TITLE: Axonal transcriptome of human stem cell derived neurons

AUTHORS: Rebecca L. Bigler[1,2], Joyce W. Kamande[2], Raluca Dumitru[3], Mark Niedringhaus[2,4], Anne Marion Taylor[2,4,5]*

AFFILIATIONS:

[1]Curriculum in Genetics and Molecular Biology, UNC-Chapel Hill, Chapel Hill, NC 27599 USA

[2]UNC/NC State Joint Department of Biomedical Engineering, UNC-Chapel Hill, Chapel Hill, NC 27599 USA

[3]UNC Human Pluripotent Stem Cell Core Facility, UNC-Chapel Hill, Chapel Hill, NC 24599 USA

[4]Neuroscience Center, UNC-Chapel Hill, Chapel Hill, NC 27599 USA

[5]Carolina Institute for Developmental Disabilities, Chapel Hill, NC 27599 USA

*To whom correspondence should be addressed:

amtaylor@unc.edu

UNC-Chapel Hill

UNC/NC State Joint Department of Biomedical Engineering CB#7575

Chapel Hill, North Carolina 27599

Phone: 919-843-8156

The identification of axonal mRNAs in model organisms has led to the discovery of multiple proteins synthesized within axons that are required for axon guidance and injury response. The extent to which these axonal mRNAs are conserved in humans is unknown. Here we report on the axonal transcriptome of glutamatergic neurons derived from human embryonic stem cells (hESC-neurons) grown in axon isolating microfluidic chambers. We identified mRNAs proportionally enriched in axons, representing a functionally unique local transcriptome as compared to the transcriptome of whole neurons inclusive of somata, dendrites, and axons. Additionally, we found that the most abundant axonal mRNAs in hESC-neurons functionally resemble the most abundant mRNAs in rat cortical neurons. The main functional categories of transcripts common to both datasets being “translational elongation”, “intracellular” and “synapse”. Pairwise comparison of our list of abundant human axonal transcripts to five similar previously published datasets generated from rat and mouse axons revealed hundreds of conserved axonal mRNAs. This new evaluation of mRNA within human axons provides an important resource for studying local mRNA translation in human neurons and has the potential to reveal both conserved and unique axonal mechanisms across species and neuronal types.

INTRODUCTION

Intra-axonal translation is involved in axon maintenance, guidance, synaptogenesis, and response to injury[1-3]. Axonal mRNAs implicated in these neuronal functions have been experimentally identified in multiple model systems, including *Aplysia* sensory neurites, rodent dorsal root ganglia, superior cervical ganglia, cortical and hippocampal neurons[2, 4-7]. In axons and synapses, remote from the nucleus and cell body, fine tuning of the local proteome by local translation can be critical for spatially restricted and rapid responses to extracellular signals such as glutamate, brain derived neurotrophic factor, neurotrophin-3 and netrin-1[6, 8-11].

Comparing axonal transcriptomes of various neuron types has the potential to not only highlight conserved roles of intra-axonal translation and the associated axonally synthesized proteins but also reveal unique functions of axonal translation and target proteins within each neuron type. Multiple vertebrate and invertebrate model systems have been used to evaluate axonally localized mRNAs but this work has not been done in human neurons[12-17]. Direct evaluation of primary human neurons is severely restricted due to the limited ability to obtain and culture these cells. Human stem cell derived neurons are a biologically relevant and tractable *in vitro* system of human neurons and protocols have been developed to differentiate stem cells into specific neuronal subtypes, such as motor, striatal, dopaminergic and glutamatergic[18-23]. Here we sought to identify the axonal transcriptome of human embryonic stem cells differentiated into glutamatergic neurons[23, 24] (hESC-neurons) and compare this transcriptome to axonal transcriptomes from rodent model systems.

RESULTS

Differentiation of human embryonic stem cell derived neurons in microfluidic chambers. To identify transcripts within human axons we first modified an existing microfluidic chamber design for compartmentalizing and harvesting pure axons, originally developed for murine neurons[25]. To optimize the growth of hESC-neurons, which have an increased nutrient demand compared to primary rodent neurons, we increased the height of the fluidically isolated compartments to provide greater access to media. In these microfluidic chambers, neurons were seeded in the soma compartment and stochastic axon growth resulted in some axons passing through the microgrooves to the axon compartment (Fig. 1A). Similar chamber designs have been used for biochemical analysis of pure axons and to characterize the axonal transcriptome of primary rat cortical neurons and mouse motoneurons[12, 13].

Differentiation of hESC-neurons began in traditional tissue culture dishes on days *in vitro* (DIV) 0 (Fig. 1A) according to an established protocol verified to produce glutamatergic neurons[23, 24]. On DIV24 neural progenitor cells were transferred to our custom axon isolating microfluidic chambers and matured to hESC-neurons over the next 30 days. By DIV58 many long, fine processes, resembling axons, could be seen within the second compartment. Microtubule-associated protein 2 (MAP2) and β-tubulin III staining, markers for cell bodies and axons respectively, confirmed that these fine processes were axons. After approximately DIV35, as is typical with long term hESC-neuron cultures, many of the soma clustered and formed balls preventing efficient antibody penetration. β-tubulin III staining revealed extensive axonal growth within the axon compartment (Fig. 1B and C) as well as large growth cones (Fig. 1D). MAP2 positive somata and dendrites were restricted to the soma compartment (Fig. 1C). The

cells in monolayer demonstrated canonical dendritic arborization as revealed by MAP2 staining (Fig. 1E). We found that 94.7% of cells within the microfluidic chamber differentiated into neurons. Four microfluidic chambers from two independent rounds of differentiation were evaluated and 198 of 209 DAPI-labeled nuclei were positive for MAP2 and/ or β-tubulin III. The remaining cells were likely neural precursor cells or astrocytes as we have observed these cell types in traditional well-plate cultures. The glutamatergic lineage of these hESC-neurons was substantiated by staining for vesicular glutamate transporter 1 (VGLUT1), a marker of neuronal glutamatergic lineage (Fig. 1F).

As additional verification that these processes were functional axons, we tested whether a modified rabies virus would infect these neurons via their isolated axons. Rabies virus infection is specific to axons, requiring endocytosis and retrograde transport of viral particles following attachment to one of three viral receptors: nicotinic acetylcholine receptor, neuronal cell adhesion molecule or p75 neurotrophin receptor[26]. A modified rabies virus incapable of trans-synaptic transmission and carrying the mCherry gene[27] was exclusively added to the axonal compartment of hESC-neuron cultures for 2 hours. Once infected through the axon, fluorescent protein expression can be detected throughout the cell, including the axons and dendrites, within 48 hours. Live DIC and fluorescent imaging was performed to visualize the fine processes and fluorescent protein expression throughout the hESC-neurons cultured within the microfluidic chamber (Fig. 1G). Many of the processes were mCherry positive indicating successful rabies viral infection, further confirming the axonal identity of the vast majority of processes.

Differential gene expression between axons and neurons derived from hESCs. To identify mRNAs within human axons we harvested total RNA from isolated axons and neurons, including somata, dendrites and axons, grown in microfluidic chambers, referred to as "axonal" and "neuronal" samples. The amount of axonal RNA obtained from microfluidic chambers was below current detection limits, therefore one round of linear cDNA amplification, optimized for very small biological samples, was performed on a fixed volume of axonal RNA and a fixed mass of neuronal RNA. As a control reaction 10 pg of neuronal RNA was subjected to linear amplification in parallel. The amount of cDNA generated in this reaction suggested that the axonal RNA yield was approximately 2 pg/µL. Equivalent amounts of cDNA for all samples were processed for microarray expression profiling (Supplemental Table S1), including the 10 pg neuronal sample. The potential effect of chamber-to-chamber variability on gene expression as well as possible variation in sample collection lead us to evaluate the specificity and robustness of our expression profiling. Pearson's correlation analysis demonstrated a high degree of correlation between replicates, confirming the robustness of our methods (Fig. 2A). Further, this replicability suggests little variation in the differentiation efficiency between chambers. The 10 pg neuronal sample expression profile more closely resembled the neuronal samples than the axonal samples suggesting that the linear amplification step broadly preserved the gene expression profile of the low concentration samples (data not shown).

