M., Wang, Q., Beers, J., Zou, J., Nguyen, M.Q., Solinski, H.J., et al. (2020). Transcriptional Programming of Human Mechanosensory Neuron Subtypes from Pluripotent Stem Cells. *Cell Rep* 30, 932-946 e937. 10.1016/j.celrep.2019.12.062.
61. Nijssen, J., Comley, L.H., and Hedlund, E. (2017). Motor neuron vulnerability and resistance in amyotrophic lateral sclerosis. *Acta Neuropathol* 133, 863-885. 10.1007/s00401-017-1708-8.
62. Maury, Y., Come, J., Piskrowski, R.A., Salah-Mohellibi, N., Chevalleyre, V., Peschanski, M., Martinat, C., and Nedelec, S. (2015). Combinatorial analysis of developmental cues efficiently converts human pluripotent stem cells into multiple neuronal subtypes. *Nat Biotechnol* 33, 89-96. 10.1038/nbt.3049.
63. Jones, R.A., Harrison, C., Eaton, S.L., Llaverro Hurtado, M., Graham, L.C., Alkhamash, L., Oladiran, O.A., Gale, A., Lamont, D.J., Simpson, H., et al. (2017). Cellular and Molecular Anatomy of the Human Neuromuscular Junction. *Cell Rep* 21, 2348-2356. 10.1016/j.celrep.2017.11.008.
64. Fukuda, A., Hazelbaker, D.Z., Motosugi, N., Hao, J., Limone, F., Beccard, A., Mazzucato, P., Messana, A., Okada, C., San Juan, I.G., et al. (2021). De novo DNA methyltransferases DNMT3A and DNMT3B are essential for XIST silencing for erosion of dosage compensation in pluripotent stem cells. *Stem Cell Reports* 16, 2138-2148. 10.1016/j.stemcr.2021.07.015.
65. Mordes, D.A., Prudencio, M., Goodman, L.D., Klim, J.R., Moccia, R., Limone, F., Pietilainen, O., Chowdhary, K., Dickson, D.W., Rademakers, R., et al. (2018). Dipeptide repeat proteins activate a

- heat shock response found in C9ORF72-ALS/FTLD patients. *Acta Neuropathol Commun* 6, 55. 10.1186/s40478-018-0555-8.
66. Limone, F., Mordes, D.A., Couto, A., Pietiläinen, O., Joseph, B.J., Burberry, A., Mitchell, J.M., Ghosh, S., Meyer, D., Therrien, M., et al. (2021). Single-nucleus sequencing reveals enriched expression of genetic risk factors sensitises Motor Neurons to degeneration in ALS. *bioRxiv* [PREPRINT]. doi.org/10.1101/2021.07.12.452054.
 67. Shahini, A., Vydiam, K., Choudhury, D., Rajabian, N., Nguyen, T., Lei, P., and Andreadis, S.T. (2018). Efficient and high yield isolation of myoblasts from skeletal muscle. *Stem Cell Res* 30, 122-129. 10.1016/j.scr.2018.05.017.
 68. Hao, J., Wells, M.F., Niu, G., Guerra San Juan, I., Limone, F., Fukuda, A., Lyeton-Jaimes, M.F., Joseph, B.J., Qian, M., Mordes, D.A., et al. (2021). Loss of TBK1 activity leads to TDP-43 proteinopathy through lysosomal dysfunction in human motor neurons. *bioRxiv* [PREPRINT]. doi.org/10.1101/2021.10.11.464011.
 69. Merkle, F.T., Ghosh, S., Kamitaki, N., Mitchell, J., Avoir, Y., Mello, C., Kashin, S., Mekhoubad, S., Ilic, D., Charlton, M., et al. (2017). Human pluripotent stem cells recurrently acquire and expand dominant negative P53 mutations. *Nature* 545, 229-233. 10.1038/nature22312.
 70. Macosko, E.Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A.R., Kamitaki, N., Martersteck, E.M., et al. (2015). Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell* 161, 1202-1214. 10.1016/j.cell.2015.05.002.
 71. Limone, F., Mitchell, J.M., Guerra San Juan, I., Smith, J.L.M., Raghunathan, K., Couto, A., Ghosh, S., Meyer, D., Mello, C.J., Nemesh, J., et al. (2022). Efficient generation of lower induced Motor Neurons by coupling Ngn2 expression with developmental cues. *bioRxiv* [PREPRINT]. doi.org/10.1101/2022.01.12.476020.
 72. Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., 3rd, Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive Integration of Single-Cell Data. *Cell* 177, 1888-1902 e1821. 10.1016/j.cell.2019.05.031.
 73. Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W.M., 3rd, Zheng, S., Butler, A., Lee, M.J., Wilk, A.J., Darby, C., Zager, M., et al. (2021). Integrated analysis of multimodal single-cell data. *Cell* 184, 3573-3587 e3529. 10.1016/j.cell.2021.04.048.
 74. Hafemeister, C., and Satija, R. (2019). Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol* 20, 296. 10.1186/s13059-019-1874-1.

ACKNOWLEDGMENTS

We would like to thank members of the Eggan, McCarroll, Nehme and Pietiläinen groups for insights and useful discussions and Lee Rubin and members of the Rubin lab for support in the final steps of this project. We are grateful to the Harvard Stem Cell and Regenerative Biology Department (HSCRB) for the use of shared equipment, microscopy core and FACS facility. We are also grateful to the Stanley Center at the Broad Institute of Harvard and MIT for support with bioinformatics and cloud space. We thank BioRender.com for useful diagrams and images used in figures throughout this publication. This work was supported by U01MH115727 to S.A.M., R.N. and K.E. as well as support from Stanley Center for Psychiatric Research at Broad Institute of MIT and Harvard and Harvard Stem Cell and Regenerative Biology Department at Harvard University.

AUTHOR CONTRIBUTION

Conception and study design F.L., R.N. Data analysis and interpretation F.L., O.P., R.N. Manuscript writing and editing F.L., O.P., R.N., M.V. “Villages” design and multi-cell line experiments J.M., J.L.M.S., K.R. Immunofluorescence analysis and co-culture studies F.L., I.G.S.J., A.C., B.M.S. Conventional motor neuron experiments F.L., B.M.S., I.G.S.J., J.R.K. Electrophysiology analysis B.J.J., J.G, F.L. Bioinformatics and scRNA-seq F.L., S.M., O.P., S.D.G., D.M., C.J.M., J.M.

DECLARATION OF INTERESTS

K.E. is cofounder of Q-State Biosciences, Quralis, Enclear Therapies, and is group vice president at BioMarin Pharmaceutical.

INCLUSION AND DIVERSITY

We worked to ensure diversity in experimental samples through the selection of the cell lines. We worked to ensure diversity in experimental samples through the selection of the genomic datasets. One or more of the authors of this paper self-identifies as an underrepresented ethnic minority in their field of research or within their geographical location. One or more of the authors of this paper self-identifies as a gender minority in their field of research. One or more of the authors of this paper self-identifies as a member of the LGBTQIA+ community.

STAR METHODS

LEAD CONTACT

Kevin Eggan (kevin.eggan@bmrn.com).

MATERIALS AVAILABILITY

This study did not generate new unique reagents.

DATA AND CODE AVAILABILITY

All codes and algorithms necessary for re-analysis of the single-cell RNA-sequencing data are publicly available and can be found in other publications^{39,40}. Raw sequencing data and count matrices have been deposited in GEO and DUOS and can be requested using the ID GSE219112 and DUOS-000121 (<https://duos.broadinstitute.org/>). This paper does not report original code. Further information requests can be directed to Kevin Eggan (kevin.eggan@bmrn.com) or Francesco Limone (francesco_limone@fas.harvard.edu).

