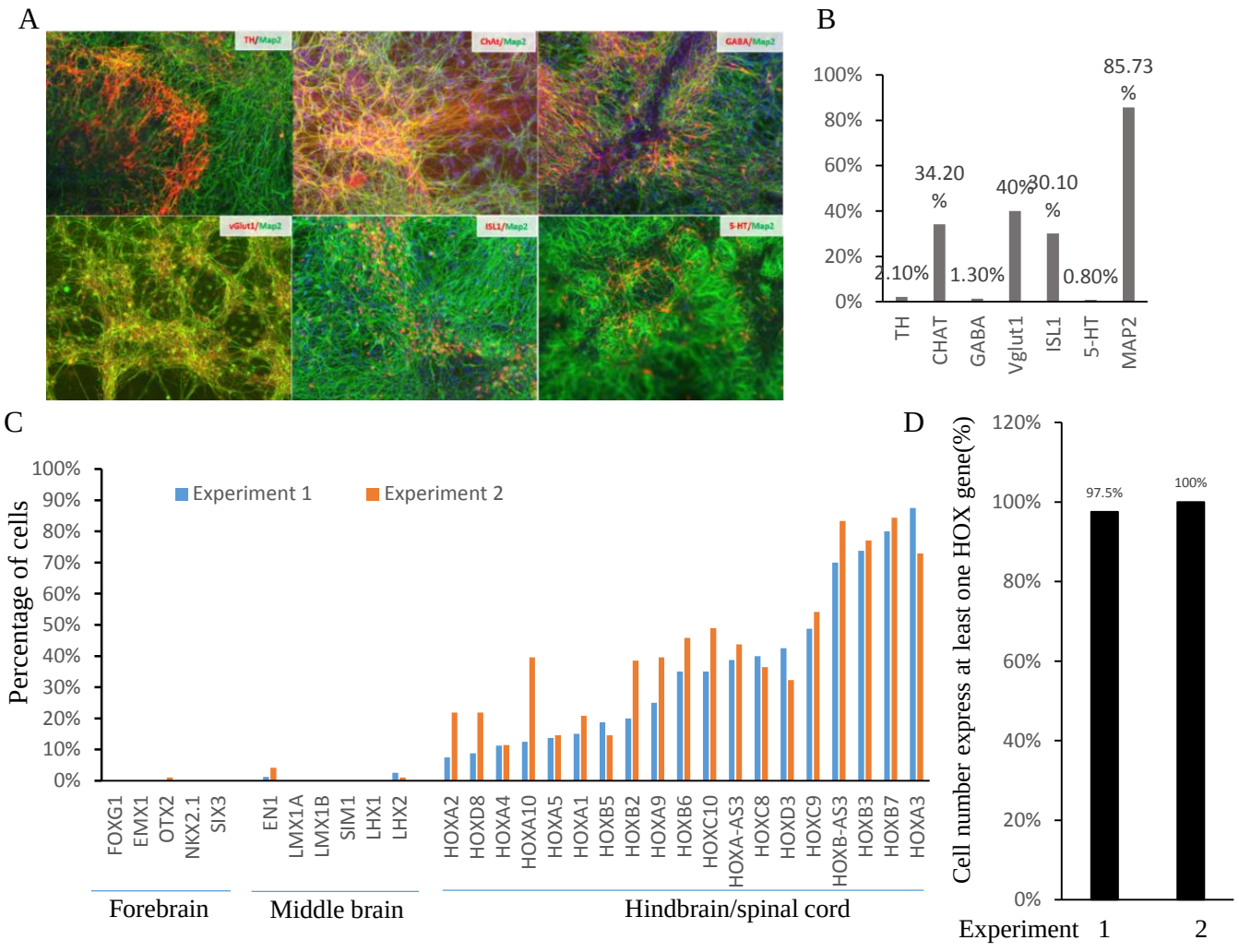
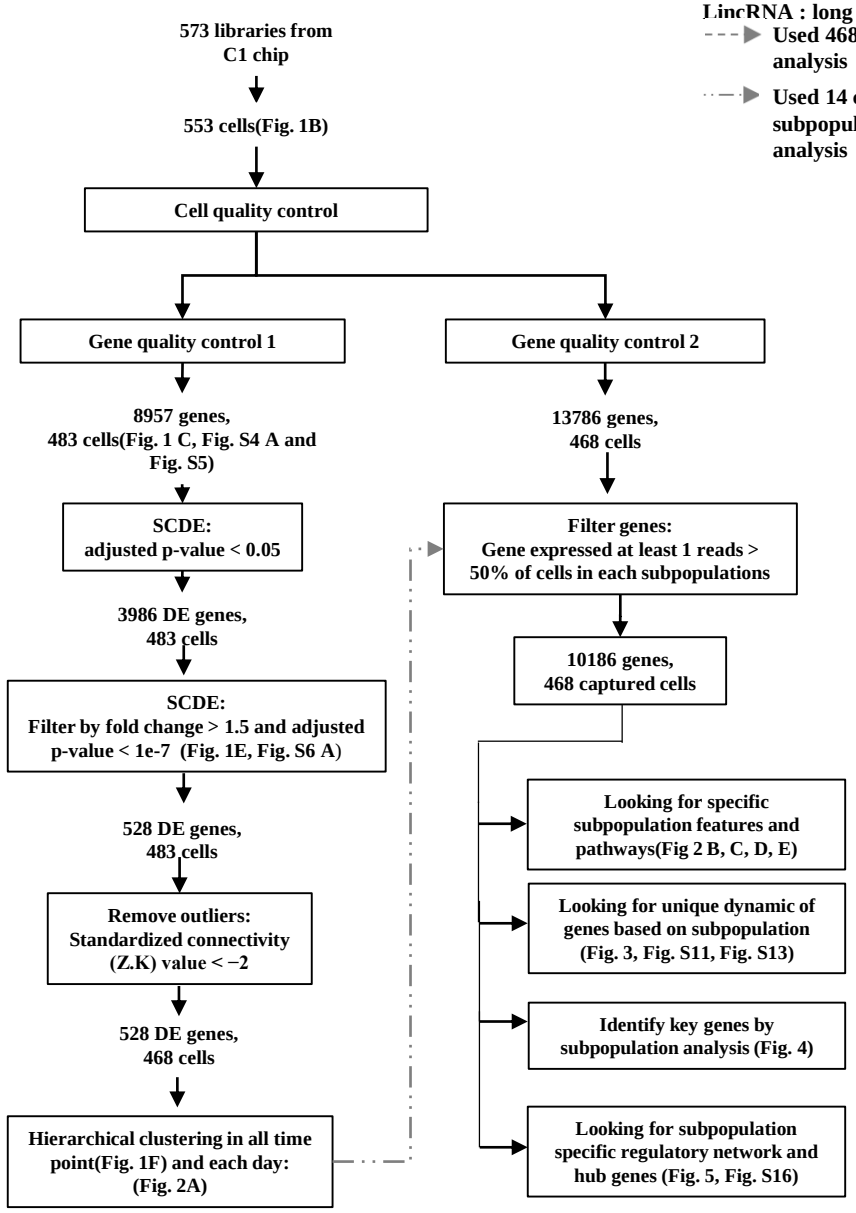
Fig_S1

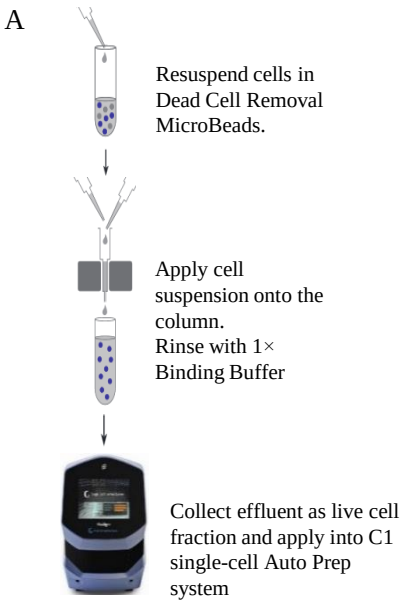


A) Immunostaining shows the expression of a general neuronal marker (Map2) and markers that are specific for neuronal subtypes: dopaminergic neurons (*TH*); motor neurons (*ChAt, ISL1*); glutamatergic neurons (*vGlut1*); and serotonergic neurons (*5-HT*). Cultured cells were stained 30 days after NPC were subjected to non-specific neural differentiation conditions. B) Immunostained images as shown in (A) were quantified by Columbus Analysis System to determine the efficieincy of neuron differentiation. C) Percentage of neuorns expressing the indicated markers of forebrain, midbrain and hindbrain/spinal cord neural cells were determined for differentiaiton experiments 1 and 2. (D) Percentage of cells expressing at least one *HOX* gene in experiment 1 and 2.