Increased proportional expression of mRNAs within the axon likely supports biologically relevant axonal functions that rely on local translation. Detection of these transcripts may reveal genes that have more important roles within the axon than the somatodendritic compartment. We evaluated the proportionally enriched fraction of mRNAs within our axonal

dataset by DAVID Gene Functional Classification[28, 29]. For comparison we also evaluated the proportionally depleted mRNAs within the axonal dataset. Figure 2B is a scatterplot of expression levels highlighting these fractions [3943 transcripts in the proportionally enriched fraction (Supplemental Table S2), 3631 transcripts in the proportionally depleted fraction (Supplemental Table S3)]. DAVID Gene Functional Classification of these two data subsets revealed unique functional categories which were assigned enrichment scores (Fig. 2C, D). The DAVID Gene Functional Classification enrichment score is a relative measure of the proportion of input genes that fall into a given functional term category. The proportionally enriched fraction was statistically associated with the terms “secreted” and “extracellular” proteins, as well as “neurofilament” proteins and “voltage-gated” and “cation” channels (Fig. 2C). This suggests that these protein classes have a more prominent role within the specialized axonal subcellular compartment than within the soma. The proportionally depleted fraction was significantly associated with the terms “intracellular” and “nuclear” proteins and proteins that function in RNA splicing, protein degradation and maintaining genome organization and integrity (Fig 2D), functions representative of the somatic identity of this dataset. Taken together these data suggest local translation of mRNAs differentially expressed within human axons is poised to support many axonal functions.

Abundant axonally localized transcripts of hESC-neurons functionally resemble rat cortical neuron axonal transcripts. While the axonally enriched fraction of mRNAs revealed a group of transcripts specific to axons, many of the mRNAs highly abundant within axons are not necessarily enriched within axons, for example β-actin. To assess the degree of functional similarity between the axonally abundant transcripts within both hESC-neurons and previously

published primary rodent neurons, we first created a subset of our axonal expression data containing 10% of the microarray probe sets with the highest average axonal signal intensity, representing the highest expressed transcripts within the axonal samples (Fig. 3A) [3696 transcripts, Supplemental Table S4)]. This threshold excluded known dendritic transcripts, such as ARC, CAMK2A and AMPA receptor subunits and included the well characterized axonal transcript β-actin. We compared this subset with a published dataset of mRNAs reliably localized to axons of primary embryonic rat cortical neurons grown in similar axon isolating microfluidic chambers[12]. Transcripts encoding proteins involved in translation and ribosomal proteins were significantly over-represented in the axonally abundant transcripts from both species as well as proteins of the mitochondrial respiratory chain (Fig. 3B, C). Overall, these data demonstrate that the axonally abundant transcriptome of hESC-neurons and primary rat cortical neurons are similar, despite differences in species of origin and differentiation conditions.

Conserved axonally localized mRNAs in neurons of the central and peripheral nervous systems. The extent of axonal transcriptome conservation between neuron types and species is unclear. We performed a comparative analysis of orthologous transcripts reported to be abundant within *in vitro* isolated axons of rat cortical neurons, mouse motoneurons, mouse peripheral sensory axons from dorsal root ganglia (DRG) and mixed peripheral and central axons from embryonic and adult DRGs[12-15]. We compared each previously published dataset to our subset of 3696 axonally abundant transcripts (Fig. 4A). DAVID Gene Functional Classification of the orthologous transcripts from pairwise comparison (Fig. 4B) revealed that “Translational elongation” is over-represented in each dataset, supporting a common role for translation in

axons of multiple species and neuronal types. “Intracellular”, “intracellular organelle”, and “transport” are also common between the multiple preparations. Interestingly, “synapse” proteins are enriched within the glutamatergic axon data sets from both human and rats. It is important to note that these analyses of overlapped transcripts likely under-reports the number of shared transcripts because at least 35% of the genes detected by the Affymetrix human expression arrays do not have orthologous/homologous counterparts on the Affymetrix rodent expression arrays.

DISCUSSION

Axons, once thought to be devoid of mRNA and ribosomes, are proving to contain transcripts encoding thousands of proteins and ribosomes have been positively identified in axons with the advent of sensitive detection tools[30, 31]. Functionally, intra-axonal translation is necessary for growth cone guidance, axon maintenance, injury response and may be involved in presynaptic plasticity. The field is beginning to identify locally translated mRNAs involved in these cellular events[4, 5, 7, 11, 32-42]. Identifying axonal transcripts in glutamatergic hESC-neurons has the potential to expand and deepen our knowledge of locally translated proteins and the role of local translation in axon morphology and maintenance. Many of the published axonally translated mRNAs (Supplemental Table S5) are excluded from our proportionally enriched fraction within the axon and our axonally abundant transcript group. This suggests that intra-axonal translation of moderate to low abundance axonal transcripts may also be functionally relevant.

Differential expression of the proportionally enriched and proportionally depleted fractions may be the cumulative result of differential trafficking of RNA to the axonal compartment as well as differential mRNA stability within this compartment. Nonsense mediated decay of mRNA is active in axons of mouse commissural neurons and microRNA mediated mRNA regulation is documented in axons[39, 43, 44]. The functional classification of these differentially enriched mRNAs reveals that local translation of retrogradely transported transcription factors (Fig 2C, "Sequence-specific DNA binding") is likely a common mechanism by which distal axonal events trigger a transcriptional response. Axonal translation of the transcription factors ATF4 and STAT3 are induced following injury to CNS and PNS neurons, respectively[38, 40]. Retrograde transport of these proteins to the soma mediate the transcriptional response to axon injury. Alternatively, these annotated transcription factors could have transcription-independent roles in the axon. For example, axonally synthesized β-catenin, a Wnt signal transducer, functions as a presynaptic scaffold protein[7, 38, 40, 45]. The highly enriched category of "secreted" protein transcripts suggests a significant demand on axons to dynamically modulate the extracellular environment, likely as part of axon guidance, synaptogenesis and possibly synaptic plasticity.

Our comparison of the most abundant axonal mRNAs within hESC-neurons with that of primary embryonic rat cortical neurons revealed functional similarities, specifically in mRNAs encoding proteins required for translation, ribosomal proteins and nuclear-encoded mitochondrial proteins necessary for ATP production (Fig. 3B and C). Axonal translation of nuclear encoded mitochondrial respiratory chain proteins has been described previously[46, 47]. We speculate that it is more efficient to replenish critical mitochondrial proteins via *de novo* and *in situ* translation than for the neuron to support a continual cycle of soma to axon mitochondrial transport. The

conservation of ribosomal protein transcripts in axons suggest that either the protein components of the ribosome have a shorter half-life than the RNA components or local translation of ribosomal proteins is a mechanism to dynamically regulate ribosomes and translation. There are over 100 human ribosomal proteins but only a few are constitutive ribosome components and the functions of many of these proteins are unknown. It is interesting to speculate that the unique repertoire of ribosomal proteins associated with a ribosome could confer target mRNA specificity. Ribosomes containing Ribosomal Protein L38 have been shown to preferentially interact with and translate Hox genes in mouse[48]. With the large number of ribosomal proteins it is possible that this is a common, under characterized, mechanism of translational regulation. Additionally, ribosomal proteins may mediate differential ribosome localization by regulating ribosome interaction with transmembrane proteins of organelles, such as the axonal endoplasmic reticulum, or growth factor receptors, such as the netrin-1 receptor Deleted in Colorectal Cancer (DCC)[49].