EXPERIMENTAL MODELS AND SUBJECT DETAILS

NGN2-based differentiations

Stem cells were grown in mTeSR1 (Stem Cell Technologies, 05850) and grown on Matrigel (Corning) coated plates at 37°C and 5% CO₂. hPSCs were infected with TetO-Ngn2-Puro, TetO-GFP and rtTA lentiviral constructs⁹ produced by Alstem in mTeSR medium with 1 μM RoCK inhibitor Y-27632 for 24 hours. hPSs were then passaged and differentiation was started when cells reached 70-80% confluency. For the first four days of differentiation cells were grown in induction medium: DMEM/F12 (Life Technologies, 11320-033), N2 supplement (0.5%v/v, Gibco), 1X GlutaMAX (Gibco), 0.1mM Non-essential amino acid (Gibco), 0.5% glucose, doxycycline hyclate (2 μg/mL). Small molecules added: day 1 – DOX: none; LSB: 10 μM SB431543 (Custom Synthesis), 200 nM LDN193189 (Custom Synthesis); piNs: 10 μM SB431543 (Custom Synthesis), 200 nM LDN193189 (Custom Synthesis), 4 μM XAV939 (Stemgent, 04-00046); liMNs: 10 μM SB431543 (Custom Synthesis), 200 nM LDN193189 (Custom Synthesis), 2 μM retinoic acid (Sigma) and 2 μM Smoothed agonist (Custom Synthesis). Day 2 to 4 - DOX: puromycin (5 μg/mL); LSB: puromycin (5 μg/mL), 10 μM SB431543 (Custom Synthesis), 100 nM LDN193189 (Custom Synthesis); piNs: puromycin (5 μg/mL), 10 μM SB431543 (Custom Synthesis), 100 nM LDN193189 (Custom Synthesis), 2 μM XAV939 (Stemgent, 04-00046); liMNs: puromycin (5 μg/mL), 10 μM SB431543 (Custom Synthesis), 100 nM LDN193189 (Custom Synthesis), 1 μM retinoic acid (Sigma) and 1 μM Smoothed agonist (Custom Synthesis). On day 4 cells were dissociated using Accutase (Gibco) and replated in a 1:2 dilution to ensure puromycin selection of uninfected cells. For day 4 to 7, DOX, LSB and piNs cells were grown in neuronally supportive medium supplemented with small molecules as described above: Neurobasal (Life Technologies 21103049) supplemented with B27 supplement w/o vitA (2%v/v, Gibco), 1X GlutaMAX (Gibco), 0.1mM Non-essential amino acid (Gibco), 0.5% glucose with the addition of 10 ng/ml of BDNF, CNTF and GDNF (R&D Systems). For day 7 to 10, liMNs were grown with small molecules as described above in neuronally supportive medium: Neurobasal (Life Technologies 21103049) supplemented with B27 supplement (2%v/v, Gibco), N2 supplement (0.5%v/v, Gibco), 1X GlutaMAX (Gibco), 0.1mM Non-essential amino acid (Gibco), 0.5% glucose with the addition of 10 ng/ml of BDNF, CNTF and GDNF (R&D Systems). On day 7, cells were dissociated using accutase and replated on glial co-cultures as described previously²⁶ in medium described above. From this time onwards, half-media change was performed every 2-3 days in neuronally supportive media described above with the only addition of 10 ng/ml of BDNF, CNTF and GDNF (R&D Systems). For most experiments, neurons were co-cultured with murine glial cells (50,000 cells/cm²) derived from postnatal brains (P0-2) as previously described²⁶, neurons were mixed with glia when replating day 7 cells 30,00 cells/cm².

2D MN differentiation

Stem cells were grown in mTeSR1 (Stem Cell Technologies, 05850) and grown on Matrigel (Corning) coated plates at 37°C and 5% CO₂. Stem cells were differentiated to bona fide 2D Motor Neurons as previously described^{36,64,65}. This protocol based on the principle of neuralization by dual-Smad inhibition

followed by the inhibition of NOTCH/FGF pathway both under the patterning capability of retinoids and Sonic Hedgehog. Briefly, once 90-95% confluent, stem cell medium was switched to differentiation medium: 1:1 mix of Neurobasal (Life Technologies 21103049) and DMEM/F12 (Life Technologies, 11320-033) supplemented with B27 supplement (2%v/v, Gibco), N2 supplement (0.5%v/v, Gibco), 1X GlutaMAX (Gibco), 0.1mM Non-essential amino acid (Gibco). For the first six days, differentiation medium was supplemented with 10 μ M SB431543 (Custom Synthesis), 100nM LDN193189 (Custom Synthesis), 1 μ M retinoic acid (Sigma) and 1 μ M Smoothed agonist (Custom Synthesis). For the second week, differentiation medium was supplemented with: 5 μ M DAPT (Custom Synthesis), 4 μ M SU-5402 (Custom Synthesis), 1 μ M retinoic acid (Sigma) and 1 μ M Smoothed agonist (Custom Synthesis). To isolate neurons from mixed cultures we utilised an immune-panning based method previously described^{36,66}. At day 14, monolayers were dissociated with Accutase (Gibco) for 1 hour at 37°C. After gentle, repeated pipetting, cells were collected, spun down and resuspended in sorting buffer and filtered. Single cell suspensions were incubated with antibody against NCAM (BD Bioscience, 557919, 1:200) for 25 minutes, washed and NCAM⁺ cells were sorted with an BD FACS Aria II cell sorter. Sorted 2D MN were plated on mouse glial cultures in motor neuron medium (Neurobasal (Life Technologies 21103049) supplemented with B27 supplement (2%v/v, Gibco), N2 supplement (0.5%v/v, Gibco), 1X GlutaMAX (Gibco), 0.1mM Non-essential amino acid (Gibco), 0.5% glucose) with the addition of 10 ng/ml of BDNF, CNTF and GDNF (R&D Systems). Neurons were co-cultured with murine glial cells (150,000 cells/cm²) derived from postnatal brains (P0-2) as previously described²⁶.

Co-culture of Ngn2 motor neurons and mouse myoblasts in microfluidic devices

Mouse myoblasts from hindlimb skeletal muscles of young adult mice and mouse glia from neonatal mouse brains were isolated and cultured as previously described^{26,67}. Microfluidic device chips (XC450, XONA Microfluidics) were designated a motor neuron compartment and a muscle compartment. The motor neuron compartment was coated with 0.1 mg/ml poly-L-ornithine (Sigma-Aldrich) in 50 mM Borate buffer, pH = 8.5 and 5 μ g/ml laminin (Invitrogen), while the muscle compartment was coated with Matrigel (Corning). Day 7 Ngn2 motor neurons and mouse glia were seeded at a concentration of 100,000 neurons-200,000 glia/device. Myoblasts were seeded at a concentration of 150,000 device. Motor neurons were seeded in the motor neuron media described above with the addition of 10 ng/ml of BDNF, CNTF and GDNF (R&D Systems). For seeding and culturing the first 2 days, myoblasts were maintained in Myoblast media (DMEM/F12, 20% Foetal Bovine Serum and 10% heat-inactivated Horse Serum, and 10 ng/ml bFGF), after that, differentiation was initiated by adding myoblast differentiation media (DMEM high glucose, 5% heat-inactivated Horse Serum). Myoblast were sustained in differentiation medium for 3 days and then switched to motor neuron medium with the addition of 10 ng/ml of BDNF, CNTF and GDNF (R&D Systems) and while medium in motor neuron compared contained no neurotrophic factors to start recruitment of motor neuron axons to the muscle compartment by generation of a volumetric gradient (50 μ l difference in volume between the compartments) in the device. Volumetric gradient was kept for every medium change, done every other day. Co-cultures were fixed at day 21 post-seeding for visualization of motor axon-muscle synaptic contacts.

METHOD DETAILS

FACS analyses

We used an *Hb9::GFP* reporter stem cell line previously described infected with the Ngn2 lentiviral constructs as described above²⁶. Briefly, cells were differentiated in 24 well plates and subjected to different patterning molecules. At each time point, cells were dissociated with Accutase (Gibco) as previously described, each replicate was frozen in Cryostor® CS10 (STEMCELL Technologies). After all samples were collected, cells were thawed in separated tubes and resuspended in sorting buffer as described by others³⁶. The BD FACS Aria II cell sorter was used to quantify the percentage of *Hb9::GFP*⁺ cells in each sample after using DAPI signal to determine cell viability.

Immunofluorescence assays

Cells were washed once with PBS, fixed with 4% PFA for 20 minutes, washed again in PBS and blocked for one hour in 0.1% Triton in PBS with 10% donkey serum. Fixed cells were then washed and incubated overnight with primary antibodies at 4°C. Primary antibody solution was washed and cells were subsequently incubated with secondary antibodies (1:2000, Alexa Fluor, Life Technologies) at room temperature for 1 hour, washed with PBS and stained with DAPI. Primary antibodies used: Tuj1 (R&D, MAB1195), Islet1 (Abcam, ab178400), MAP2 (Abcam, ab5392), Synapsin (Millipore, AB1543), SMI-32 (BioLegend, 801702), Chat (Millipore, AB144P), Foxp1 (Abcam, ab16645), AnkyrinG (Millipore, MABN466), Synaptophysin (Synaptic Systems, 101 004), PSD-95 (Abcam, ab2723), STMN2 (Novus NBP49461). Images were analysed using FIJI.