Fig_S2


Note:
DE: differential expressed
SCDE: single cell differential expression
FC: fold change
TFs: transcription factors
LincRNA : long –noncoding RNA
Used 468 cells in further analysis
Used 14 clusters as cell subpopulations in further analysis
573 libraries from C1 chip
553 cells(Fig. 1B)
Cell quality control
Gene quality control 1
Gene quality control 2
8957 genes, 483 cells(Fig. 1 C, Fig. S4 A and Fig. S5)
13786 genes, 468 cells
SCDE: adjusted p-value < 0.05
Filter genes: Gene expressed at least 1 reads > 50% of cells in each subpopulations
3986 DE genes, 483 cells
10186 genes, 468 captured cells
SCDE: Filter by fold change > 1.5 and adjusted p-value < 1e-7 (Fig. 1E, Fig. S6 A)
528 DE genes, 483 cells
Remove outliers: Standardized connectivity (Z.K) value < −2
528 DE genes, 468 cells
Hierarchical clustering in all time point(Fig. 1F) and each day: (Fig. 2A)
Looking for specific subpopulation features and pathways(Fig 2 B, C, D, E)
Looking for unique dynamic of genes based on subpopulation (Fig. 3, Fig. S11, Fig. S13)
Identify key genes by subpopulation analysis (Fig. 4)
Looking for subpopulation specific regulatory network and hub genes (Fig. 5, Fig. S16)

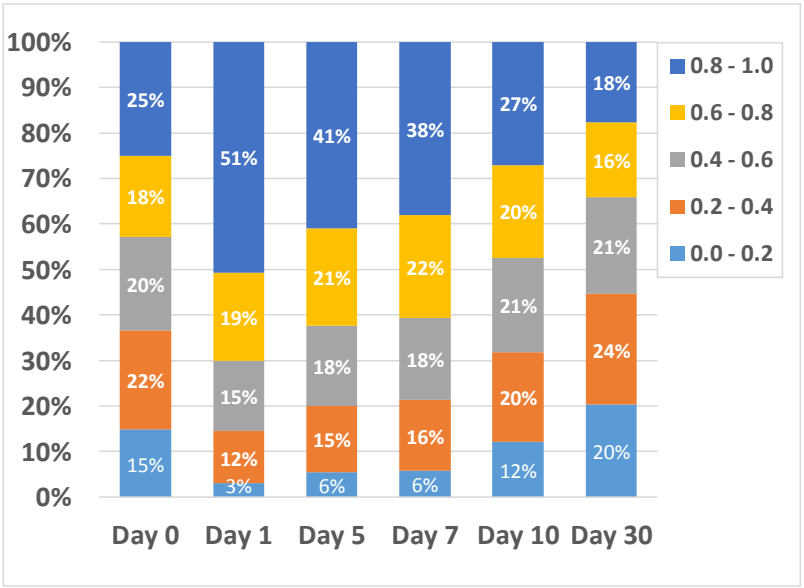
Fig_S3



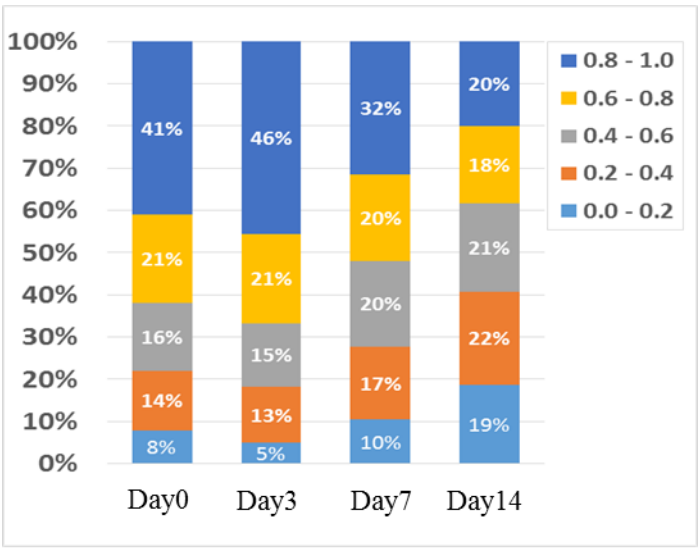
A) Outline of the workflow for experiment 2. Harvested cells at day 0, day 3, day 7 and day 14 were filtered of dead cells and then loaded on the C1 Fluidigm system for single cell capture. This is a second independent differentiation experiment that provided a replicate set of single cell RNA sequence data. B) After sequencing, similar analysis as in Fig. 1C was performed to determine the total number of genes expressed per cell at the indicated time points. Each dot represent one cell.