Methods to identify axonal mRNAs generate a “snap shot” of a dynamic molecular environment and with each axonal transcriptome dataset we gain a better understanding of the RNA within axons[12, 14]. Our comparison of axonal mRNAs from hESC-neurons and five primary rodent neuron types demonstrates that there are potentially thousands of transcripts uniquely abundant within axons of each (Fig. 4A). These differences could represent unique axonal requirements to support the distinctive characteristics of each neural identity. We predict that there are more axonal mRNAs in common than we detected in our analysis. First, in the five datasets generated from primary rodent tissue β-catenin mRNA (Ctnnb1) was identified as axonally abundant while in hESC-neuron axons CTNNB1 expression was below the threshold we

selected for axonal abundance (Supplemental Table S5). We suspect this is only one example of many such conserved transcripts. Secondly, the platforms used for gene expression quantification are not capable of detecting the same population of transcripts between species. The mouse arrays used by Saal *et al.* (2014) with motoneurons have 64.1% gene overlap with the human arrays used here. The rat arrays of Taylor *et al.* (2009) and Gumy *et al.* (2011) have 48.7% gene overlap with the human arrays. 91.9% of the peripheral sensory axonal transcriptome generated by Minis *et al.* (2014) using RNA-seq can be detected by the human arrays. It is possible that if the various platforms were in concordance the overlap of axonally abundant transcripts would be greater and the number of transcripts unique to hESC-neurons would be smaller. Our analysis of the transcripts in common between datasets highlights "Translational elongation" as a mechanism that likely depends on intra-axonal translation across neuron types. This category includes translational elongation factors as well as ribosomal proteins.

The question of the maturation state of stem cell derived neurons is an important one in neurodevelopmental and neurodegenerative models. Recent work by Patani *et al.* (2012)[50] evaluated the gene expression profile of H9 hESC derived midbrain dopaminergic neurons in culture for 59 days, approximately the same length of time as our cultures. They compared this to gene expression profiles of hESC derived neural precursors as well as previously published gene expression data from human fetal brain and adult brain tissue. They found that their hESC-neuron expression profile most closely resembled the human fetal brain tissue. In addition, developmentally regulated gene splicing was a significant differentiating factor between the datasets and they suggest that splice variant profiling might be a more sensitive

measure of maturation than gene expression profiling. As the techniques and algorithms necessary to quantitatively evaluate single cell transcriptomes by RNA-seq become robustly established they can be applied to axonal mRNA to evaluate the localization and relative quantities of splice variants, possibly revealing additional levels of complexity within axonal transcriptomes.

The ability to generate patient derived induced pluripotent stem cells and differentiate these cells to neurons in axon isolating microfluidic chambers will facilitate studying axon function in complex genetic diseases such as schizophrenia and autism spectrum disorders. CRISPR-mediated engineered stem cells, such as targeted deletion of RNA binding proteins associated with neurological diseases, applied to our system would provide further insight into the role of axonal translation in human neurons by allowing reverse genetic studies of the axonal transcriptome. Further, mechanistic examination of axonal translation in human axons has the potential to reveal relevant axonal mRNAs in disease and normal neuron function.

METHODS

Microfluidic chambers. Custom microfluidic chambers were fabricated by soft lithography using an established protocol[25] with the following modifications. The soma and axon compartments of these chambers were 1.5 mm by 7 mm by 450 µm tall. These compartments were connected by microgrooves of 450 µm by 10 µm by 3 µm tall. The only dimensions different from previously published microfluidic chambers[12] was the microgroove length and compartment height, which was increased to facilitate the daily media changes necessary to meet the nutrient demands of hESC-neurons. To create the tall cell compartments a thick layer of

photoresist (SU-8-2050; Microchem) was spun on a clean 3 inch silicon wafer in 2 coatings. Each coating was spun at 800 rpm for 45 s, the first coating was baked on a leveled 95°C hot plate for 3 h and the second coating was baked for 5 h. Wafers were UV-exposed (1000 mJ total over 3 sessions with at least 45 s between exposures) and baked for 1 h in a leveled 95 °C oven. Finally wafers were gradually cooled to room temperature over 30 min and developed in PGMEA.

The resulting masters were used to cast microfluidic chambers using poly(dimethylsiloxane) (PDMS) (Sylgard 184 Silicon Elastomer, Dow Corning) as described previously[51]. German glass coverslips were sterilized and coated overnight with a solution of 500-550 kDa poly-D-lysine/laminin (80 µg/ml and 10 µg/ml, respectively), washed and dried. Microfluidic chambers were assembled from PDMS devices and coated coverslips.

Maintenance and Differentiation of human ESC-neurons. Stem cell differentiation into neural progenitor cells (neuroepithelial cells) was performed according to previously published modifications of the original protocol[23, 24, 52]. Days *in vitro* (DIV) numbering began when H9 human embryonic stem cells were plated in embroid body media to induce neuronal differentiation.

Maturation of hESC-neurons in microfluidic chambers. On DIV24 neural progenitor cells were manually lifted and dissociated manually or with Accutase (Life Technologies). Approximately $5x10^3$ cells were seeded into the soma compartment of microfluidic chambers in N2B27 media supplemented with 100 ng/ml human recombinant brain derived neurotrophic factor (BDNF). Half of the media (100 µl) from each compartment was changed daily and cells were matured for at least 35 days (DIV59). This protocol has been demonstrated to generate glutamatergic

neurons as early as DIV25, as verified by staining for the telencephalic transcription factor FOXG1, the transcription factor CTIP specific for subcerebral projection neurons, the glutamatergic transcription factor TBR1 and the glutamatergic marker protein VGLUT1[23, 24].

Immunocytochemistry. Human ESC-neurons grown in axon isolating microfluidic chambers were fixed, stained and mounted within the microfluidic chamber because the cells were susceptible to damage when the chambers were removed. Using this method the neuronal processes within the microgrooves were not exposed to antibodies due to the fluidic isolation of the two compartments. Chambers were fixed with freshly prepared 4% paraformaldehyde in PBS containing 40 mg/ml sucrose, 1 µm MgCl2, and 0.1 µm CaCl2 for 45 min. Neurons were permeabilized in 0.25% Triton X-100 for 15 min then blocked in PBS containing 10% goat serum for 15 min. Primary antibodies to β-tubulin (1:2000; chicken; Aves Labs), MAP2 (1:1000; rabbit; Millipore) and VGLUT1 (1:200; mouse; NeuroMab) were diluted in PBS with 1% goat serum and incubated overnight at 4 °C. AlexaFluor goat secondary antibodies conjugated to fluorophores with 488 nm, 568 nm, or 647 nm excitation wavelengths (1:1000; Invitrogen) were diluted in PBS and incubated for 1 h at room temperature. Cells were counterstained with DAPI.