RNA extraction and RT-qPCR analyses

RNA was extracted with the miRNeasy Mini Kit (Qiagen, 217004). cDNA was produced with iScript kit (BioRad) using 50 ng of RNA. RT-qPCR reactions were performed in triplicates using 20 ng of cDNA with SYBR Green (BioRad) and were run on a CFX96 Touch™ PCR Machine for 39 cycles at: 95°C for 15s, 60°C for 30s, 55°C for 30s.

Western blots

For WB analyses, cells were lysed in RIPA buffer with protease inhibitors (Roche). After protein quantification by BCA assay (ThermoFisher), ten micrograms of protein were preheated in Laemmli's buffer (BioRad), loaded in 4-20% mini-PROTEAN® TGX™ precast protein gels (BioRad) and gels were transferred to a PDVF membrane. Membranes were blocked in Odyssey Blocking Buffer (Li-Cor) and incubated overnight at 4°C with primary antibodies (1:1000 dilution). After washing with TBS-T, membranes were incubated with IRDye® secondary antibodies (Li-Cor) for one hour and imaged with Odyssey® CLx imaging system (Li-Cor). Primary antibodies used: Tuj1 (R&D, MAB1195), Synapsin (Millipore, AB1543), PSD-95 (Neuromab, 75-028), GAPDH (Millipore, MAB374).

Multi Electrode Array analysis

Electrophysiological recordings were obtained by Axion Biosystems Multi-Electrode Array (MEA) plate system (Axion Biosystems, 12 wells or 48 wells formats) that recorded extracellular spike potential. On day 7 of differentiation, cells were detached and counted and mixed with murine glia as described above. MEA plates were previously coated with Matrigel (Corning) and cells were seeded in Neurobasal medium supplemented with ROCK inhibitor for 24 hours. Recordings were performed every 2-3 days and medium was changed after recordings. Analysis was performed with AxIS (Axion Biosystems – Neuronal Metric Tool) as described by others^{11,68}.

QUANTIFICATION AND STATISTICAL ANALYSIS

Stem cell lines, villages, single-cell RNA-sequencing, Census-seq and Dropulation

Methods for Census-seq and Dropulation are described elsewhere^{39,40}, brief description below:

Human pluripotent cell lines and village generation³⁹

Human ESC lines used in this study were part of a collection previously described⁶⁹. These lines were exome sequenced and whole genome sequenced after minimal passaging and cultured as described. Individual lines were cultured and differentiated into neurons as described. At day 6 after doxycycline induction, when cells are postmitotic, cultures were dissociated with Accutase (Gibco) and resuspended in mTeSR medium with 1 μ M ROCK inhibitor Y-27632. To generate balanced “villages”, cell suspensions were counted using a Scepter 2.0 Handheld Cell Counter (Millipore Sigma) with 60 μ M Scepter Cell Counting Sensor (Millipore Sigma), 0.5M viable cells from each donor cell line were mixed. At this timepoint 0.5M cells were harvested for Census-seq analysis and ensure balanced representation, the rest was plated for subsequent experiments.

DNA isolation and library preparation³⁹

Every seven days, pellets were harvested from separate wells of the “liMNs village” after dissociation with Accutase (Gibco). Pellets were lysed and DNA precipitated and DNA was used to generate libraries using TruSeq Nano DNA Library Prep Kit (Illumina), libraries were sequenced using NextSeq 500 Sequencing System (Illumina). Generated libraries were aligned to human genome using BWA, reference genome was selected to match the genomes used to generate VCF files containing the whole-genome sequenced genotypes of each donor cell line. To exclude confounding mouse DNA from glia, a multi-organism reference was used, reads competitively aligned to both genomes and only high quality (MQ $\geq$ 10) were used for assignment.

Census-seq analysis³⁹

The algorithms used to assign donor contribution to villages are extensively described elsewhere and their validation is outside the scope of this publication. However, briefly the aim of Census-seq algorithms is to accurately detect and precisely quantify the contribution of donors in a mixed DNA sample to monitor population dynamics over time and/or conditions. This can be achieved systematically and inexpensively by lightly sequencing genomic DNA, the algorithms attempt to determine the donors' mixture by determining the ratio of alleles present at every SNP. The gradient-descent algorithm can then use this data to identify the donor-mix that maximizes the likelihood of any observed sequence data. Once the best ratio is identified, the algorithms compare the computed “most likely donor mix” to a VCF file that contains whole genome-sequencing data from all stem cell lines in the collection. These VCF files contain a filtered and refined matrix with alternate alleles at each variant for every donor's genotype. Census-seq can use this data to find a vector of donor-specific contribution (to the mix) that can explain the allele counts detected at each site in the sequencing data provided. For each site, the allele frequency is inferred using the VCF reference files and its proportion of donor in the pool of DNA can then be calculated over the total counts for that specific site. The algorithms are

then able to sum the proportion of each donor's representation at every specific site and calculate total representation of each genotype, a.k.a. donor, in the pooled DNA, providing us an estimate of the ratio of donors in the village.

Dropulation: scRNA-sequencing and donor assignment⁴⁰

For single-cell analyses, cells were harvested and prepared with 10X Chromium Single Cell 3' Reagents V3 and sequenced on a NovaSeq 6000 (Illumina) using a S2 flow cell at 2 x 100bp. Raw sequence files were then aligned and prepared following previous Drop-seq workflow⁷⁰. Human reads were aligned to GRCh18 and filtered for high quality mapped reads ($MQ \geq 10$). In order to identify donor identity of each droplet, variants were filtered through several quality controls as described previously be included in the VCF files^{40,71}, to summarise the goal is to only use sites that unambiguously and unequivocally can be detected as A/T or G/C. Once both the sequenced single-cell libraries and VCF reference files are filtered and QC'ed, the Dropulation algorithm is run. Dropulation analyses each droplet, hence a cell, independently and for each cell generates a number representing the likely provenance of each droplet from one donor. Each variant site is assigned a probability score for a given allele in the sequenced unique molecular identifier (UMI) calculated as the probability of the base observed compared to expected based, and $1 - \text{probability that those reads disagree with the base sequenced}$. Donor identity is then assigned as the computed diploid likelihood at each UMI summed up across all sites.

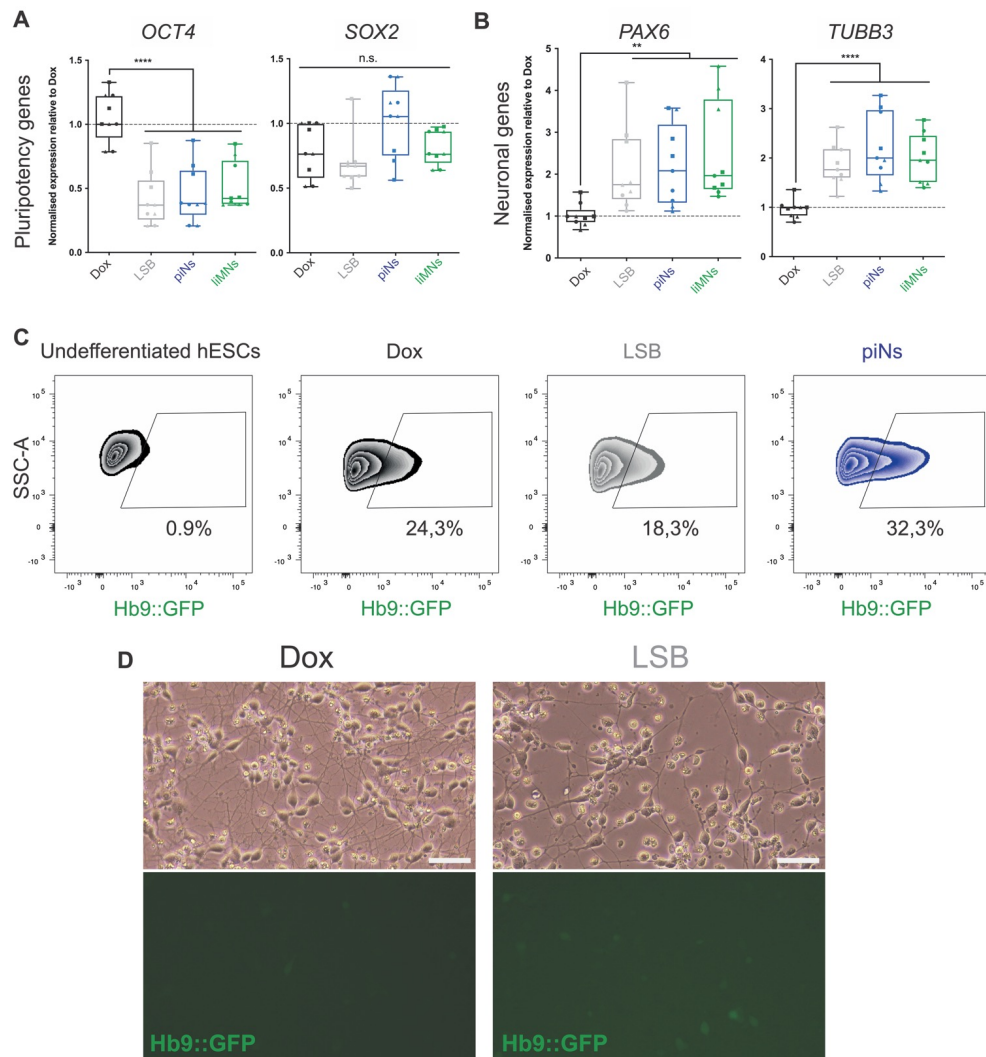
This probability-based analysis allows to increase confidence in donor detection per barcode by increase the numbers of individuals in the VCF files: more individuals, more UMIs with site variants, more confident scores, higher quality donor assignments. After assigning a "likelihood score", sites where only few donors have detected reads are ignored and scores are adjusted to allow only high confidence variant sites to be included. This second computer score is then added to the original likelihood as a weighted average score, this mixed coefficient defines the proportion of the population that presents each genotype and in adds to 1. Based on this mixed coefficient that takes into account reads mapped to each donors and the confidence to which each site can be used for this assignment, Dropulation then contains algorithms able to detect "doublets", barcoded droplets with genetic DNAs assigned to two different donors, to avoid analysing barcodes with admixed identity but also to avoid excluding barcoded droplets with unclear donor assignment based on the coefficient previously calculated⁴⁰.

