(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
30 June 2016 (30.06.2016)

WIPO | PCT

(10) International Publication Number
WO 2016/103269 A1

(51) **International Patent Classification**:
C12N 5/0793 (2010.01) *C12N 5/079* (2010.01)

(21) **International Application Number**:
PCT/IL2015/051253

(22) **International Filing Date**:
23 December 2015 (23.12.2015)

(25) **Filing Language**: English

(26) **Publication Language**: English

(30) **Priority Data**:
62/096,184 23 December 2014 (23.12.2014) US

(71) **Applicant**: **RAMOT AT TEL-AVIV UNIVERSITY LTD.** [IL/IL]; P.O. Box 39296, 6139201 Tel-Aviv (IL).

(72) **Inventors**: **ELKABETZ, Yechiel**; 16/3 Dror Street, 6813520 Tel-Aviv (IL). **EDRI, Reuven**; c/o Ramot at Tel-Aviv University Ltd., P.O. Box 39296, 6139201 (IL). **YAFFE, Yakey**; 6 Ruchama Street, 5235944 Ramat-Gan (IL). **MEISSNER, Alexander**; 134 Larch Road, Cambridge, Massachusetts 02138 (US). **ZILLER, Michel J.**; 89 Museum Street, Cambridge, Massachusetts 02138 (US).

(74) **Agents**: **EHRLICH, Gal** et al.; G.E. EHRLICH (1995) LTD., 11 Menachem Begin Road, 5268104 Ramat Gan (IL).

(81) **Designated States** *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) **Designated States** *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

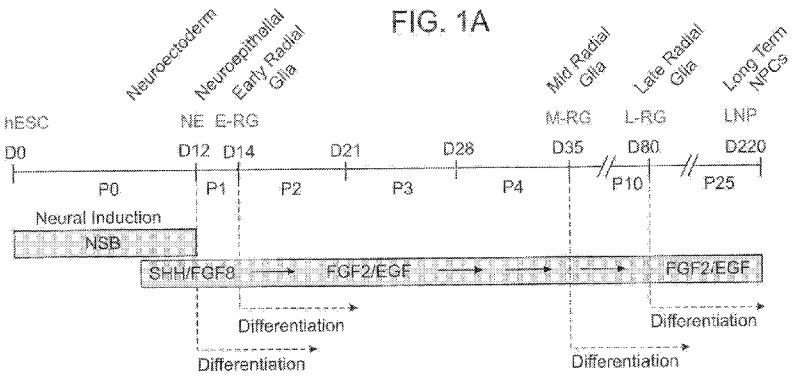
— *with international search report (Art. 21(3))*

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

— *with sequence listing part of description (Rule 5.2(a))*

(54) **Title**: POPULATIONS OF NEURAL PROGENITOR CELLS AND METHODS OF PRODUCING AND USING SAME

FIG. 1A



(57) **Abstract**: An isolated population of cells is provided the cells comprise at least 10 % HES5+ cells, wherein said HES5+ cells are: (i) early radial glial (E -RG) cells; (ii) mid radial glial (M-RG) cells; (iii) late radial glial (L-RG) cells; or (iv) long term neural progenitor (LNP) cells. Also provided are additional populations of neural cells and use of the populations and methods of producing same.

WO 2016/103269 A1

POPULATIONS OF NEURAL PROGENITOR CELLS AND METHODS OF PRODUCING AND USING SAME

FIELD AND BACKGROUND OF THE INVENTION

The present invention, in some embodiments thereof, relates to populations of neural progenitor cells and methods of producing and using same.

Human pluripotent stem cell derived models that accurately recapitulate neural development *in vitro* and allow for the generation of specific neuronal subtypes are of major interest to the stem cell and biomedical community. To date, the process of neural differentiation has not been elucidated and therefore any attempts to accurately mimic it or manipulate it have failed.

The identification of neural stem cells (NSCs) in the developing and adult brain has transformed the way the scientists understand central nervous system (CNS) development and regeneration. However, long following their isolation from the CNS or the derivation of neural progenitors from pluripotent stem cells (PSCs), the ability to address the dynamic changes in self-renewal and potency of distinct NSC types *in vitro* has remained poor. The exceptionally pioneering studies done in the NSC field *in vivo* have led to the identification of fundamental NSC types populating the germinal zones - neuroepithelial (NE) cells, radial glial (RG) cells, and adult NSCs (aNSCs). These studies provided the basis for understanding of the dynamic nature and lineage relationship of these distinct NSC types *in vivo*, describing the unique timing mechanism of neuronal cell type generation. However, in depth *in vitro* dissection of the molecular characteristics of each stage, particularly within the RG compartment, has been stalled mainly by the heterogeneity of NSC cultures and the lack of stage specific markers. In fact, despite being highly heterogeneous, distinct RG cell types as well as aNSCs are known to share similar RG cell markers rather than distinctive ones. The reporter gene- and surface marker-based prospective isolation of acute mouse aNSCs serves as a great example for a more in depth analysis of aNSC characteristics. However, applying such a study to human CNS derived RG cells is limited due to obvious shortage in early human CNS tissue. Thus, in depth understanding on human NSC ontogeny and dynamics in culture is still elusive.

The advent of PSCs has brought the ability to direct early neural progenitors

towards a range of neuronal cell fates including midbrain dopaminergic neurons, spinal motoneurons and telencephalic cortical neurons. One remarkable study by Knoblich and co-workers allows monitoring early to mid gestation cerebral morphogenesis and neurogenesis, making up an attractive approach to model development and disease of the human brain (Lancaster, M. A. et al. Cerebral organoids model human brain development and microcephaly. Nature 501, 373-379, doi:10.1038/nature12517, 2013). Another recently published comprehensive work delineates the temporal transcriptome analysis of cerebral cortex neuronal subtypes derived from PSCs [van de Leemput, J. et al. CORTECON: A Temporal Transcriptome Analysis of In Vitro Human Cerebral Cortex Development from Human Embryonic Stem Cells. Neuron 83, 51-68, doi:10.1016/j.neuron.2014.05.013 (2014)]. These two latter advancements have significantly helped to demonstrate the capability of hESC differentiation strategies to recapitulate major hallmarks of *in vivo* neural development and serve as a valuable resource for modeling development and disease of the human brain. Further to these important findings, however, there is a need to better understand how different types of progenitors emerge and exert their full potential while progressing through distinct competences during development. Addressing such an aim requires employing differentiation culture strategies that allow distinguishing primary progenitor cells holding extensive proliferation capacity and broad differentiation potential from the bulk of accompanying progenitors that lack these abilities. The present inventors have previously isolated an early progenitor cell type from PSCs that exhibits considerable self-renewal capacity (termed rosette-neural stem cells (R-NSCs)), and showed their developmental potential and distinct molecular signature [Elkabetz, Y. et al. Human ES cell-derived neural rosettes reveal a functionally distinct early neural stem cell stage. Genes & Development 22, 152-165 (2008)]. However, also the R-NSC stage exhibits high heterogeneity with respect to NSC potential and corresponds to a transient stage *in vitro*. Currently there is no list of genes at high confidence that are known for specific types of neural progenitors emerging in culture, stressing the need to unravel generalized networks and pathways involved in the extensively changing dynamics of early neuroepithelial (NE) cells. Taken together, despite many years of NSC research, the heterogeneity and rapid transition through distinct neural stem and progenitor cell types still impedes the understanding of origin, lineage transitions, and the key factors

that maintain or alter the epigenetic stability of early NE cells.

SUMMARY OF THE INVENTION

According to an aspect of some embodiments of the present invention there is provided a method of isolating neural progenitor cells, the method comprising:

(a) culturing pluripotent stem cells having been transformed to express a Notch-activated reporter under culture conditions suitable for differentiation of the pluripotent stem cells into neural progenitor cells; and

(b) successively isolating progenitor cells of interest based on activation of the Notch-activated reporter.

According to an aspect of some embodiments of the present invention there is provided an isolated population of cells comprising at least 10 % HES5+ cells, wherein the HES5+ cells are:

(i) early radial glial (E-RG) cells;

(ii) mid radial glial (M-RG) cells;

(iii) late radial glial (L-RG) cells; or

(iv) long term neural progenitor (LNP) cells.

According to an aspect of some embodiments of the present invention there is provided an isolated population of cells comprising at least 10 % HES5- cells, wherein the HES5- cells are:

(i) non-CNS cells comprising neural crest cells, placodal cells, non-neuroepithelial cells; and CNS cells which exhibit an NEUROD4+/NGN1+/NGN2+/TBR2+/DCX+ expression signature and which form neurons of layers 1 and 6 of the brain cortex;

(ii) neural progenitor cells which belong to the CNS, having a lower proliferative capacity as compared to the HES5+ ERG cells, which form layers 1, 5 and 6 of the brain cortex;

(iii) intermediate progenitor cells (INPs) which belong to the CNS, and which are capable of differentiating into the neurons forming layers 4, and 2 of the brain cortex;

(iv) HES5- neurons and astrocytes, wherein the neurons form layers 2, 4 and 3 of the brain cortex; or

(v) neurons, oligodendrocyte and astrocytes, wherein the neurons comprise neurons reaching the olfactory bulb.

According to an aspect of some embodiments of the present invention there is provided a culture medium for neuroepithelial differentiation comprising noggin, LDN-193189and SB-431542.

According to some embodiments of the invention, the isolated population of cells of some embodiments of the invention further comprising HES5+ neuroepithelial (NE) cells.

According to some embodiments of the invention, the HES5+ NE cells are capable of differentiating into HES5+ E-RG cells and into HES5- central nervous system cells neurons.

According to some embodiments of the invention, the HES5+ E-RG cells are capable of differentiating into the HES5+ M-RG cells and into HES5- neural progenitor cells.

According to some embodiments of the invention, the HES5+ M-RG cells are capable of differentiating into the HES5+ L-RG cells and into HES5- intermediate progenitor cells (INPs).

According to some embodiments of the invention, the HES5+ L-RG cells are capable of differentiating into the HES5+ LNP cells and into HES5- neurons and astrocytes.

According to some embodiments of the invention, the HES5+ LNP cells which comprise HES5+ adult neural stem cells (aNSCs) and into HES5- neurons, oligodendrocyte and astrocytes.

According to some embodiments of the invention, the HES5+ neuroepithelial cells exhibit an HES5+/SOX1+/PAX6+/SOX2+/Nestin+ expression signature.

According to some embodiments of the invention, the HES5+ neuroepithelial cells further exhibit a CDC6+/CDX1+/CENPH+/TOP2A+ expression signature.

According to some embodiments of the invention, the HES5+ neuroepithelial cells are characterized by a higher expression level of at least one gene selected from the group consisting of: TOP2A, HIST1H4C, TRIM71, PPIG, MLLT4, TNC, CDK1, OIP5, GDF15, MCM6, TP53TG1, FAM83D, FANCI, GINS2, KDM5A, GSTM3, FAM64A, LIMS1, CENPH, KIF2C, ATAD2, DTL, CDCA5, ARHGEF6, LIPA,

POLE2, RRM2, MAD2L1, CKS1B, TTK, DHFR, S100A4, NUP37, PMAIP1, CENPN, RNASEH2A, BST2, MCM10, MAF, KIAA0101, C8ORF4, E2F7, CENPA, UBE2T, RAB13, TMEM126A, MAGT1, CDC6, C6ORF211, RFC5, PSMD1, HMMR, UNG, UBE2C, GINS1, AURKB, LEPREL1, SBNO1, ZWINT, MKI67, CCAR1, FKBP5, PVRL3, CCNB1, NOP58, COL4A1, GGH, LSM6, EID1, GPX8, STC2, CD276, HS2ST1, EIF5B, HDGF, and NOL7 as compared to the expression level of the at least one gene in HES5- differentiated cells obtained by culturing undifferentiated pluripotent stem cells (PSCs) under culture conditions suitable for differentiation the PSCs into HES5+ neuroepithelial cells.

According to some embodiments of the invention, the culture conditions suitable for differentiation the PSCs into neural cell lineage comprise a culture medium which comprises Noggin, SB-431542 and LDN-193189.

According to some embodiments of the invention, the early radial glial cells exhibit an PAX6+/SOX1+/SOX2+/Nestin+ expression signature.

According to some embodiments of the invention, the early radial glial cells further exhibit an HES5+/ARX+/FEZF2+/NR2E1+ expression signature.

According to some embodiments of the invention, the HES5+ early radial glial cells are characterized by a higher expression level of at least one gene selected from the group consisting of: NR2E1, HES5, ARX, C1ORF61, FRZB, GRM3, EPHA3, NAV3, EGR2, RGMA, NRXN3, FAM107A, FABP7, EGR3, ZNF385B, TTYH1, SNCAIP, NRARP, PLP1, LIX1, LFNG, HES4, CD82, HS6ST1, PTPRZ1, CACHD1, DACH1, FEZF2, DTX4, FUT9, WNT5B, ENPP2, POU3F3, EMX2, MECOM, XYLT1, ARMCX2, FOS, PPAP2B, NOS2, LRP2, SOX9, NLGN3, TMEM2, CXCR7, EPHA7, SMOC1, TBC1D9, FAT4, SCUBE3, FUT8, CSPG5, DLL1, BOC, ID4, EGR1, ALPL, RFX4, GALNT12, CBX2, FHOD3, SORBS2, GUCY1B3, MBIP, FBXO16, SHISA2, DAB1, GLI3, FZD3, SEMA5B, LGALS3BP, SFRP1, C1QL1, RING1, GPRC5B, ZNF710, WSCD1, VPS37B, ZIC2, SDK2, DOCK11, GAS1, ZNF436, TMSB15A, IER2, FEZ1, CELF2, SFT2D3, NCALD, AKAP7, MYADM, NEDD4L, PHC2, PI4KAP2, STARD3, and CAMK1D as compared to the expression level of the at least one gene in HES5- differentiated cells obtained by culturing HES5+ neuroepithelial cells under culture conditions suitable for differentiation the HES5+ neuroepithelial cells into HES5+ early radial glial cells.

According to some embodiments of the invention, the HES5+ mid radial glial cells exhibit an HES5+/POU3F2+ expression signature.

According to some embodiments of the invention, the HES5+ mid radial glial cells further exhibit an GLAST+/FABP7+ expression signature.

According to some embodiments of the invention, the HES5+ mid radial glial cells are characterized by a higher expression level of at least one gene selected from the group consisting of: FZD10, ZEB2, EN2, ST20, CDKN2C, RAB10, WASF1, ZBED4, EZH2, PPA2, H1F0, CCNJ, ITGB8, SH3BGRL3, IRX2, KIF23, PEG10, SMC3, NUSAP1, APLP1, ADAMTS3, RACGAP1, LIMCH1, ETNK1, RNF13, ARID1B, TRIM28, CNOT8, CRNDE, TWSG1, NT5DC2, NAA50, NUF2, ABCE1, PLTP, FBRSL1, DCAF16, OGT, ZFYVE16, FOXM1, PM20D2, POU3F2, MCM4, HERPUD2, VRK1, TRIM41, SATB1, HOMER1, CCNG1, ATF2, AP1AR, GABPA, STXBP3, SMC5, CDKN1B, NUPL1, UBA1, CYTH2, FXYD6, ISYNA1, DOCK1, 41527, LPHN2, IDI1, PXMP2, U2AF2, ARHGAP12, KLHL24, CKAP2, ZNF238, PARP6, NHSL1, PBRM1, BAZ1A, MAP4K5, TSPAN12, SH3GLB1, ASPM, ANKLE2, SPG20, MAP4K4, CASC4, FUBP3, ARSB, BTAF1, SIKE1, VEZT, PBX3, CBL, EIF2AK4, API5, MTSS1, NET1, CHD6, ZNF117, PNMA1, PTPN13, MTIF2, SSFA2, KIAA1279, STRN3, WIPF1, MEIS2, ZC3H4, DYNC1I2, RTN4, TAF2, RASA1, OSBPL8, SKA2, IGF1R, RNF6, SGTB, TMEM131, HIATL1, TGIF1, TMEM170B, PSAT1, ACBD5, HECTD2, ASF1A, LAMB1, GLS, DDX39, DGCR2, EIF1AX, SALL1, GOLPH3, PTBP2, GRIP1, PNPLA8, VASH2, SUDS3, PFDN4, BAZ2A, PRKDC, GLYR1, DAZAP2, PCMTD1, SENP6, CLINT1, RECQL, CNTNAP2, CTBP1, C10ORF18, CDON, B4GALT6, CSNK1G3, STAT5B, TMEM60, HNRNPH2, TACC2, CCNG2, FSCN1, CCNA2, C21ORF45, PLRG1, ZFHX3, UBE2A, DMTF1, TRA2A, MYO5A, FAM96A, IFT80, VPS26A, MRPL50, ACYP1, WDR11, PLDN, RPRD1A, MEAF6, CKAP5, YTHDC2, GABARAPL2, IPO5, PGAP1, C14ORF147, CD200, MST4, PPT1, ANKRD50, HPS3, CCNC, THRAP3, TWF1, CYP51A1, PSPC1, WDR75, CAST, SEPW1, C21ORF59, PIK3C2A, GNG5, MED4, GIPC1, STK39, KIAA1715, PHF6, PPTC7, SOCS4, PPM1B, UQCRB, C10ORF84, SLAIN1, RAB6A, SOS2, KLF10, RNF4, C3ORF63, INSIG1, CPSF1, DNAJC4, ATP2B4, PPP2R1A, TRIM22, SDC3, TSNAX, PPIL4, ZDHHC2, ZBTB44, ANO6, PPP2CB, UBA2, BBS2, ZNF423, RNF5, C1ORF31, IFT81, CPSF6, KLHL9,

FAM164A, TTC35, CCDC90B, TM9SF4, SEC24B, SMARCC2, CAP2, SAR1A, THY1, RBPMS2, EIF3A, DZIP1, ARL6IP1, SACM1L, PAPD4, SCG2, TCF3, EFHA1, HNRNPA2B1, EWSR1, STAG2, YEATS4, PAQR3, GAR1, FTH1, C19ORF43, TMEM14C, CCDC104, PSMD12, DCTD, SSR1, HMGCS1, HMGB3, KIF3A, TMEM128, PATZ1, RBL2, ARFGAP3, DNAJB5, TMED7, G3BP2, BMPR1A, FMR1, TPST2, TMSB4X, RP2, CEP170, KLHL23, RNF7, HNRNPH1, MARCKS, HNRNPD, TOB1, UTP11L, RFK, DHX36, LCOR, WBP5, PHLDB2, USP33, EFNB2, C6ORF62, MEX3B, ABCD3, ATG3, ARID4B, C7ORF11, EPB41, TCF12, CDK8, CMIP, ATG12, CETN3, ZNF217, TMEM55A and UBE2N as compared to the expression level of the at least one gene in HES5- differentiated cells obtained by culturing HES5+ early radial glial cells under culture conditions suitable for differentiation the HES5+ early radial glial cells into HES5+ mid radial glial cells.

According to some embodiments of the invention, the HES5+ late radial glial cells exhibit an HES5+/OLIG1+/PDGFRA+ expression signature.

According to some embodiments of the invention, the HES5+ late radial glial cells further exhibit an CUX1+/CUX2+/POU3F2+ expression signature.

According to some embodiments of the invention, the HES5+ late radial glial cells further exhibit an S100B+/EGFR+ expression signature.

According to some embodiments of the invention, the HES5+ late radial glial cells are characterized by a higher expression level of at least one gene selected from the group consisting of: PMP2, GABBR2, BCAN, LUZP2, SALL3, SYNM, DCT, OLIG1, SPON1, PDGFRA, COL22A1, KIAA1239, PCDH10, LPAR4, VAV3, CADM2, SOX6, SLC6A1, DPP6, FGFR3, PDE3B, MOXD1, TNFRSF19, PYGL, GPC6, COL11A1, TRIM9, GABRB3, TFPI, CREB5, RAB3GAP2, NCAN, EFHD1, SLITRK2, PAX6, SLC1A4, GPR155, GPD2, CHST11, PAQR8, MT2A, GPC3, TMEM51, CHST3, PAG1, MYO5C, CACNB2, NDRG2, ST3GAL5, TPD52L1, TRIB1, PRKCA, BCKDHB, GLT25D2, LITAF, PLCB1, TIMP3, ZBTB46, OPCML, CTDSPL, MDGA2, MEGF10, EYA2, KANK1, RAB31, TRIL, FAM171B, ALCAM, RAB6B, PGM2L1, LARGE, HPCAL1, HTRA1, TRPS1, TRIB3, IGF2BP2, PITPNC1, CMTM4, IAH1, DHTKD1, SNAP29, CTNNBIP1, NQO2, MAP4, CBR1, LTBP1, C5ORF32, MARK1, AASS, CISD1, DSC2, SLC25A33, RIMS3, ZIC3, EGF, SRGAP2, RANGAP1, SCRG1, PRCP, CA12, HEATR5A, ZNF503, GYG2, ANAPC1,

C19ORF63, ASAP1, C1ORF96, DHX33, FASTKD1, STAU2, MAML2, RRAS2, GLTP, VPS13B, GPT2, NKAIN4 and ZC3HAV1 as compared to the expression level of the at least one gene in HES5- differentiated cells obtained by culturing HES5+ mid radial glial cells under culture conditions suitable for differentiation the HES5+ mid radial glial cells into HES5+ late radial glial cells.

According to some embodiments of the invention, the HES5+ long term neural progenitor cells exhibit an HES5+/ANXA2+/LGALS1+ expression signature.

According to some embodiments of the invention, the HES5+ long term neural progenitor cells further exhibit EGFR+/ S100B+ expression signature.

According to some embodiments of the invention, the HES5+ long term neural progenitor cells are characterized by a higher expression level of at least one gene selected from the group consisting of: ANXA2P2, ANXA2, FRAS1, SPOCK1, PCDHB15, SLC10A4, TPBG, C5ORF39, MMP14, TNFRSF10D, S100A6, RNF182, LGALS1, ISL1, SPINK5, DOCK10, LECT1, LYPD1, ARMCX1, NAP1L2, COL4A6, GSN, PLAG1, MMD, PTGR1, PDP1, COL18A1, ZIC4, BASP1, AHNAK, REC8, KLHDC8B, FRMD6, MYL9, RBMS1, TNFRSF21, and FAM38A as compared to the expression level of the at least one gene in HES5- differentiated cells obtained by culturing HES5+ late radial glial cells under culture conditions suitable for differentiation the HES5+ late radial glial cells into HES5+ long term neural progenitor cells.