Modified rabies virus infection. The axonal compartment of DIV61 cultures were incubated with 100,000 viral units of modified rabies virus carrying the mCherry gene[27] in a total of 50 µl media for 2 hours at 37 °C, washed twice with fresh media and live imaging was performed 48 hours later. We have observed within 48 hours this viral load results in detectable mCherry expression within 80-85% of hESC-neuron axons in microfluidic chambers. Similar observations were made with primary embryonic rat hippocampal neurons grown in axon isolating microfluidic chambers (data not shown). The presence of mCherry negative axons within the

axon compartment could arise from variability in the time course of mCherry expression. In primary rat hippocampal cultures the minimum amount to time to detect fluorescent protein is 48 hours but some cells do not express detectable mCherry until 4 days after transfection.

Confocal imaging. Z-stack fluorescent images and single plane differential contrast (DIC) images were acquired using a spinning disk confocal imaging system (Revolution XD, Andor Technology) configured for an Olympus IX81 zero-drift microscope and spinning disk unit (CSU-X1, Yokogawa). Light excitation was provided by 50 mW, 488 nm; 50 mW, 561 nm; and 100 mW, 640 nm lasers. The following bandpass emission filters (BrightLine, Semrock) were used: 525/30 nm (TR-F525-030), 607/36 nm (TR-F607-036), 685/40 nm (TR-F685-040). All slides within a series were imaged during a single session in which image capture settings were identical for all fields.

Image analysis. Z-stack fluorescent images were MAX projected. 16-bit images were manually thresholded to the same values within each series of stained chambers.

Determination of differentiation efficiency. Seven 20x fields of hESC-neurons in monolayer from 4 independently stained chambers of at least DIV39 were used to estimate the efficiency of differentiation to mature hESC-neurons as detected by colocalization of DAPI signal with MAP2 or β-tubulin III positive staining.

RNA isolation. On DIV59 isolation of RNA from microfluidic cultures using the RNAqueous-Micro Kit (Ambion) was performed as previously described[12, 25]. To collect pure axonal RNA continuous aspiration was applied to the soma compartment to prevent somatic RNA from entering the sample. Most of the axonal media was removed and lysis buffer was added to one

well, the reagent that flowed to the other axonal well through the axonal compartment was collected as the axonal RNA sample. Subsequent RNA purification was performed according to manufacturer's instructions and RNA was eluted in a final volume of 10 µl. Neuronal RNA was collected from the soma compartment of hESC-neurons grown in microfluidic chambers not used for axonal RNA collection, this compartment contained somata, dendrites and axons which did not pass through the microgrooves. The majority of the media from the soma compartment was removed and the same RNA isolation procedure was followed, omitting constant aspiration. Each sample was obtained from one microfluidic chamber. Samples were stored at -80°C until further processing by the UNC Lineberger Comprehensive Cancer Center Genomics Core and the UNC School of Medicine Functional Genomics Core. Three axonal and three neuronal samples were collected and analyzed.

RNA amplification and microarray. Axonal RNA harvested from microfluidic chambers is below the detection limit of currently available technology. A fixed volume of each axonal RNA sample (5 µL) and a fixed mass of each neuronal RNA sample (250 pg) were subjected to one round of linear amplification using the Ovation One-Direct System (NuGEN). This system initiates amplification at the 3' end as well as randomly throughout the transcript and is optimized for very small biological samples such as single cell transcriptomics. The integrity and concentration of the resulting cDNA samples were quantified with an Agilent Bioanalyzer 2100 at the UNC Lineberger Comprehensive Cancer Center Genomics Core. After amplification axonal samples were 230.5 ng/µl, 305.0 ng/µl and 199.0 ng/µl. Neuronal samples were 519.7 ng/µl, 497.3 ng/µl and 469.9 ng/µl. As a control reaction, 10 pg of neuronal RNA was amplified resulting in a concentration of 252.1 ng/µl, suggesting our original axonal RNA yield ranged from

approximately 1.6 to 2.4 pg/µl. A fixed concentration of amplified cDNA from all samples, including the 10 pg neuronal sample, were processed in parallel for microarray analysis using the Encore Biotin Module (NuGEN) at the UNC School of Medicine Functional Genomics Core.

Human Gene 2.0 ST Arrays (Affymetrix) were used and scanned with an Affymetrix GeneChip Scanner 3000 7G Plus with Autoloader. Affymetrix Expression Console software was used to analyze microarray CEL files. Affymetrix normalization controls, built into the microarray chips and analysis software, were used for raw data normalization. The Robust Multichip Analysis (RMA) algorithm was used for global background adjustment, quantile normalization and gene expression summarization between samples. This method is sensitive to small changes between samples without negatively affecting signal variance. Affymetrix Transcriptome Analysis Console v2.0 software algorithms were used to determine differential expression and statistical analysis using one-way between-subject ANOVA of normalized intensities (Supplemental Table S1 online).

Data analysis. Microarray probe sets with fold change linear (Axon vs Neuron) greater than or equal to 1.5 and an ANOVA p-value (Axon vs Neuron) less than 0.05 were designated the axonally enriched fraction (3943 probe sets). Probe sets with fold change linear (Axon vs Neuron) less than or equal to -1.5 and an ANOVA p-value (Axon vs Neuron) less than 0.05 were designated the axonally depleted fraction (3631 probe sets). The fold change threshold of 1.5 was selected to counter balance the compressed calculated fold change values obtained by the RMA method. This threshold generated gene symbol lists of an appropriate size for DAVID Functional Annotation Clustering (<3000). Official gene symbols were extracted, duplicates were removed and gene symbol lists (3297 for axonally enriched fraction and 3638 for axonally

depleted fraction, Supplemental Tables S2 and S3 online, respectively) were submitted to DAVID Bioinformatics Resources for Gene ID Conversion and Functional Annotation Clustering with the single species *Homo sapiens* selected and the default human background list[28, 29].

To obtain a gene symbol list of axonally abundant transcripts within hESC-neuron axons microarray probe sets with Axon bi-weight average signal (log2) equal to or greater than 6.48 were selected. This threshold excluded known dendritic transcripts and included β-actin, a well characterized axonal mRNA. These criteria were similar to those used by Taylor *et al.* (2009) to determine axonal mRNAs within primary rat cortical neurons. Official gene symbols were extracted, duplicates were removed and the gene symbol list (3696 gene symbols, Supplemental Table S4 online) was submitted to DAVID Bioinformatics Resources for Gene ID Conversion and Functional Annotation Clustering with the single species *Homo sapiens* selected and the default human background list[28, 29].

This list of 3696 human gene symbols was compared to published gene symbols of axonally localized transcripts from previous rodent derived axonal transcriptomes[12-15] obtained through the NCBI Gene Expression Omnibus or in published supplementary data. GEO numbers of datasets used are: GSE11730 Taylor *et al.* (2009), GSE59506 Saal *et al.* (2014), GSE51572 Minis *et al.* (2014) and GSE22638 Gumy *et al.* (2011). Axonally abundant transcripts as reported by Taylor *et al.* (2009) and Gumy *et al.* (2011) were used. We designated the 4000 highest expressed axonal RNAs from fully processed RNA-seq data by Minis *et al.* (2014) as axonally abundant. Processed microarray CEL and chp files from Saal *et al.* (2014) were analyzed using Affymetrix Transcriptome Analysis Console v2.0 software algorithms to determine differential expression and statistical analysis using one-way between-subject ANOVA of normalized

intensities. Axon abundant transcript thresholding of this mouse motoneuron data was performed using the same criteria as for the hESC-neuron data generating a list of 3751 unique gene symbols.