Fig_S4

A

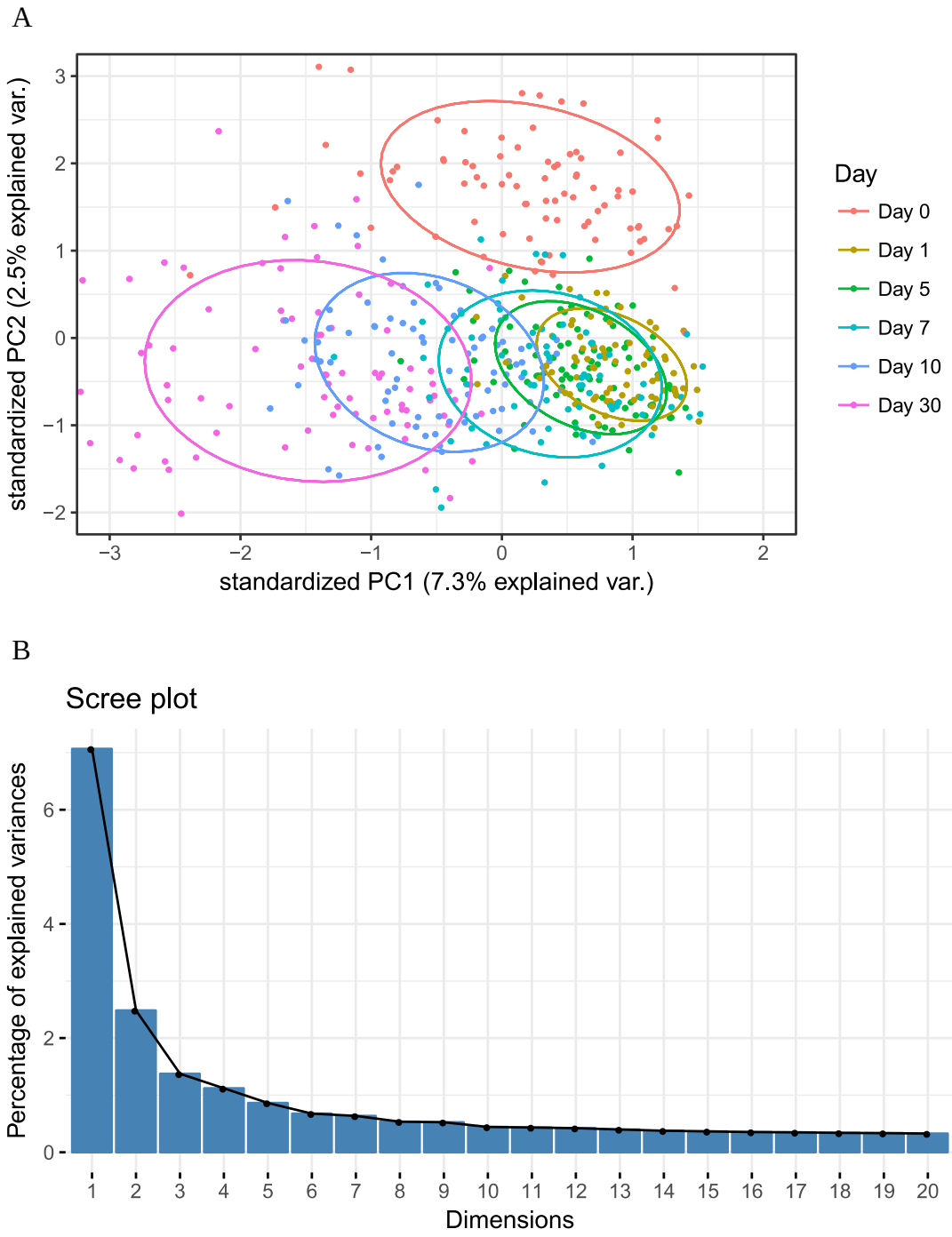


B



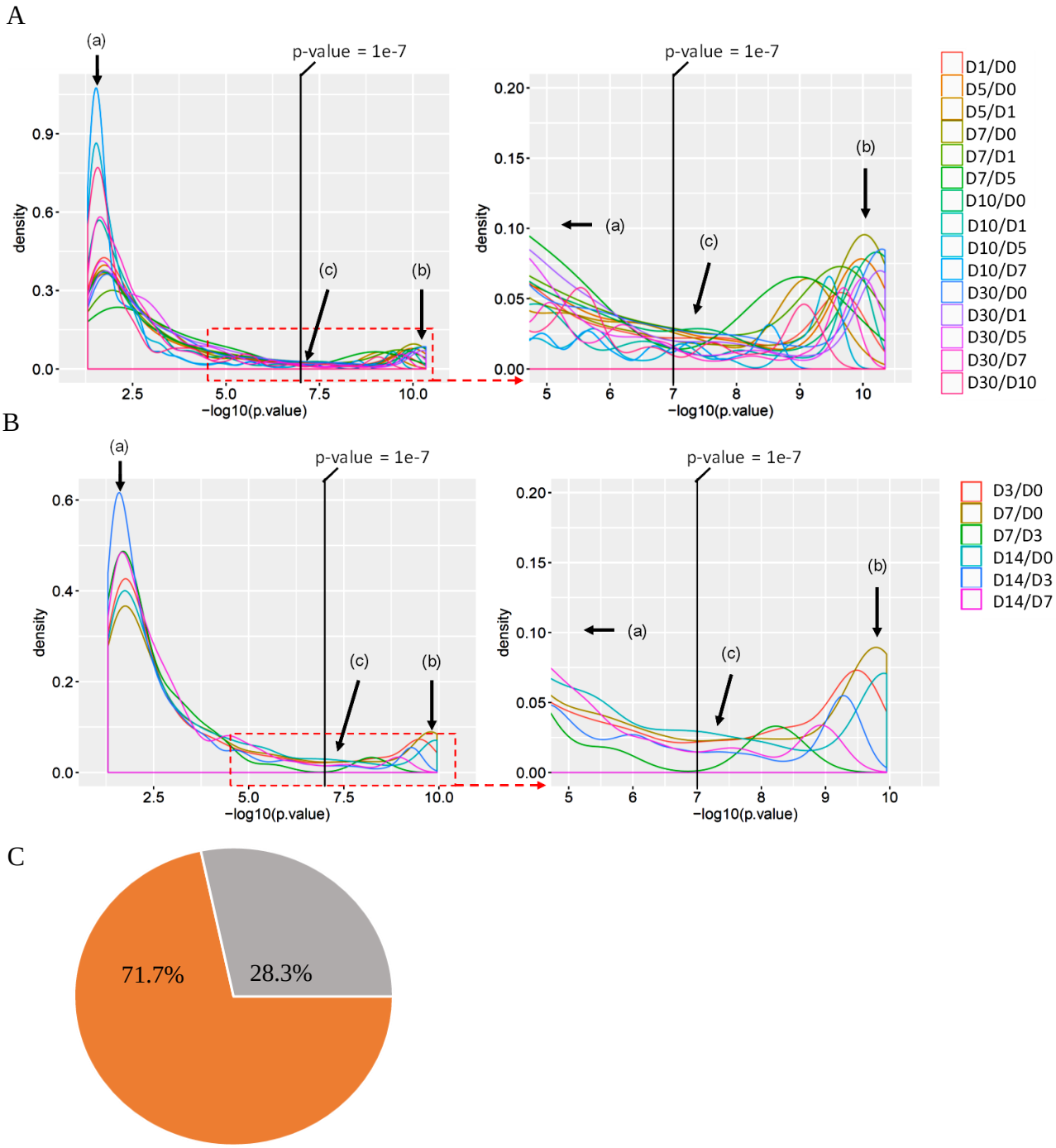
A) The gene penetrance across the six time points of experiment 1 are shown. High, medium and low penetrance genes are distributed across five quintiles as shown for each time point. For example: the 0.8-1 quintile indicates that more than 80% of the cells in a given time point express this gene. B) A similar analysis was performed for experiment 2 (day 0 day 3 day 7 and day 14). Both of these experiments show that high penetrance genes (80%-100%) increased early and decreased later as neurons developed from NPC.

Fig_S5



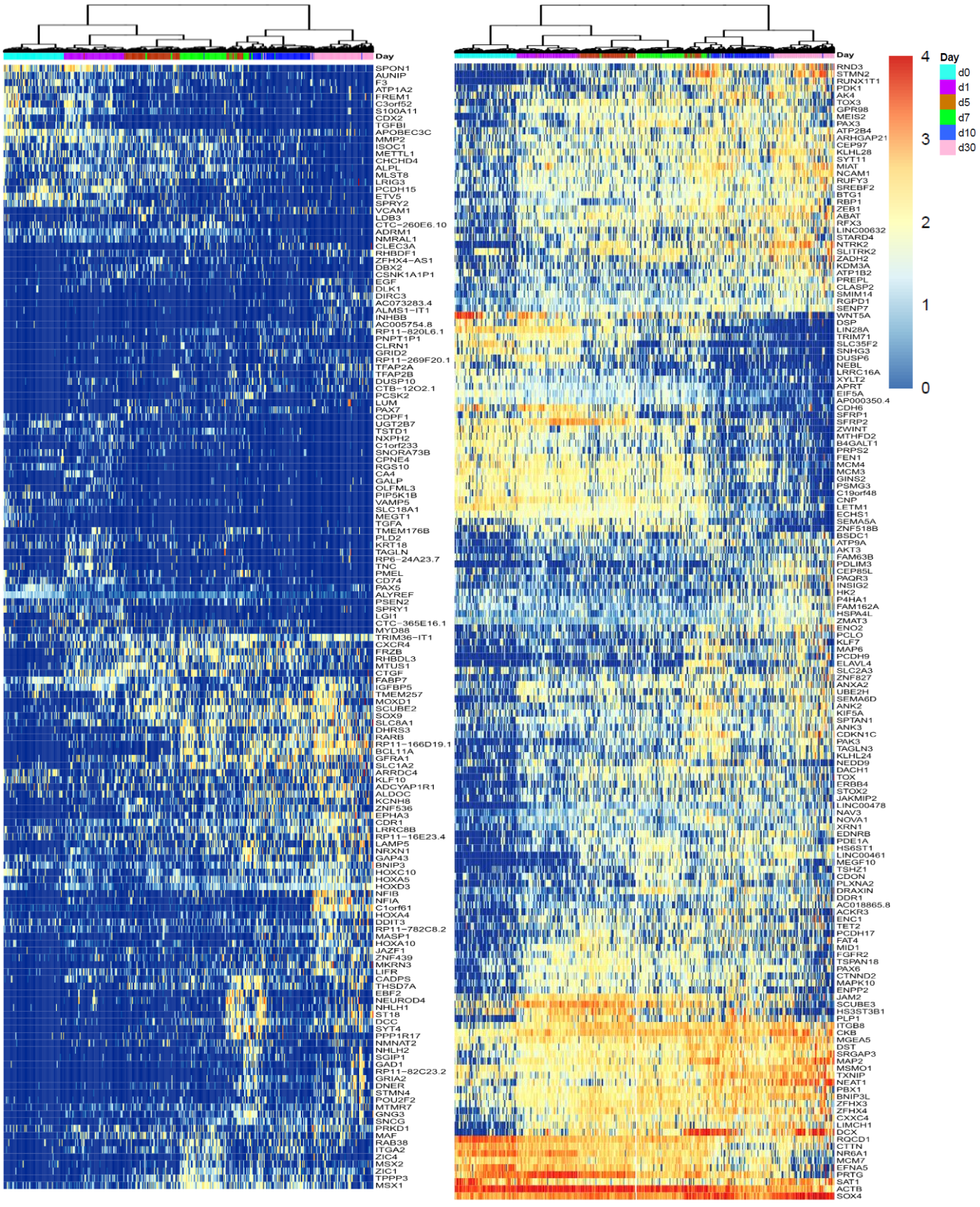
Principal component analysis (PCA) of a scRNA-seq data set. (A) The dot plots show the groups of cells corresponding to sampling time points (day 0 to day 30) based on expression of filtered genes with each dot representing a cell. Cells are labeled with different colors based upon time points as shown. Ellipses shows clustering of the samples. (B) The scree plot indicates the explained variance for each of the top 20 principal component dimensions.