According to some embodiments of the invention, the HES5+ cells are neuroepithelial cells (NE) which constitute at least about 80 % of the isolated population of cells.

According to some embodiments of the invention, the HES5+ cells are early radial glial cells (E-RG) which constitute at least about 70-80 % of the isolated population of cells.

According to some embodiments of the invention, the HES5+ cells are mid radial glial cells (M-RG) which constitute at least about 30 % of the isolated population of cells.

According to some embodiments of the invention, the HES5+ cells are late radial glial cells (L-RG) which constitute at least about 10-15 % of the isolated population of cells.

According to some embodiments of the invention, the HES5+ cells are long term neural progenitors (LNP); which constitute at least about 7-10 % of the isolated population of cells.

According to some embodiments of the invention, the HES5+ cells are genetically modified.

According to some embodiments of the invention, the cells are human cells.

According to some embodiments of the invention, the cells are derived from a subject having a CNS disease or disorder.

According to some embodiments of the invention, the cells having been subjected to an ex-vivo differentiation protocol.

According to some embodiments of the invention, the HES5+ neuroepithelial (NE) cells are capable of differentiating into E-RG, M-RG, L-RG and LNP cells.

According to some embodiments of the invention, the HES5+ early radial glial (E-RG) cells are capable of differentiating into M-RG, L-RG and LNP cells.

According to some embodiments of the invention, the HES5+ mid radial glial (M-RG) cells are capable of differentiating into L-RG and LNP cells.

According to some embodiments of the invention, the HES5+ late radial glial (L-RG) cells are capable of differentiating into LNP cells.

According to some embodiments of the invention, the isolated population of cells of some embodiments of the invention is for use in the treatment of a CNS disease or disorder.

According to some embodiments of the invention, the successive isolation comprises at least two isolation steps following at least two culturing steps, wherein a first isolation of the at least two isolation steps is effected up to 12 days of a first culturing of the at least two culturing steps, and wherein a second isolation of the at least two isolation steps is effected up to 5 days of a second culturing of the at least two culturing steps.

According to some embodiments of the invention, the successive isolation comprises at least three isolation steps following at least three culturing steps, wherein a third isolation of the at least three isolation steps is effected up to 21 days of a third culturing of the at least three culturing steps.

According to some embodiments of the invention, the successive isolation comprises at least four isolation steps following at least four culturing steps, wherein a fourth isolation of the at least four isolation steps is effected up to 45 days of a fourth culturing of the at least four culturing steps.

According to some embodiments of the invention, the successive isolation comprises at least five isolation steps following at least five culturing steps, wherein a fifth isolation of the at least five isolation steps is effected up to 140 days of a fifth culturing of the at least five culturing steps.

According to some embodiments of the invention, the first isolation results in a population of cells comprising HES5+ neuroepithelial cells.

According to some embodiments of the invention, the second isolation results in a population of cells comprising HES5+ early radial glial cells.

According to some embodiments of the invention, the third isolation results in a population of cells comprising HES5+ mid radial glial cells.

According to some embodiments of the invention, the fourth isolation results in a population of cells comprising HES5+ late radial glial cells.

According to some embodiments of the invention, the fifth isolation results in a population of cells comprising HES5+ long term neural progenitor cells.

According to some embodiments of the invention, the first culturing is performed on an extracellular matrix or a feeder cell layer.

According to some embodiments of the invention, the first culturing is performed in the presence of a culture medium which comprises Noggin, SB-431542 and LDN-193189.

According to some embodiments of the invention, the second culturing is performed on an extracellular matrix.

According to some embodiments of the invention, the second culturing is performed in the presence of a culture medium which comprises fibroblast growth factor 8 (FGF8) and brain-derived neurotrophic factor (BDNF).

According to some embodiments of the invention, the third culturing is performed on an extracellular matrix.

According to some embodiments of the invention, the third culturing is performed in the presence of a culture medium which comprises fibroblast growth factor 2 (FGF2) and epidermal growth factor (EGF).

According to some embodiments of the invention, the fourth culturing is performed on an extracellular matrix.

According to some embodiments of the invention, the fourth culturing is performed in the presence of a culture medium which comprises fibroblast growth factor 2 (FGF2) and epidermal growth factor (EGF).

According to some embodiments of the invention, the fifth culturing is performed on an extracellular matrix.

According to some embodiments of the invention, the fifth culturing is performed in the presence of a culture medium which comprises fibroblast growth factor 2 (FGF2) and epidermal growth factor (EGF).

According to some embodiments of the invention, the method of isolating neural progenitor cells further comprising qualifying presence of a neural progenitor cell of interest according to at least one marker comprised in an expression signature of the neural progenitor cells, wherein:

(i) an expression signature of HES5+ neuroepithelial cells comprises HES5+/SOX1+/PAX6+/SOX2+/Nestin+/CDC6+/CDX1+/CENPH+/TOP2A+;

(ii) an expression signature of HES5+ early radial glial cells comprises HES5+/ARX+/FEZF2+/NR2E1+;

(iii) an expression signature of HES5+ mid radial glial cells comprises HES5+/POU3F2+/GLAST+/FABP7+;

(iv) an expression signature of HES5+ late radial glial cells comprise HES5+/OLIG1+/PDGFRA+/CUX1+/CUX2+/POU3F2+/S100B+/EGFR+;

(v) an expression signature of HES5+ long term neural progenitor cells comprise HES5+/ANXA2+/LGALS1+/EGFR+/ S100B+.

According to some embodiments of the invention, the method further comprising qualifying presence of a neural progenitor cell of interest according to epigenetic analysis functional phenotype and/or morphological phenotype.

According to some embodiments of the invention, the stem cells are derived from a subject having a CNS disease or disorder.

According to some embodiments of the invention, the CNS disease or disorder comprises a motor-neuron disease.

According to some embodiments of the invention, the CNS disease or disorder is characterized by cortex damage.

According to some embodiments of the invention, the culture medium further comprising sonic hedgehog.

According to some embodiments of the invention, the HES5- cells of (i) are characterized by a higher expression level of at least one gene selected from the group consisting of: LHX1, CNTN2, ST18, EBF3, NFASC, FSTL5, ONECUT2, SLC17A6, EBF1, SLIT1, SYT4, NEFM, NEUROD1, PARM1, CHN2, DNER, HMP19, TFAP2B, DCX, KLHL35, PAPPA, OLFM1, NHLH1, RTN1, GAP43, GFRA1, CHL1, FNDC5, SCN3A, NPTX2, EOMES, CADPS, NHLH2, TMEM163, STMN3, LRRN3, NEFL, ROBO2, INA, PHLDA1, GRIA1, GRIA2, DCLK1, CRABP1, OLIG2, SCG3, TMEM158, FBXL16, FAM123A, SYP, KIF21B, PCDH9, CDKN1C, IGFBPL1, RSPO3, GABRB3, TAGLN3, KCNH2, EPB41L3, EYA2, TMOD2, NCAN, GABBR2, D4S234E, PI15, ANK2, SLC1A2, NRCAM, CBLB, CAMK2N1, ZIC3, PTPRN2, SORBS1, NXPH1, BAALC, CLASP2, DPP6, MAP6, FIGNL2, KIAA0802, KIF1A, TMEM170B, SLC22A23, TMEM178, CTNND2, CADM1, LGALS3, TCEAL7, CD9, GLB1L2, NFIA, YPEL3, TNFRSF19, SPINK5, PNMA2, IRX5, SIN3B, STK38, NR3C1, SOX8, SLC6A8, SYNPR, SGK223, BASP1 and APC as compared to the expression level of the at least one gene in HES5+ neuroepithelial cells.

According to some embodiments of the invention, the HES5- cells of (ii) are characterized by a higher expression level of at least one gene selected from the group consisting of: GREM1, COL3A1, PCDH8, SEMA3C, BMP4, NID2, TNC, COL1A2, ANKRD1, ANXA1, TMEFF2, PDZRN3, ANXA3, KRT8, LEPREL1, NOX4, LAMB1, FLNC, FST, IMMP2L, S100A4, GDF15, PHACTR2, METTL7A, MAMDC2, DDIT4, BCHE, OCIAD2, TNFRSF10D, BBS9, ELOVL2, TUBA1C, CHST7, RBM47, TFPI, NEBL and LHFP as compared to the expression level of the at least one gene in HES5+ early radial glial cells.

According to some embodiments of the invention, the HES5- cells of (iii) are characterized by a higher expression level of the ACSS1 gene as compared to the expression level of the gene in HES5+ mid radial glial cells.

According to some embodiments of the invention, the HES5- cells of (iv) are characterized by a higher expression level of at least one gene selected from the group consisting of: THBS1, KLHL4, A2M, EN2, SLC6A6, ACTA2, ST6GAL1, SLC7A8, GRM3, FAM65B, CALB1, MYLK, TNNT1, PTX3, MFAP2 and HMGA2 as compared to the expression level of the at least one gene in HES5+ late radial glial cells.

According to some embodiments of the invention, the HES5- cells of (v) are characterized by a higher expression level of at least one gene selected from the group consisting of: FBN2, NELL2, KAL1, PCDHB5, ST8SIA4, DCN, SLC6A1, CADM2, BCL11A, DDB2, ANXA11, PAK1, ID3, IGF2BP1, ANK3, ZEB2 and CREB5 as compared to the expression level of the at least one gene in HES5+ long term neural progenitor cells.

According to some embodiments of the invention, the HES5+ neuroepithelial cells are characterized by an increased transcriptional epigenetic activity of at least one transcription factor encoded by the gene selected from group consisting of: RFX4, PAX6, TCF4, SOX9, LEF1, POU3F2, SOX8, SOX21, TEAD1, NFATC1, SOX5, TGIF1, MEIS1, TCF4, MEIS2, OTX2, TEF, ZBTB16, MSX1, RFX1, NR4A2, MEIS2, SOX15, STAT5B, SATB1, RBPJ, FOXK1, MYBL2, DMRT3, NFIA, CUX1, TFAP4, MSX1, CDC5L, RFX1, FOXJ2, POU6F1, TEF, RBPJ, PKNOX2, BCL6, PRRX1, STAT1, POU3F1, FOXB1, CTNNB1, PBX1, ZNF143, NFATC1, SOX21, TCF7L1, ARX, SOX15, RXRA, TFAP4, CUX1, OTX2, NR2E1, CUX1, ZNF232, NR2F1, SOX4, MEIS1, PBX1, CUX1, NEUROD1, MSX1, ZNF652, MEF2A, OLIG1, POU6F2, IKZF2, MECOM, STAT1, ESRRA, IRF7, STAT1, MYBL2, BCL6, ELK1, ATF2, SMAD3, ATF4, DLX1, MEF2A, DBP, MAF, MEF2A, TEAD2, SMAD3, SOX15, POU6F1, BARHL2, FOXG1, LHX9, MECOM, ARNTL, MYC, ZNF75A, NFIA, VAX1, GBX2, HOMEZ, FOXO4, FOXO4, FOXB1, ZSCAN16, ELK1, ATF2, CREB1, USF1, ESRRA, ZNF282, NEUROG2, NFYA, NR4A1, CTF1, ELK1, POU3F2, ELK1, HSF1, E2F3, CUX1, CREB1, ELF2, MYBL2, HMGA2, SRF, ZNF410, JDP2, NR2F1, PAX3, NRF1, SMAD4, ZNF85, ZNF628, NFATC1, CREB1, E4F1, NR2F6, NHLH1, IRF2, PBX1, FOXJ3, RORA, IRF7, NR6A1, LHX2, PAX3, NR2E1, POU3F1, ZFP42, E2F4, ETV5, ELF3, USF1, ATF6, TFAP2B, CUX1, IRF3, RXRA, PEBP1, LHX2, GZF1, MEF2A, MEF2A, IRF9, MGA, VBP1, GMEB2, YY1, ELF1, POU3F3, GTF2IRD1, IRF3, SRF, XBP1, ESRRA, HEY1, NFKB2, IRF2,

EOMES, FOXB1, NR2F1, NR4A2, STAT3, SP1, RARA, CREB1, NR2F1, FOXJ3, HSF1, MYBL2, SRF, ETV2, MECP2, E2F1, FOXJ2, JUN, SCRT2, DLX1, E2F1, E2F1 and ATF2as compared to the transcriptional epigenetic activity of the at least one transcription factor in human ESC cells.

According to some embodiments of the invention, the HES5+ early radial glial cells are characterized by an increased transcriptional epigenetic activity of at least one transcription factor encoded by the gene selected from group consisting of: NFIA, MZF1, SMAD4, CTF1, SMAD3, RFX4, MAF, TCF12, NFYC, ZNF263, MECP2, ZFP42, ZIC1, YY1, ETS2, NR2C2, SREBF2, SREBF1, MEIS2, NR4A1, REST, SF1, ZBTB7A, STAT3, E2F1, NR1H2, NFKB1, NR2F6, GLIS3, MAZ, STAT1, TGIF1, SOX9, HES1, THRA, GLIS3, MEIS1, ESRRA, ZBTB7A, NFYA, MECP2, PKNOX2, EP300, TFAP2B, NR2F1, MZF1, ESRRA, TFCP2, NR2F1, ESRRA, TERF1, KLF3, XBP1, RORA, PBX1, MYC, SNAI2, TEAD2, CENPB, PEBP1, HINFP, SREBF1, YY1, E2F1, HSF2, CNOT3, MEIS2, MEF2A, RXRA, CREB3L1, MYBL2, ZNF524, TFAP4, RFX1, NFYA, ZBTB33, RREB1, NR6A1, HES5, TFAP2B, HIF1A, INSM1, ZNF524, ELF1, SMAD4, STAT3, TFAP2B, USF1, GLI3, GLI2, ETV2, E2F1, CREB3, NFKB1, MYC, SRF, TFCP2, ATF3, ELK1, TP53, E2F4, ELF2, DEAF1, RXRA, ZNF423, FOXG1, ZNF628, NEUROD1, ARNTL, STAT1, PAX3, NR6A1, ESRRA, TCF4, YY1, PBX1, VAX1, NRF1, ETV1, FOXB1, ZNF85, CREB1, PRDM4, MYBL2, NHLH1, CREB1, KLF13, TRIM26, ZNF148, E4F1, USF1, JDP2, OLIG1, GZF1, ATF2, CREB1, FOXO4, SMAD3, CLOCK, ZNF282, ZSCAN16, CUX1, NFKB1, RFX1, ETV5, MEF2A, GTF2IRD1, TFAP2B, NHLH1, ELF1, PAX3, MNT, ZNF740, CREB1, ESRRA, EOMES, EGR2, NR4A2, BCL6, ZNF143, MAFF, CUX1, NR2F1, TFAP2B, NEUROG2, TCF3, ATF3, PLAG1, ESRRA, E2F1, KLF13, MSX1, SCRT2, GABPB1, SOX9, NFKB2, EOMES, SIX3, MECOM, BCL6, POU3F3, HINFP, NFATC1, TBR1, MYBL2, USF1, SP1, ATF2, LHX2, MYBL2, POU3F3, HEY1, MEF2A, MGA, SOX9, E2F7, IRF3, SRF, POU3F2, SREBF2, NFKB1, MAF, ZNF784, NR1H2, MSX1, RORA, LMO2, HINFP, ZNF23, FOXO4, NR4A2, POU6F2, ID4, RARA, KLF13, LEF1, PATZ1, ELK1, DLX1, E2F4, E2F4, ELK1, MAZ, ATF6, TFAP4, RARA, PAX6, EGR2, HOMEZ, ELK1, E2F1, GMEB2, ATF4, ATF2, MAX, MGA, NFKB2, ESRRA, E2F4, STAT3, TCF3, NR2E1, ELF3, SOX4, USF1 and

ZNF652 as compared to the transcriptional epigenetic activity of the at least one transcription factor in HES5+ neuroepithelial cells.

According to some embodiments of the invention, the HES5+ mid radial glial cells are characterized by an increased transcriptional epigenetic activity of at least one transcription factor encoded by the gene selected from group consisting of: RFX4, NFIA, NFATC1, CTF1, NEUROD1, RFX1, TGIF1, ETS2, TFAP4, TCF12, TFAP4, SMAD3, NFIA, SOX9, RFX1, TCF4, MZF1, STAT1, MECP2, MEIS2, ZBTB33, NFYA, ELF1, MYBL2, LEF1, NFYC, MAFF, ZNF263, YY1, POU3F3, TGIF1, STAT3, SMAD4, NR6A1, TGIF1, MEIS1, ZNF628, ZFP42, FOXK1, PRDM4, STAT1, MAF, SCRT2, CREB1, GZF1, CREB1, VAX1, MECP2, NHLH1, ETV1, SOX9, PEBP1, SMAD4, XBP1, USF1, POU3F2, CREB3, EP300, PBX1, STAT3, TFCP2, POU2F2, IRF3, FOXB1, MSX1, POU3F3, ELK1, DBP, CUX1, MEF2A, POU6F2, ARNTL, ZSCAN16, MEF2A, ETV4, OLIG1, HOMEZ, DLX1, PRRX1, MSX1, MYC, FOXO4, MEF2A, MZF1, ATF2, GMEB2, NFYA, ESRRA, SOX9, PBX1, POU3F2, MECOM, SMAD3, MAZ, ELK1, BCL6, ELF3, PAX3, ELF1, SF1, BCL6, EMX2, STAT1, E2F1, ELF2, CLOCK, ELK1, STAT3, ATF5, THRA, SOX9, SRF, TCF4, ATF3, CTNNB1, USF1, FOXJ3, USF1, ZNF282, NEUROG2, ESRRA, REST, E2F3, ZBTB7A, MYBL2, HSF2, MAX, ZNF143, MYBL2, SRF, FOXO4, NR4A2, CUX1, E2F4, MSX1, EOMES, MAF, MNT, POU6F1, NFATC1, STAT3, CREB3, SOX9, ZNF85, CREB3L1, TCF3, ELK1, IRF2, YY1, SOX15, PAX6, E4F1, MEF2A, ATF2, NFE2L1, NFATC1, ATF2, YY1, SRF, ARX, ETV2, HINFP, MAZ, NR4A1, INSM1, ZNF652, USF1, NFKB1, GBX2, POU3F4, IRF9, POU6F1, LHX2, NR4A2, POU6F2, FOXJ2, CEBPG, VBP1, TERF1, ESRRA, PAX6, ZIC3, IRF2, MAF, SOX21, CREB1, CREB3, IRF7, POU3F3, NFKB2, ATF4, MYBL2, SREBF2, SOX9, YY1, RBPJ, FOXJ3, HSF1, HMGA2, CUX1, POU3F2, EOMES, ZNF423, ESRRA and KLF13 as compared to the transcriptional epigenetic activity of the at least one transcription factor in HES5+ early radial glial cells.

Unless otherwise defined, all technical and/or scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which the invention pertains. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of embodiments of the invention, exemplary methods and/or materials are described below. In case of conflict, the patent

specification, including definitions, will control. In addition, the materials, methods, and examples are illustrative only and are not intended to be necessarily limiting.

BRIEF DESCRIPTION OF THE SEVERAL VIEWS OF THE DRAWINGS

Some embodiments of the invention are herein described, by way of example only, with reference to the accompanying drawings. With specific reference now to the drawings in detail, it is stressed that the particulars shown are by way of example and for purposes of illustrative discussion of embodiments of the invention. In this regard, the description taken with the drawings makes apparent to those skilled in the art how embodiments of the invention may be practiced.

In the drawings:

FIGs. 1A-C depict a schematic illustration and microscopy images showing that notch activation links major neural lineage transitions in hESC derived neural stem and progenitor cells. Figure 1A - Neural differentiation scheme showing the differentiation of human stem cells into the neocortex. Neural induction was performed by a dual SMAD inhibition protocol followed by long-term propagation with the factors indicated for 220 days ("D" = day(s)). Naming conventions representing neuroepithelial (NE), early radial glial (E-RG), mid-radial glial (M-RG), late radial glial (L-RG) and long term cultured progenitors (LNP) are indicated. Figure 1B - Bright field microscopy of progenitor cells during long-term differentiation shows dynamic morphological features. Scale bar: 50 μm (valid for all images in Figure 1B). GFP (green fluorescent protein) matched images can be seen in Figure 8A. Figures 1C-E - Combined *HES5::eGFP* reporter expression and immunostainings of stem/ progenitor as well as differentiation markers throughout the progression period. Figure 1C - PAX6, SOX1 and HES5 induction during early stages (See also Figures 8A-B for *HES5::eGFP* percentages).

FIG. 1D – SOX1 expression and DAPI nuclear counterstaining;

FIG. 1E - NESTIN and DCX expression. Scale bar: 50 μm (valid for all images in Figures 1C-E). Individual qPCR analyses for all genes tested at all stages are shown in Figure 8D.

FIGs. 2A-C demonstrate that early Notch activation in NE cells confers amenability to neural patterning cues. Figure 2A - Neural patterning paradigm scheme. PSCs (pluripotent stem cells) were subjected to neural induction and were exposed to

patterning cues directing differentiation into forebrain, midbrain and spinal cord cell fates with the morphogenes indicated. Region specific progenitors were sorted to high, medium or low *HES5::eGFP* expressing populations followed by neuronal differentiation. Figure 2B - Immunostaining for respective neuronal progeny derived from HES5+ (top) or HES5- (bottom) on day 12 of progression. Cortical neurons are marked by TBR1 (left top image, in red), midbrain dopaminergic neurons marked by FOXA2/TH (middle top image, in red) and spinal cord motoneurons are marked by HB9 (right top image, in red) are shown. Scale bar: 50 μm. Figures 2C-E – Histograms depicting quantitative PCR analysis of transcript levels of HES5 as well as selected regional markers in high (++, dark green bars), medium (+, light green bars) and low (-, gray bars) HES5 expressing progenitors, in their proliferative state (Day 12 or Day 14) and following terminal neuronal differentiation (Day 19, Day 26, or Day 28). Figure 2C – Midbrain dopamine neurons;

FIG. 2D – Forebrain cortical neurons;

FIG. 2E – Spinal cord Motoneurons. All transcript levels shown are normalized to respective HPRT levels in each sample. Values were obtained from three technical replicates. Statistical analysis: mean ± SEM; T- Test: (***) $P < 0.001$; (**) $P < 0.01$; (*) $P < 0.05$. Individual qPCR analyses for additional regional or neuronal markers are shown in Figure 9A.