Evaluation of microarray probe set similarity. The degree of concordance between microarray platforms (Human Gene 2, Mouse Genome 430 2.0 and Rat Expression 230 2.0) was determined using the Affymetrix Expression Array Comparison Tool available at www.affymetrix.com. To determine the concordance between the Human Gene 2 microarray and the RNA-seq data of Minis et al. (2014), the list of 4000 axonally abundant transcripts within these peripheral sensory axons described above was submitted as a batch query against the Human Gene 2 array. The Rat Expression 230 2.0 array has an orthologous/homologous gene overlap with the Human Gene 2 array of 48.7%. The Mouse Genome 230 2.0 array has an orthologous/homologous gene overlap with the Human Gene 2 array of 64.1%. The majority of the genes that are not shared between human and rat or human and mouse are unique to the Human Gene 2.0 array though there are a few thousand transcripts detected by each of the rodent arrays which are not represented on the Human Gene 2.0 array.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the CHANL facility at UNC-CH. R.L.B. was supported in part by a grant from the National Institute of General Medical Sciences under award 5T32 GM007092. A.M.T. is an Alfred P. Sloan Research Fellow. A.M.T. acknowledges support from the NIH (R41MH097377) and the Simons Foundation (SFARI #236390).

AUTHOR CONTRIBUTIONS

R.L.B., J.W.K., R.D., M.N. and A.M.T. designed and performed experiments. R.L.B. and A.M.T. prepared the figures and wrote the manuscript.

ADDITIONAL INFORMATION

Competing financial interests: Yes there is potential competing financial interest. A.M.T. is an inventor of the microfluidic chambers (US 7419822 B2) and has financial interest in Xona Microfluidics, LLC. R.B., J.K., R.D., and M.N. declare no competing financial interest.

Accession codes: All microarray data have been deposited to the NCBI Gene Expression Omnibus under accession number X.

REFERENCES

1. Holt, C. E. & Schuman, E. M. The central dogma decentralized: new perspectives on RNA function and local translation in neurons. *Neuron* **80**, 648-657 (2013).

2. Lyles, V., Zhao, Y. & Martin, K. C. Synapse formation and mRNA localization in cultured Aplysia neurons. *Neuron* **49**, 349-356 (2006).

3. Donnelly, C. J., Fainzilber, M. & Twiss, J. L. Subcellular communication through RNA transport and localized protein synthesis. *Traffic* **11**, 1498-1505 (2010).

4. Merianda, T. T. *et al*. Axonal amphoterin mRNA is regulated by translational control and enhances axon outgrowth. *J. Neurosci.* **35**, 5693-5706 (2015).

5. Aschrafi, A. *et al*. MicroRNA-338 regulates local cytochrome c oxidase IV mRNA levels and oxidative phosphorylation in the axons of sympathetic neurons. *J. Neurosci.* **28**, 12581-12590 (2008).

6. Goldman, J. S. *et al*. Netrin-1 promotes excitatory synaptogenesis between cortical neurons by initiating synapse assembly. *J. Neurosci.* **33**, 17278-17289 (2013).

7. Taylor, A. M., Wu, J., Tai, H. C. & Schuman, E. M. Axonal translation of beta-catenin regulates synaptic vesicle dynamics. *J. Neurosci.* **33**, 5584-5589 (2013).

8. Hsu, W. L. *et al*. Glutamate Stimulates Local Protein Synthesis in the Axons of Rat Cortical Neurons by Activating AMPA Receptors and Metabotropic Glutamate Receptors. *J. Biol. Chem.* (2015).

9. Kang, H. & Schuman, E. M. A requirement for local protein synthesis in neurotrophin-induced hippocampal synaptic plasticity. *Science* **273**, 1402-1406 (1996).

10. Genheden, M. *et al*. BDNF stimulation of protein synthesis in cortical neurons requires the MAP kinase-interacting kinase MNK1. *J. Neurosci.* **35**, 972-984 (2015).

11. Welshhans, K. & Bassell, G. J. Netrin-1-induced local beta-actin synthesis and growth cone guidance requires zipcode binding protein 1. *J. Neurosci.* **31**, 9800-9813 (2011).

12. Taylor, A. M. *et al*. Axonal mRNA in uninjured and regenerating cortical mammalian axons. *J. Neurosci.* **29**, 4697-4707 (2009).

13. Saal, L., Briese, M., Kneitz, S., Glinka, M. & Sendtner, M. Subcellular transcriptome alterations in a cell culture model of spinal muscular atrophy point to widespread defects in axonal growth and presynaptic differentiation. *RNA* **20**, 1789-1802 (2014).

14. Gumy, L. F. *et al*. Transcriptome analysis of embryonic and adult sensory axons reveals changes in mRNA repertoire localization. *RNA* **17**, 85-98 (2011).

15. Minis, A. *et al*. Subcellular transcriptomics-dissection of the mRNA composition in the axonal compartment of sensory neurons. *Dev. Neurobiol.* **74**, 365-381 (2014).

16. Moccia, R. *et al*. An unbiased cDNA library prepared from isolated Aplysia sensory neuron processes is enriched for cytoskeletal and translational mRNAs. *J. Neurosci.* **23**, 9409-9417 (2003).

17. Cajigas, I. J. *et al*. The local transcriptome in the synaptic neuropil revealed by deep sequencing and high-resolution imaging. *Neuron* **74**, 453-466 (2012).

18. Hu, W. *et al*. Derivation, Expansion, and Motor Neuron Differentiation of Human-Induced Pluripotent Stem Cells with Non-Integrating Episomal Vectors and a Defined Xenogeneic-free Culture System. *Mol. Neurobiol.* (2015).

19. Reddington, A. E., Rosser, A. E. & Dunnett, S. B. Differentiation of pluripotent stem cells into striatal projection neurons: a pure MSN fate may not be sufficient. *Front. Cell. Neurosci.* **8**, 398 (2014).

20. Duan, L., Peng, C. Y., Pan, L. & Kessler, J. A. Human pluripotent stem cell-derived radial glia recapitulate developmental events and provide real-time access to cortical neurons and astrocytes. *Stem Cells Transl. Med.* **4**, 437-447 (2015).

21. Baronchelli, S. *et al*. Investigating DNA methylation dynamics and safety of human embryonic stem cell differentiation towards striatal neurons. *Stem Cells Dev.* (2015).

22. Sagal, J. *et al*. Proneural transcription factor Atoh1 drives highly efficient differentiation of human pluripotent stem cells into dopaminergic neurons. *Stem Cells Transl. Med.* **3**, 888-898 (2014).

23. Li, X. J. *et al*. Coordination of sonic hedgehog and Wnt signaling determines ventral and dorsal telencephalic neuron types from human embryonic stem cells. *Development* **136**, 4055-4063 (2009).

24. Zeng, H. *et al*. Specification of region-specific neurons including forebrain glutamatergic neurons from human induced pluripotent stem cells. *PLoS One* **5**, e11853 (2010).

25. Taylor, A. M. *et al*. A microfluidic culture platform for CNS axonal injury, regeneration and transport. *Nat. Methods* **2**, 599-605 (2005).

26. Lafon, M. Rabies virus receptors. *J. Neurovirol.* **11**, 82-87 (2005).

27. Wickersham, I. R., Finke, S., Conzelmann, K. K. & Callaway, E. M. Retrograde neuronal tracing with a deletion-mutant rabies virus. *Nat. Methods* **4**, 47-49 (2007).

28. Huang da, W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* **37**, 1-13 (2009).

29. Huang da, W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44-57 (2009).

30. Koenig, E. & Martin, R. Cortical plaque-like structures identify ribosome-containing domains in the Mauthner cell axon. *J. Neurosci.* **16**, 1400-1411 (1996).

31. Koenig, E., Martin, R., Titmus, M. & Sotelo-Silveira, J. R. Cryptic peripheral ribosomal domains distributed intermittently along mammalian myelinated axons. *J. Neurosci.* **20**, 8390-8400 (2000).