Fig_S6



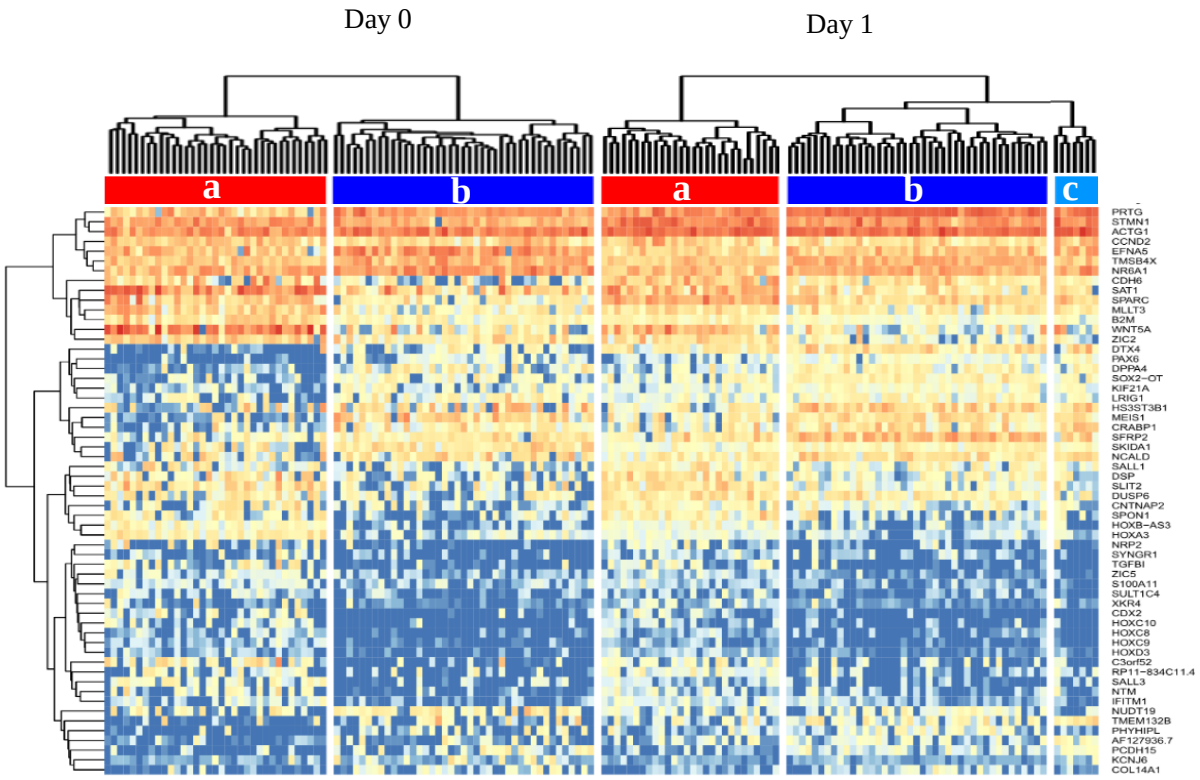
A) After comparing any two time points, as shown in Fig. 1E, 3986 significantly ($p<0.05$) differentially expressed genes were obtained(Table S3). To enhance pattern detection of significantly differentially expressed genes for downstream analysis the curve was drawn based on the gene frequency and $-\log_{10}$ (p-value) for all 3986 genes. Representative p-values (0.05, 10^{-11} and 10^{-7}) are labeled by (a), (b) and (c), and 528 genes were obtained(Table S4) with fold change > 1.5 and $p < 10^{-7}$. Each line with different colour represents one comparison of 2 time points as shown. D0, D1, D5, D7, D10 and D30 represents Day 0, Day 1, Day 5, Day 7, Day 10 and Day 30 respectively. 279 genes were obtained with fold change >1.5 and $p<10^{-7}$ when similar analysis was performed for experiment 2 as indicated(B). C) Pie chart showing that 71.7% of the 279 genes in experiment 2 were also significantly differentially expressed in experiment 1.

Fig_S7



Heat map shows the hierarchical clustering of cells based on differential expression of representative genes from Fig1. F during neuron differentiation using a statistical cutoff of $p<10^{-7}$

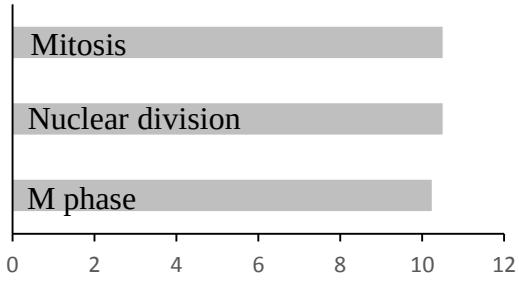
Fig_S8



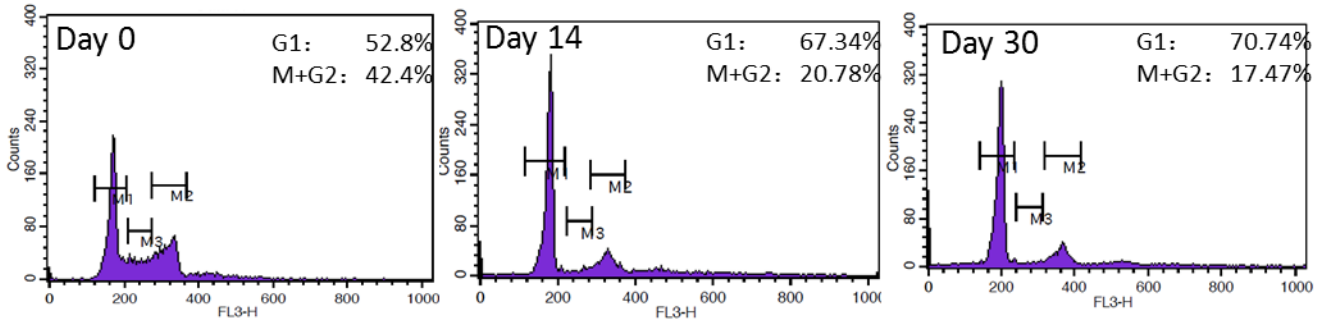
After comparing subpopulations ‘a’, ‘b’ and ‘c’ on day 0 and day 1, there were 58 differentially expressed genes (Table S6). A heat map was generated to show the distinct pattern of gene expression for all the 58 genes on day 0 and day 1.

Fig_S9

A



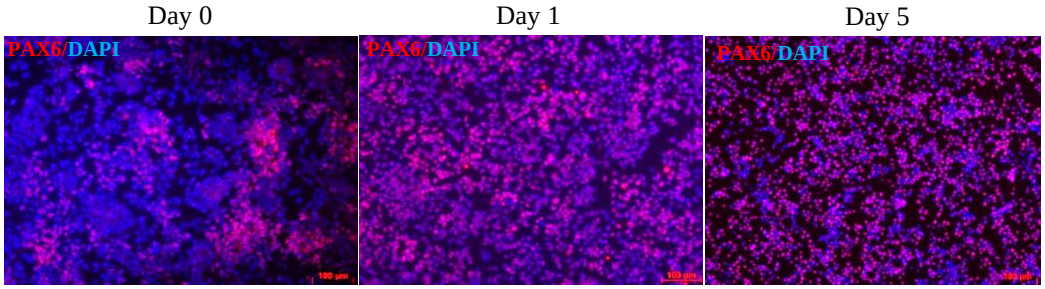
B



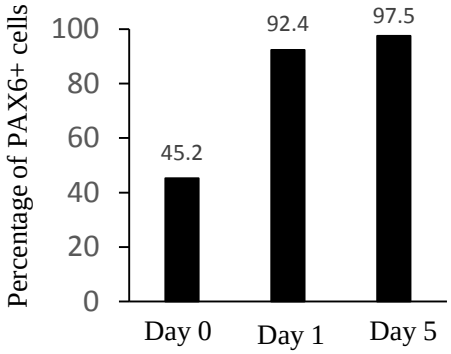
A) GO analysis of ‘ab’ subpopulation higher expressed genes compared with ‘c’ subpopulation at day 30, The X-axis indicates the –log (Benjamini p-value). B) Flow cytometry was performed for day 0, day 14 and day 30, and the percentage of cells in different cell cycle stages was calculated to ascertain the fraction of dividing (M+G2 phase) and non-dividing cells (G1 phase). M1= G1, M2+M3=M+G2.