FIGs. 3A-E demonstrate that consecutive isolation of Notch active progenitors recapitulates cortical lamination and glial fates.

Figure 3A - Combined *HES5::eGFP* reporter expression and immunostainings of cortical layer specific neuronal markers. Early born neurons expressing TBR1, RELN and CTIP2 (top two panels), and late derived neurons expressing SATB2, BRN2 and CUX1 (bottom two panels), are shown for NE, M-RG and L-RG progenitors that were subjected to neuronal differentiation. Insets for RELN/TBR1 and SATB2/BRN2 show magnified areas within the image. Inset for CTIP/TUJ1 show same magnification but a different view of neuronal axons. Images of HES5+ derived neurons are shown. Scale bars: 50 μm for images, 25 μm for Insets. Images of HES5- derived neurons and percentages of all cortical subtypes derived from both HES5+ and HES5- cells are presented in Figure 12B.

Figure 3B - Distribution of relative transcript abundance based on qPCR for selected stage specific marker gene groups for either deep or upper layer neuronal progeny. Contributions of HES5+ (marked as +) and HES5- (marked as –) populations per each respective stage are shown. Marker gene groups for each progenitor stage were created by collapsing the normalized values of TBR1/RELN, CTIP2/FEZF2 and CUX1/CUX2/SATB2 (see Experimental Methods in the Examples section which follows for details). Individual qPCR analyses for all genes tested at all stages are shown in Figures 11A-K.

Figures 3C-D – A histogram depicting the cumulative neuronal marker levels based on absolute transcript levels (Figure 3C) and a graph depicting the sum of neuronal transcripts in HES5+ or GES5- cells (Figure 3D). Note the decrease in total neuronal progeny as shown in Figure 3D.

Figure 3E - Distribution of relative transcript abundance based on qPCR for selected stage specific marker genes for indicated progenitor or neuronal cell markers. Contributions of HES5+ and HES5- populations per each respective stage from either untreated or DAPT treated cells are shown. Expression levels relative to HPRT of all four conditions (color coded) were summed per each gene and plotted as a single bar.

FIGs. 3F-G – An image (Figure 3F) and a histogram (Figure 3G) depicting combined *HES5::eGFP* reporter expression and immunostaining of the glial marker GFAP following differentiation of distinct progenitor stages (Figure 3F). Scale bar: 50 μm. Figure 3G - GFAP transcript level as assessed by qPCR. Values were obtained from three technical replicates. Statistical analysis: mean ± SEM. Note that the glial marker GFAP is upregulated (Figure 3G).

FIGs. 4A-E demonstrate that transition through progenitor cell stages demarcates developing rosettes as VZ and SVZ equivalents. Figure 4A - Differential expression levels for selected genes that are most differentially expressed between HES5+ and HES5- cells in a stage specific fashion. Selected gene members are indicated on the left, developmental stages are indicated on the bottom, and gene categories classified by stage are indicated on the right. Values plotted on the heatmap represent ratios of expression levels relative to ES cells. Figures 4B-C - Relative expression levels (z-scores) based on microarray expression data for the entire differentiation time course for selected germinal zone marker genes. Expression levels are shown for HES5+ (Figure

4B) and HES5- (Figure 4C) samples separately. Genes are ordered from VZ to SVZ and from neurogenic to gliogenic markers. Individual qPCR analyses for all genes tested at all stages are shown in Figure 13C. Note that the apparently high GFAP expression in HES5+ cells at the L-RG stage has in fact low absolute expression values, and only appear high relatively to expression in other stages (all stages per each gene are normalized to 1; *i.e.* highest red intensity). To compare GFAP transcript levels during proliferation and serum induced astrocytic differentiation, see Figures 5D and 3G, respectively. Figure 4D - Combined *HES5::eGFP* reporter expression and immunostainings of neural stem / progenitor markers, RG markers, and proliferation markers throughout the progression period. From Top: PAX6 marking the VZ and TBR2 marking the SVZ are shown. Middle: CUX1 marking SVZ is shown. Bottom: the mainly SVZ marker POU3F2 is shown. Scale bar: 50 μm (valid for all images in Figure 4D). Figure 4E - High power magnification of E-RG and M-RG images shown in Figure 4D. Dashed lines demarcate proposed VZ, SVZ and OSVZ regions, containing apical RG, INPs and basal RG, respectively. Scale bar: 25 μm (valid for all images in Figure 4E).

FIGs. 5A-D - Glial transformation with respect to Notch activation. Figure 5A - Combined *HES5::eGFP* reporter expression and immunostainings of the RG markers GLAST (top) and FABP7 (bottom). Scale bar: 50 μm (valid for all images in Figure 5A). Figure 5B - High power magnification of E-RG and M-RG images for selected genes shown in Figure 5A. Scale bar: 25 μm (valid for all images in Figure 5B). Figure 5C - EGFR expression percentages by FACS analysis for L-RG (purple) and LNP (turquoise) stages is shown. Average of 2 independent experiments is shown. Statistical analysis: mean ± SEM. Figure 5D - Relative GFAP expression levels based on qPCR data for the entire progression period. Relative expression levels are shown for HES5+ and HES5- samples during progenitor proliferation. Values were obtained from three technical replicates. Statistical analysis: mean ± SEM. Compare the very low absolute levels of GFAP during proliferation (Day 80 HES5+ cells) to GFAP levels at the same progenitor type following astrocytic differentiation in Figures 3F-G.

FIGs. 6A-B Global gene expression cluster analysis for stage specifically expressed genes. Figure 6A - Selected trends of global gene expression clustering comprising 496 genes using k- means (k=100) for all differentially expressed genes across the differentiation time course of HES5+ cells. Gene expression levels were log2

transformed and normalized to hESC by subtracting their respective hESC level. For all stages expression levels in HES5+ and HES5- cells, are shown. Selected gene members of each cluster are indicated on the right while the cluster naming conventions are indicated on the left. Arrays were obtained from single replicates first used as a discovery tool, and then extensively validated by qPCR and immunostainings from independent experiments throughout the manuscript. Figure 6B - Gene set enrichment analysis results (using IPA, p-Values are calculated using right-tailed Fisher Exact Test) of gene sets selected from the top 10 categories for each cluster are shown. Color code indicates –log10 p-Value.

FIG. 7 - Schematic model for NSC progression. Neuroectodermal cells yield the earliest NE cells of the CNS by launching Notch activation and HES5 expression, while non-CNS neuroectodermal cells lack this activation. Under proliferation conditions, HES5+ NE cells yield consecutive radial glial progenitor cell types and their corresponding neuronal and glial progeny, hence considered as primary NSCs generating CNS neural diversity. Following mitogen withdrawal, HES5+ NE cells exert their competence towards deep layer specific neuronal types (RELN, TBR1 but also FEZF2 and CTIP2; blue-to-red wave, bottom panel) and do so in a Notch dependent manner. In addition, they also upregulate SVZ progenitor markers such as TBR2 and CUX1, CUX2 at the RNA level (Red font, light brown early wave; bottom panel) and in a Notch dependent manner, implying on their future potential to generate these progenitors at later stages. In contrast to NE cells, HES5+ E-RG cells are already committed to early dorsocaudal cortical identity, based on their elongated polarized cell morphology, rosette formation capacity and FEZF2 and EMX2 expression. Hence, they exhibit competence towards deep layer neurons (CTIP2, FEZF2; blue-to-red wave, bottom panel). M-RG stage cells are characterized by lower HES5 percentages, reduced rosette organization, substantial accumulation of HES5+ derived HES5- progenitors expressing CUX1, CUX2 and TBR2 at the protein level, and competence for yielding upper layer neuronal fates (CUX1, CUX2, SATB2; brown wave, bottom panel) in a Notch independent manner. HES5+ L-RG cells are able to give rise to astrocytes in a Notch dependent manner (GFAP; light blue wave, bottom panel), yet both HES5+ and HES5- cells at that stage continue to contribute to neurogenesis. Ultimately, L-RG cells transform to long-term progenitors (LNP) associated with adult NSC progeny (purple

wave, bottom panel). Horizontal green and black arrows mark transition in a Notch dependent and independent manner, respectively. Diagonal green and black arrows mark HES5+ and HES5- cells, respectively, subjected to differentiation following FACS-based separation. When indicated, Notch active pathways were confirmed by DAPT (red bar-headed lines). Top panel shows cell types and developmental potential. Bottom panel shows temporal phases of neuronal and glial markers derived by the stages indicated above. BP, basal progenitors.

FIGs. 8A-F demonstrate HES5 expression dynamics in PSC derived neural progenitors. Figure 8A - Fluorescent microscopy of *HES5::eGFP* during long-term differentiation shows dynamic morphological features through neural progenitor cell progression *in vitro*. Scale bar: 50 μm. Figure 8B - Top: FACS charts depicting ES cells purified for pluripotency markers. SSEA-4 and TRA-1-60 surface markers are presented (right). Unstained cells are shown on the left. Bottom: *HES5::eGFP* percentages for all stages are shown. Percentages indicated are representative of three independent experiments. Figure 8C - SOX1 and PAX6 expression in neuroectodermal cells at days 5 and 8, prior onset of *HES5::eGFP*. Scale bar: 25 μm. Figure 8D - Quantitative PCR analysis of transcript levels of neural stem and progenitor cell markers (for whom immunostainings is shown in Figures 1C-E. Relative expression levels for HES5+ and HES5- samples across the entire progression period are shown. Values were obtained from three technical replicates. Statistical analysis: mean ± SEM. Figure 8E - Top: Expression dynamics of early rosette markers (PLZF) and late glial progenitor markers (S100B) during progression in vitro are shown. Scale bar: 50 μm. Bottom: qPCR analysis of PLZF, S100B and EGFR are shown. Figure 8F - Combined *HES5::eGFP* reporter intensity and immunostainings for multipotency marker expression of progeny derived from LNP progenitor stages. GFAP, TUJ1 and O4 are shown from left to right. Scale bar: 50 μm.

FIGs. 9A-F demonstrate CNS fate specification and regional patterning potential in HES5+ and HES5- progenitors. Figure 9A-Cs - Quantitative PCR analysis of transcript levels of the regional marker HOXB4 (Figure 9B) and the neuronal marker TUJ1 (Figure 9A) during motoneuron differentiation (Figures 9A-B), or TUJ1 expression during dopamine neuron differentiation (Figure 9C). High (++, dark green bars), medium (+, light green bars) and low (-, gray bars) HES5 expressing progenitors,

in their proliferative state (Day 12 or Day 14) and following terminal neuronal differentiation (Day 28 for motoneurons, Day 19 and Day 26 for dopamine neurons) are shown. All transcript levels shown are normalized to respective HPRT levels in each sample. Values were obtained from three technical replicates. Statistical analysis: mean ± SEM; T-Test: (***) $P < 0.001$; (**) $P < 0.01$; (*) $P < 0.05$. Figures 9D-F - CNS fate specification required Notch activation. Neurally induced cells were treated with or without the Notch inhibitor DAPT on either day 2 or day 6 of neural induction and harvested for qPCR analysis on Day 9. Relative expression (compared to HPRT) of early appearing CNS markers PAX6 (Figure 9D) and HES5 (Figure 9E) and the neural crest/placodal marker SIX1 (Figure 9F) on Day 9 is shown. Values were obtained from three technical replicates. Statistical analysis: mean ± SEM.

FIGs. 10A-E depict cell fate and proliferation marker segregation in consecutively sorted HES5+ and HES5- cells. Figures 10A-C - Immunostainings for CNS and non-CNS markers PAX6 and AP2a, respectively, in NE stage HES5+ and HES5- cells and their directly and consecutively derived E-RG stage HES5+ and HES5- progeny (see Experimental Methods in the Examples section which follows for details). Acute fixation and staining following sorting is shown (Figure 10A). Quantification of PAX6 (Figure 10B) and AP2a (Figure 10C) cell ratios reflecting segregation of CNS and non-CNS cell fates via Notch activation and inactivation, respectively, is shown. Lineage relations for each stage analyzed (NE, E-RG) are indicated by vertical arrows on the x-axis. Scale bar for images and insets: 50 μm. Figure 10D - Additional anterior CNS and NSC markers OTX2 and SOX2, as well as the neuronal marker DCX are shown for NE stage HES5+ and HES5- progenitors acutely sorted and analyzed. Scale bar: 50 μm. Figure 10E - Immunostainings for the S-Phase marker BrdU are shown in an experiment performed similar to the one presented in Figure 10A. Immunostainings were performed immediately after sorting, re-plating, and 1 hour of BrdU labeling are shown. Quantification of BrdU+ cell ratios is shown through stages examined on the right. Quantifications in Figures 10A through 10E are representative of at least 2 independent experiments. Scale bar: 25 μm.

FIGs. 11A-K depict transcript validation of cortical lamination by PSC derived consecutively appearing neural progenitors. Individual qPCR analyses of laminar markers in neuronal progeny derived from HES5+ and HES5- progenitor populations

from NE, E-RG, M-RG and L-RG stages. Figure 11A – RELN expression levels in marginal zone; Figure 11B – TBR1 expression levels in subplate, layer VI; Figure 1C – CTIP2 expression levels in subcerebral, layer V; Figure 11D – FEZF expression levels in subcerebral, layer V; Figure 11E – NR2F1 expression levels in subcerebral, layer V; Figure 11F – PCP4 expression levels in subcerebral, layer V;; Figure 11G – TBR2 expression levels in SVZ, callosal layer, II-IV; Figure 11H – CUX2 expression levels in SVZ, callosal layer, II-IV; Figure 11I – SATB2 expression levels in SVZ, callosal layer, II-IV; Figure 11J – BRN2 expression levels in SVZ, callosal layer, II-IV; Figure 11K – CUX1 expression levels in SVZ, callosal layer, II-IV; Values from selected genes were collapsed together for generating the pie charts and bars shown in Figures 3B, 3C, and 3E. All transcript levels shown are normalized to respective HPRT levels in each sample. Values shown were obtained from three technical replicates of a representative experiment. Statistical analysis: mean ± SEM; T-Test: (***) $P < 0.001$; (**) $P < 0.01$; (*) $P < 0.05$.

FIGs. 12A-G depict differentiation capacity of HES5- progenitor cell stages. Figure 12A - Combined *HES5::eGFP* reporter expression and immunostainings of cortical layer specific neuronal markers for neuronal progeny derived from HES5- cells across stages is shown. Panels and stainings are ordered identically to the ones shown for HES5+ progenitor stages in Figure 3A. Insets show compressed magnification of a matched DAPI image for the entire corresponding image. Scale bars: 50 µm for images, 100 µm for Insets. Figures 12B-G - Quantification of marker immunofluorescence intensity of the neuronal progeny shown in Figure 12A. Figure 12B – RELN expression; Figure 12C – TBR1; Figure 12D – CTIP2; Figure 12E – SATB2; Figure 12F – BRN2; Figure 12G – CUX1. Entire image cell counting relative to DAPI of at least 2 independently taken images for one representative experiment is shown. Also shown in each of the charts is the quantification of cell ratios expressing these specific neuronal markers from HES5+ progenitors (for which images are shown in Figure 3A). Statistical analysis: mean ± SEM.

FIGs. 13A-J depict spatiotemporal progenitor marker expression during progression in vitro: rosettes as VZ and SVZ equivalents. Figure 13A - Combined *HES5::eGFP* reporter expression and immunostainings for the mitotic (M-Phase) marker PHH3 and the cell cycle marker KI67. Scale bar: 50 µm. Right panel shows high

power magnification of E-RG and M-RG rosettes shown on the left. Scale bar: 2 μm. Figure 13B - Separate channel presentation for high power magnification images of E-RG and M-RG rosettes shown in Figure 4E. POU3F2, TBR2, and CUX1 are shown. Dashed lines demarcate proposed VZ, SVZ (also referred to as inner SVZ) and OSVZ (outer SVZ) regions, containing apical RG (radial glial), Intermediate neural progenitor (INP)s and putative basal RG, respectively. Scale bar: 25 μm. Figures 13C-J - Quantitative PCR validations of transcript levels for all VZ and SVZ markers whose gene array levels are represented by heatmaps on Figure 4B are shown. Figure 13C – FEZF; Figure 13D – TBR2; Figure 13E – POU3F2; Figure 13F – POU3F3; Figure 13G – CUX1; Figure 13H – CUX2; Figure 13I – FABP7; Figure 13J – GLAST. Relative expression levels for HES5+ and HES5- samples across the entire progression period are shown. Values were obtained from three technical replicates of a representative experiment. Statistical analysis: mean ± SEM.

FIGs. 14A-G depict stage specific marker validation. Figure 14A-E - qPCR validation of transcript levels for selected markers across all stages and across HES5+ and HES5- populations. Figure 14A – LRG5; Figure 14B – NR2E1; Figure 14C – EZH2; Figure 14D – DCN; Figure 14E – LGALS1. Relative expression (compared to HPRT) is shown. Values were obtained from three technical replicates. Statistical analysis: mean ± SEM. Figure 14F-G - Combined *HES5::eGFP* reporter expression and immunostainings for selected stage specific identified markers. Figure 14F – images, Scale bar: 50 μm; Figure 14G - Statistical analysis for qPCR, shown are the Mean ± SEM.

FIGs. 15A-D show that consecutive stages of ES cell derived neural progenitors are characterized by distinct epigenetic states. Figure 15A - Left: Schematic illustration of the cell system. Middle: Normalized read-count level for H3K27ac over a 1.4 mega base (mb) region around the *SOX2* locus (chr3:180,854,252-182,259,543). ChIP-Seq read counts were normalized to 1 million reads and scaled to the same level (1.5) for all tracks shown. Right: Additional tracks for H3K4me3, H3K4me1 and H3K27me3 as well as DNAme (scale 0-100%), OTX2 and expression covering a 100 kilo base (kb) sub-region (chr3:181,389,523-181,490,148) of this locus. Histone and RNA-Seq data were normalized to 1 million reads and are shown on distinct scales. Figure 15B - Maximum gene set activity levels shown as z-scores for genes expressed in defined brain structures

(left) and developmental time points (right) based on the mouse Allen Brain Atlas. Gene set activity was defined as average expression level of all member genes followed by z-score computation across all nine cell types. Abbreviations used: Rostral secondary prosencephalone (RSP), Telencephalon (Tel), peduncular (caudal) hypothalamus (PHy), Hypothalamus (p3), pre-thalamus (p2), pre-tectum (p1), midbrain (M), prepontine hindbrain (PPH), pontine hindbrain (PH), pontomedullary hindbrain (PMH), medullary hindbrain (MH); and embryonic (E)11.5, E13.5, E15.5, E18.5 as well postnatal P4, P14 and P28. Figure 15C - Distribution of DNAme levels for differentially methylated regions (delta meth$\geq$0.2, p$\leq$0.01) across state transitions. For instance, distributions for regions gaining methylation in the transition from ES cell to NE (top left) at all stages of differentiation. Distinct methylation level trace plots are shown for regions gaining methylation (left) during the specific transitions (indicated on the side) and loss of methylation (right). Black labeled samples are based on WGBS data and grey color samples (LRG and LNP) were profiled by RRBS. Figure 15D – Bar plot of the frequency and associated mark of epigenetic changes for all cell state transitions broken up into gain and loss for consecutive differentiation stages.

FIGs. 16A-B show that distinct transcription factor modules are associated with stage specific epigenetic transitions. Figure 16A - Illustration of epigenomic footprinting across the *PAX6* locus (chr11:31,780,014-31,842,503) for dips in H3K27ac regions (right). Black boxes highlight footprints (FP) determined for H3K27ac peaks that harbor various putative transcription factor (TF) binding sites based on motif matching. Figure 16B - The 40 top ranked TFs predicted to be activated during the cell state transition indicated on the bottom. Color-coding represents normalized TF epigenetic remodeling scores, averaging over all TERAs based on 7H3K4me3, H3K4me1, H3K27ac and DNAme. In addition, predictions were filtered for factors expressed at least at the stage of predicted induction.

FIGs. 17A-D show that a pooled shRNA (short hairpin RNA) screen recovers predicted regulators of *in vitro* NPC differentiation. Figure 17A – Simplified schematic of the pooled shRNA screen (Figures 22A-G). Figure 17B - Depletion scores for all genes that are significantly reduced (q-value$\leq$ 0.05 for at least 2 different shRNAs per gene) in at least one stage for FACS purified HES5+ cells 6 days after knockdown compared to FACS sorted HES5- obtained from the same infection or compared to cells

collected 24h after infection (Figure 22A). Depletion score indicates the extent to which shRNAs targeting a particular gene were lost during the knockdown period relative to the control, indicating potential relevance of a particular gene for HES5+ maintenance, NPC state progression and proliferation or cell survival. Higher depletion scores (red) indicate stronger reduction in shRNA presence; scores were capped at 1.5 and computed based on at least three technical replicates per condition. Figure 17C - Overlap of genes detected to be significantly depleted in the HES5+ population relative to at least one of the control conditions. Figure 17D - Performance of combined regulator predictions based on TERA ranking averaged over H3K4me3, H3K4me1, H3K27ac and DNAme. Performance is measured as percentage of the top 20 predicted activating or repressing motifs for each stage mapping to TFs included in the shRNA library.

FIGs. 18A-E show that a set of core TFs dynamically associates with stage-specific factors to modulate NPC identity and differentiation potential. Figure 18A - Predicted top 10 significant ($p \leq 0.01$, odds ratio ≥ 1.5) co-binding relationships in dynamically regulated H3K27ac footprints for a set of 10 TFs (bold) essential for HES5+ cells at each stage. Stage-specific predicted co-binding relationships are indicated in blue (NE), red (ERG) and grey (MRG). All predicted relations are supported by a knockdown effect of each gene at the relevant stage. Figure 18B - Gene expression patterns shown as z-scores for the core network TFs as well as all predicted co-binding partners across ES cells, all NPCs and more mature cellular states. Figure 18C - Venn diagram showing the overlap of *OTX2* binding sites determined by ChIP-Seq in early NE and MRG cells. Figure 18D - Gene set enrichment analysis results for *OTX2* binding sites in early NE and MRG cells. Figure 18E - Median expression patterns for ES cells, all NPCs and more mature cell populations shown as z-scores for putative downstream target genes of *OTX2* binding sites.