32. Eng, H., Lund, K. & Campenot, R. B. Synthesis of beta-tubulin, actin, and other proteins in axons of sympathetic neurons in compartmented cultures. *J. Neurosci.* **19**, 1-9 (1999).

33. Perlson, E. *et al*. Vimentin-dependent spatial translocation of an activated MAP kinase in injured nerve. *Neuron* **45**, 715-726 (2005).

34. Wu, K. Y. *et al*. Local translation of RhoA regulates growth cone collapse. *Nature* **436**, 1020-1024 (2005).

35. Yudin, D. *et al*. Localized regulation of axonal RanGTPase controls retrograde injury signaling in peripheral nerve. *Neuron* **59**, 241-252 (2008).

36. Andreassi, C. *et al*. An NGF-responsive element targets myo-inositol monophosphatase-1 mRNA to sympathetic neuron axons. *Nat. Neurosci.* **13**, 291-301 (2010).

37. Perry, R. B. *et al*. Subcellular knockout of importin beta1 perturbs axonal retrograde signaling. *Neuron* **75**, 294-305 (2012).

38. Ben-Yaakov, K. *et al*. Axonal transcription factors signal retrogradely in lesioned peripheral nerve. *EMBO J.* **31**, 1350-1363 (2012).

39. Colak, D., Ji, S. J., Porse, B. T. & Jaffrey, S. R. Regulation of axon guidance by compartmentalized nonsense-mediated mRNA decay. *Cell* **153**, 1252-1265 (2013).

40. Baleriola, J. *et al*. Axonally synthesized ATF4 transmits a neurodegenerative signal across brain regions. *Cell* **158**, 1159-1172 (2014).

41. Moretti, F. *et al*. Growth Cone Localization of the mRNA Encoding the Chromatin Regulator HMGN5 Modulates Neurite Outgrowth. *Mol. Cell. Biol.* **35**, 2035-2050 (2015).

42. Calliari, A. *et al*. Myosin Va is locally synthesized following nerve injury. *Cell Motil. Cytoskeleton* **51**, 169-176 (2002).

43. Hengst, U., Cox, L. J., Macosko, E. Z. & Jaffrey, S. R. Functional and selective RNA interference in developing axons and growth cones. *J. Neurosci.* **26**, 5727-5732 (2006).

44. Natera-Naranjo, O., Aschrafi, A., Gioio, A. E. & Kaplan, B. B. Identification and quantitative analyses of microRNAs located in the distal axons of sympathetic neurons. *RNA* **16**, 1516-1529 (2010).

45. Bamji, S. X. *et al*. Role of beta-catenin in synaptic vesicle localization and presynaptic assembly. *Neuron* **40**, 719-731 (2003).

46. Gioio, A. E. *et al*. Local synthesis of nuclear-encoded mitochondrial proteins in the presynaptic nerve terminal. *J. Neurosci. Res.* **64**, 447-453 (2001).

47. Hillefors, M., Gioio, A. E., Mameza, M. G. & Kaplan, B. B. Axon viability and mitochondrial function are dependent on local protein synthesis in sympathetic neurons. *Cell. Mol. Neurobiol.* **27**, 701-716 (2007).

48. Kondrashov, N. *et al*. Ribosome-mediated specificity in Hox mRNA translation and vertebrate tissue patterning. *Cell* **145**, 383-397 (2011).

49. Tcherkezian, J., Brittis, P. A., Thomas, F., Roux, P. P. & Flanagan, J. G. Transmembrane receptor DCC associates with protein synthesis machinery and regulates translation. *Cell* **141**, 632-644 (2010).

50. Patani, R. *et al*. Investigating the utility of human embryonic stem cell-derived neurons to model ageing and neurodegenerative disease using whole-genome gene expression and splicing analysis. *J. Neurochem.* **122**, 738-751 (2012).

51. Taylor, A. M. *et al*. Microfluidic Multicompartment Device for Neuroscience Research. *Langmuir* **19**, 1551-1556 (2003).

52. Niedringhaus, M. *et al*. Transferable neuronal mini-cultures to accelerate screening in primary and induced pluripotent stem cell-derived neurons. *Sci. Rep.* **5**, 8353 (2015).