Fig_S10

A

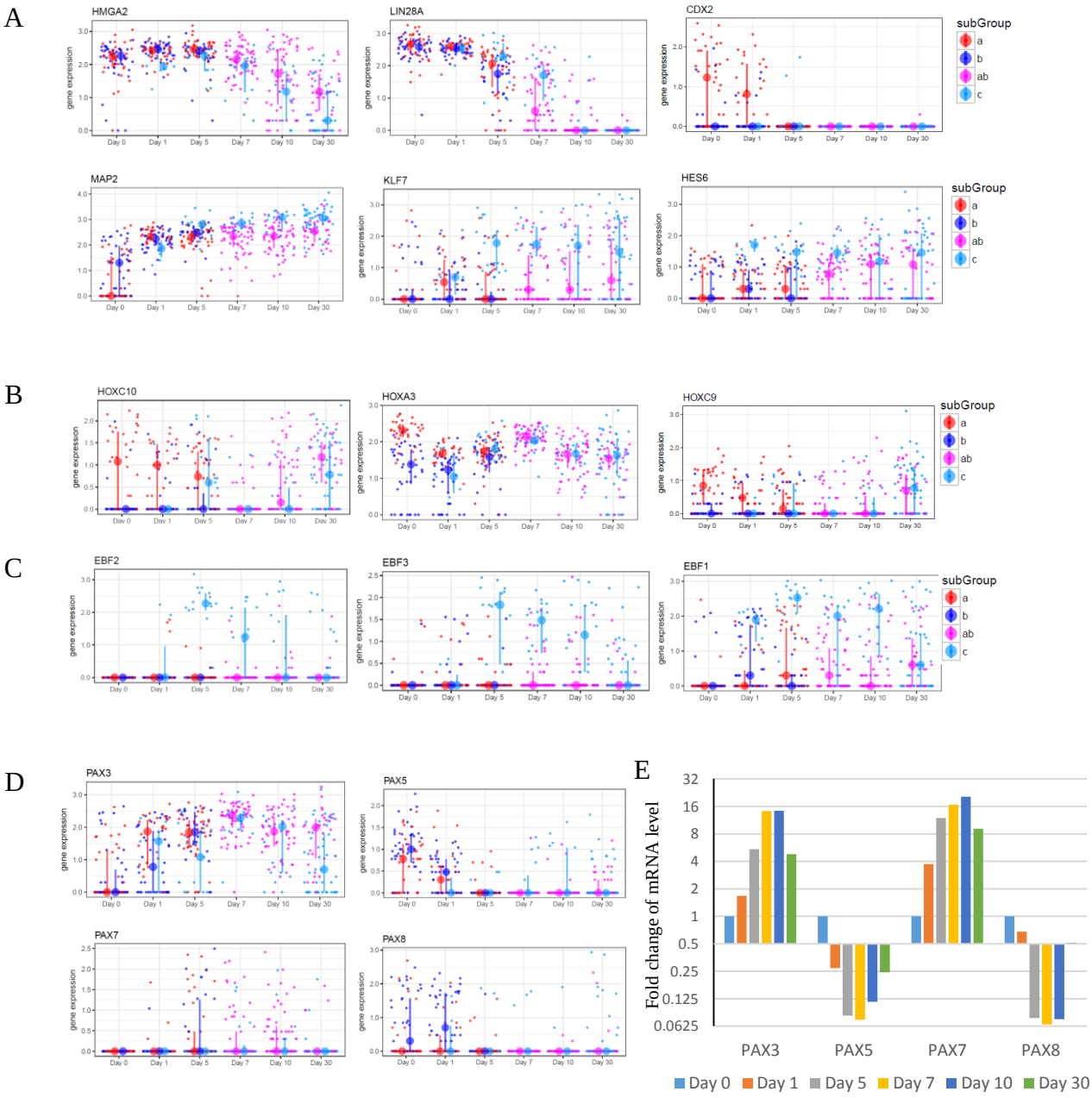


B



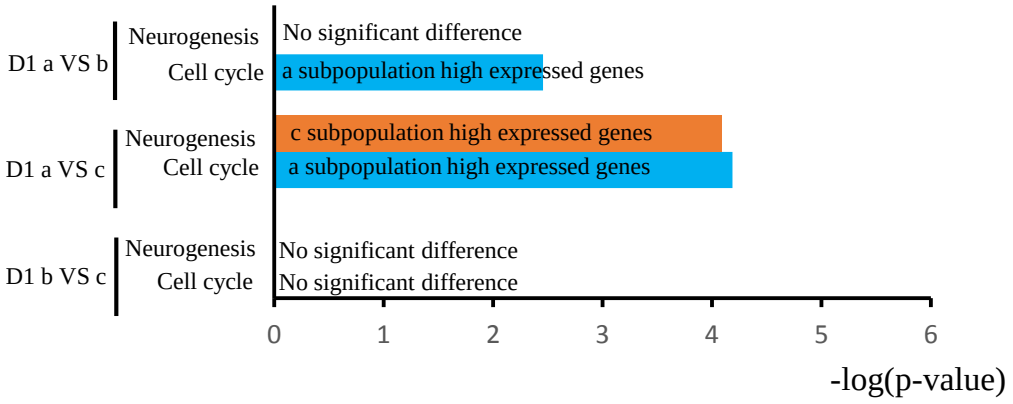
A) Immunostaining analysis was performed for neural cell marker PAX6 on day 0, day 1 and day 5 as shown. Nuclei were stained with DAPI(blue). B) Immunostained images were quantified by Columbus Analysis System to determine the percentage of *PAX6+* cells.

Fig_S11



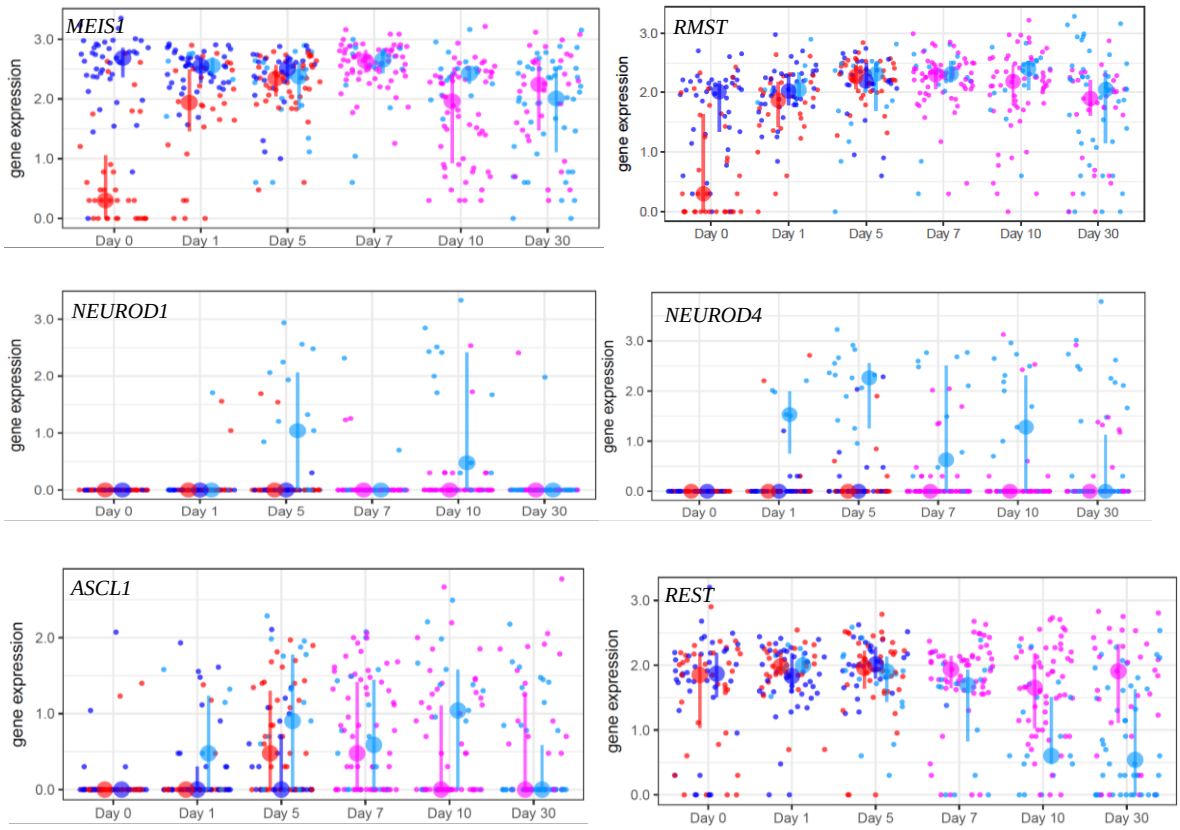
Gene expression of selected genes across time course of neuronal differentiaton. Dot plots of expression in individual cells in different subpopulation as indicated of all 6 time points: A) cell cycle genes (*HMGA2*), stem cell markers (*LIN28A*, *CDX2*) and neuronal markers (*DCX*, *STMN2*); B) *HOX* genes(*HOXC10*, *HOXA3*, *HOXC9*); C) early B cell factor (*EBF1*, *EBF2*, *EBF3*); and D) *PAX* family(*PAX3*, *PAX5*, *PAX7* and *PAX8*) . Y-axis represents the gene expression level. Each small dot represents a cell and the large dot indicates the medium expression of a given cell subpopulation. Error bar represents variation of the given gene in a given cell subpopulation. E) Q-RT-PCR performed for *PAX* genes at 6 time point to confirm the developmental dynamics in (D).

Fig_S12



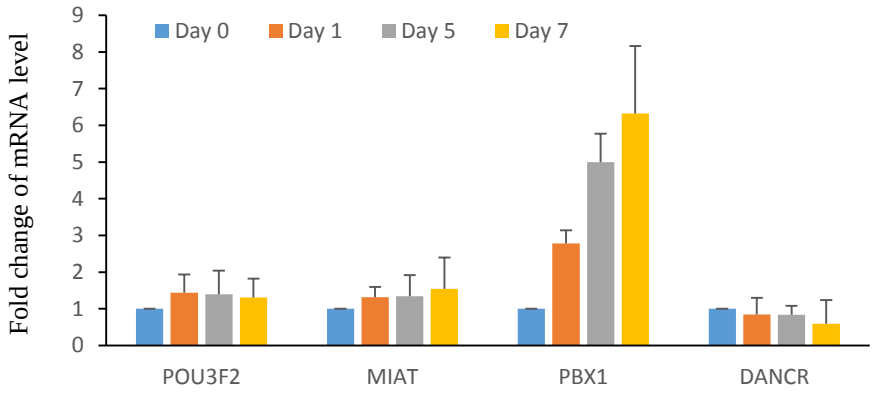
Gene ontology was performed by David 6.7 for differentially expressed genes between 'a', 'b' and 'c' subpopulations of day 1 respectively. The X-axis indicates the –log (Benjamini p-value). This figure shows that b and c subpopulations are similar at day 1.