Once scores are calculated, the algorithm assigns donors to single droplets. Then runs the double detection and cells that are likely doublets are filtered out. After that, donor identities are confirmed only if $p\text{-value} < 0.05$. These cells are then validated by crossing proportions of each donors as known inputs in the village and excluding any unexpected identity. Donors composing less then 0.2% of the libraries are excluded from the experiment⁴⁰.

More details on the preparation of libraries and donor identification can be found in published work⁴⁰.

scRNAseq analysis of villages and integrated datasets

Matrices from neuronal villages were built from 12 separate runs of 10X Chromium Single Cell 3' Reagents V3 as described above. Any barcode with less than 400 genes and combined UMI matrices were used for downstream analysis using Seurat (v3.0.2)⁷². After that, barcodes were further filtered by number of genes detected $1500 < n\text{Feature_RNA} < 7000$ and percent of mitochondrial and ribosomal genes to reduced the number of dying cells/debris: $\text{percent.mito} < 20$, $3 < \text{percent.RPS} < 15$, $5 < \text{percent.RPS} < 10$. The matrix was then processed via the Seurat pipeline: log-normalized by a factor of 10,000, followed by regressing out UMI counts, mitochondrial and ribosomal genes, scaled for gene expression. After quality filtering, barcodes were used to compute SNN graphs and *t*-SNE projections using numbers of Principal Components based on ElbowPlot analysis. SNN-graphed *t*-SNE projection was used to determine minimum number of clusters obtain at resolution=0.2 (FindClusters) as described previously⁶⁶. Integration with Rayon et al. 2021 was performed using matrices and metadata available at https://github.com/briscoelab/human_single_cell. Only barcodes with available metadata concerning their cellular identity from Rayon et al. were selected to use identities assigned by peer review publication⁴⁴. The available barcodes were then loaded into Seurat v4.0.1⁷³. Integration with libraries previously generated from villages of liMNs was achieved using SCTransform on a merged object running the PreSCTIntegration() function according to the sctransform integration pipeline⁷⁴. Analysis of MN alone was conducted as described above by comparing liMNs generated in this study with barcodes identified as "pMN" and "MN" by Rayon et al.

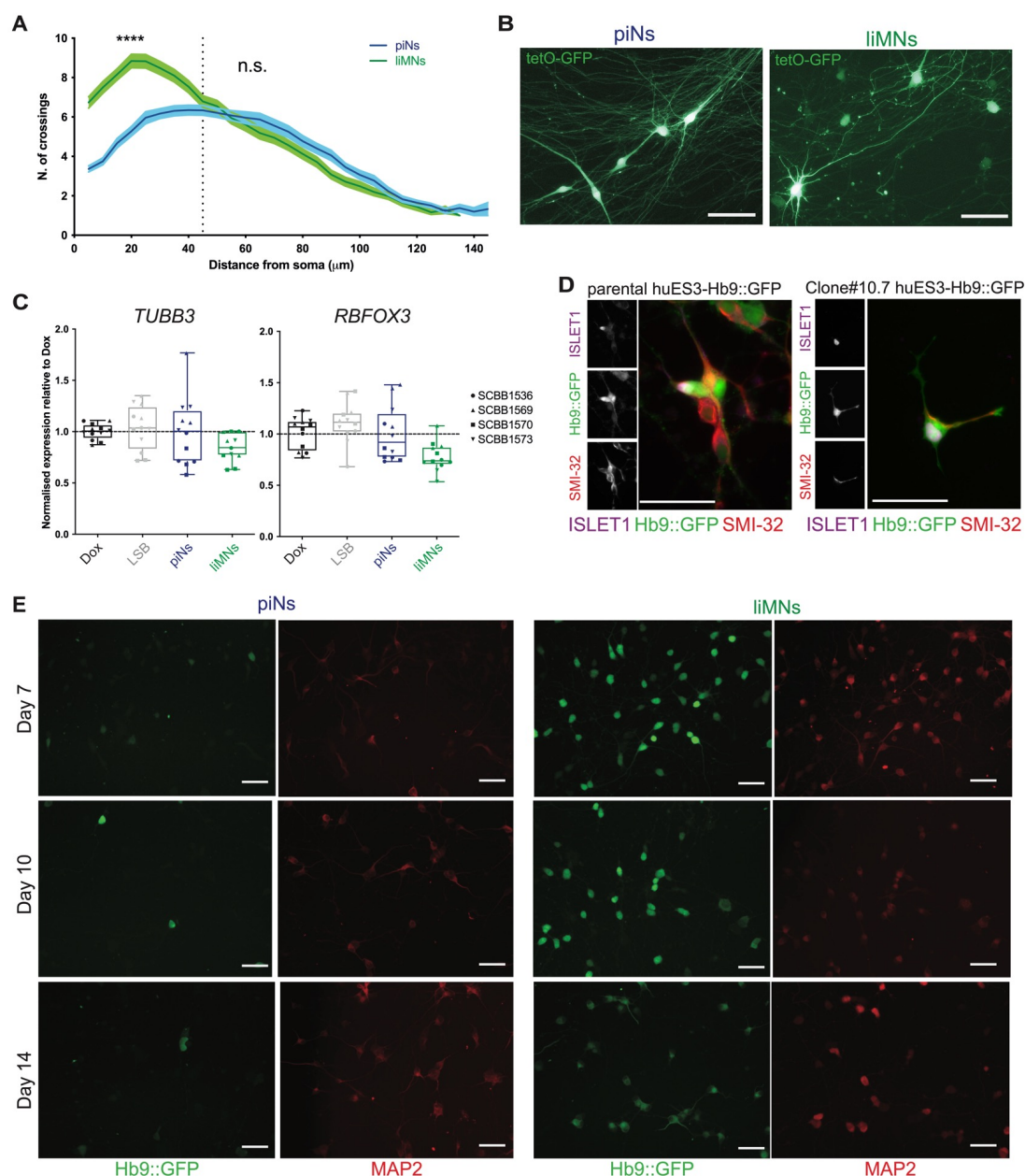


Supplementary Figure 1. Ngn2 neuronal patterning can be directed to different neuronal fates by small molecules patterning

(A-B) RT-qPCR quantification of pluripotency genes and genes involved in pan-neuronal development (p-values from one-way ANOVA).

(C) Flow cytometry quantification of Hb9::GFP positive cells by day 4 for the other conditions.

(D) Hb9::GFP expression at day 7 post-induction in original Ngn2-induced Dox and LSB conditions (scale bar 50 μ m).



Supplementary Figure 2. Patterned Ngn2-induced neuronal fate is maintained throughout the differentiation.