FIGs. 19A-D show that binding of core and stage-specific NPC TFs is associated with epigenetic priming of pro-neural genes. Figure 19A - Characterization of TFs associated with motifs gaining H3K4me1 or losing DNAme at the NE stage prior to their expression at a later or more differentiated cell state as determined by high TERA scores (bold), termed priming. In addition, significant ($p \leq 0.01$, odds ratio ≥ 1.5) co-binding relationships with factors expressed at the NE are indicated by colored lines. For each TF (from outer to inner circles, see example below for *NEUROD4*) heatmaps

indicating the relative expression level as z-score in all cell types as well as normalized TERA scores for H3K27ac, H3K4me3, H3K4me1 and DNAme. Figure 19B - Heatmaps depicting the H3K27ac (upper) and the H3K4me1 (second from top) enrichment level for predicted NEUROD binding sites at each NPC stage for 5 distinct dynamic patterns. Here, none of the NEUROD family proteins is expressed (<2.5 FPKM (fragments per kilobase of exon per million fragments mapped)). Figure 19C - Heatmap showing the z-scores of the median gene expression levels for predicted NEUROD downstream target genes for each of the 5 dynamic patterns in the more mature neuron and astrocyte-like populations. Figure 19D – Schematic illustration of the TERA and expression analyses.

FIGs. 20A-D depict isolation and characterization of ES cell derived neural progenitor cells. Figure 20A – A schematic illustration of the differentiation model including the specific days of sample collection. Human ES cells were differentiated into neuroepithelial (NE) cells using dual inhibition of TGFβ (transforming growth factor beta) and BMP (bone morphogenic protein) followed by the transition to neural base media. Subsequently, sonic hedgehog and FGF8 (fibroblast growth factor 8), are used to transition to the early radial glial stage (ERG). For the rest of the differentiation experiment the cells were constantly maintained in FGF2 (fibroblast growth factor 2) and EGF2 (epidermal growth factor 2) neural base media to reach the mid radial glia (MRG) stage after 35 days, the late radial glia (LRG) stage after 80 days and the long term neural progenitor (LNP) stage after about 200 days of *in vitro* culture. Cell type names indicated in red were profiled for gene expression, histone modifications as well as DNAme by WGBS, while names shown in grey for gene expression only and names in black for DNAme by RRBS only. Figure 20B - Hierarchical clustering for all RNA-Seq datasets collapsing replicates using the Jensen-Shannon divergence as metric. Figure 20C - Gene expression patterns (RNA sequencing (RNAseq)) shown as z-scores for all differentially expressed genes (q-value$\leq$ 0.1) across ES cells and four HES5+ NE, HES5+ ERG, HES5+ MRG and HES5+ LRG expressed with at $\geq$ 2 FPKM in at least one stage (n=20,306). Genes were grouped into 18 clusters (described in Example 11 herein below) based on minimal average silhouette width using PAM clustering and Jensen-Shannon divergence based metric. The cluster number is shown on the left from 1 to 18. Pie charts below indicate fraction of up (red) and down-regulated (green) genes during each transition. Figure 20D - Gene expression patterns shown as z-scores for all

significantly differentially expressed genes (q-value≤ 0.1) across four more mature cell populations (corresponding to the sequentially generated cortical layers) obtained through differentiation of NE, ERG or MRG cells to neuronal like cells (NE/ERG/MRGdN) and astrocyte like cells (LRGdA) derived from the LRG stage. "NEdn" = terminally differentiated neurons from HES5+ NE cells; "ERGdn" = terminally differentiated neurons from HES5+ ERG cells; "MRGdn" = terminally differentiated neurons from HES5+ MRG cells; "LRGdn" = terminally differentiated neurons and glial cells from HES5+ LRG cells. Genes were grouped into 12 clusters (described in Example 12 herein below) based on minimal average silhouette width using PAM clustering and Jensen- Shannon divergence based metric. The cluster number is shown on the left from 1 to 12. Raw data for all RNAseq presented in Figures 20C and 20D is provided in Supplementary data 7, which is fully incorporated herein by reference in its entirety.

FIGs. 21A-E depict epigenetic dynamics and TF footprints. Figure 21A - Median TPR (True Positive Rate) (red), FPR (False Positive Rate) (blue) and PPV (Positive Predictive Values) (black) for n=46 TFs with matching motif for H3K27ac footprints (n=27,292) in K562 cells as a function of confidence in predicted binding (-log10 p-value). True positives were defined as predicted binding events overlapping with peaks determined by ChIP-Seq and false positives accordingly. The entire set of positives was defined as all TF ChIP-Seq peaks for a particular factor that overlapped with any H3K27ac footprint. Figure 21B - ROC curve of the median TPR/FPR values from Figure 21A. Figure 21C - Epigenetic dynamics across the *APOE* locus (chr19:45,391kb – 45,414kb) for ES cells and three stages of the NPCs. H3K4me3 read counts 10 normalized to 1 million reads are shown on a scale of 0 to 2 (green). DNAme levels for single CpGs are indicated as blue dots on a scale of 0 to 100% of methylation (y-axis). H3K27ac read counts normalized to 1 million reads are shown on a scale of 0 to 1 (purple). For reference footprints (FP) and CpG islands (CGIs) are indicated as blue boxes (bottom). Shaded gray box indicates the position of the putative enhancer element overlapping with the Alzheimer related SNP rs157580. Figure 21D - Top: Decomposition of H3K27ac dynamics into 7 distinct modules based on PLS regression. Colors indicate median epigenetic enrichment level of gene regulatory elements assigned to each module for each cellular state for H3K27ac. Bottom: Gene set enrichment

analysis results for gene regulatory elements associated with each module. Figure 21E – Connectivity matrix showing the association strength of each of the factors listed in Figure 16B with each of the 7 modules identified by the partial least square (PLS) regression.

FIGs. 22A-G depict functional validation using a pooled shRNA screen. Figure 22A - Detailed outline of the pooled shRNA screen. Each stage (NE, ERG and MRG) was infected with an optimized virus titer aiming for an average of one shRNA integration per cell. Immediately after infection, cells were subjected to puromycin (puro) selection and bulk population material was collected 24 hours after infection and prior to efficient shRNA knockdown. Five days after infection and selection, cells were FACS sorted for HES5-GFP and both GFP+ and GFP- were collected for analysis. Subsequently, genomic DNA was extracted and all integrated shRNAs were amplified by PCR for each population separately. The resulting material was then used to construct libraries for next generation sequencing to count the number of shRNA integrations for each shRNA in each cell population. Figure 22B - Overlap of genes identified to facilitate HES5+ cell maintenance, progression or proliferation determined by genes with at least two shRNAs significantly ($q \leq 0.05$) overrepresented in the HES5+ population with respect to the 24 hours or HES5- control. Figure 22C - Regulator predictions based on differential gene expression. Performance is measured as percentage of the top 20 differentially expressed factors for each stage linked to the TF included in the shRNA library. Figure 22D - Regulator predictions based on TERA ranking for H3K4me3, H3K4me1, H3K27ac or DNAme. Performance is measured as percentage of the top 20 predicted activating or repressive motifs for each stage mapping to a TF included in the shRNA library. Figure 22E - Detailed heatmap showing the top 20 predicted motifs and corresponding TFs differentially active between the ES cell and NE stage based on the combined TERA scores for H3K27ac, H3K4me3, H3K4me1 and DNAme. In addition, knockdown results as depletion scores (green-red heatmap) obtained at each stage are shown on the right. Figure 22F - Heatmap showing the pairwise pearson-correlation coefficient (PCC) of the log2 read-count normalized shRNA libraries across all conditions and replicates. Figure 22G - Individual validation for shRNAs against *OTX2* and *PAX6* at the NE stage, which showed no effect in the pooled screening approach at any stage. Shown are qPCR levels for *OTX2* or *PAX6*,

HES5 and Puromycin relative to *HPRT*. Each gene was measured in an independent knockdown experiment for a pool of the 5 shRNAs against *PAX6* (blue), *OTX2* (green), *lacZ* (orange) as well as the uninfected control (red).

FIGs. 23A-D depict co-binding analysis. Figure 23A - Gene expression levels reported as z-scores for core network TFs and epigenetic modifiers with and without a known DNA binding motif. Figure 23B - Illustration of predicted significant co-binding relationships ($p \leq 0.01$, odds ratio ≥ 1.5) of core factors (rows) with more stage-specific or pro- neuronal/glial factors (columns). Color-coding indicates whether binding is stage-specific or occurs at multiple stages. Figure 23C - Overlap of predicted binding sites in dynamic putative enhancer regions based on H3K27ac for *OTX2* in NE and ERG. Figure 23D - Gene set enrichment analysis results for predicted *OTX2* binding sites in dynamic putative enhancer regions at the NE and MRG stage.

FIGs. 24A-E depict epigenetic priming. Figure 24A - TERA scores for H3K27ac, H3K4me3, H3K4me1 and DNAme for TFs showing evidence of priming (top bold) and TFs predicted to significantly co-occur in these primed binding sites. Figure 24B - Gene expression levels shown as z-scores for primed and co-binding TFs from Figure 24A. Figure 24C - Detailed predicted co-binding relationship ($p \leq 0.01$, odds ratio ≥ 1.5) of primed TFs (columns) with significantly associated co-binding factors (rows). Figure 24D - Illustration of a potential priming event and the associated predicted target gene at the *ATOH1* locus (chr4:94,740-94,800). For each stage, H3K27ac, H3K27me3 and DNAme patterns are shown along with predicted NEUROD binding sites (black boxes) in putative gene regulatory elements marked by a loss of DNAme (highlighted by the grey bars). Figure 24E - Gene set enrichment analysis results for predicted NEUROD binding sites split up by dynamic patterns defined in Figure 19B (top). Binding sites in patterns 3 and 4 showed no significant enrichment.

FIG. 25 is a schematic illustration depicting the transcriptional codes for the generation of human stem cells of the neocortex.

FIGs. 26A-D depict formulas used in the "GENERAL MATERIALS AND EXPERIMENTAL METHODS" section. Figure 26A – Formula 1; Figure 26B – Formula 2; Figure 26C – Formula 3; Figure 26D – Formula 4.

DESCRIPTION OF SPECIFIC EMBODIMENTS OF THE INVENTION

The present invention, in some embodiments thereof, relates to populations of neural progenitor cells and methods of producing and using same.

Before explaining at least one embodiment of the invention in detail, it is to be understood that the invention is not necessarily limited in its application to the details set forth in the following description or exemplified by the Examples. The invention is capable of other embodiments or of being practiced or carried out in various ways.

Decoding heterogeneity of pluripotent stem cell (PSC)-derived neural progeny is fundamental for revealing the origin of diverse progenitors, for defining their lineages, and for identifying fate determinants driving transition through distinct potencies.

The present inventors establish a long-term neural differentiation system from PSCs using the *HES5::eGFP* reporter human embryonic stem cell (hESC) line. HES5 is a major and direct downstream target of Notch activation pathway (Kageyama, R. & Ohtsuka, T. The Notch-Hes pathway in mammalian neural development. Cell Res. 9, 179-188, 1999). This allows the prospective isolation and characterization of primary progenitors retaining low proneural transcriptional activity and broad developmental potential and thus serving as the primary progenitors - or NSCs - that generate neural cellular diversity. The stepwise isolation of Notch active NSCs during neural differentiation of PSCs enables a systematic investigation of human NSC ontogeny and proposes a controlled module-based platform for understanding the development of normal and pathogenic NSCs and their progeny.

The present inventors isolated consecutively appearing PSC-derived primary progenitors based on their Notch activation state. As shown in Examples 1 and 2 of the Examples section which follows, the present inventors isolated early neuroepithelial cells and show their broad Notch-dependent developmental and proliferative potential. Neuroepithelial cells further yield successive Notch-dependent functional primary progenitors, from early and mid neurogenic radial glia and their derived basal progenitors, to gliogenic radial glia and adult- like neural progenitors, together recapitulating hallmarks of neural stem cell (NSC) ontogeny. Gene expression profiling reveals dynamic stage specific transcriptional patterns that may link development of distinct progenitor identities through Notch activation. The present observations provide a platform for characterization and manipulation of distinct progenitor cell types

amenable for developing streamlined neural lineage specification paradigms for modeling development in health and disease.

Human pluripotent stem cell derived models that accurately recapitulate neural development *in vitro* and allow for the generation of specific neuronal subtypes are of major interest to the stem cell and biomedical community. Notch signaling, particularly through the Notch effector HES5, is a major pathway critical for the onset and maintenance of neural progenitor cells (NPCs) in the embryonic and adult nervous system (Imayoshi, I. et al., 2010; Shimojo, H., et al., 2011; Carlen, M. *et al. 2009*). This can be exploited to isolate distinct populations of human embryonic stem (ES) cell derived NPCs (Edri, R. *et al. 2015*). Here, the present inventors report the transcriptional and epigenomic analysis of six consecutive stages derived from a HES5-GFP reporter ES cell line (Placantonakis, D. G. *et al. 2009*) differentiated along the neural trajectory aimed at modeling key cell fate decisions including specification, expansion and patterning during the ontogeny of cortical neural stem and progenitor cells. In order to dissect the regulatory mechanisms that orchestrate the stage-specific differentiation process, the present inventors developed a computational framework to infer key regulators of each cell state transition based on the progressive remodeling of the epigenetic landscape and then validated these through a pooled shRNA screen. The present inventors were also able to refine the previous observations on epigenetic priming at transcription factor binding sites and show here that they are mediated by combinations of core and stage-specific factors. Taken together, the present inventors demonstrate the utility of the system and outline a general framework, not limited to the context of the neural lineage, to dissect regulatory circuits of differentiation.

According to an aspect of some embodiments of the invention, there is provided an isolated population of cells comprising at least 5% HES5+ cells, e.g., at least 6% HES5+ cells, e.g., at least 7% HES5+ cells, e.g., at least 8% HES5+ cells, e.g., at least 9% HES5+ cells, e.g., at least 10% HES5+ cells, wherein the HES5+ cells are:

(i) early radial glial (E-RG) cells;

(ii) mid radial glial (M-RG) cells;

(iii) late radial glial (L-RG) cells; or

(iv) long term neural progenitor (LNP) cells.

According to some embodiments of the invention, the isolated population of cells further comprising HES5+ neuroepithelial (NE) cells.

As used herein the term "HES5" refers to the hes family basic helix-loop-helix (bHLH) transcription factor 5. The HES5 gene encodes a protein, which is activated downstream of the Notch pathway and regulates cell differentiation in multiple tissues. Disruptions in the normal expression of this gene have been associated with developmental diseases and cancer. The coding sequence for human HES5 can be found in GenBank Accession No. NM_001010926.3 (SEQ ID NO:41 for the nucleic acid sequence encoding HES5), and the HES5 encoded protein can be found in GenBank Accession No. NP_001010926.1 (SEQ ID NO:42 for the amino acid sequence encoding HES5).

As used herein the term "isolated" refers to at least partially separated from the natural environment e.g., the human body.

According to some embodiments of the invention the isolated population of cells is positive for one or more markers. Positive is also abbreviated by (+) or simply "+". Positive for a marker means that at least about 10 %, 20 %, 30, 40 %, 50 %, 60 %, or even at least about 70 %, 80 %, 85 %, 90 %, 95 %, or 100 % of the cells in the population present detectable levels of the marker assayed by a method known to those of skill in the art. It should be noted that cells which are positive to one or more markers can be negative for expression of other marker(s).

According to some embodiments of the invention the isolated population of cells is negative for one or more markers. Negative is also abbreviated by (-) or simply "-". Negative for a marker means that no more than about 5 %, 10 %, 20 %, 25 %, or 30 % of the cells in the population present detectable levels of the marker.

According to some embodiments of the invention, the marker is an expressed product, e.g., RNA or a polypeptide encoded by an endogenous gene. Alternatively or additionally, the marker is exogenous (heterologous) to the cell such as in the case of a reporter molecule which expression in a cell is under the control of a promoter sequence of a gene-of-interest (also referred to as a promoter-driven reporter).

It should be noted that each of the genes encompassed by some embodiments of the invention is designated using the accepted nomenclature of the gene symbol and/or by the gene identification (ID) number as available via the National Center for

Biotechnology Information (NCBI). Sequence information of the genes used according to some embodiments of the invention, which includes representative, non-limiting sequences of the polynucleotides transcribed from - and polypeptides encoded by- the genes is provided in the below Tables and in the accompanying sequence listing.

The presentation of the marker on the cell can be detected by directly monitoring expression of the marker (i.e., the RNA and/or protein encoded by the gene) using RNA or protein detection methods, or it can be monitored by means of detecting the reporter molecule (RNA or protein) driven by the promoter of the gene.

For example, "Notch+" means presence of detectable levels within (or on) the cell of the mRNA encoded by Notch signaling pathway genes and/or of detectable levels of the protein encoded by Notch signaling pathway genes, and/or of a reporter molecule which expression is under the control of Notch signaling pathway regulatory sequence (e.g., promoters, enhancers, and other regulatory sequences being upstream and/or downstream of the Notch signaling pathways coding sequence).

Non-limiting examples of Notch signaling pathway genes include genes which are downstream of Notch activation such as the HES family of genes (e.g., HES1, HES2, HES3, HES4, HES5, HES6, HES7) and RBDJ.

For example, "HES5+" means presence of detectable levels of the mRNA encoded by HES5 and/or of detectable levels of the protein encoded by HES5, and/or of a reporter molecule which expression is under the control of HES5 promoter within the cell.

Methods of detecting the expression level of RNA include, but are not limited to Northern Blot analysis, RT-PCR analysis, quantitative RT-PCR or quantitative PCR, RNA in situ hybridization stain, In situ RT-PCR stain, DNA microarrays/DNA chips, and Oligonucleotide microarray (all of which are further described hereinunder).

Methods of detecting expression and/or activity of proteins include but are not limited to Enzyme linked immunosorbent assay (ELISA), Western blot, Radio-immunoassay (RIA), Fluorescence activated cell sorting (FACS), Immunohistochemical analysis, and In situ activity assay, In vitro activity assays (all of which are further described hereinunder).

As used herein the phrase "at least one gene (or marker)" encompasses any combination of genes or markers higher than one, e.g., at least 2, 3, 4, 5, 6, 7, 8, 9, 10,

11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, e.g., at least 35, 40, 45, 50, 55, 60, 65, 70, 75, 80 and more genes or markers. Accordingly, a collection of markers (i.e., more than one) in a single cell is also referred to as a signature.

As used herein the phrase "HES5+ neuroepithelial (NE) cells" refers to a population of cells which express HES5 which present an NE phenotype characterized by an epithelial morphology and character (e.g., tight junctions), symmetrically dividing, and characterized by the SOX1+/PAX6+/SOX2+/Nestin+ expression signature, and which has the potential of differentiating ex-vivo into HES5+ early radial glial (E-RG) cells and into HES5- central nervous system neurons.

It should be noted that the HES5- central nervous system neurons which are formed (differentiated) from the HES5+ neuroepithelial cells comprise the earliest neurons capable of forming layers 1 and 6 of the cortex.

According to some embodiments of the invention the HES5+ NE cells exhibit an HES5+/SOX1+/PAX6+/SOX2+/Nestin+ expression signature.

According to some embodiments of the invention the HES5+ NE cells further exhibit a CDC6+/CDX1+/CENPH+/TOP2A+ expression signature.

According to some embodiments of the invention the HES5+ NE cells are characterized by a higher expression level of at least one gene selected from the group consisting of: TOP2A, HIST1H4C, TRIM71, PPIG, MLLT4, TNC, CDK1, OIP5, GDF15, MCM6, TP53TG1, FAM83D, FANCI, GINS2, KDM5A, GSTM3, FAM64A, LIMS1, CENPH, KIF2C, ATAD2, DTL, CDCA5, ARHGEF6, LIPA, POLE2, RRM2, MAD2L1, CKS1B, TTK, DHFR, S100A4, NUP37, PMAIP1, CENPN, RNASEH2A, BST2, MCM10, MAF, KIAA0101, C8ORF4, E2F7, CENPA, UBE2T, RAB13, TMEM126A, MAGT1, CDC6, C6ORF211, RFC5, PSMD1, HMMR, UNG, UBE2C, GINS1, AURKB, LEPREL1, SBNO1, ZWINT, MKI67, CCAR1, FKBP5, PVRL3, CCNB1, NOP58, COL4A1, GGH, LSM6, EID1, GPX8, STC2, CD276, HS2ST1, EIF5B, HDGF, and NOL7 as compared to the expression level of the at least one gene in HES5- differentiated cells obtained by culturing undifferentiated pluripotent stem cells (PSCs) under culture conditions suitable for differentiation the PSCs into HES5+ neuroepithelial cells.

According to some embodiments of the invention, the culture conditions suitable for differentiation the PSCs into HES5+ neuroepithelial cells comprise a culture medium which comprises Noggin, SB-431542 and LDN-193189.

According to some embodiments of the invention, the HES5+ NE cells are characterized by presence of at least one active transcription factor selected from the group consisting of: RFX4, NR2F2, REST, CDC6, CDX1, CENPH, and TOP2A.

According to some embodiments of the invention, the isolated population of cells comprises at least 50%, e.g., at least 60%, e.g., at least 70%, e.g., at least 75%, e.g., at least 80%, 81%, 82%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, e.g., 90% of HES5+ cells neuroepithelial cells (NE).

According to some embodiments of the invention the HES5+ NE cells are capable of differentiating into E-RG, M-RG, L-RG and LNP cells.