Fig_S13



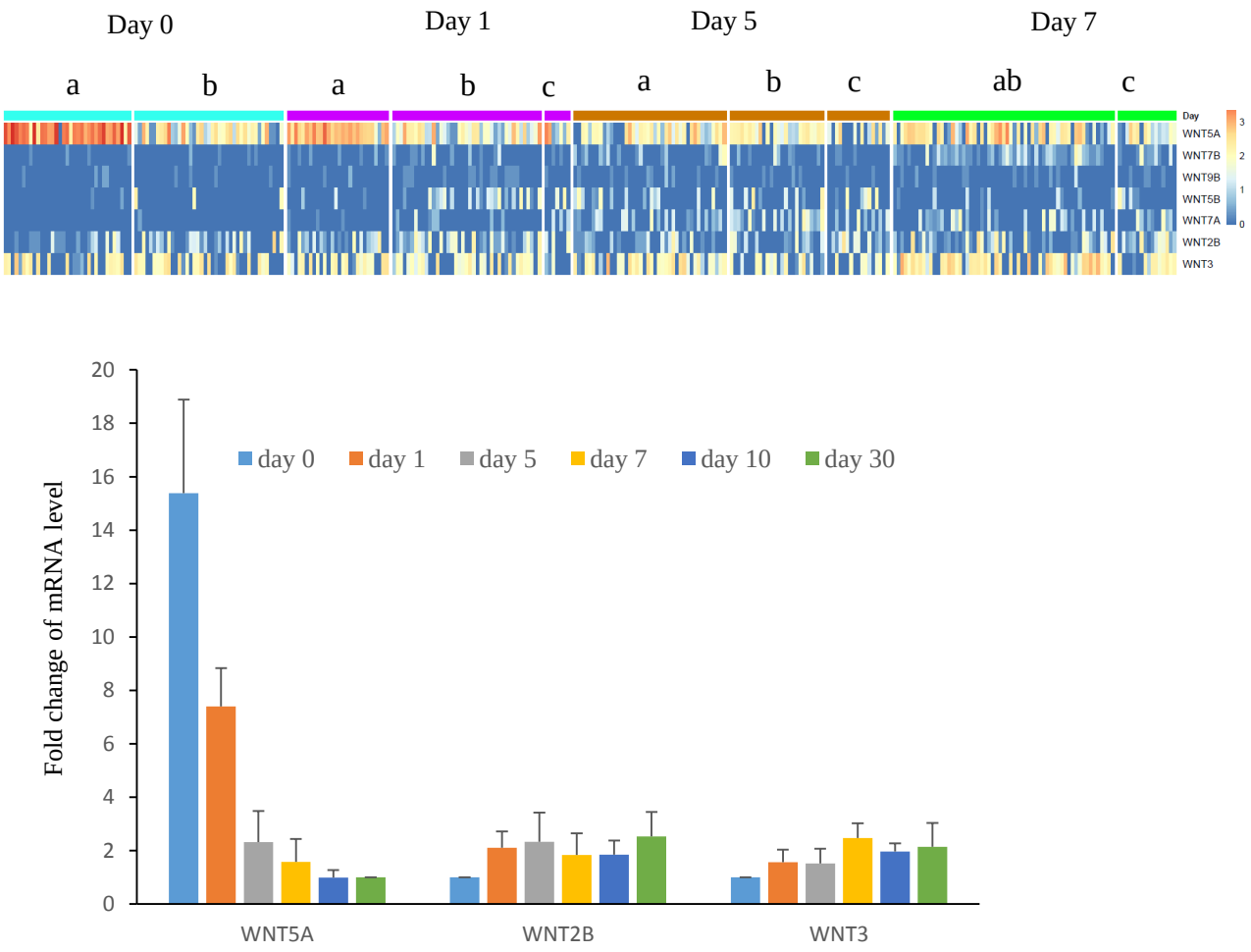
Dot plots shows the gene expression of well-known key genes during neural cell development in different subpopulations as indicated for all 6 time points. Y-axis represents the gene expression level. Each small dot represents a cell and the large dot indicates the medium expression of a given cell subpopulation. Error bar represents variation of the given gene in a given cell subpopulation.

Fig_S14



Q-RT-PCR performed for *POU3F2*, *MIAT*, *PBX1* and *DANCR* on days 0, 1, 5 and 7.

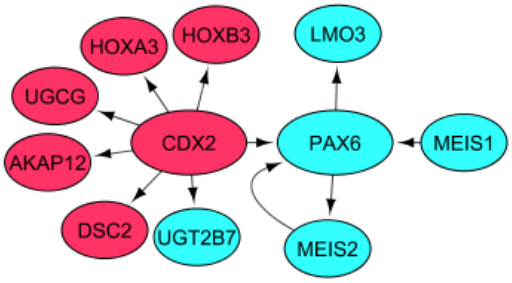
Fig_S15



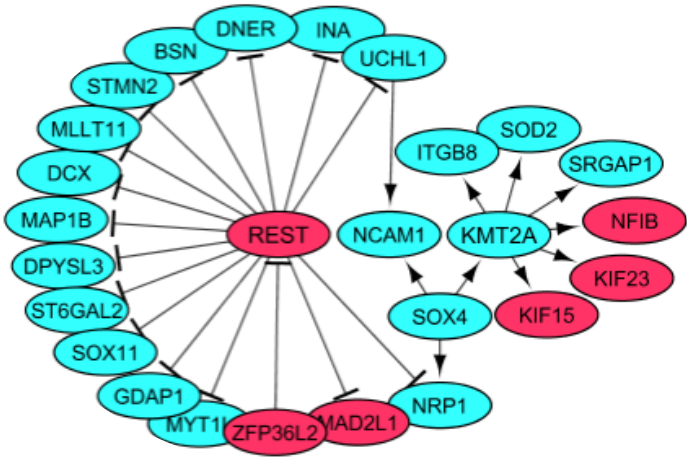
A) Heat map showing the expression of *WNT* genes in subpopulations at different time points. Each column represents a cell, each row represents the indicated gene. B)Q-RT-PCR shows the expression levels of *WNT5A, WNT2B, WNT3* in all 6 time points as indicated. These 3 members showed relatively high expression in our single cell RNA-seq data.