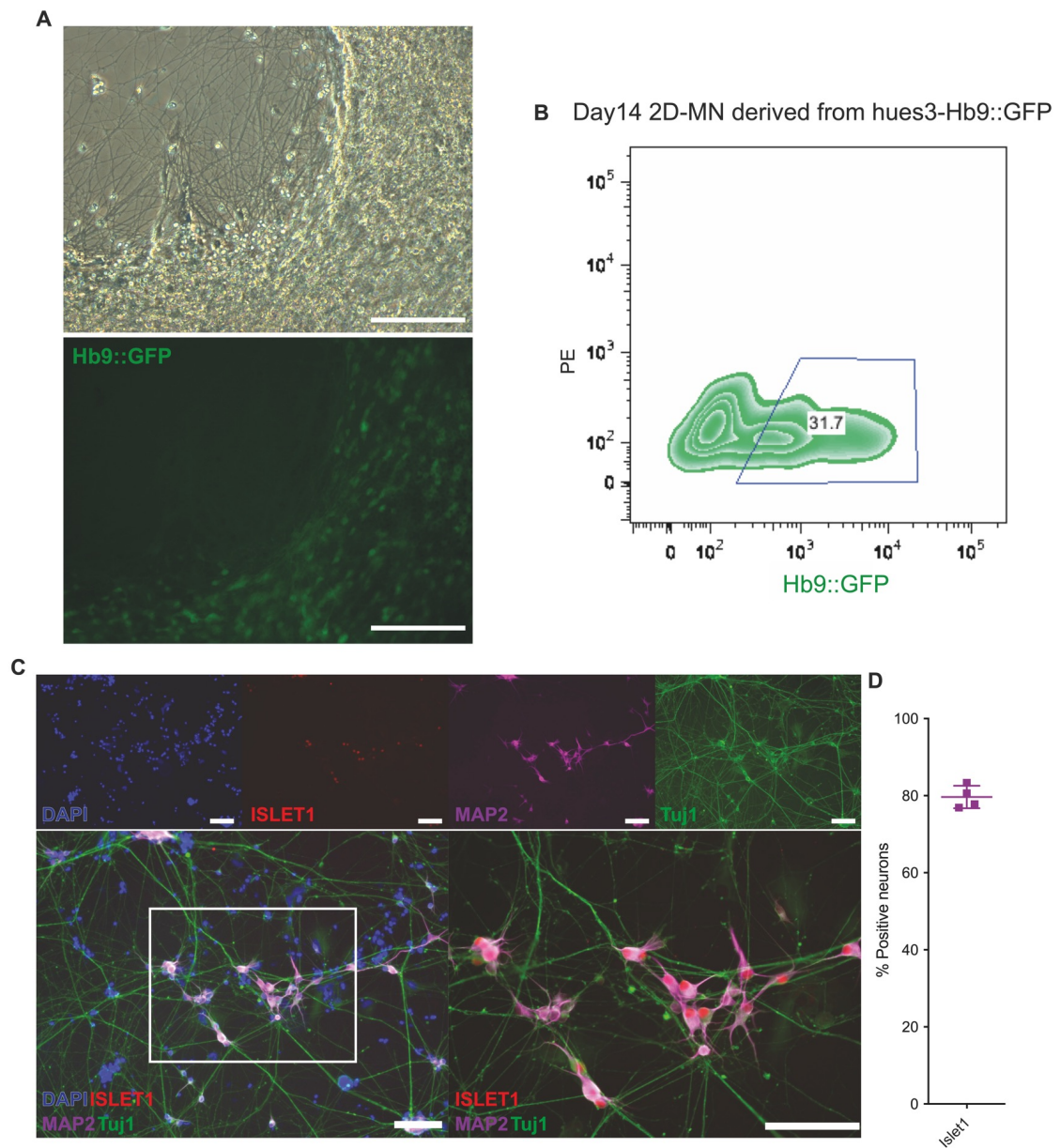
(A) Quantification of arborization of piNs and liMNs.

(B) Viral tetO-GFP imaging at day 30 in piNs and liMNs, showing different cell morphology (scale bar $50 \mu\text{m}$).

(C) RT-qPCR quantification of pan-neuronal markers (p-values from one-way ANOVA).

(D) IF analysis for pan-MN SMI-32, Islet1 and Hb9::GFP reporter expression at day 7 post-induction in two clones of the same reporter (scale bar $50 \mu\text{m}$).

(E) Images of Hb9::GFP expression at day 7, 10 and 14 post-induction in piNs and liMNs by immunofluorescence.



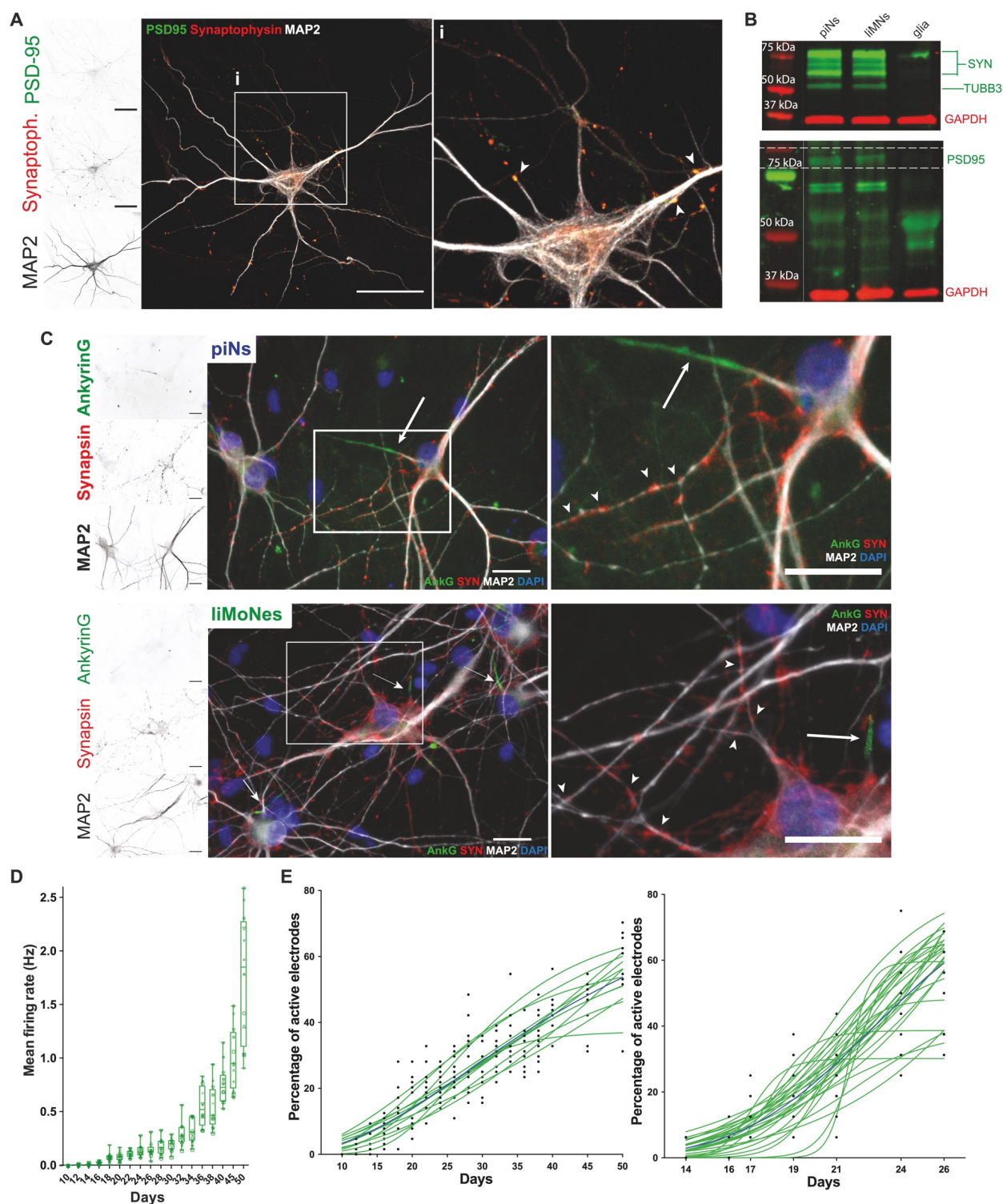
Supplementary Figure 3. Bona fide, hPSC-derived 2D MN express similar MN markers as liMNs.

(A) Hb9::GFP expression day 14 of differentiation in 2D-MN (scale bar 50 μ m).

(B) Flow cytometry quantification of Hb9::GFP positive cells in day 14 2D-MN.

(C) Immunofluorescence analysis for cholinergic transcription factor Islet1 and neuronal cytoskeletal proteins MAP2 and TUBB3 (Tuj1) in sorted 2D MN (scale bar 50 μ m).

(D) Quantification for cells in A.