According to some embodiments of the invention the HES5+ neuroepithelial cells are characterized by an increased transcriptional epigenetic activity of at least one transcription factor encoded by the gene selected from group consisting of: RFX4, PAX6, TCF4, SOX9, LEF1, POU3F2, SOX8, SOX21, TEAD1, NFATC1, SOX5, TGIF1, MEIS1, TCF4, MEIS2, OTX2, TEF, ZBTB16, MSX1, RFX1, NR4A2, MEIS2, SOX15, STAT5B, SATB1, RBPJ, FOXK1, MYBL2, DMRT3, NFIA, CUX1, TFAP4, MSX1, CDC5L, RFX1, FOXJ2, POU6F1, TEF, RBPJ, PKNOX2, BCL6, PRRX1, STAT1, POU3F1, FOXB1, CTNNB1, PBX1, ZNF143, NFATC1, TCF7L1, ARX, RXRA, TFAP4, CUX1, OTX2, NR2E1, CUX1, ZNF232, NR2F1, SOX4, MEIS1, PBX1, CUX1, NEUROD1, MSX1, ZNF652, MEF2A, OLIG1, POU6F2, IKZF2, MECOM, STAT1, ESRRA, IRF7, STAT1, MYBL2, BCL6, ELK1, ATF2, SMAD3, ATF4, DLX1, MEF2A, DBP, MAF, MEF2A, TEAD2, SMAD3, POU6F1, BARHL2, FOXG1, LHX9, MECOM, ARNTL, MYC, ZNF75A, NFIA, VAX1, GBX2, HOMEZ, FOXO4, FOXO4, FOXB1, ZSCAN16, ELK1, ATF2, CREB1, USF1, ESRRA, ZNF282, NEUROG2, NFYA, NR4A1, CTF1, ELK1, POU3F2, ELK1, HSF1, E2F3, CUX1, CREB1, ELF2, MYBL2, HMGA2, SRF, ZNF410, JDP2, NR2F1, PAX3, NRF1, SMAD4, ZNF85, ZNF628, NFATC1, CREB1, E4F1, NR2F6, NHLH1, IRF2, PBX1, FOXJ3, RORA, IRF7, NR6A1, LHX2, PAX3, NR2E1, POU3F1, ZFP42, E2F4, ETV5, ELF3, USF1, ATF6, TFAP2B, CUX1, IRF3, RXRA, PEBP1, LHX2, GZF1, MEF2A, MEF2A, IRF9, MGA, VBP1, GMEB2, YY1, ELF1, POU3F3, GTF2IRD1,

IRF3, SRF, XBP1, ESRRA, HEY1, NFKB2, IRF2, EOMES, FOXB1, NR2F1, NR4A2, STAT3, SP1, RARA, CREB1, NR2F1, FOXJ3, HSF1, MYBL2, SRF, ETV2, MECP2, E2F1, FOXJ2, JUN, SCRT2, DLX1, E2F1, E2F1 and ATF2 as compared to the transcriptional epigenetic activity of the at least one transcription factor in human ESC cells.

According to some embodiments of the invention the HES5+ neuroepithelial cells are characterized by an increased transcriptional epigenetic activity of at least one transcription factor encoded by the gene selected from group consisting of: RFX4, PAX6, TCF4, SOX9, LEF1, POU3F2, SOX8, SOX21, TEAD1, NFATC1, SOX5, TGIF1, MEIS1, TCF4, MEIS2, OTX2, TEF, ZBTB16, SOX9, MSX1, RFX1, SOX8, NR4A2, MEIS2, SOX15, STAT5B, SOX8, SATB1, RBPJ, FOXK1, MYBL2, DMRT3, NFIA, CUX1, TFAP4, MSX1, CDC5L, RFX1, FOXJ2, POU6F1, TEF, RBPJ, PKNOX2, BCL6, PRRX1, STAT1, POU3F1, FOXB1, CTNNB1, PBX1, ZNF143, NFATC1, TCF7L1, ARX, RXRA, TFAP4, CUX1, OTX2, NR2E1, CUX1, ZNF232, NR2F1, SOX4, MEIS1, PBX1, CUX1, NEUROD1, MSX1, ZNF652, MEF2A, OLIG1, POU6F2, IKZF2, MECOM, STAT1, ESRRA, IRF7, STAT1, MYBL2, BCL6, ELK1, ATF2, SMAD3, ATF4, DLX1, MEF2A, DBP, MAF, MEF2A, TEAD2, SMAD3, POU6F1, BARHL2, FOXG1, LHX9, MECOM, ARNTL, MYC, ZNF75A, NFIA, VAX1, GBX2, HOMEZ, FOXO4, FOXO4, FOXB1, ZSCAN16, ELK1, ATF2, SOX9, CREB1, USF1, ESRRA, ZNF282, NEUROG2, NFYA, NR4A1, CTF1, ELK1, POU3F2, ELK1, HSF1, E2F3 and CUX1, as compared to the transcriptional epigenetic activity of the at least one transcription factor in human ESC cells.

According to some embodiments of the invention the HES5+ neuroepithelial cells are characterized by an increased transcriptional epigenetic activity of at least one transcription factor encoded by the gene selected from group consisting of: RFX4, PAX6, TCF4, SOX9, LEF1, POU3F2, SOX8, SOX21, TEAD1, NFATC1, SOX5, TGIF1, MEIS1, TCF4, MEIS2, OTX2, TEF, ZBTB16, SOX9, MSX1, RFX1, SOX8, NR4A2, MEIS2, SOX15, STAT5B, SOX8, SATB1, RBPJ, FOXK1, MYBL2, DMRT3, NFIA, CUX1, TFAP4, MSX1, CDC5L, RFX1, FOXJ2, POU6F1, TEF, RBPJ, PKNOX2, BCL6, PRRX1, STAT1, POU3F1, FOXB1, CTNNB1, PBX1, ZNF143, NFATC1, TCF7L1, ARX, RXRA, TFAP4, CUX1, OTX2, NR2E1, CUX1, ZNF232, NR2F1, SOX4, MEIS1, PBX1, CUX1, NEUROD1, MSX1, ZNF652, MEF2A, OLIG1,

POU6F2, IKZF2, MECOM and STAT1 as compared to the transcriptional epigenetic activity of the at least one transcription factor in human ESC cells.

According to some embodiments of the invention the HES5+ neuroepithelial cells are characterized by an increased transcriptional epigenetic activity of at least one transcription factor encoded by the gene selected from group consisting of: RFX4, PAX6, TCF4, SOX9, LEF1, POU3F2, SOX8, SOX21, TEAD1, NFATC1, SOX5, TGIF1, MEIS1, TCF4, MEIS2, OTX2, TEF, ZBTB16, MSX1, RFX1, NR4A2, MEIS2, SOX15, STAT5B, SATB1, RBPJ, FOXK1, MYBL2, DMRT3, NFIA, CUX1, TFAP4, MSX1, CDC5L, RFX1, FOXJ2, POU6F1, TEF, RBPJ, PKNOX2, BCL6, PRRX1, STAT1, POU3F1, FOXB1, CTNNB1 and PBX1 as compared to the transcriptional epigenetic activity of the at least one transcription factor in human ESC cells.

According to some embodiments of the invention the HES5+ neuroepithelial cells are characterized by an increased transcriptional epigenetic activity of at least one transcription factor encoded by the gene selected from group consisting of: RFX4, PAX6, TCF4, SOX9, LEF1, POU3F2, SOX8, SOX21, TEAD1, NFATC1, SOX5, TGIF1, MEIS1, TCF4, MEIS2, OTX2, TEF, ZBTB16, and MSX1 as compared to the transcriptional epigenetic activity of the at least one transcription factor in human ESC cells.

According to some embodiments of the invention the HES5+ neuroepithelial cells are characterized by an increased transcriptional epigenetic activity of at least one transcription factor encoded by the gene selected from group consisting of: RFX4, PAX6, TCF4, SOX9, LEF1, POU3F2 and SOX8 as compared to the transcriptional epigenetic activity of the at least one transcription factor in human ESC cells.

According to some embodiments of the invention the HES5+ neuroepithelial cells are characterized by an increased transcriptional epigenetic activity of at least one transcription factor encoded by the gene selected from group consisting of: RFX4 and PAX6 as compared to the transcriptional epigenetic activity of the at least one transcription factor in human ESC cells.

As used herein the phrase "HES5+ early radial glial (E-RG) cells" refers to an isolated population of cells which express HES5, having an elongated morphology, express PAX6 and form neural rosettes (highly polarized structures containing radially

organized columnar cells) and have the potential of differentiating ex-vivo into HES5+ M-RG cells and into HES5- neural progenitor cells.

It should be noted that the HES5+ ERG cells loose some of the morphology of epithelial cells that was present in the HES5+ NE cells (e.g., HES5+ ERG cells loose the tight junctions morphology as compared to HES5+ NE cells) and gain some astro glial characters, such as expression of S100B, EGFR, GLAST and FABP7 and also to some extent expression of GFAP (shown by RNAseq data in Supplementary data 7, which is fully incorporated herein by reference in its entirety).

It should be noted that the HES5- neural progenitor cells belong to the CNS. These cells are non-stem cells but rather are limited progenitor cells, which upon the immediate differentiation form the earliest neurons which form layer 1 of the cortex, and eventually the HES5- neural progenitor cells can also differentiate to the neurons forming layers 5 and 6 of the cortex.

According to some embodiments of the invention, the HES5+ early radial glial cells exhibit an PAX6+/SOX1+/SOX2+/Nestin+ expression signature.

According to some embodiments of the invention, the HES5+ early radial glial cells further exhibit an ARX+/FEZF2+/NR2E1+ expression signature.

According to some embodiments of the invention, the HES5+ early radial glial cells are characterized by a higher expression level of at least one gene selected from the group consisting of: NR2E1, HES5, ARX, C1ORF61, FRZB, GRM3, EPHA3, NAV3, EGR2, RGMA, NRXN3, FAM107A, FABP7, EGR3, ZNF385B, TTYH1, SNCAIP, NRARP, PLP1, LIX1, LFNG, HES4, CD82, HS6ST1, PTPRZ1, CACHD1, DACH1, FEZF2, DTX4, FUT9, WNT5B, ENPP2, POU3F3, EMX2, MECOM, XYLT1, ARMCX2, FOS, PPAP2B, NOS2, LRP2, SOX9, NLGN3, TMEM2, CXCR7, EPHA7, SMOC1, TBC1D9, FAT4, SCUBE3, FUT8, CSPG5, DLL1, BOC, ID4, EGR1, ALPL, RFX4, GALNT12, CBX2, FHOD3, SORBS2, GUCY1B3, MBIP, FBXO16, SHISA2, DAB1, GLI3, FZD3, SEMA5B, LGALS3BP, SFRP1, C1QL1, RING1, GPRC5B, ZNF710, WSCD1, VPS37B, ZIC2, SDK2, DOCK11, GAS1, ZNF436, TMSB15A, IER2, FEZ1, CELF2, SFT2D3, NCALD, AKAP7, MYADM, NEDD4L, PHC2, PI4KAP2, STARD3, and CAMK1D as compared to the expression level of the at least one gene in HES5- differentiated cells obtained by culturing HES5+

neuroepithelial cells under culture conditions suitable for differentiation the HES5+ neuroepithelial cells into HES5+ early radial glial cells.

According to some embodiments of the invention, the HES5+ E-RG cells are characterized by presence of at least one active transcription factor selected from the group consisting of: ARX, NR2E1, FEZF2 and EMX2.

According to some embodiments of the invention, the HES5+ cells are early radial glial cells (E-RG) which constitute at least about 60%, e.g., at least about 65%, e.g., at least about 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, e.g., at least 90% or more of the isolated population of cells.

According to some embodiments of the invention, the HES5+ cells early radial glial cells (E-RG) are capable of differentiating into M-RG, L-RG and LNP cells.

According to some embodiments of the invention, the HES5+ early radial glial cells are characterized by an increased transcriptional epigenetic activity of at least one transcription factor encoded by the gene selected from group consisting of: NFIA, MZF1, SMAD4, CTF1, SMAD3, RFX4, MAF, TCF12, NFYC, ZNF263, MECP2, ZFP42, ZIC1, YY1, ETS2, NR2C2, SREBF2, SREBF1, MEIS2, NR4A1, REST, SF1, ZBTB7A, STAT3, E2F1, NR1H2, NFKB1, NR2F6, GLIS3, MAZ, STAT1, TGIF1, SOX9, HES1, THRA, GLIS3, MEIS1, ESRRA, ZBTB7A, NFYA, MECP2, PKNOX2, EP300, TFAP2B, NR2F1, MZF1, ESRRA, TFCP2, NR2F1, ESRRA, TERF1, KLF3, XBP1, RORA, PBX1, MYC, SNAI2, TEAD2, CENPB, PEBP1, HINFP, SREBF1, YY1, E2F1, HSF2, CNOT3, MEIS2, MEF2A, RXRA, CREB3L1, MYBL2, ZNF524, TFAP4, RFX1, NFYA, ZBTB33, RREB1, NR6A1, HES5, TFAP2B, HIF1A, INSM1, ZNF524, ELF1, SMAD4, STAT3, TFAP2B, USF1, GLI3, GLI2, ETV2, E2F1, CREB3, NFKB1, MYC, SRF, TFCP2, ATF3, ELK1, TP53, E2F4, ELF2, DEAF1, RXRA, ZNF423, FOXG1, ZNF628, NEUROD1, ARNTL, STAT1, PAX3, NR6A1, ESRRA, TCF4, YY1, PBX1, VAX1, NRF1, ETV1, FOXB1, ZNF85, CREB1, PRDM4, MYBL2, NHLH1, CREB1, KLF13, TRIM26, ZNF148, E4F1, USF1, JDP2, OLIG1, GZF1, ATF2, CREB1, FOXO4, SMAD3, CLOCK, ZNF282, ZSCAN16, CUX1, NFKB1, RFX1, ETV5, MEF2A, GTF2IRD1, TFAP2B, NHLH1, ELF1, PAX3, MNT, ZNF740, CREB1, ESRRA, EOMES, EGR2, NR4A2, BCL6, ZNF143, MAFF, CUX1, NR2F1, TFAP2B, NEUROG2, TCF3, ATF3, PLAG1, ESRRA, E2F1, KLF13, MSX1,

SCRT2, GABPB1, SOX9, NFKB2, EOMES, SIX3, MECOM, BCL6, POU3F3, HINFP, NFATC1, TBR1, MYBL2, USF1, SP1, ATF2, LHX2, MYBL2, POU3F3, HEY1, MEF2A, MGA, SOX9, E2F7, IRF3, SRF, POU3F2, SREBF2, NFKB1, MAF, ZNF784, NR1H2, MSX1, RORA, LMO2, HINFP, ZNF23, FOXO4, NR4A2, POU6F2, ID4, RARA, KLF13, LEF1, PATZ1, ELK1, DLX1, E2F4, E2F4, ELK1, MAZ, ATF6, TFAP4, RARA, PAX6, EGR2, HOMEZ, ELK1, E2F1, GMEB2, ATF4, ATF2, MAX, MGA, NFKB2, ESRRA, E2F4, STAT3, TCF3, NR2E1, ELF3, SOX4, USF1 and ZNF652 as compared to the transcriptional epigenetic activity of the at least one transcription factor in HES5+ neuroepithelial cells.

According to some embodiments of the invention, the HES5+ early radial glial cells are characterized by an increased transcriptional epigenetic activity of at least one transcription factor encoded by the gene selected from group consisting of: NFIA, MZF1, SMAD4, CTF1, SMAD3, RFX4, MAF, TCF12, NFYC, ZNF263, MECP2, ZFP42, ZIC1, YY1, ETS2, NR2C2, SREBF2, SREBF1, MEIS2, NR4A1, REST, SF1, ZBTB7A, STAT3, E2F1, NR1H2, NFKB1, NR2F6, GLIS3, MAZ, STAT1, TGIF1, SOX9, HES1, THRA, GLIS3, MEIS1, ESRRA, ZBTB7A, NFYA, MECP2, PKNOX2, EP300, TFAP2B, NFIA, NR2F1, MZF1, ESRRA, TFCP2, NR2F1, TERF1, KLF3, XBP1, RORA, PBX1, MYC, SNAI2, TEAD2, CENPB, PEBP1, HINFP, SREBF1, YY1, E2F1, HSF2, CNOT3, MEIS2, MEF2A, RXRA, CREB3L1, MYBL2, ZNF524, TFAP4, RFX1, NFYA, ZBTB33, RREB1, NR6A1, HES5, TFAP2B, HIF1A, INSM1, ZNF524, ELF1, SMAD4, STAT3, TFAP2B, USF1, GLI3, GLI2, ETV2, E2F1, CREB3, MYC, SRF, TFCP2, ATF3, ELK1, TP53, E2F4, ELF2, DEAF1, RXRA, ZNF423, FOXG1, ZNF628, NEUROD1, ARNTL, STAT1, PAX3, NR6A1 and ESRRA as compared to the transcriptional epigenetic activity of the at least one transcription factor in HES5+ neuroepithelial cells.

According to some embodiments of the invention, the HES5+ early radial glial cells are characterized by an increased transcriptional epigenetic activity of at least one transcription factor encoded by the gene selected from group consisting of: NFIA, MZF1, SMAD4, CTF1, SMAD3, RFX4, MAF, TCF12, NFYC, ZNF263, MECP2, ZFP42, ZIC1, YY1, ETS2, NR2C2, SREBF2, SREBF1, MEIS2, NR4A1, REST, SF1, ZBTB7A, STAT3, E2F1, NR1H2, NR2F6, GLIS3, MAZ, STAT1, TGIF1, SOX9, HES1, THRA, GLIS3, MEIS1, ESRRA, ZBTB7A, NFYA, MECP2, PKNOX2, EP300,

TFAP2B, NFIA, NR2F1, MZF1, ESRRA, TFCP2, NR2F1, TERF1, KLF3, XBP1, RORA, PBX1, MYC, SNAI2, TEAD2, CENPB, PEBP1, HINFP, SREBF1, YY1, E2F1, HSF2, CNOT3, MEIS2, MEF2A, RXRA, CREB3L1, MYBL2, ZNF524, TFAP4, RFX1, NFYA, ZBTB33, RREB1, NR6A1, HES5, TFAP2B, HIF1A, INSM1, ZNF524, ELF1, SMAD4, STAT3, TFAP2B, USF1, GLI3, GLI2, ETV2, E2F1, CREB3 and NFKB1 as compared to the transcriptional epigenetic activity of the at least one transcription factor in HES5+ neuroepithelial cells.

According to some embodiments of the invention, the HES5+ early radial glial cells are characterized by an increased transcriptional epigenetic activity of at least one transcription factor encoded by the gene selected from group consisting of: NFIA, MZF1, SMAD4, CTF1, SMAD3, RFX4, MAF, TCF12, NFYC, ZNF263, MECP2, ZFP42, ZIC1, YY1, ETS2, NR2C2, SREBF2, SREBF1, MEIS2, NR4A1, REST, SF1, ZBTB7A, STAT3, E2F1, NR1H2, NFKB1, NR2F6, GLIS3, MAZ, STAT1, TGIF1, SOX9, HES1, THRA, GLIS3, MEIS1, ESRRA, ZBTB7A, NFYA, MECP2, PKNOX2, EP300, TFAP2B and NFIA as compared to the transcriptional epigenetic activity of the at least one transcription factor in HES5+ neuroepithelial cells.

According to some embodiments of the invention, the HES5+ early radial glial cells are characterized by an increased transcriptional epigenetic activity of at least one transcription factor encoded by the gene selected from group consisting of: NFIA, MZF1, SMAD4, CTF1, SMAD3, RFX4, MAF, TCF12, NFYC, ZNF263, MECP2, ZFP42, ZIC1, YY1, ETS2, NR2C2, SREBF2, SREBF1, MEIS2, NR4A1, REST, SF1, ZBTB7A, STAT3, E2F1, NR1H2, NFKB1, NR2F6, GLIS3, MAZ and STAT1 as compared to the transcriptional epigenetic activity of the at least one transcription factor in HES5+ neuroepithelial cells.

According to some embodiments of the invention, the HES5+ early radial glial cells are characterized by an increased transcriptional epigenetic activity of at least one transcription factor encoded by the gene selected from group consisting of: NFIA, MZF1, SMAD4, CTF1, SMAD3, RFX4, MAF, TCF12, NFYC and ZNF263 as compared to the transcriptional epigenetic activity of the at least one transcription factor in HES5+ neuroepithelial cells.

As used herein the phrase "HES5+ mid radial glial (M-RG) cells" refers to an isolated population of cells expressing HES5, which form neural rosettes and being

capable of differentiating into the HES5+ late radial glial (L-RG) cells and into HES5- intermediate progenitor cells (INPs).

It should be noted that the HES5+ M-RG cells are comprised in the ventricular zone (VZ) of the brain, and include HES5+ basal radial progenitors.

It should be noted that the HES5- intermediate progenitor cells (INPs) belong to the CNS and are capable of differentiating into the neurons forming mainly layers 4 and 3 of the brain cortex, and also a small fraction forming layer 2 of the brain cortex. The INPs constitute about 80% of the SVZ (80%), exhibit TBR2+ expression pattern, wherein each INP can divide on average 3 times to create neurons of layers 4 and 3 of the brain cortex, and a small fraction forming layer 2.

According to some embodiments of the invention, the HES5+ mid radial glial cells exhibit an HES5+/PAX6+/Nestin+ expression signature.

According to some embodiments of the invention, the HES5+ mid radial glial cells exhibit an HES5+/POU3F2+ expression signature.

According to some embodiments of the invention, the HES5+ mid radial glial cells further exhibit an GLAST+/FABP7+ expression signature.