Fig_S16

A

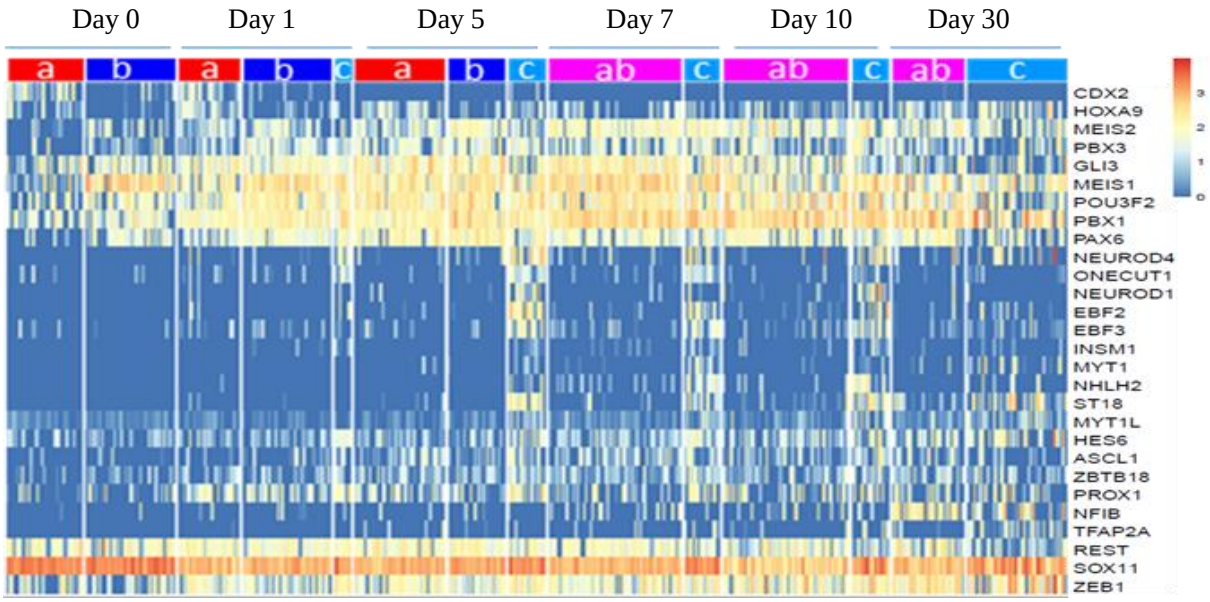


B



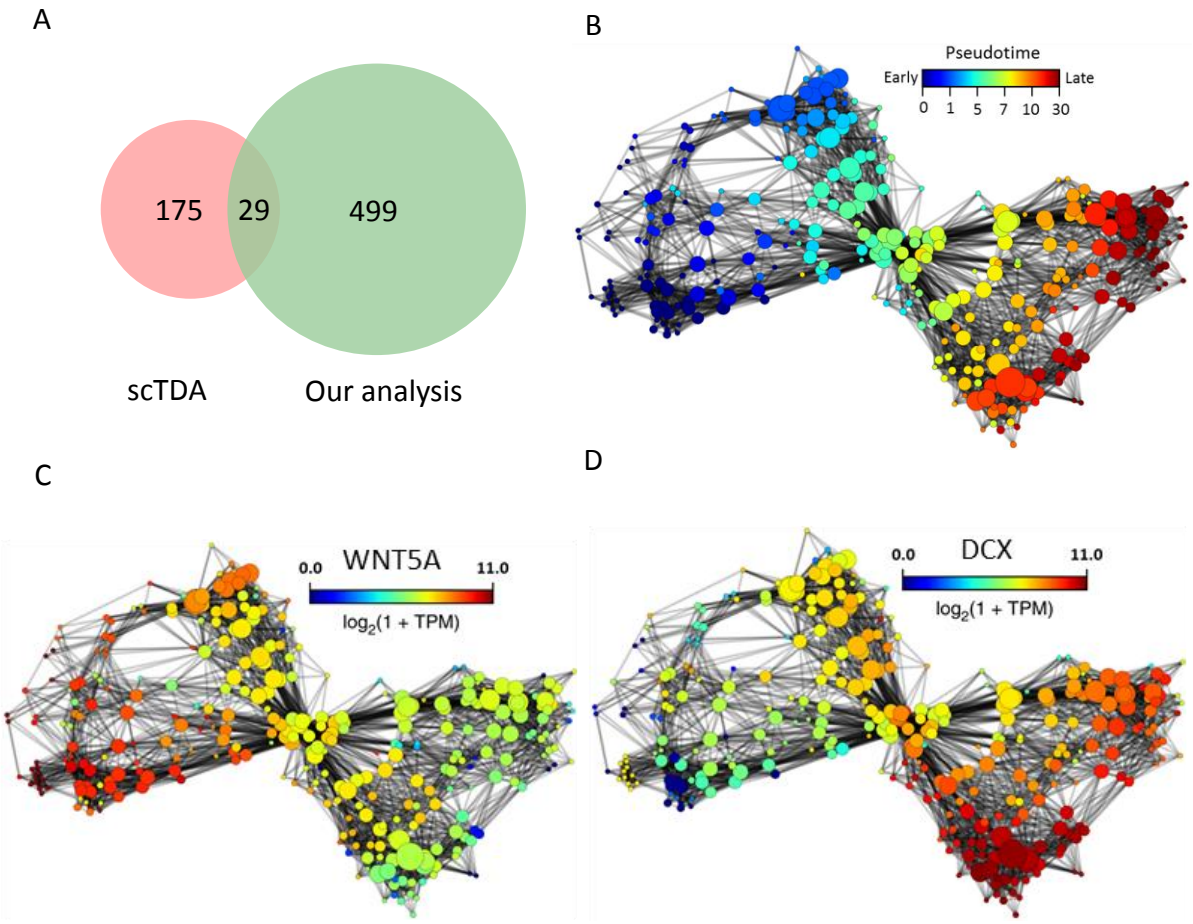
Similar regulatory network analysis as in Fig. 5A was performed for all differentially expressed genes of the subpopulations on day 0(A) and day 30(B).

Fig_S17



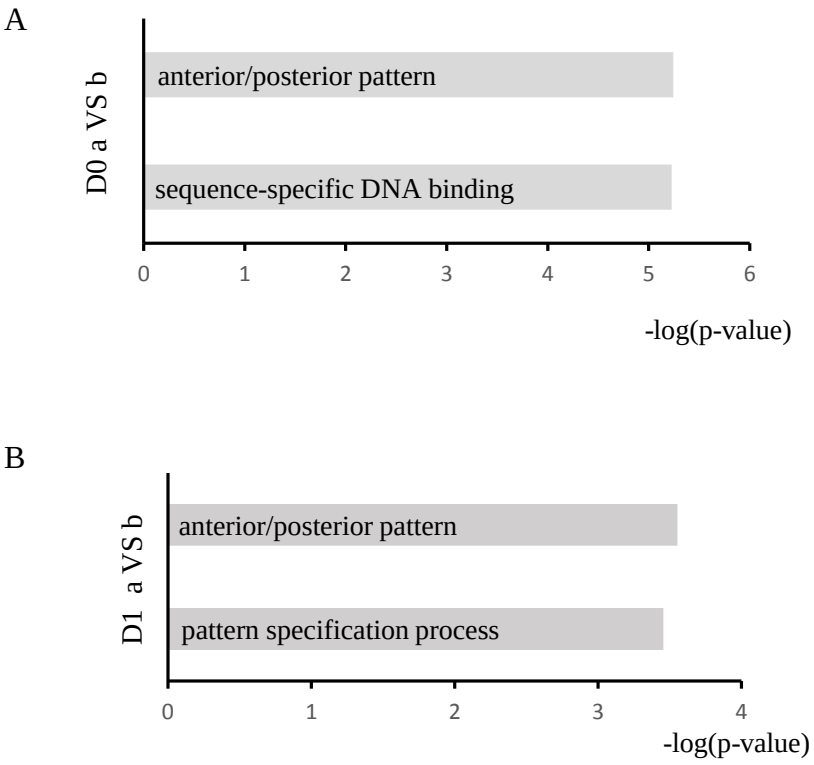
A heat map was generated for the expression of TFs shown in the derived network of Fig. 5D for different cell subpopulation at different time points. Each column represents a cell and each row represents the indicated gene.

Fig_S18



Our scRNA-seq data was analyzed by the published scTDA method (Rizvi et al. 2017). A) We selected genes that have an average of 2 or more counts across the cells and a minimum z-score of +2.0 with respect to the sigmoidal fit. 204 genes were obtained by scTDA, 29 of which are included in our 528 gene list (see Fig. 1E). B) scTDA reconstructed differentiation trajectories based on expression of all cells from NPCs to neurons (colored by sampling times). The root nodes correspond to sets of cells sharing similar global transcriptional profiles, with the node sizes proportional to the number of cells in the node (i.e. larger node size representing more cells). Nodes that are connected by an edge have at least one cell in common. Topological representation (labelled by mRNA levels) of WNT5A(C) and DCX(D) show concordant expression of these transcripts during the development of neurons.

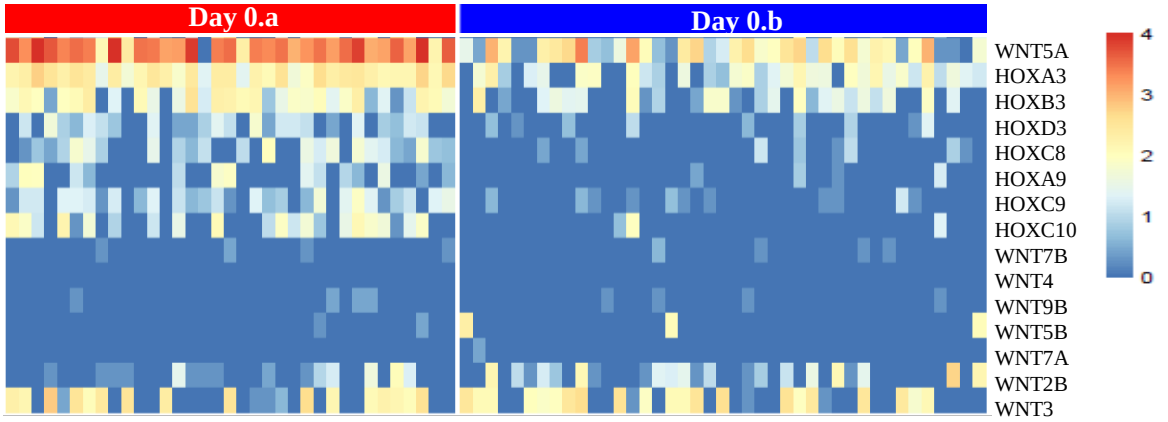
Fig_S19



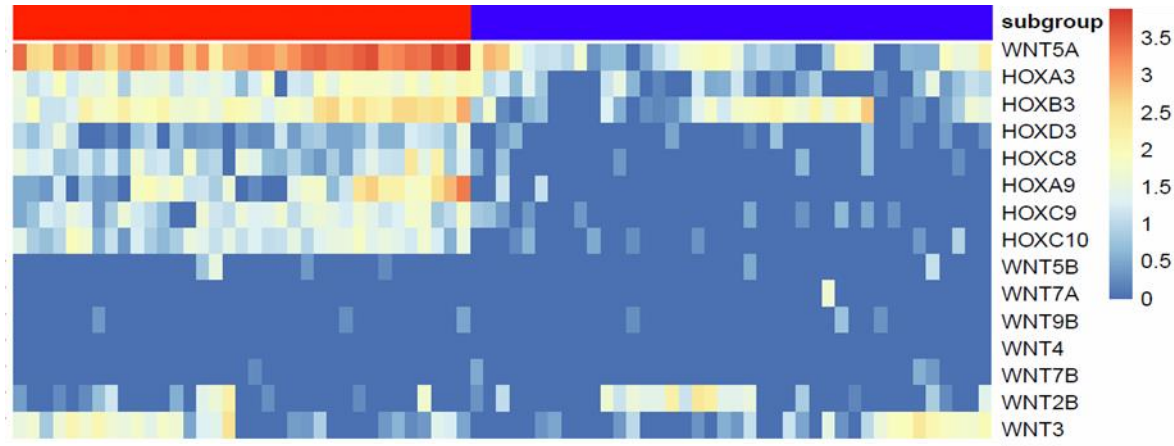
A) Gene ontology was performed by David 6.7 for all the 192 differentially expressed genes (Table S5) between 'a' and 'b' subpopulations of day 0. The X-axis indicates the –log (Benjamini p-value). No cell cycle differences were apparent in 'a' and 'b' subpopulations on day 0. B) Similar analysis was performed for differentially expressed genes between ‘a’ and ’b’ subpopulation of day 1.