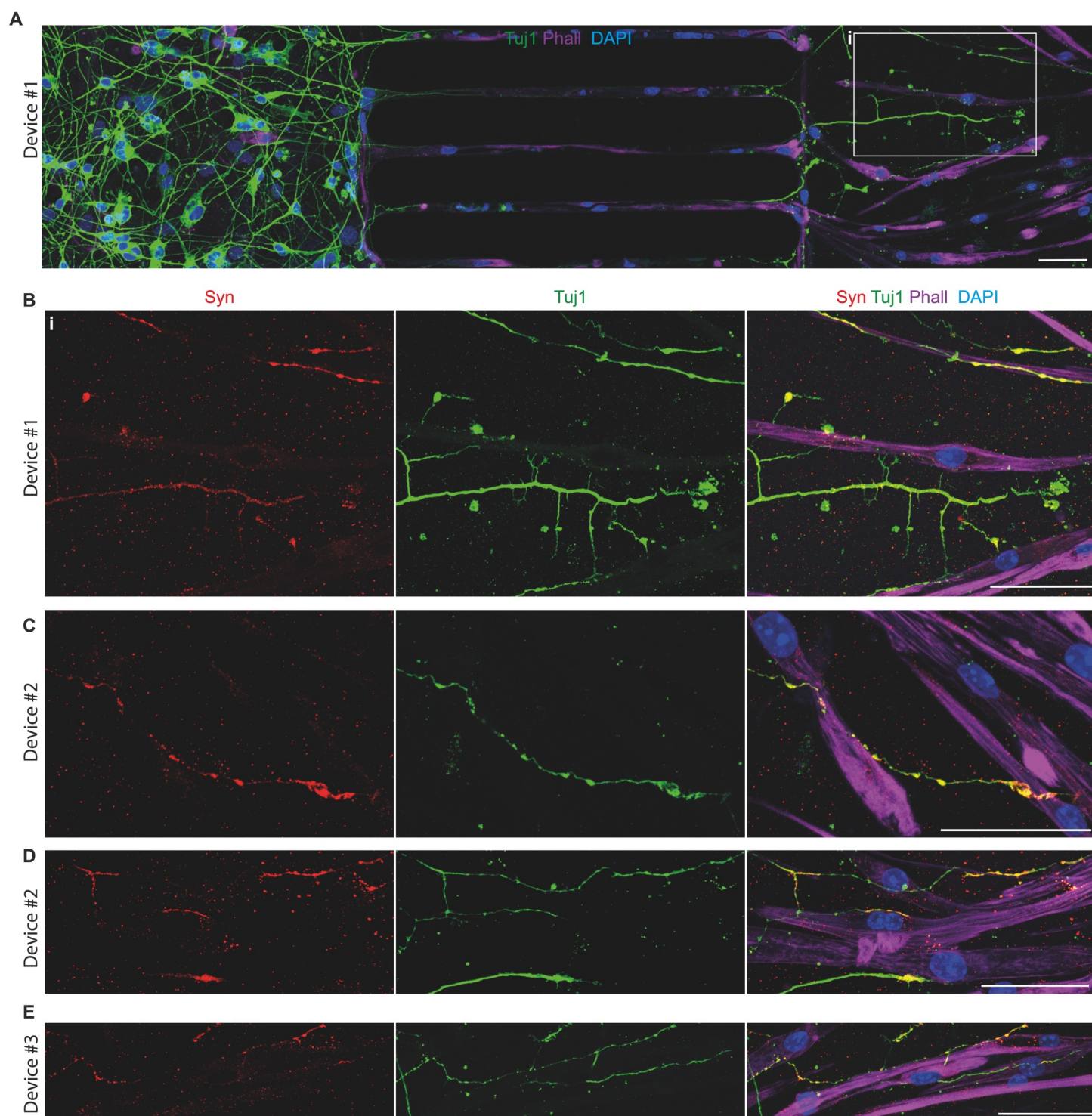
Supplementary Figure 4. liMNs can form active synaptic contacts comparable to previously characterised piNs.

(A) Day 50 liMNs express pre- and post-synaptic density proteins (glial co-cultures - scale bar 50 μ m).

(B) Western blot analysis shows expression of pre- and post-synaptic density molecules in both cell types.

(C) Immunofluorescence for proteins involved in the formation of functional axons and synaptic structure in piNs and liMNs (glial co-cultures - scale bar 50 μ m).

(C-D) Network activity of liMNs. (C) Mean number of spikes in 10-s period in liMNs co-cultured with murine cortical glial preparations. (D) Proportion of active electrodes detecting spontaneous activity throughout the differentiation (days). Data fit by sigmoidal function (green), median sigmoidal in black.



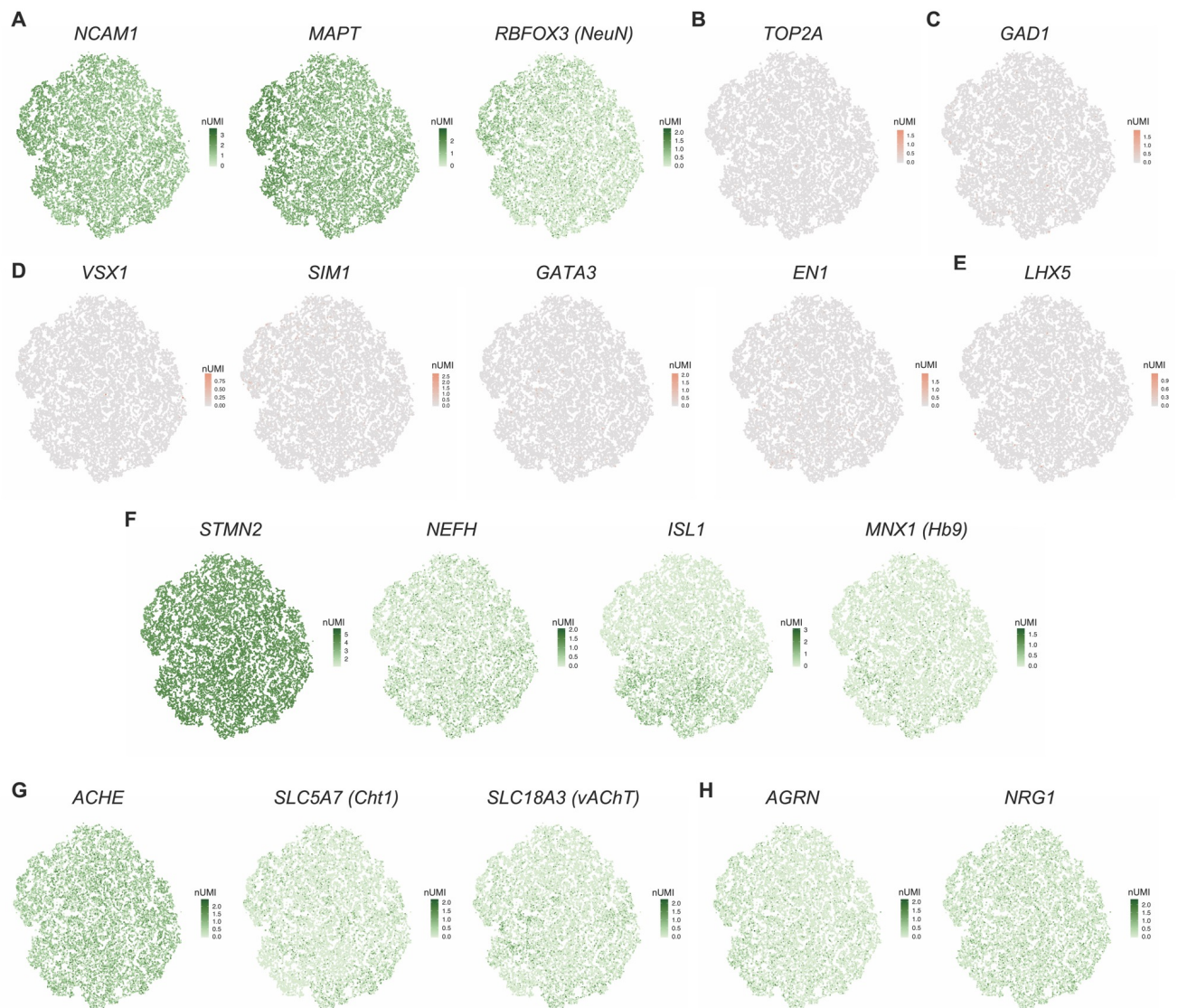
Supplementary Figure 5. liMoNs can form NMJ-like structures *in vitro*.

Immunofluorescence of co-cultures of liMoNs and primary murine myoblasts from three devices from separate rounds of differentiation.

(A) Representative image of neurons extending axons through the channels (middle), contacting primary muscle cells (right).