According to some embodiments of the invention, the HES5+ mid radial glial cells are characterized by a higher expression level of at least one gene selected from the group consisting of: FZD10, ZEB2, EN2, ST20, CDKN2C, RAB10, WASF1, ZBED4, EZH2, PPA2, H1F0, CCNJ, ITGB8, SH3BGRL3, IRX2, KIF23, PEG10, SMC3, NUSAP1, APLP1, ADAMTS3, RACGAP1, LIMCH1, ETNK1, RNF13, ARID1B, TRIM28, CNOT8, CRNDE, TWSG1, NT5DC2, NAA50, NUF2, ABCE1, PLTP, FBRSL1, DCAF16, OGT, ZFYVE16, FOXM1, PM20D2, POU3F2, MCM4, HERPUD2, VRK1, TRIM41, SATB1, HOMER1, CCNG1, ATF2, AP1AR, GABPA, STXBP3, SMC5, CDKN1B, NUPL1, UBA1, CYTH2, FXYD6, ISYNA1, DOCK1, 41527, LPHN2, IDI1, PXMP2, U2AF2, ARHGAP12, KLHL24, CKAP2, ZNF238, PARP6, NHSL1, PBRM1, BAZ1A, MAP4K5, TSPAN12, SH3GLB1, ASPM, ANKLE2, SPG20, MAP4K4, CASC4, FUBP3, ARSB, BTAF1, SIKE1, VEZT, PBX3, CBL, EIF2AK4, API5, MTSS1, NET1, CHD6, ZNF117, PNMA1, PTPN13, MTIF2, SSFA2, KIAA1279, STRN3, WIPF1, MEIS2, ZC3H4, DYNC1I2, RTN4, TAF2, RASA1, OSBPL8, SKA2, IGF1R, RNF6, SGTB, TMEM131, HIATL1, TGIF1, TMEM170B, PSAT1, ACBD5, HECTD2, ASF1A, LAMB1, GLS, DDX39, DGCR2,

EIF1AX, SALL1, GOLPH3, PTBP2, GRIP1, PNPLA8, VASH2, SUDS3, PFDN4, BAZ2A, PRKDC, GLYR1, DAZAP2, PCMTD1, SENP6, CLINT1, RECQL, CNTNAP2, CTBP1, C10ORF18, CDON, B4GALT6, CSNK1G3, STAT5B, TMEM60, HNRNPH2, TACC2, CCNG2, FSCN1, CCNA2, C21ORF45, PLRG1, ZFHX3, UBE2A, DMTF1, TRA2A, MYO5A, FAM96A, IFT80, VPS26A, MRPL50, ACYP1, WDR11, PLDN, RPRD1A, MEAF6, CKAP5, YTHDC2, GABARAPL2, IPO5, PGAP1, C14ORF147, CD200, MST4, PPT1, ANKRD50, HPS3, CCNC, THRAP3, TWF1, CYP51A1, PSPC1, WDR75, CAST, SEPW1, C21ORF59, PIK3C2A, GNG5, MED4, GIPC1, STK39, KIAA1715, PHF6, PPTC7, SOCS4, PPM1B, UQCRB, C10ORF84, SLAIN1, RAB6A, SOS2, KLF10, RNF4, C3ORF63, INSIG1, CPSF1, DNAJC4, ATP2B4, PPP2R1A, TRIM22, SDC3, TSNAX, PPIL4, ZDHHC2, ZBTB44, ANO6, PPP2CB, UBA2, BBS2, ZNF423, RNF5, C1ORF31, IFT81, CPSF6, KLHL9, FAM164A, TTC35, CCDC90B, TM9SF4, SEC24B, SMARCC2, CAP2, SAR1A, THY1, RBPMS2, EIF3A, DZIP1, ARL6IP1, SACM1L, PAPD4, SCG2, TCF3, EFHA1, HNRNPA2B1, EWSR1, STAG2, YEATS4, PAQR3, GAR1, FTH1, C19ORF43, TMEM14C, CCDC104, PSMD12, DCTD, SSR1, HMGCS1, HMGB3, KIF3A, TMEM128, PATZ1, RBL2, ARFGAP3, DNAJB5, TMED7, G3BP2, BMPR1A, FMR1, TPST2, TMSB4X, RP2, CEP170, KLHL23, RNF7, HNRNPH1, MARCKS, HNRNPD, TOB1, UTP11L, RFK, DHX36, LCOR, WBP5, PHLDB2, USP33, EFNB2, C6ORF62, MEX3B, ABCD3, ATG3, ARID4B, C7ORF11, EPB41, TCF12, CDK8, CMIP, ATG12, CETN3, ZNF217, TMEM55A and UBE2N as compared to the expression level of the at least one gene in HES5- differentiated cells obtained by culturing HES5+ early radial glial cells under culture conditions suitable for differentiation the HES5+ early radial glial cells into HES5+ mid radial glial cells.

According to some embodiments of the invention, the HES5+ M-RG cells are characterized by presence of at least one active transcription factor selected from the group consisting of: NFIA, NFIB, REST, CDKN1B, SALL1, and POU3F2.

According to some embodiments of the invention, the HES5+ cells are mid radial glial cells (M-RG) which constitute at least about 20%, e.g., at least about 25%, e.g., at least about 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, e.g., 40% of the isolated population of cells.

According to some embodiments of the invention, the HES5+ mid radial glial (M-RG) cells are capable of differentiating into L-RG and LNP cells.

According to some embodiments of the invention, the HES5+ mid radial glial cells are characterized by an increased transcriptional epigenetic activity of at least one transcription factor encoded by the gene selected from group consisting of: RFX4, NFIA, NFATC1, CTF1, NEUROD1, RFX1, TGIF1, ETS2, TFAP4, TCF12, TFAP4, SMAD3, NFIA, SOX9, RFX1, TCF4, MZF1, STAT1, MECP2, MEIS2, ZBTB33, NFYA, ELF1, MYBL2, LEF1, NFYC, MAFF, ZNF263, YY1, POU3F3, TGIF1, STAT3, SMAD4, NR6A1, TGIF1, MEIS1, ZNF628, ZFP42, FOXK1, PRDM4, STAT1, MAF, SCRT2, CREB1, GZF1, CREB1, VAX1, MECP2, NHLH1, ETV1, SOX9, PEBP1, SMAD4, XBP1, USF1, POU3F2, CREB3, EP300, PBX1, STAT3, TFCP2, POU2F2, IRF3, FOXB1, MSX1, POU3F3, ELK1, DBP, CUX1, MEF2A, POU6F2, ARNTL, ZSCAN16, MEF2A, ETV4, OLIG1, HOMEZ, DLX1, PRRX1, MSX1, MYC, FOXO4, MEF2A, MZF1, ATF2, GMEB2, NFYA, ESRRA, SOX9, PBX1, POU3F2, MECOM, SMAD3, MAZ, ELK1, BCL6, ELF3, PAX3, ELF1, SF1, BCL6, EMX2, STAT1, E2F1, ELF2, CLOCK, ELK1, STAT3, ATF5, THRA, SOX9, SRF, TCF4, ATF3, CTNNB1, USF1, FOXJ3, USF1, ZNF282, NEUROG2, ESRRA, REST, E2F3, ZBTB7A, MYBL2, HSF2, MAX, ZNF143, MYBL2, SRF, FOXO4, NR4A2, CUX1, E2F4, MSX1, EOMES, MAF, MNT, POU6F1, NFATC1, STAT3, CREB3, SOX9, ZNF85, CREB3L1, TCF3, ELK1, IRF2, YY1, SOX15, PAX6, E4F1, MEF2A, ATF2, NFE2L1, NFATC1, ATF2, YY1, SRF, ARX, ETV2, HINFP, MAZ, NR4A1, INSM1, ZNF652, USF1, NFKB1, GBX2, POU3F4, IRF9, POU6F1, LHX2, NR4A2, POU6F2, FOXJ2, CEBPG, VBP1, TERF1, ESRRA, PAX6, ZIC3, IRF2, MAF, SOX21, CREB1, CREB3, IRF7, POU3F3, NFKB2, ATF4, MYBL2, SREBF2, SOX9, YY1, RBPJ, FOXJ3, HSF1, HMGA2, CUX1, POU3F2, EOMES, ZNF423, ESRRA and KLF13 as compared to the transcriptional epigenetic activity of the at least one transcription factor in HES5+ early radial glial cells.

According to some embodiments of the invention, the HES5+ mid radial glial cells are characterized by an increased transcriptional epigenetic activity of at least one transcription factor encoded by the gene selected from group consisting of: RFX4, NFIA, NFATC1, CTF1, NEUROD1, RFX1, TGIF1, ETS2, TFAP4, TCF12, TFAP4, SMAD3, NFIA, SOX9, RFX1, TCF4, MZF1, STAT1, MECP2, MEIS2, ZBTB33,

NFYA, ELF1, MYBL2, LEF1, NFYC, MAFF, ZNF263, YY1, POU3F3, TGIF1, STAT3, SMAD4, NR6A1, TGIF1, MEIS1, ZNF628, ZFP42, FOXK1, PRDM4, STAT1, MAF, SCRT2, CREB1, GZF1, CREB1, VAX1, MECP2, NHLH1, ETV1, SOX9, PEBP1, SMAD4, XBP1, USF1, POU3F2, CREB3, EP300, PBX1, STAT3, TFCP2, POU2F2, IRF3, FOXB1, MSX1, POU3F3, ELK1, DBP, CUX1, MEF2A, POU6F2, ARNTL, ZSCAN16, MEF2A, ETV4, OLIG1, HOMEZ, DLX1, PRRX1, MSX1, MYC, FOXO4, MEF2A, MZF1, ATF2, GMEB2, NFYA, ESRRA, SOX9, PBX1, POU3F2, MECOM, SMAD3, MAZ, ELK1, BCL6, ELF3, PAX3, ELF1, SF1, BCL6, EMX2, STAT1, E2F1, ELF2, CLOCK, ELK1, STAT3, ATF5, THRA, SOX9, SRF, TCF4, ATF3, CTNNB1, USF1, FOXJ3, USF1, ZNF282, NEUROG2, ESRRA, REST, E2F3, ZBTB7A, MYBL2, HSF2, MAX, ZNF143, MYBL2, SRF, FOXO4, NR4A2, CUX1, E2F4, MSX1 and EOMES as compared to the transcriptional epigenetic activity of the at least one transcription factor in HES5+ early radial glial cells.

According to some embodiments of the invention, the HES5+ mid radial glial cells are characterized by an increased transcriptional epigenetic activity of at least one transcription factor encoded by the gene selected from group consisting of: RFX4, NFIA, NFATC1, CTF1, NEUROD1, RFX1, TGIF1, ETS2, TFAP4, TCF12, TFAP4, SMAD3, NFIA, SOX9, RFX1, TCF4, MZF1, STAT1, MECP2, MEIS2, ZBTB33, NFYA, ELF1, MYBL2, LEF1, NFYC, MAFF, ZNF263, YY1, POU3F3, TGIF1, STAT3, SMAD4, NR6A1, TGIF1, MEIS1, ZNF628, ZFP42, FOXK1, PRDM4, STAT1, MAF, SCRT2, CREB1, GZF1, CREB1, VAX1, MECP2, NHLH1, ETV1, SOX9, PEBP1, SMAD4, XBP1, USF1, POU3F2, CREB3, EP300, PBX1, STAT3, TFCP2, POU2F2, IRF3, FOXB1, MSX1, POU3F3, ELK1, DBP and CUX1 as compared to the transcriptional epigenetic activity of the at least one transcription factor in HES5+ early radial glial cells.

According to some embodiments of the invention, the HES5+ mid radial glial cells are characterized by an increased transcriptional epigenetic activity of at least one transcription factor encoded by the gene selected from group consisting of: RFX4, NFIA, NFATC1, CTF1, NEUROD1, RFX1, TGIF1, ETS2, TFAP4, TCF12, TFAP4, SMAD3, NFIA, SOX9, RFX1, TCF4, MZF1, STAT1, MECP2, MEIS2, ZBTB33, NFYA, ELF1, MYBL2, LEF1, NFYC, MAFF, ZNF263, YY1, POU3F3, TGIF1,

STAT3, SMAD4, NR6A1, TGIF1, MEIS1 and ZNF628 as compared to the transcriptional epigenetic activity of the at least one transcription factor in HES5+ early radial glial cells.

According to some embodiments of the invention, the HES5+ mid radial glial cells are characterized by an increased transcriptional epigenetic activity of at least one transcription factor encoded by the gene selected from group consisting of: RFX4, NFIA, NFATC1, CTF1, NEUROD1, RFX1, TGIF1, ETS2, TFAP4, TCF12, TFAP4, SMAD3, NFIA, SOX9, RFX1, TCF4, MZF1, STAT1, MECP2, MEIS2, ZBTB33, NFYA, ELF1, MYBL2, LEF1 and NFYC as compared to the transcriptional epigenetic activity of the at least one transcription factor in HES5+ early radial glial cells.

According to some embodiments of the invention, the HES5+ mid radial glial cells are characterized by an increased transcriptional epigenetic activity of at least one transcription factor encoded by the gene selected from group consisting of:RFX4, NFIA, NFATC1, CTF1, NEUROD1, RFX1, TGIF1, ETS2, TFAP4, TCF12, TFAP4, SMAD3, NFIA and SOX9 as compared to the transcriptional epigenetic activity of the at least one transcription factor in HES5+ early radial glial cells.

As used herein the phrase "HES5+ late radial glial (L-RG) cells" refers to HES5+ cells characterized by downregulation of at least one rosette marker and upregulation of at least one glial marker as compared to HES5+ MRG cells, and being capable of differentiating into the HES5+ LNP cells and into HES5- neurons and astrocytes.

According to some embodiments of the invention, the at least one rosette marker comprises PLZF.

According to some embodiments of the invention, the at least one glial marker comprises epidermal growth factor receptor (EGFR) and/or S100B.

It should be noted that the HES5- neurons which are differentiated from the HES5+ L-RG form mainly layer 2 of the brain cortex, but also layers 4 and 3 of the brain cortex.

According to some embodiments of the invention, the HES5+ late radial glial cells exhibit an HES5+/OLIG1+/PDGFRA+ expression signature.

According to some embodiments of the invention, the HES5+ late radial glial cells further exhibit an CUX1+/CUX2+/POU3F2+ expression signature.

According to some embodiments of the invention, the HES5+ late radial glial cells further exhibit an S100B+/EGFR+ expression signature.

According to some embodiments of the invention, the HES5+ late radial glial cells are characterized by a higher expression level of at least one gene selected from the group consisting of: PMP2, GABBR2, BCAN, LUZP2, SALL3, SYNM, DCT, OLIG1, SPON1, PDGFRA, COL22A1, KIAA1239, PCDH10, LPAR4, VAV3, CADM2, SOX6, SLC6A1, DPP6, FGFR3, PDE3B, MOXD1, TNFRSF19, PYGL, GPC6, COL11A1, TRIM9, GABRB3, TFPI, CREB5, RAB3GAP2, NCAN, EFHD1, SLITRK2, PAX6, SLC1A4, GPR155, GPD2, CHST11, PAQR8, MT2A, GPC3, TMEM51, CHST3, PAG1, MYO5C, CACNB2, NDRG2, ST3GAL5, TPD52L1, TRIB1, PRKCA, BCKDHB, GLT25D2, LITAF, PLCB1, TIMP3, ZBTB46, OPCML, CTDSPL, MDGA2, MEGF10, EYA2, KANK1, RAB31, TRIL, FAM171B, ALCAM, RAB6B, PGM2L1, LARGE, HPCAL1, HTRA1, TRPS1, TRIB3, IGF2BP2, PITPNC1, CMTM4, IAH1, DHTKD1, SNAP29, CTNNBIP1, NQO2, MAP4, CBR1, LTBP1, C5ORF32, MARK1, AASS, CISD1, DSC2, SLC25A33, RIMS3, ZIC3, EGF, SRGAP2, RANGAP1, SCRG1, PRCP, CA12, HEATR5A, ZNF503, GYG2, ANAPC1, C19ORF63, ASAP1, C1ORF96, DHX33, FASTKD1, STAU2, MAML2, RRAS2, GLTP, VPS13B, GPT2, NKAIN4 and ZC3HAV1 as compared to the expression level of the at least one gene in HES5- differentiated cells obtained by culturing HES5+ mid radial glial cells under culture conditions suitable for differentiation the HES5+ mid radial glial cells into HES5+ late radial glial cells.

According to some embodiments of the invention, the HES5+ L-RG cells are characterized by presence of the active transcription factor GFAP.

According to some embodiments of the invention, the HES5+ cells are late radial glial cells (L-RG) which constitute at least about 2%, e.g., at least 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20% of the isolated population of cells, e.g., about 10-15% of the isolated population of cells.

According to some embodiments of the invention, the HES5+ late radial glial (L-RG) cells are capable of differentiating into LNP cells.

As used herein the phrase "HES5+ LNP cells" refers to HES5+ adult neural stem cells (aNSCs) characterized by a higher expression level of EGFR and S100B as

compared to HES5+ L-RG cells, and which are capable of differentiation into HES5- neurons, oligodendrocyte and astrocytes.

It should be noted that the HES5- neurons which are differentiated from the HES5+ adult neural stem cells (aNSCs) comprise limited types of neurons, mainly those reaching the olfactory bulb.

According to some embodiments of the invention, the long term neural progenitor cells exhibit an HES5+/ANXA2+/LGALS1+ expression signature.

According to some embodiments of the invention, the long term neural progenitor cells further exhibit EGFR+/ S100B+ expression signature.

According to some embodiments of the invention, the HES5+ long term neural progenitor cells are characterized by a higher expression level of at least one gene selected from the group consisting of: ANXA2P2, ANXA2, FRAS1, SPOCK1, PCDHB15, SLC10A4, TPBG, C5ORF39, MMP14, TNFRSF10D, S100A6, RNF182, LGALS1, ISL1, SPINK5, DOCK10, LECT1, LYPD1, ARMCX1, NAP1L2, COL4A6, GSN, PLAG1, MMD, PTGR1, PDP1, COL18A1, ZIC4, BASP1, AHNAK, REC8, KLHDC8B, FRMD6, MYL9, RBMS1, TNFRSF21, and FAM38A as compared to the expression level of the at least one gene in HES5- differentiated cells obtained by culturing HES5+ late radial glial cells under culture conditions suitable for differentiation the HES5+ late radial glial cells into HES5+ long term neural progenitor cells.

According to some embodiments of the invention, the HES5+ **LNP** cells are characterized by presence of at least one active transcription factor selected from the group consisting of: ANXA2, LGALS1, S100B, and FABP7.

According to some embodiments of the invention, the HES5+ cells are long term neural progenitors (LNP) which constitute at least about 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, e.g., about 7-10 % of the isolated population of cells.

According to some embodiments of the invention, the HES5+ cells (at all stage of NE, E-RG, M-RG, L-RG and LNP) are characterized by presence of at least one active transcription factor selected from the group consisting of: E2F4, PAX6, RFX5, CREB3, OTX2, NR2F2, THRA, PBX2, ATF5, and CREM.

According to some embodiments of the invention, the HES5+ cells are genetically modified.

As used herein the phrase "genetically modified" refers to having been transformed with an exogenous polynucleotide or a nucleic acid construct.

According to some embodiments of the invention, the genetic modification of the cells comprise transforming the cells with a nucleic acid construct which comprises a Notch-driven reporter as is further described hereinunder.

According to an aspect of some embodiments of the invention, there is provided an isolated population of cells comprising at least 5% HES5- cells, e.g., at least 6% HES5- cells, e.g., at least 7% HES5- cells, e.g., at least 8% HES5- cells, e.g., at least 9% HES5- cells, e.g., at least 10% HES5- cells, wherein the HES5- cells are:

(i) non-CNS (CNS = central nervous system) cells comprising neural crest cells, placodal cells, non-neuroepithelial cells; and CNS cells which exhibit an NEUROD4+/NGN1+/NGN2+/TBR2+/DCX+ expression signature and which form neurons of layers 1 and 6;

(ii) neural progenitor cells which belong to the CNS, having a limited differentiation potential and are characterized by a lower proliferative capacity as compared to the HES5+ ERG cells. Upon the immediate differentiation of these cells, the cells form the earliest neurons which form layer 1 of the cortex, and eventually can also differentiate to neurons of layers 5 and 6 of the cortex. These neural progenitor cells also constitute up to 20% of the early emerging SVZ, which are also termed "intermediate progenitor cells (INPs)" or "basal progenitor cells".

(iii) intermediate progenitor cells (INPs) which belong to the CNS, and which are capable of differentiating into the neurons forming mainly layers 4 and 3 of the brain cortex, and also a small fraction forming layer 2 of the brain cortex. The INPs constitute about 80% of the SVZ (80%), exhibit TBR2+ expression pattern, wherein each INP can divide on average 3 times to create neurons of layers 4 and 3 of the brain cortex, and a small fraction forming layer 2.

(iv) HES5- neurons and some astrocytes, wherein the neurons form mainly layer 2 of the brain cortex, but also layers 4 and 3 of the brain cortex; or

(v) neurons, oligodendrocyte and astrocytes, wherein the neurons of this group comprise limited types of neurons, mainly those reaching the olfactory bulb.

According to some embodiments of the invention, HES5- cells of (i) (which comprise non-CNS cells and CNS cells), are characterized by a higher expression level of at least one gene selected from the group consisting of: LHX1, CNTN2, ST18, EBF3, NFASC, FSTL5, ONECUT2, SLC17A6, EBF1, SLIT1, SYT4, NEFM, NEUROD1, PARM1, CHN2, DNER, HMP19, TFAP2B, DCX, KLHL35, PAPPA, OLFM1, NHLH1, RTN1, GAP43, GFRA1, CHL1, FNDC5, SCN3A, NPTX2, EOMES, CADPS, NHLH2, TMEM163, STMN3, LRRN3, NEFL, ROBO2, INA, PHLDA1, GRIA1, GRIA2, DCLK1, CRABP1, OLIG2, SCG3, TMEM158, FBXL16, FAM123A, SYP, KIF21B, PCDH9, CDKN1C, IGFBPL1, RSPO3, GABRB3, TAGLN3, KCNH2, EPB41L3, EYA2, TMOD2, NCAN, GABBR2, D4S234E, PI15, ANK2, SLC1A2, NRCAM, CBLB, CAMK2N1, ZIC3, PTPRN2, SORBS1, NXPH1, BAALC, CLASP2, DPP6, MAP6, FIGNL2, KIAA0802, KIF1A, TMEM170B, SLC22A23, TMEM178, CTNND2, CADM1, LGALS3, TCEAL7, CD9, GLB1L2, NFIA, YPEL3, TNFRSF19, SPINK5, PNMA2, IRX5, SIN3B, STK38, NR3C1, SOX8, SLC6A8, SYNPR, SGK223, BASP1 and APC as compared to the expression level of the at least one gene in HES5+ neuroepithelial cells.