FIGURE 1

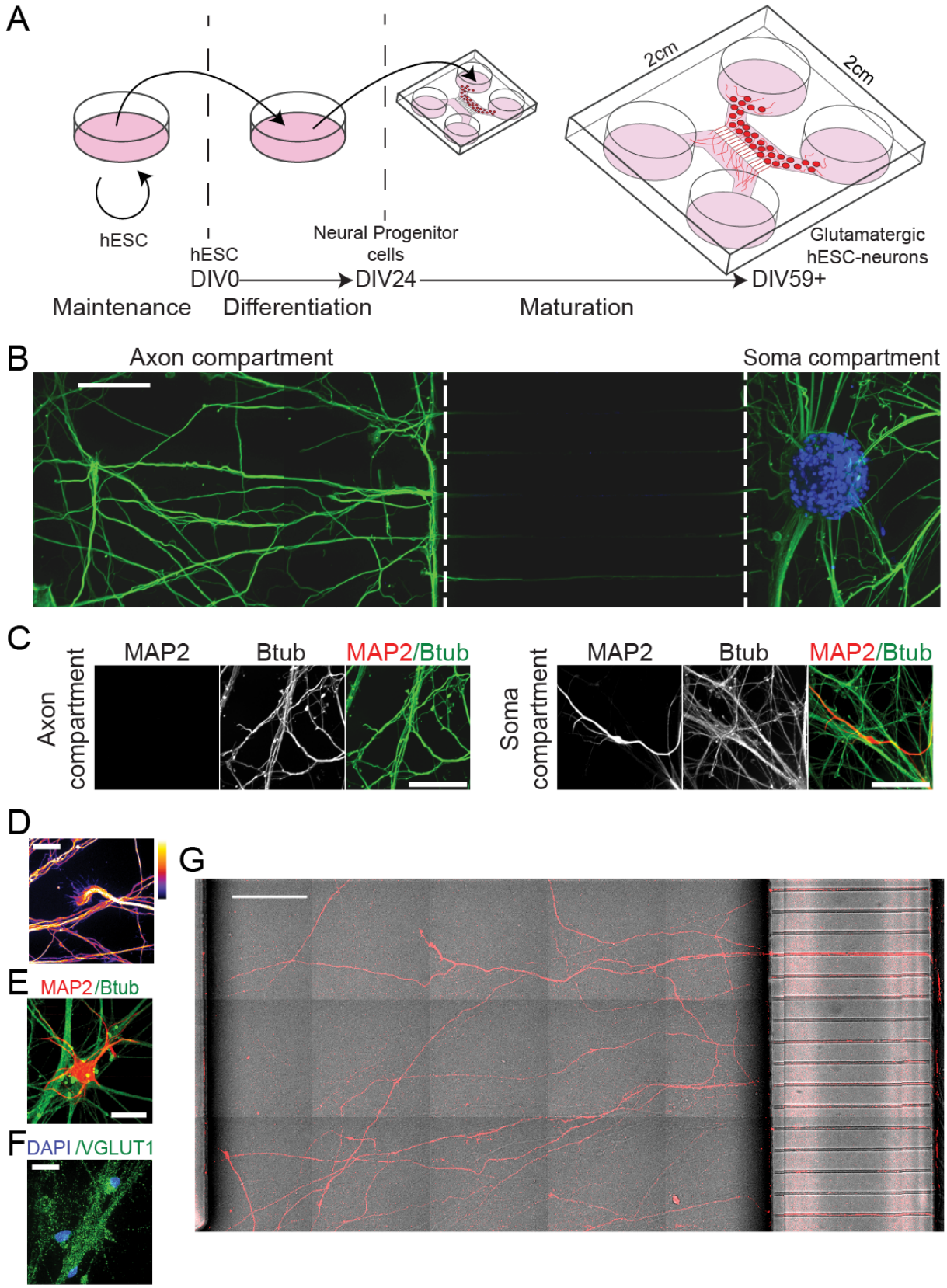

Figure 1. Human embryonic stem cell derived neurons (hESC-neurons) matured in axon isolating microfluidic chambers. (A) Schematic of hESC-neuron differentiation and maturation in microfluidic chambers. (B) Representative montage image of β-tubulin III (Btub, green) immunostaining spanning the soma and axon compartments of DIV58 hESC-neurons cultured within a microfluidic chamber. Nuclei are counterstained with DAPI (blue). hESC-neurons were fixed, stained and mounted within the microfluidic chamber limiting the penetration of antibodies into the microgrooves due to the fluidic isolation between the two compartments. White dashed bars delineate the boundaries of the microgroove barrier. (C) Representative fluorescent images taken within the axon and soma compartments show that MAP2 positive somata and dendrites were restricted to the soma compartment while Btub axons were detected in both compartments. (D) Representative image of a Btub stained hESC-neuron growth cone within the axon compartment, “Fire” LUT pseudocolor. (E) hESC-neurons exhibited canonical dendritic arborization as revealed by MAP2 immunostaining. (F) Mature hESC-neurons were positive for the glutamatergic marker VGLUT1. (G) Axons of live hESC-neurons were infected with a modified rabies virus encoding the mCherry fluorescent protein and imaged 2 days after infection (DIV63). The presence of a small percentage of mCherry negative axons within the axon compartment could arise from variability in the time course of mCherry expression. Scale bars are 100 µm (B, C), 25 µm (D, E, F) and 200 µm (G).

FIGURE 2

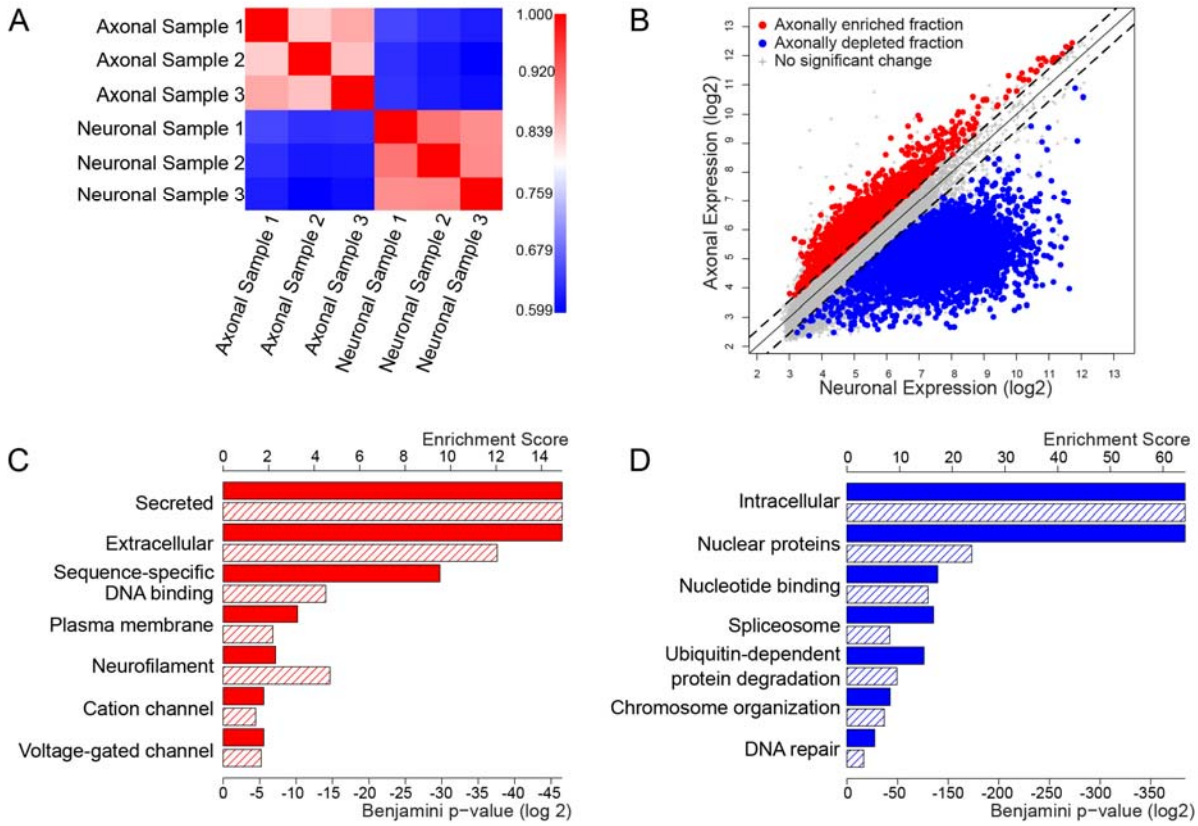

Figure 2. Differential gene expression between the axonal transcriptome and the neuronal transcriptome of hESC-neurons. Axonal and neuronal mRNAs from hESC-neurons grown in axon isolating microfluidic chambers were evaluated using Affymetrix microarrays. (A) Pearson correlation coefficient of microarray data from axonal and neuronal samples (B) Gene expression scatterplot comparing expression levels in hESC-neurons and axons. Threshold for proportional enrichment or depletion within the axonal fraction was set at ±1.5 fold change, ANOVA p-value < 0.05. The proportionally enriched fraction is shown in red. The proportionally depleted fraction is shown in blue. (C) Gene ontology (GO) enrichment scores (solid red bars, scale along the top) and Benjamini p-values (hashed red bars, scale along the bottom) of the proportionally enriched fraction as determined by DAVID Gene Functional Classification. (D) GO enrichment scores (solid blue bars, scale along the top) and Benjamini p-values (hashed blue bars, scale along the bottom) of the proportionally depleted fraction as determined by DAVID Gene Functional Classification.