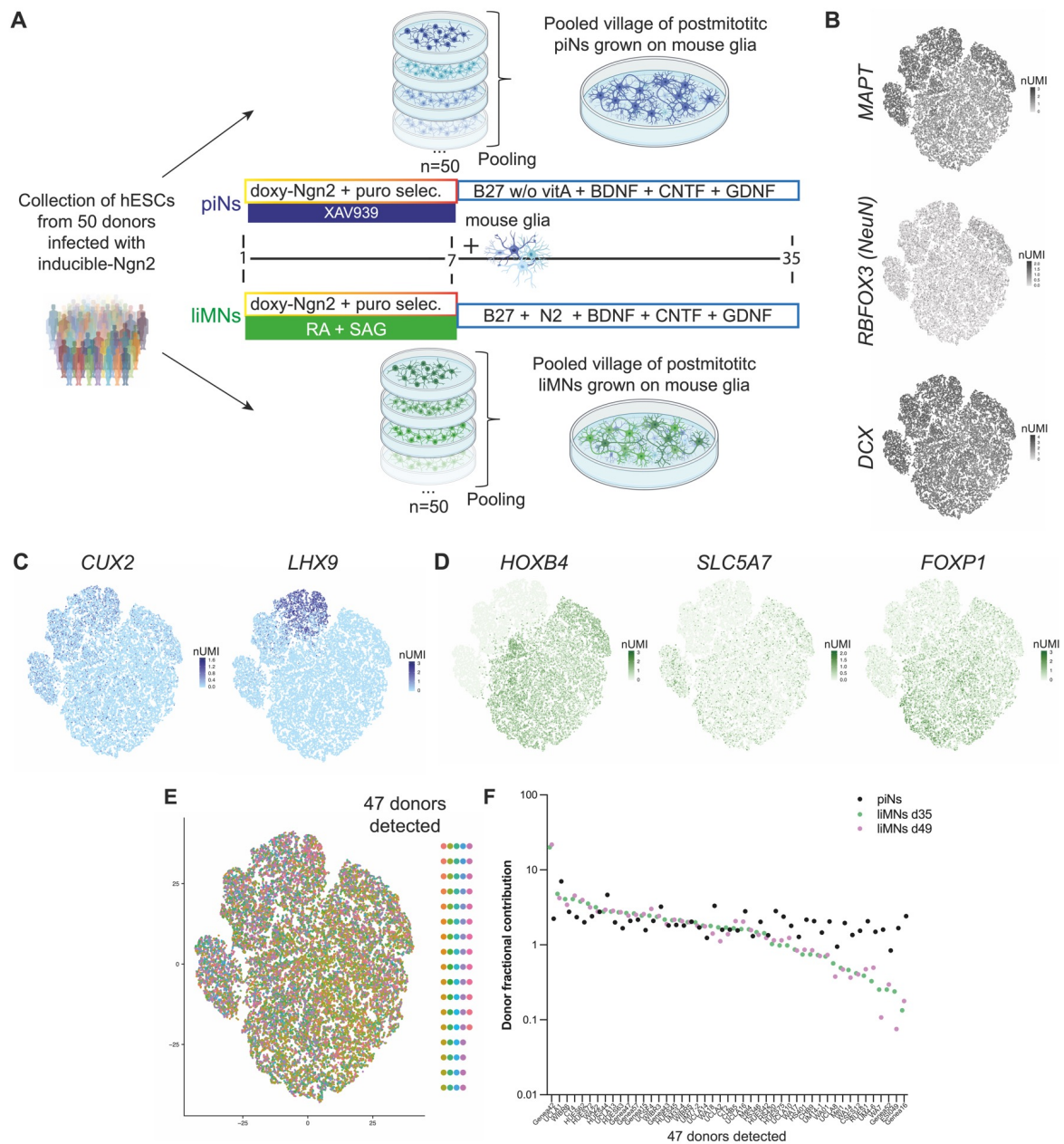
(B) Insets of (A) showing liMoNs forming synaptic-like contacts with muscle cells.

(C-E) Representative images from separate devices showing liMoNs forming synaptic-like contacts with muscle cells ($n=10$ devices).



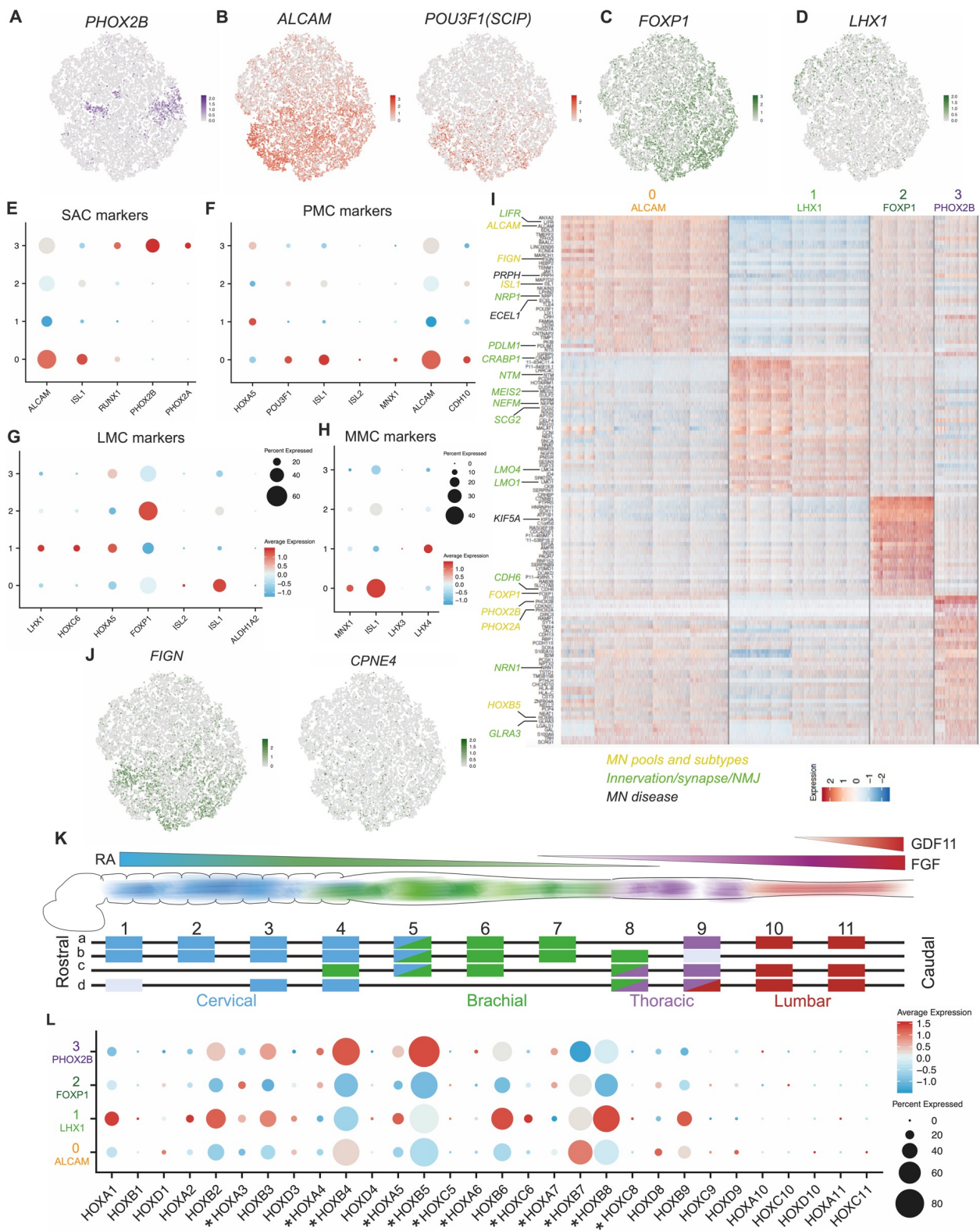
Supplementary Figure 6. scRNAseq confirms expression of MN markers.

- (A) *t*-SNE projection with expression of markers specific for post-mitotic neurons.
 (B) *t*-SNE projection with expression of cycling cells markers.
 (C) *t*-SNE projection with expression of inhibitory neurons.
 (D) *t*-SNE projection with expression of spinal ventral inhibitory V1, V2a, V2b, V3.
 (E) *t*-SNE projection with expression of markers specific for mid-spinal neurons V0.
 (F) *t*-SNE projection with expression of MN-specific markers.
 (G) *t*-SNE projection with expression of genes involved in cholinergic machinery.
 (H) *t*-SNE projection with expression of genes involved NMJ formation.