Fig_S20

A

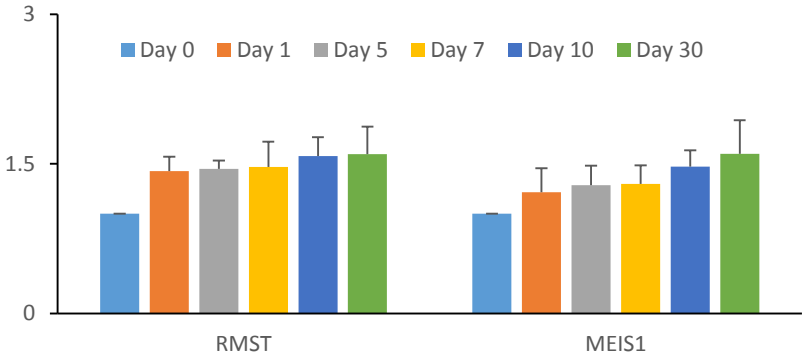


B



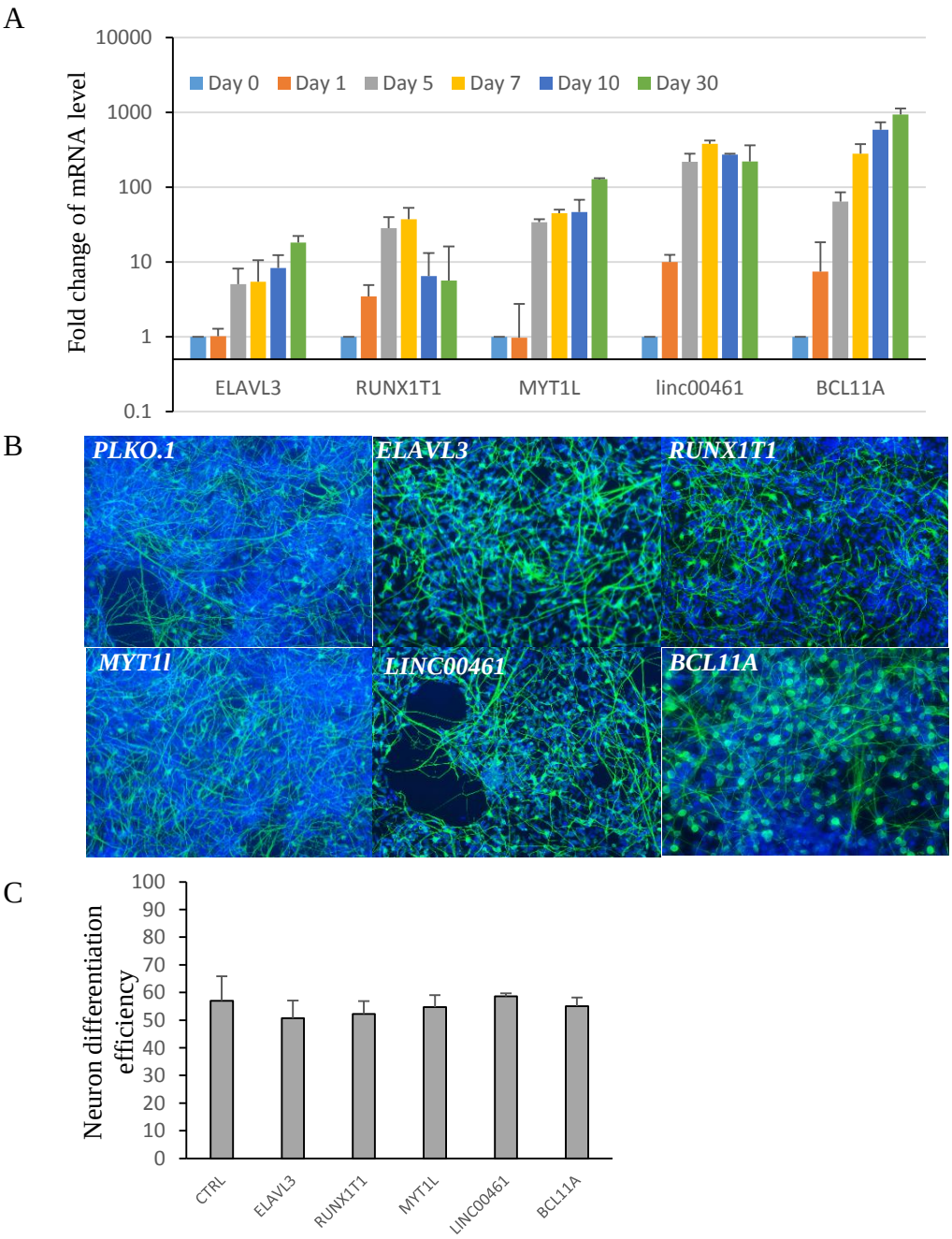
The heat map shows the expression level of the same genes on day 0 of experiment 1(A) and 2(B). Each column represents a cell and each row represents the indicated gene.

Fig_S21



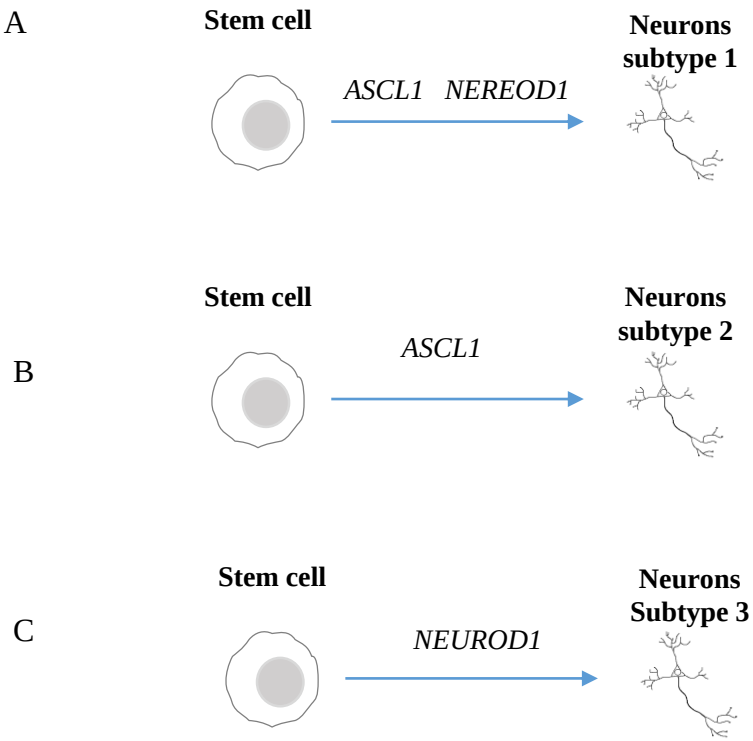
Q-RT-PCR shows the expression level of *RMST* and *MEIS1* for bulk cells at all 6 time points. These genes were not significantly up-regulated during neurogenesis.

Fig_S22



A) Q-RT-PCR shows the expression level in bulk cells for *ELAVL3*, *RUNX1T1*, *MYT1l*, *LINC00461* and *BCL11A* at all 6 time points. These genes were significantly up-regulated during neurogenesis. B) Knock down gene expression experiments were performed *for ELAVL3*, *RUNX1T1*, *MYT1l*, *LINC00461* and *BCL11A* during neuronal differentiation (the empty plko.1 vector served as a negative control). Neuronal marker *TUJ1*(Green) and nuclei were stained with DAPI (blue) after 7 days of knockdown of these candidates. C) Immunostained images were quantified by Columbus Analysis System to determine the efficiency of neuronal differentiation.

Fig_S23



Simple cell differentiation model. A) Some NPC require sequential activation of *ASCL1* and *NEUROD1* to differentiate into neurons. B, C) Other neurons develop with activation of either *ASCL1* or *NEUROD1*, which may in turn give rise to alternative neuronal subtypes.