FIGURE 3

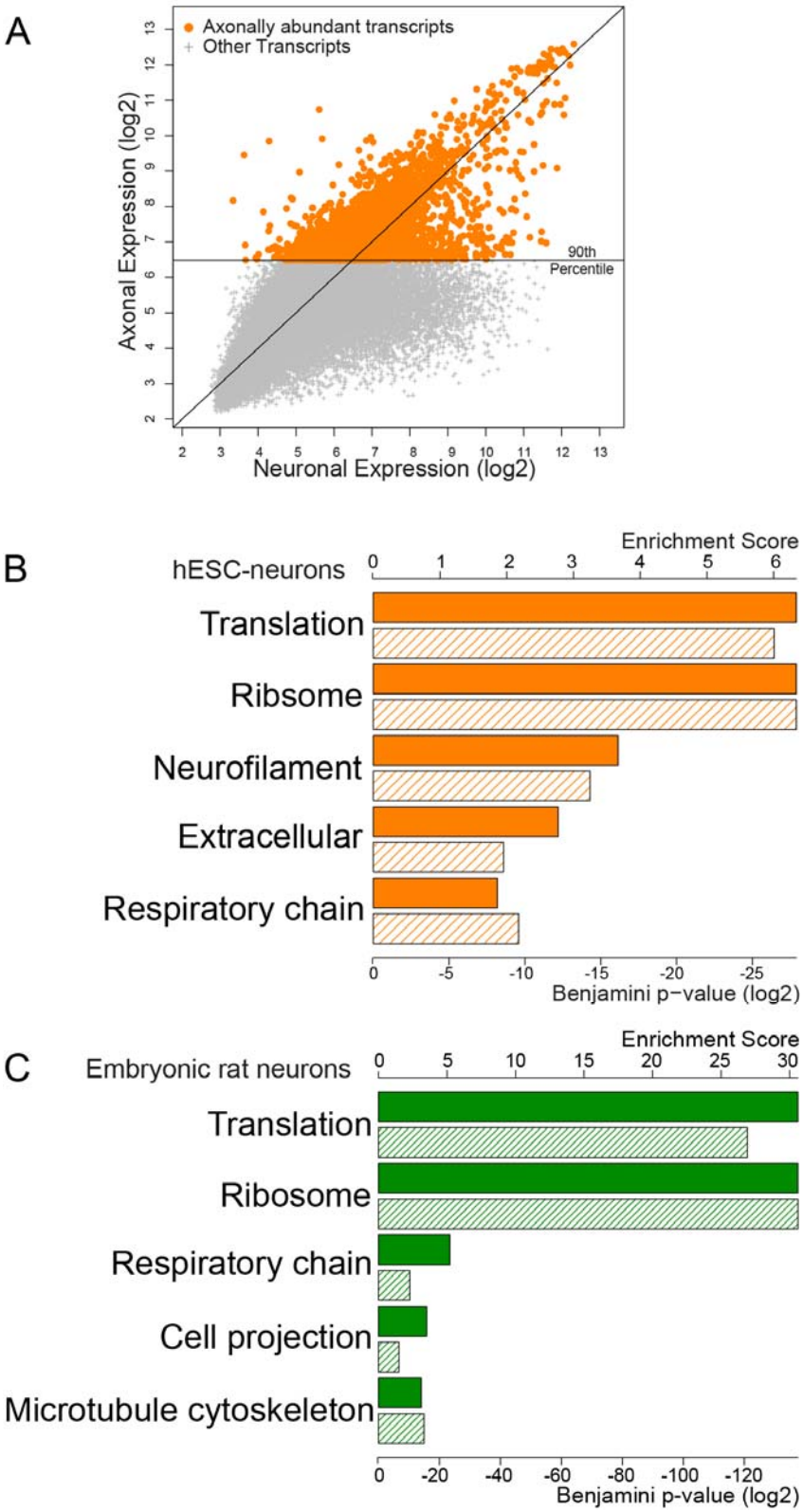

Figure 3. The axonal transcriptome of hESC-neurons is similar to embryonic rat cortical neurons. (A) hESC-neuron gene expression scatterplot highlighting axonally abundant transcripts. (B) GO enrichment scores (solid orange bars, scale along the top) and p-values (hashed orange bars, scale along the bottom) of axonally abundant transcripts as determined by DAVID Gene Functional Classification. (C) GO enrichment scores (solid green bars, scale along the top) and p-values (hashed green bars, scale along the bottom) of axonal transcripts from rat cortical neurons[12] as determined by DAVID Gene Functional Classification.

FIGURE 4

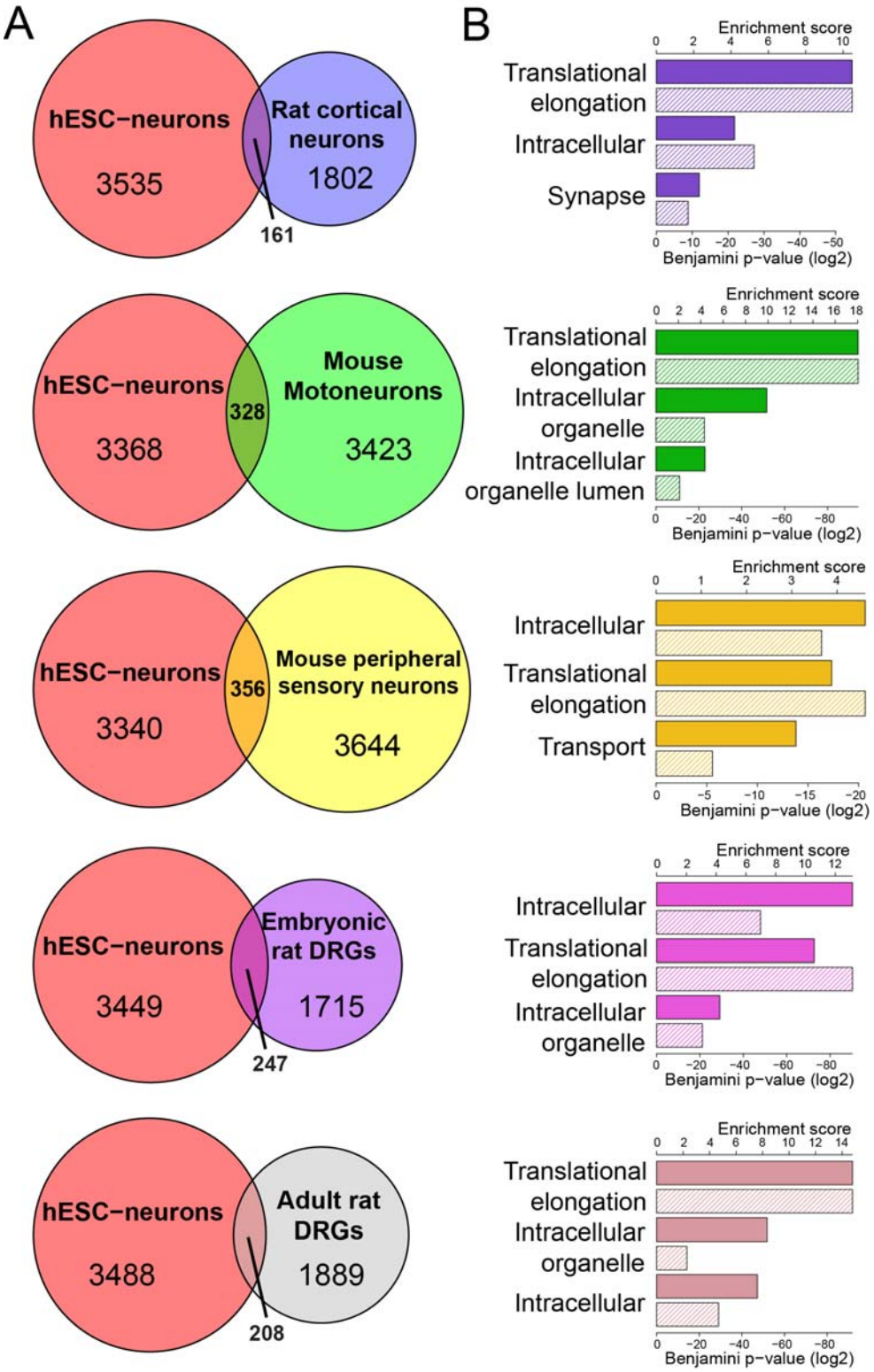

Figure 4. Conservation of axonal transcripts across species and neuron types. (A) Pairwise Venn diagrams of orthologous transcripts between hESC-neuron axonally abundant transcripts and axonal transcripts from five primary rodent axon types[12-15]. Values within the non-overlap areas indicate the number of unique genes in each pairwise comparison. Values within the overlap area indicate the number of axonally abundant transcripts in common. (B) GO enrichment scores (solid bars, scale along the top) and Benjamini p-values (hashed bars, scale along the bottom) as determined by DAVID Gene Functional Classification of common transcripts from each pairwise comparison.