According to some embodiments of the invention, the HES5- cells of (ii) (which comprise neural progenitor cells which form layers 1 5 and 6 of the brain cortex) are characterized by a higher expression level of at least one gene selected from the group consisting of: GREM1, COL3A1, PCDH8, SEMA3C, BMP4, NID2, TNC, COL1A2, ANKRD1, ANXA1, TMEFF2, PDZRN3, ANXA3, KRT8, LEPREL1, NOX4, LAMB1, FLNC, FST, IMMP2L, S100A4, GDF15, PHACTR2, METTL7A, MAMDC2, DDIT4, BCHE, OCIAD2, TNFRSF10D, BBS9, ELOVL2, TUBA1C, CHST7, RBM47, TFPI, NEBL and LHFP as compared to the expression level of the at least one gene in HES5+ early radial glial cells.

According to some embodiments of the invention, the HES5- cells of (iii) (which comprise intermediate progenitor cells (INPs) capable of differentiating into the neurons forming layers 4, and 2 of the brain cortex) are characterized by a higher expression level of the ACSS1 gene as compared to the expression level of the gene in HES5+ mid radial glial cells.

According to some embodiments of the invention, the HES5- cells of (iii) are characterized by presence of the active transcription factor TBR2.

According to some embodiments of the invention, the HES5- cells of (iv) (which comprise neurons and some astrocytes, wherein the neurons form layers 2, 4 and 3 of the brain cortex) are characterized by a higher expression level of at least one gene selected from the group consisting of: THBS1, KLHL4, A2M, EN2, SLC6A6, ACTA2, ST6GAL1, SLC7A8, GRM3, FAM65B, CALB1, MYLK, TNNT1, PTX3, MFAP2 and HMGA2 as compared to the expression level of the at least one gene in HES5+ late radial glial cells.

According to some embodiments of the invention, the HES5- cells of (iv) are characterized by presence of the active transcription factor(s) POU3F3 and/or POU3F2.

According to some embodiments of the invention, the HES5- cells of (v) (which comprise neurons, oligodendrocyte and astrocytes, wherein the neurons comprise neurons reaching the olfactory bulb) are characterized by a higher expression level of at least one gene selected from the group consisting of: FBN2, NELL2, KAL1, PCDHB5, ST8SIA4, DCN, SLC6A1, CADM2, BCL11A, DDB2, ANXA11, PAK1, ID3, IGF2BP1, ANK3, ZEB2 and CREB5 as compared to the expression level of the at least one gene in HES5+ long term neural progenitor cells.

According to some embodiments of the invention, the cells are human cells.

According to some embodiments of the invention, the cells are derived from a subject having a CNS disease or disorder.

According to some embodiments of the invention, the cells having been subjected to an ex-vivo differentiation protocol.

According to some embodiments of the invention, the isolated population of cells can be used in the treatment of a CNS disease or disorder.

According to an aspect of some embodiments of the invention, there is provided a method of isolating neural progenitor cells, the method comprising:

(a) culturing pluripotent stem cells having been transformed to express a Notch-activated reporter under culture conditions suitable for differentiation of the pluripotent stem cells into neural progenitor cells; and

(b) successively isolating progenitor cells of interest based on activation of the Notch-activated reporter.

As used herein the term "successively" means more than once in a time dependent sequential manner.

According to some embodiments of the invention a successive isolation of the cells is an isolation of cell(s) derived from a previously isolated cell(s) based on activation state of the Notch reporter. For example, the first isolation is based on HES5 expression such that only HES5+ cells are isolated and further cultured. Then the cultured cells (which were previously isolated based on HES5+) are further isolated based on HES5+ expression (showing Notch activation).

As used herein the term "isolating" refers to the enrichment of a mixed population of cells (e.g., in a cell culture) with cells predominantly displaying at least one characteristic associated with a specific phenotype.

The specific phenotype can be for example, presence or absence of Notch activation. Methods of determining the status of Notch activation are known in the art. For example, the cells can be evaluated for expression of markers which are activated in the Notch pathway, such as presence or absence of expression of HES5, which is activated downstream of the Notch pathway and regulates cell differentiation in multiple tissues. The expression of the markers can be detected by monitoring presence of a reporter protein driven by the regulation of the Notch pathway related promoter, e.g., the HES5 promoter.

According to some embodiments of the invention, the method further comprising monitoring expression of the reporter, wherein cells exhibiting negative expression of the reporter are more mature, terminally differentiated cells.

The isolation of the specific cells can be performed using methods known in the art such as by fluorescence activated cell sorter (FACS), magnetically-labeled antibodies and magnetic separation columns (MACS, Miltenyi) as described by Kaufman, D.S. et al., (Hematopoietic colony-forming cells derived from human embryonic stem cells. Proc. Natl. Acad. Sci. USA. 2001, 98: 10716–10721).

To genetically modify a cell to express a reporter driven by a Notch signaling promoter, a nucleic acid construct, which comprises the regulatory sequences of a Notch signaling pathway gene (e.g., promoters, enhancers, and other regulatory sequences being upstream and/or downstream of the Notch signaling pathways coding sequence, e.g., the HES5 promoter) ligated to a coding sequence of a reporter molecule, is introduced into the cell (integration into the genome). Thus, signals affecting native

Notch signaling pathway gene expression within the cell will result in expression of the reporter molecule, which will further be detected within the cell.

Examples of the regulatory sequences of a Notch signaling pathway gene include, the regulatory sequences of genes which are downstream of Notch activation such as the HES family of genes, such as of HES1, HES2, HES3, HES4, HES5, HES6, HES7 and other Notch signaling pathway genes such as RBDJ. Such regulatory sequences are known in the art and can be obtained from the database, e.g., via the NCBI web site.

Non-limiting examples of suitable reporters include, green fluorescent protein [e.g., the enhanced green fluorescent protein from Mycobacterium tuberculosis H37Rv, depicted by polynucleotide set forth by SEQ ID NO:43; and the polypeptide set forth by GenBank Accession No. YP_009062989.1, SEQ ID NO:44], blue fluorescent protein (BFP), red fluorescent protein (RFP) and yellow fluorescent protein (YFP). The reporter molecule can be ligated under the control of a suitable promoter, e.g., the promoter of the gene-of-interest.

A non-limiting example of such a HES5-driven reporter is bacterial artificial chromosome (BAC) RP24-341I10, in which the coding sequence for enhanced green fluorescent protein (EGFP), followed by a polyadenylation signal, was inserted into the mouse genomic bacterial artificial chromosome (BAC) RP24-341I10 at the ATG transcription initiation codon of the Hes5 gene so that expression of the reporter mRNA/protein is driven by the regulatory sequences of the HES5 gene, essentially as described in Placantonakis DG, et al., 2009, which is fully incorporated herein by reference in its entirety.

According to some embodiments of the invention, the successive isolation comprises at least two isolation steps following at least two culturing steps, wherein a first isolation of the at least two isolation steps is effected up to 12 days of a first culturing of the at least two culturing steps, and wherein a second isolation of the at least two isolation steps is effected up to 5 days of a second culturing of the at least two culturing steps.

According to some embodiments of the invention, the first isolation of the at least two isolation steps is effected up to 8, 9, 10, 11 or 12 days of a first culturing of the at least two culturing steps.

According to some embodiments of the invention, the first isolation results in a population of cells comprising HES5+ neuroepithelial cells.

According to some embodiments of the invention, the first culturing is performed on an extracellular matrix or a feeder cell layer.

According to some embodiments of the invention, the first culturing is performed on an extracellular matrix.

Non-limiting examples of suitable extracellular matrixes include matrixes composed of laminin, fibronectin, collagen and the like.

MATRIGEL™ (BD Biosciences) is a solubilized basement membrane preparation extracted from the Engelbreth-Holm-Swarm (EHS) mouse sarcoma, a tumor rich in extracellular matrix proteins which include laminin (a major component), collagen IV, heparan sulfate proteoglycans, and entactin/nidogen.

According to some embodiments of the invention, the first culturing is performed on a feeder cell layer.

The feeder cell layers can be from various sources, such as fibroblasts or stroma cells, and can be from human (in order to avoid animal contamination) or from non-human source (e.g., mouse feeder cells layers).

A non-limiting example of suitable feeder cells is MS5 stromal cells. MS5 stromal cells is a murine stromal cell line established after irradiation of the adherent cells in long-term bone marrow culture; the cells produce extracellular matrix proteins such as fibronectin, laminin, and collagen type 1.

According to some embodiments of the invention, the first culturing is performed without passaging of the cells.

According to some embodiments of the invention, the first culturing is performed in the presence of a culture medium which comprises Noggin, SB-431542 and LDN-193189.

Noggin (gene name NOG, Gene ID 9241, GenBank Accession No. NP_005441.1; SEQ ID NO:45) is a secreted polypeptide, which binds and inactivates members of the transforming growth factor-beta (TGF-beta) superfamily signaling proteins, such as bone morphogenetic protein-4 (BMP4). Noggin can be obtained R&D systems (e.g., Catalogue No. 6057-NG).

According to some embodiments of the invention, the concentration of Noggin in the culture medium is between from 200-300 ng/ml, e.g., about 250 ng/ml (nanograms per milliliter).

SB-431542 is a potent and specific inhibitor of transforming growth factor-beta superfamily type I activin receptor-like kinase (ALK) receptors ALK4, ALK5, and ALK7, and can be obtained, for example, from TOCRIS (a biotechne brand, e.g., Cat. No. 1614).

According to some embodiments of the invention, the concentration of SB-431542 in the culture medium is from 5-20 μM, e.g., about 10 μM (micromolar).

LDN-193189 is a cell permeable small molecule inhibitor of bone morphogenetic protein (BMP) type I receptors ALK2 and ALK3 (IC_{50} = 5 nM and 30 nM respectively), and can be obtained from STEMGENT (e.g., Catalogue Number 04-0074).

According to some embodiments of the invention, the concentration of LDN-193189 in the culture medium is from 50-200 nM, e.g., about 100 nM (nanomolar).

Following is non-limiting protocol suitable for producing HES5+ neuroepithelial cell:

Pluripotent stem cells such as human embryonic stem cells (ESCs) are cultured on an extracellular matrix [e.g., MATRIGEL™ (BD Biosciences)] or on a feeder cell layer [e.g., MS5 stromal cells] in the presence of a culture medium which induce the cells to differentiation into the neural lineage. The culture medium can comprise: Noggin (e.g., from 200-300 ng/ml, e.g., about 250 ng/ml) and SB-431542 (e.g., from 5-20 μM, e.g., about 10 μM, Tocris), and LDN-193189 (e.g., from 50-200 nM, e.g., about 100 nM, Stemgent) for up to day 9-12. It should be noted that "day 0" (also referred to as "D0" herein) is the first day of culturing in which the pluripotent stem cells are transferred to the differentiation conditions towards the neural lineage.

According to some embodiments of the invention, the second isolation results in a population of cells comprising HES5+ early radial glial cells.

According to some embodiments of the invention, the second culturing is performed on an extracellular matrix.

For example, as shown in the Examples section which follows, the second culturing (reaching to HES5+ ERG cells) was performed on moist MATRIGEL™ drops.

According to some embodiments of the invention, the second culturing is performed in the presence of a culture medium which comprises fibroblast growth factor 8 (FGF8; GenBank Accession Nos: NP_149355.1, NP_149354.1, NP_149353.1, NP_006110.1, NP_001193318.1; SEQ ID NOs:46-50) and brain-derived neurotrophic factor (BDNF, GenBank Accession NOs. NP_001137277.1, NP_001137278.1, NP_001137279.1, NP_001137280.1, NP_001137281.1, NP_001137282.1, NP_001137283.1, NP_001137284.1, NP_001137285.1, NP_001137286.1, NP_001137288.1, NP_001700.2, NP_733927.1, NP_733928.1, NP_733929.1, NP_733930.1, NP_733931.1; SEQ ID NOs:51-67).

According to some embodiments of the invention, the medium used in the second culturing further comprises Sonic hedgehog protein.

Sonic hedgehog (gene symbol: SHH; Gene ID 6469; GenBank Accession Nos. NP_000184.1 SEQ ID NO:68; and NP_001297391.1 SEQ ID NO:69) is instrumental in patterning the early embryo. It has been implicated as the key inductive signal in patterning of the ventral neural tube, the anterior-posterior limb axis, and the ventral somites.

According to some embodiments of the invention, the concentration of Sonic hedgehog in the culture medium is about 10-50 ng/ml, e.g., about 30 ng/ml.

A non-limiting example of such a culture medium includes the DMEM/F12 medium with the N2 supplement containing Insulin, Apo-transferin, Sodium Selenite, Putrecine and Progesterone, which is further supplemented FGF8 (e.g., between 50-150 ng/ml, e.g., about 100 ng/ml) and BDNF (e.g., between 1-20 ng/ml, e.g., between 1-10 ng/ml, e.g., about 5 ng/ml).

Following is a non-limiting description of a protocol suitable for generating HES5+ E-RG cells. HES5+ neuroepithelial cells are isolated from about day 9-12 of the first culturing (preferably from day 12), and are replated at a high density (e.g., 500,000 cells/cm^2) on moist Matrigel™ drops in a DMEM/F12 and N2 supplement containing Insulin, Apo-transferin, Sodium Selenite, Putrecine and Progesterone), and further

supplemented FGF8 (100 ng/ml) and BDNF (5 ng/ml) until rosettes appeared (E-RG stage), which is about 2 days of culturing.

According to some embodiments of the invention, the successive isolation comprises at least three isolation steps following at least three culturing steps, wherein a third isolation of the at least three isolation steps is effected up to 21 days of a third culturing of the at least three culturing steps.

According to some embodiments of the invention, the third isolation results in a population of cells comprising HES5+ mid radial glial cells.

According to some embodiments of the invention, the third culturing is performed on an extracellular matrix.

According to some embodiments of the invention, the extracellular matrix which is used for the third culturing comprises polyornithine, Laminin and Fibronectin.

It should be noted that the matrix can be prepared as a solution with the extracellular matrix proteins at specific concentrations, and the solution is then poured over a surface of a culture vessel in order to form a layer of such extracellular matrix on the vessel.

According to some embodiments of the invention, the concentration of polyornithine in the extracellular matrix is in the range of 5-30 μg/ml, e.g., about 15 μg/ml.

According to some embodiments of the invention, the concentration of Laminin in the extracellular matrix is in the range of 0.5-3 μg/ml, e.g., about 1 μg/ml.

According to some embodiments of the invention, the concentration of Fibronectin in the extracellular matrix is in the range of 0.5-3 μg/ml, e.g., about 1 μg/ml.

According to some embodiments of the invention, the concentration of the polyornithine in the matrix is about 15 μg/ml, a concentration of the Laminin in the matrix is about 1 μg/ml and a concentration of the Fibronectin in the matrix is about 1 μg/ml.

According to some embodiments of the invention, the third culturing is performed in the presence of a culture medium which comprises fibroblast growth factor 2 (FGF2) and epidermal growth factor (EGF).

According to some embodiments of the invention, the successive isolation comprises at least four isolation steps following at least four culturing steps, wherein a fourth isolation of the at least four isolation steps is effected up to 45 days of a fourth culturing of the at least four culturing steps.

According to some embodiments of the invention, the fourth isolation results in a population of cells comprising HES5+ late radial glial cells.

According to some embodiments of the invention, the fourth culturing is performed on an extracellular matrix, e.g., an extracellular matrix comprises polyornithine, Laminin and Fibronectin.

According to some embodiments of the invention, the fourth culturing is performed in the presence of a culture medium which comprises fibroblast growth factor 2 (FGF2) and epidermal growth factor (EGF).

According to some embodiments of the invention, the successive isolation comprises at least five isolation steps following at least five culturing steps, wherein a fifth isolation of the at least five isolation steps is effected up to 140 days of a fifth culturing of the at least five culturing steps.

According to some embodiments of the invention, the fifth isolation results in a population of cells comprising HES5+ long term neural progenitor cells.

According to some embodiments of the invention, the fifth culturing is performed on an extracellular matrix, e.g., an extracellular matrix comprises polyornithine, Laminin and Fibronectin.

According to some embodiments of the invention, the fifth culturing is performed in the presence of a culture medium which comprises fibroblast growth factor 2 (FGF2) and epidermal growth factor (EGF).

According to some embodiments of the invention, the concentration of FGF2 in the culture medium is in the range of 1-50 ng/ml, e.g., about 20 ng/ml.

According to some embodiments of the invention, the concentration of EGF in the culture medium is in the range of 1-50 ng/ml, e.g., about 20 ng/ml.

According to some embodiments of the invention, the concentration of the FGF2 in the medium is about 20 ng/ml, and a concentration of the EGF in the medium is about 20 ng/ml.

A non-limiting example of a culture medium which can be used for the third, fourth and/or fifth culturing steps can be DMEM/F12 and N2 supplement containing Insulin, Apo-transferin, Sodium Selenite, Putrecine and Progesterone, and further supplemented with FGF2 (1-50 ng/ml, e.g., about 20 ng/ml) and EGF (1-50 ng/ml, e.g., about 20 ng/ml)

According to some embodiments of the invention, the culturing in the third, fourth and/or fifth culturing steps comprises passaging the cells every about 5-8 days, e.g., every 7 days.

Following is a non-limiting protocol of producing and isolating HES5+ M-RG cells: HES5+ E-RG cells from the second culturing step are isolated and replated on a matrix which comprises polyornithine (15 μg/ml), Laminin (1 μg/ml) and Fibronectin (1 μg/ml) in a medium which comprises: DMEM/F12 and N2 supplement containing Insulin, Apo-transferin, Sodium Selenite, Putrecine and Progesterone, and further supplemented with FGF2 (20 ng/ml) and EGF (20 ng/ml). The cells are cultured for 3 weeks (21 days), while being passaged about every 7 days, in order to maintain a proliferative (FGF and EGF responsive) neural progenitor cells state.

Following is a non-limiting protocol for producing HES5+ L-RG cells: HES5+ MRG cells isolated following the third culturing step (e.g., on day 35 of total culturing days starting from day 0) are replated on matrix which comprises polyornithine (15 μg/ml), Laminin (1 μg/ml) and Fibronectin (1 μg/ml) in the presence of a culture medium such as DMEM/F12 and N2 supplement containing Insulin, Apo-transferin, Sodium Selenite, Putrecine and Progesterone, and further supplemented with FGF2 (20 ng/ml) and EGF (20 ng/ml). The cells are cultured for up to about 45 days (thus reaching a total of about 80 days in culture from day 0), while being passaged weekly.

Following is a non-limiting protocol for producing HES5+ L-NP cells: HES5+ L-RG cells isolated following the fourth culturing step (e.g., on day 80 of the total culturing days starting form day 0) are replated on a matrix which comprises polyornithine (15 μg/ml), Laminin (1 μg/ml) and Fibronectin (1 μg/ml) with a culture medium such as DMEM/F12 and N2 supplement containing Insulin, Apo-transferin, Sodium Selenite, Putrecine and Progesterone, and further supplemented with FGF2 (20 ng/ml) and EGF (20 ng/ml) for up to 140 days (thus reaching a total of about 220 days in culture from day 0), while being passaged weekly.

According to some embodiments of the invention, the method of isolating neural progenitor cells further comprising qualifying presence of a neural progenitor cell of interest according to at least one marker comprised in an expression signature of the neural progenitor cells, wherein:

(i) an expression signature of HES5+ neuroepithelial cells comprises HES5+/SOX1+/PAX6+/SOX2+/Nestin+/CDC6+/CDX1+/CENPH+/TOP2A+;

(ii) an expression signature of HES5+ early radial glial cells comprises HES5+/ARX+/FEZF2+/NR2E1+;

(iii) an expression signature of HES5+ mid radial glial cells comprises HES5+/POU3F2+/GLAST+/FABP7+;

(iv) an expression signature of HES5+ late radial glial cells comprise HES5+/OLIG1+/PDGFRA+/CUX1+/CUX2+/POU3F2+/S100B+/EGFR+;

(v) an expression signature of HES5+ long term neural progenitor cells comprise HES5+/ANXA2+/LGALS1+/EGFR+/ S100B+.

According to some embodiments of the invention, qualifying the HES5+ neuroepithelial cells is performed by detecting an increase above a predetermined threshold in the expression level in the HES5+ cells isolated in the first isolation step of at least one gene (marker) selected from the group consisting of: TOP2A, HIST1H4C, TRIM71, PPIG, MLLT4, TNC, CDK1, OIP5, GDF15, MCM6, TP53TG1, FAM83D, FANCI, GINS2, KDM5A, GSTM3, FAM64A, LIMS1, CENPH, KIF2C, ATAD2, DTL, CDCA5, ARHGEF6, LIPA, POLE2, RRM2, MAD2L1, CKS1B, TTK, DHFR, S100A4, NUP37, PMAIP1, CENPN, RNASEH2A, BST2, MCM10, MAF, KIAA0101, C8ORF4, E2F7, CENPA, UBE2T, RAB13, TMEM126A, MAGT1, CDC6, C6ORF211, RFC5, PSMD1, HMMR, UNG, UBE2C, GINS1, AURKB, LEPREL1, SBNO1, ZWINT, MKI67, CCAR1, FKBP5, PVRL3, CCNB1, NOP58, COL4A1, GGH, LSM6, EID1, GPX8, STC2, CD276, HS2ST1, EIF5B, HDGF, and NOL7 as compared to the expression level of the at least one gene in HES5- cells which remain in the culture following the first isolation (when the culture is depleted of the HES5+ cells).

As used herein the phrase "an increase above a predetermined threshold" refers to an increase in the level of expression in of a specific gene in certain cells (e.g., the HES5+ cell) measured following a specific isolation step which is higher than the

predetermined threshold relative to the level of expression of the same gene in reference cells (e.g., the HES5- cell) measured following the same specific isolation step.