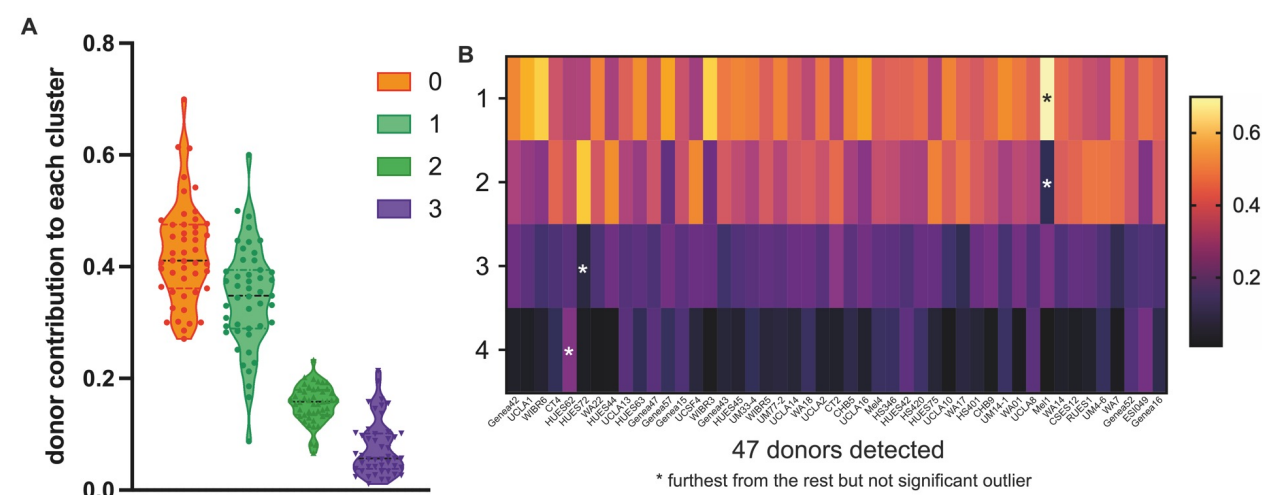
Supplementary Figure 7. piNs and liMNs transcriptomic differences.

- (A) Pooling strategy and village construction for sequencing analysis of piNs and liMNs.
- (B) *t*-SNE projection with expression of neuronal markers.
- (C) *t*-SNE projection with expression of cortical-enriched markers.
- (D) *t*-SNE projection with expression of MN-specific markers.
- (E) *t*-SNE projection of cells depicting donor's identity of each cell from 47 donors detected by *Dropulation* analysis.
- (F) Fraction representation of 47 donors in the two timepoints for liMNs and day 35 piNs.

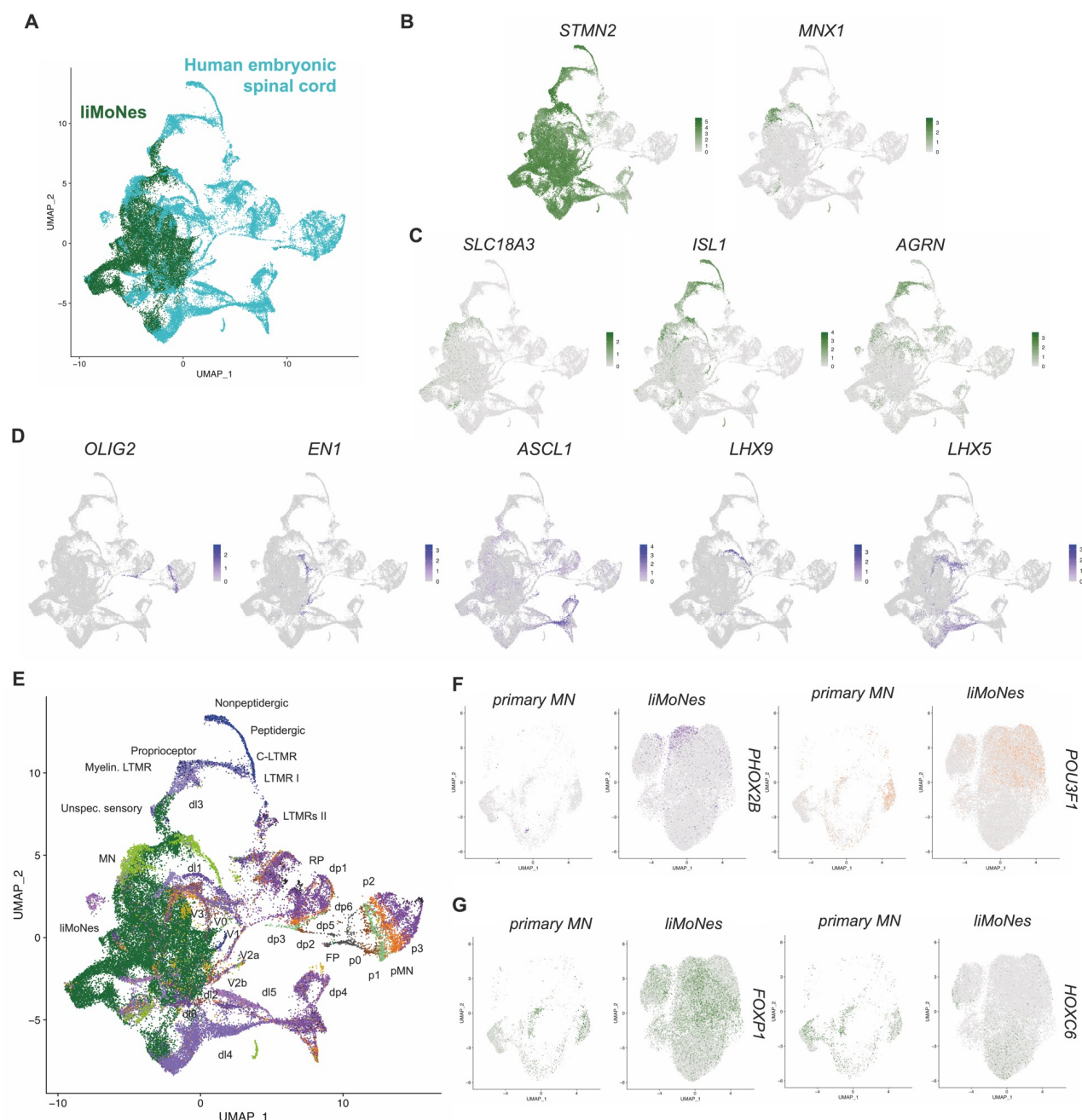


Supplementary Figure 8. Sub-columnar localization of liMNs.

- (A) *t*-SNE projection with expression of markers specific for SAC subtypes.
 (B) *t*-SNE projection with expression of markers specific for PMC subtypes.
 (C) *t*-SNE projection with expression of markers specific for mLMC subtypes.
 (D) *t*-SNE projection with expression of markers specific for lLMC subtypes.
 (E) Dotplot for expression of markers specific for SAC – Spinal Accessory Column.
 (F) Dotplot for expression of markers specific for PMC – Phrenic Motor Column.
 (G) Dotplot for expression of markers specific for LMC – Lateral Motor Column.
 (H) Dotplot for expression of markers specific for MMC – Median Motor Column.
 (I) Heatmap of genes differentially expressed in each subtype, highlighted genes involved in subtype specific MN biology.
 (J) *t*-SNE projection with expression of markers expressed by digit-innervating motor neurons.
 (K) Schematic of spinal cord HOX genes expression.
 (L) Dotplot for expression of all HOX genes detected in the four subclusters. Retinoid dependent Hox activation in vertebrates (green line) and specifically expressed in ventral spinal cord MNs (asterisks).



Supplementary Figure 9. Donor composition of liMNs.
(A) Violin plot showing distribution of each donor in the villages.
(B) Heatmap depicting donors composition of each subgroup and highlighting the absence of outliers by Grubb's test $\alpha=0.05$.



Supplementary Figure 10. Comparison of liMoNes and human embryonic spinal cord from Rayon et al. 2021.

(A) t-SNE projection of integrated datasets: liMoNes and human embryonic spinal cord Rayon et al. 2021.

(B) t-SNE projection with expression of MN-specific markers.

(C) t-SNE projection with expression of genes involved in cholinergic machinery.

(D) t-SNE projection with expression of markers specific to spinal, ventral inhibitory neurons, mid- and dorsal spinal neurons and progenitor cells.

(E) t-SNE projection of integrated datasets: liMoNes and human embryonic spinal cord with cell types identified by Rayon et al. 2021.

(F-G) t-SNE projection with expression of markers specific to hindbrain/cervical (F) and brachial (G) MN pools in primary MNs from Rayon et al. 2021 (left side of panel) and liMoNes (right side of panel).