According to some embodiments of the invention, qualifying the HES5+ early radial glial cells is performed by detecting an increase above a predetermined threshold in the expression level in the HES5+ cells isolated in the second isolation step of at least one gene (marker) selected from the group consisting of: NR2E1, HES5, ARX, C1ORF61, FRZB, GRM3, EPHA3, NAV3, EGR2, RGMA, NRXN3, FAM107A, FABP7, EGR3, ZNF385B, TTYH1, SNCAIP, NRARP, PLP1, LIX1, LFNG, HES4, CD82, HS6ST1, PTPRZ1, CACHD1, DACH1, FEZF2, DTX4, FUT9, WNT5B, ENPP2, POU3F3, EMX2, MECOM, XYLT1, ARMCX2, FOS, PPAP2B, NOS2, LRP2, SOX9, NLGN3, TMEM2, CXCR7, EPHA7, SMOC1, TBC1D9, FAT4, SCUBE3, FUT8, CSPG5, DLL1, BOC, ID4, EGR1, ALPL, RFX4, GALNT12, CBX2, FHOD3, SORBS2, GUCY1B3, MBIP, FBXO16, SHISA2, DAB1, GLI3, FZD3, SEMA5B, LGALS3BP, SFRP1, C1QL1, RING1, GPRC5B, ZNF710, WSCD1, VPS37B, ZIC2, SDK2, DOCK11, GAS1, ZNF436, TMSB15A, IER2, FEZ1, CELF2, SFT2D3, NCALD, AKAP7, MYADM, NEDD4L, PHC2, PI4KAP2, STARD3, and CAMK1D as compared to the expression level of the at least one gene in HES5- cells which remain in the culture following the second isolation (when the culture is depleted of the HES5+ cells).

According to some embodiments of the invention, qualifying the HES5+ mid radial glial cells is performed by detecting an increase above a predetermined threshold in the expression level in the HES5+ cells isolated in the third isolation step of at least one gene (marker) selected from the group consisting of: FZD10, ZEB2, EN2, ST20, CDKN2C, RAB10, WASF1, ZBED4, EZH2, PPA2, H1F0, CCNJ, ITGB8, SH3BGRL3, IRX2, KIF23, PEG10, SMC3, NUSAP1, APLP1, ADAMTS3, RACGAP1, LIMCH1, ETNK1, RNF13, ARID1B, TRIM28, CNOT8, CRNDE, TWSG1, NT5DC2, NAA50, NUF2, ABCE1, PLTP, FBRSL1, DCAF16, OGT, ZFYVE16, FOXM1, PM20D2, POU3F2, MCM4, HERPUD2, VRK1, TRIM41, SATB1, HOMER1, CCNG1, ATF2, AP1AR, GABPA, STXBP3, SMC5, CDKN1B, NUPL1, UBA1, CYTH2, FXYD6, ISYNA1, DOCK1, 41527, LPHN2, IDI1, PXMP2, U2AF2, ARHGAP12, KLHL24, CKAP2, ZNF238, PARP6, NHSL1, PBRM1, BAZ1A, MAP4K5, TSPAN12, SH3GLB1, ASPM, ANKLE2, SPG20, MAP4K4, CASC4,

FUBP3, ARSB, BTAF1, SIKE1, VEZT, PBX3, CBL, EIF2AK4, API5, MTSS1, NET1, CHD6, ZNF117, PNMA1, PTPN13, MTIF2, SSFA2, KIAA1279, STRN3, WIPF1, MEIS2, ZC3H4, DYNC1I2, RTN4, TAF2, RASA1, OSBPL8, SKA2, IGF1R, RNF6, SGTB, TMEM131, HIATL1, TGIF1, TMEM170B, PSAT1, ACBD5, HECTD2, ASF1A, LAMB1, GLS, DDX39, DGCR2, EIF1AX, SALL1, GOLPH3, PTBP2, GRIP1, PNPLA8, VASH2, SUDS3, PFDN4, BAZ2A, PRKDC, GLYR1, DAZAP2, PCMTD1, SENP6, CLINT1, RECQL, CNTNAP2, CTBP1, C10ORF18, CDON, B4GALT6, CSNK1G3, STAT5B, TMEM60, HNRNPH2, TACC2, CCNG2, FSCN1, CCNA2, C21ORF45, PLRG1, ZFHX3, UBE2A, DMTF1, TRA2A, MYO5A, FAM96A, IFT80, VPS26A, MRPL50, ACYP1, WDR11, PLDN, RPRD1A, MEAF6, CKAP5, YTHDC2, GABARAPL2, IPO5, PGAP1, C14ORF147, CD200, MST4, PPT1, ANKRD50, HPS3, CCNC, THRAP3, TWF1, CYP51A1, PSPC1, WDR75, CAST, SEPW1, C21ORF59, PIK3C2A, GNG5, MED4, GIPC1, STK39, KIAA1715, PHF6, PPTC7, SOCS4, PPM1B, UQCRB, C10ORF84, SLAIN1, RAB6A, SOS2, KLF10, RNF4, C3ORF63, INSIG1, CPSF1, DNAJC4, ATP2B4, PPP2R1A, TRIM22, SDC3, TSNAX, PPIL4, ZDHHC2, ZBTB44, ANO6, PPP2CB, UBA2, BBS2, ZNF423, RNF5, C1ORF31, IFT81, CPSF6, KLHL9, FAM164A, TTC35, CCDC90B, TM9SF4, SEC24B, SMARCC2, CAP2, SAR1A, THY1, RBPMS2, EIF3A, DZIP1, ARL6IP1, SACM1L, PAPD4, SCG2, TCF3, EFHA1, HNRNPA2B1, EWSR1, STAG2, YEATS4, PAQR3, GAR1, FTH1, C19ORF43, TMEM14C, CCDC104, PSMD12, DCTD, SSR1, HMGCS1, HMGB3, KIF3A, TMEM128, PATZ1, RBL2, ARFGAP3, DNAJB5, TMED7, G3BP2, BMPR1A, FMR1, TPST2, TMSB4X, RP2, CEP170, KLHL23, RNF7, HNRNPH1, MARCKS, HNRNPD, TOB1, UTP11L, RFK, DHX36, LCOR, WBP5, PHLDB2, USP33, EFNB2, C6ORF62, MEX3B, ABCD3, ATG3, ARID4B, C7ORF11, EPB41, TCF12, CDK8, CMIP, ATG12, CETN3, ZNF217, TMEM55A and UBE2N as compared to the expression level of the at least one gene in HES5- cells which remain in the culture following the third isolation (when the culture is depleted of the HES5+ cells).

According to some embodiments of the invention, qualifying the HES5+ late radial glial cells is performed by detecting an increase above a predetermined threshold in the expression level in the HES5+ cells isolated in the fourth isolation step of at least one gene (marker) selected from the group consisting of: PMP2, GABBR2, BCAN,

LUZP2, SALL3, SYNM, DCT, OLIG1, SPON1, PDGFRA, COL22A1, KIAA1239, PCDH10, LPAR4, VAV3, CADM2, SOX6, SLC6A1, DPP6, FGFR3, PDE3B, MOXD1, TNFRSF19, PYGL, GPC6, COL11A1, TRIM9, GABRB3, TFPI, CREB5, RAB3GAP2, NCAN, EFHD1, SLITRK2, PAX6, SLC1A4, GPR155, GPD2, CHST11, PAQR8, MT2A, GPC3, TMEM51, CHST3, PAG1, MYO5C, CACNB2, NDRG2, ST3GAL5, TPD52L1, TRIB1, PRKCA, BCKDHB, GLT25D2, LITAF, PLCB1, TIMP3, ZBTB46, OPCML, CTDSPL, MDGA2, MEGF10, EYA2, KANK1, RAB31, TRIL, FAM171B, ALCAM, RAB6B, PGM2L1, LARGE, HPCAL1, HTRA1, TRPS1, TRIB3, IGF2BP2, PITPNC1, CMTM4, IAH1, DHTKD1, SNAP29, CTNNBIP1, NQO2, MAP4, CBR1, LTBP1, C5ORF32, MARK1, AASS, CISD1, DSC2, SLC25A33, RIMS3, ZIC3, EGF, SRGAP2, RANGAP1, SCRG1, PRCP, CA12, HEATR5A, ZNF503, GYG2, ANAPC1, C19ORF63, ASAP1, C1ORF96, DHX33, FASTKD1, STAU2, MAML2, RRAS2, GLTP, VPS13B, GPT2, NKAIN4 and ZC3HAV1 as compared to the expression level of the at least one gene in HES5- cells which remain in the culture following the fourth isolation (when the culture is depleted of the HES5+ cells).

According to some embodiments of the invention, qualifying the HES5+ long term neural progenitor cells is performed by detecting an increase above a predetermined threshold in the expression level in the HES5+ cells isolated in the fifth isolation step of at least one gene (marker) selected from the group consisting of: ANXA2P2, ANXA2, FRAS1, SPOCK1, PCDHB15, SLC10A4, TPBG, C5ORF39, MMP14, TNFRSF10D, S100A6, RNF182, LGALS1, ISL1, SPINK5, DOCK10, LECT1, LYPD1, ARMCX1, NAP1L2, COL4A6, GSN, PLAG1, MMD, PTGR1, PDP1, COL18A1, ZIC4, BASP1, AHNAK, REC8, KLHDC8B, FRMD6, MYL9, RBMS1, TNFRSF21, and FAM38A as compared to the expression level of the at least one gene in HES5- cells which remain in the culture following the fifth isolation (when the culture is depleted of the HES5+ cells).

According to some embodiments of the invention, the method further comprising qualifying presence of a neural progenitor cell of interest according to epigenetic analysis (e.g., DNA methylation and histone modification) functional phenotype and/or morphological phenotype. Non-limiting examples of such analyses are described in the Examples section which follows.

The expression of several transcription factors was analyzed in cells appearing (or present) in the various stages of neural differentiation. Cells of the "NE stage" refer to the cells present following the first culturing step; Cells of the "E-RG stage" or "ERG stage" refer to the cells present following the second culturing step; Cells of the "M-RG stage" or "MRG stage" refer to cells present following the third culturing step; Cells of the "L-RG stage" or "LRG stage" refer to cells present following the fourth culturing step; and Cells of the "LNP stage" refer to cells present following the fifth culturing step.

Transcription factors which are upregulated early in neural differentiation of pluripotent stem cells include FOXG1, PAX6, ZIC1, SP8 and ARX. These factors are upregulated in NE, E-RG, M-EG, L-RG and LNP stages as compared to their level in undifferentiated pluripotent stem cells (hESC, Figure 6A).

Transcription factors which are upregulated in the middle (mid upregulated) of neural differentiation of pluripotent stem cells include NFIA, NFIB and SLITRK3. These factors are upregulated in M-RG, L-RG and LNP stages as compared to their level in undifferentiated pluripotent stem cells (hESC, Figure 6A).

Transcription factors which are upregulated late in neural differentiation of pluripotent stem cells include GABBR2, GRIA4, GRM3, DLX1, DLX2, OLIG1, OLIG2 and LGALS1. These factors are upregulated in L-RG and LNP stages as compared to their level in undifferentiated pluripotent stem cells (hESC, Figure 6A).

Transcription factors which are upregulated early but transiently during neural differentiation of pluripotent stem cells include HES5, NR2E1, DLL1, EMX2, MEIS2, LGR5, DACH1, PLAGL1 and LGI1. These factors are upregulated in NE, E-RG, and some also in the M-EG stage as compared to their level in undifferentiated pluripotent stem cells (hESC), but this upregulation is transient and reduces during the L-RG and LNP stages (Figure 6A).

Transcription factors which are unique to the HES5- cells of the NE stage include COMES and RSPO2 (activated as compared to undifferentiated hESCs). These factors are also upregulated in the M-RG stage in both HES5+ and HES5- as compared to their levels in undifferentiated hESCs.

A non-limiting example of a transcription factor which is downregulated early in neural differentiation includes POU5F1. This factor in downregulated in cells of the

NE, E-RG, M-EG, L-RG and LNP stages as compared to its level in undifferentiated pluripotent stem cells (hESC, Figure 6A).

Transcription factors which are downregulated in the middle (mid upregulated) stage of neural differentiation of pluripotent stem cells include LIN28A, HMGA2 and F11R. These factors are downregulated in M-RG, L-RG and LNP stages as compared to their level in undifferentiated pluripotent stem cells (hESC, Figure 6A).

A non-limiting example of a transcription factor which is downregulated late in neural differentiation of pluripotent stem cells includes OTX2. This factor is downregulated in L-RG and LNP stages as compared to its level in undifferentiated pluripotent stem cells (hESC, Figure 6A).

According to some embodiments of the invention, the at least one marker comprises a cell surface marker.

According to some embodiments of the invention, the at least one marker comprises a secreted marker.

According to some embodiments of the invention, the stem cells comprise pluripotent stem cells.

According to some embodiments of the invention, the pluripotent stem cells comprise embryonic stem cells.

According to some embodiments of the invention, the pluripotent stem cells comprise induced pluripotent stem (iPS) cells.

According to some embodiments of the invention, the stem cells are derived from a subject having a CNS disease or disorder.

According to some embodiments of the invention, the CNS disease or disorder comprises a motor-neuron disease.

According to some embodiments of the invention, the CNS disease or disorder is characterized by cortex damage.

According to some embodiments of the invention, the cells of some embodiments of the invention are characterized according to the expression values, epigenetic analyses and/or transcriptional activity provided by the TERA and/or DNAchip and/or RNAseq and/or microarray and/or histon modification and/or DNA methylation analyses provided in the Examples and Tables included in the application

as well as in the supplementary Data 1-7 attached herein, which are fully incorporated herein in their entirety.

According to an aspect of some embodiments of the invention there is provided a culture medium for neuroepithelial differentiation comprising noggin, LDN-193189 and SB-431542.

According to some embodiments of the invention, the concentration of Noggin in the culture medium is between from 200-300 ng/ml, e.g., about 250 ng/ml (nanograms per milliliter), the concentration of LDN-193189 in the culture medium is from 50-200 nM, e.g., about 100 nM (nanomolar), and the concentration of SB-431542 in the culture medium is from 5-20 μM, e.g., about 10 μM (micromolar).

According to specific embodiments of the invention, the culture medium comprises Noggin (250 ng/ml), LDN-193189 (100 nM) and SB-431542 (10 μM).

According to some embodiments of the invention, the culture medium further comprises sonic hedgehog.

According to some embodiments of the invention, the concentration of Sonic hedgehog in the culture medium is about 10-50 ng/ml, e.g., about 30 ng/ml.

As used herein, the phrase "stem cells" refers to cells which are capable of remaining in an undifferentiated state (*e.g.*, pluripotent or multipotent stem cells) for extended periods of time in culture until induced to differentiate into other cell types having a particular, specialized function (*e.g.*, fully differentiated cells). Preferably, the phrase "stem cells" encompasses embryonic stem cells (ESCs), induced pluripotent stem cells (iPS), adult stem cells and hematopoietic stem cells.

The phrase "embryonic stem cells" refers to embryonic cells which are capable of differentiating into cells of all three embryonic germ layers (*i.e.*, endoderm, ectoderm and mesoderm), or remaining in an undifferentiated state. The phrase "embryonic stem cells" may comprise cells which are obtained from the embryonic tissue formed after gestation (e.g., blastocyst) before implantation of the embryo (*i.e.*, a pre-implantation blastocyst), extended blastocyst cells (EBCs) which are obtained from a post-implantation/pre-gastrulation stage blastocyst (see WO2006/040763) and embryonic germ (EG) cells which are obtained from the genital tissue of a fetus any time during gestation, preferably before 10 weeks of gestation.

Induced pluripotent stem cells (iPS; embryonic-like stem cells), are cells obtained by de-differentiation of adult somatic cells which are endowed with pluripotency (*i.e.*, being capable of differentiating into the three embryonic germ cell layers, *i.e.*, endoderm, ectoderm and mesoderm). According to some embodiments of the invention, such cells are obtained from a differentiated tissue (e.g., a somatic tissue such as skin) and undergo de-differentiation by genetic manipulation which re-program the cell to acquire embryonic stem cells characteristics. According to some embodiments of the invention, the induced pluripotent stem cells are formed by inducing the expression of Oct-4, Sox2, Kfl4 and c-Myc in a somatic stem cell.

The phrase "adult stem cells" (also called "tissue stem cells" or a stem cell from a somatic tissue) refers to any stem cell derived from a somatic tissue [of either a postnatal or prenatal animal (especially the human)]. The adult stem cell is generally thought to be a multipotent stem cell, capable of differentiation into multiple cell types. Adult stem cells can be derived from any adult, neonatal or fetal tissue such as adipose tissue, skin, kidney, liver, prostate, pancreas, intestine, bone marrow and placenta.

Hematopoietic stem cells, which may also referred to as adult tissue stem cells, include stem cells obtained from blood or bone marrow tissue of an individual at any age or from cord blood of a newborn individual. Preferred stem cells according to this aspect of some embodiments of the invention are embryonic stem cells, preferably of a human or primate (e.g., monkey) origin.

Placental and cord blood stem cells may also be referred to as "young stem cells".

The embryonic stem cells of some embodiments of the invention can be obtained using well-known cell-culture methods. For example, human embryonic stem cells can be isolated from human blastocysts. Human blastocysts are typically obtained from human *in vivo* preimplantation embryos or from *in vitro* fertilized (IVF) embryos. Alternatively, a single cell human embryo can be expanded to the blastocyst stage. For the isolation of human ES cells the zona pellucida is removed from the blastocyst and the inner cell mass (ICM) is isolated by immunosurgery, in which the trophectoderm cells are lysed and removed from the intact ICM by gentle pipetting. The ICM is then plated in a tissue culture flask containing the appropriate medium which enables its outgrowth. Following 9 to 15 days, the ICM derived outgrowth is dissociated into

clumps either by a mechanical dissociation or by an enzymatic degradation and the cells are then re-plated on a fresh tissue culture medium. Colonies demonstrating undifferentiated morphology are individually selected by micropipette, mechanically dissociated into clumps, and re-plated. Resulting ES cells are then routinely split every 4-7 days. For further details on methods of preparation human ES cells see Thomson et al., [U.S. Pat. No. 5,843,780; Science 282: 1145, 1998; Curr. Top. Dev. Biol. 38: 133, 1998; Proc. Natl. Acad. Sci. USA 92: 7844, 1995]; Bongso et al., [Hum Reprod 4: 706, 1989]; and Gardner et al., [Fertil. Steril. 69: 84, 1998].

It will be appreciated that commercially available stem cells can also be used according to some embodiments of the invention. Human ES cells can be purchased from the NIH human embryonic stem cells registry [Hypertext Transfer Protocol://grants (dot) nih (dot) gov/stem_cells/registry/current (dot) htm]. Non-limiting examples of commercially available embryonic stem cell lines are BG01, BG02, BG03, BG04, CY12, CY30, CY92, CY10, TE03, TE32, CHB-4, CHB-5, CHB-6, CHB-8, CHB-9, CHB-10, CHB-11, CHB-12, HUES 1, HUES 2, HUES 3, HUES 4, HUES 5, HUES 6, HUES 7, HUES 8, HUES 9, HUES 10, HUES 11, HUES 12, HUES 13, HUES 14, HUES 15, HUES 16, HUES 17, HUES 18, HUES 19, HUES 20, HUES 21, HUES 22, HUES 23, HUES 24, HUES 25, HUES 26, HUES 27, HUES 28, CyT49, RUES3, WA01, UCSF4, NYUES1, NYUES2, NYUES3, NYUES4, NYUES5, NYUES6, NYUES7, UCLA 1, UCLA 2, UCLA 3, WA077 (H7), WA09 (H9), WA13 (H13), WA14 (H14), HUES 62, HUES 63, HUES 64, CT1, CT2, CT3, CT4, MA135, Eneavour-2, WIBR1, WIBR2, WIBR3, WIBR4, WIBR5, WIBR6, HUES 45, Shef 3, Shef 6, BJNhem19, BJNhem20, SA001, SA001.

In addition, ES cells can be obtained from other species as well, including mouse (Mills and Bradley, 2001), golden hamster [Doetschman et al., 1988, Dev Biol. 127: 224-7], rat [Iannaccone et al., 1994, Dev Biol. 163: 288-92] rabbit [Giles et al. 1993, Mol Reprod Dev. 36: 130-8; Graves & Moreadith, 1993, Mol Reprod Dev. 1993, 36: 424-33], several domestic animal species [Notarianni et al., 1991, J Reprod Fertil Suppl. 43: 255-60; Wheeler 1994, Reprod Fertil Dev. 6: 563-8; Mitalipova et al., 2001, Cloning. 3: 59-67] and non-human primate species (Rhesus monkey and marmoset) [Thomson et al., 1995, Proc Natl Acad Sci U S A. 92: 7844-8; Thomson et al., 1996, Biol Reprod. 55: 254-9].

Extended blastocyst cells (EBCs) can be obtained from a blastocyst of at least nine days post fertilization at a stage prior to gastrulation. Prior to culturing the blastocyst, the zona pellucida is digested [for example by Tyrode's acidic solution (Sigma Aldrich, St Louis, MO, USA)] so as to expose the inner cell mass. The blastocysts are then cultured as whole embryos for at least nine and no more than fourteen days post fertilization (*i.e.*, prior to the gastrulation event) *in vitro* using standard embryonic stem cell culturing methods.

EG cells are prepared from the primordial germ cells obtained from fetuses of about 8-11 weeks of gestation (in the case of a human fetus) using laboratory techniques known to anyone skilled in the arts. The genital ridges are dissociated and cut into small chunks which are thereafter disaggregated into cells by mechanical dissociation. The EG cells are then grown in tissue culture flasks with the appropriate medium. The cells are cultured with daily replacement of medium until a cell morphology consistent with EG cells is observed, typically after 7-30 days or 1-4 passages. For additional details on methods of preparation human EG cells see Shamblott et al., [Proc. Natl. Acad. Sci. USA 95: 13726, 1998] and U.S. Patent No. 6,090,622.

Induced pluripotent stem cells (iPS) (embryonic-like stem cells) can be generated from somatic cells by genetic manipulation of somatic cells, e.g., by retroviral transduction of somatic cells such as fibroblasts, hepatocytes, gastric epithelial cells with transcription factors such as Oct-3/4, Sox2, c-Myc, and KLF4 [Yamanaka S, Cell Stem Cell. 2007, 1(1):39-49; Aoi T, et al., Generation of Pluripotent Stem Cells from Adult Mouse Liver and Stomach Cells. Science. 2008 Feb 14. (Epub ahead of print); IH Park, Zhao R, West JA, et al. Reprogramming of human somatic cells to pluripotency with defined factors. Nature 2008;451:141-146; K Takahashi, Tanabe K, Ohnuki M, et al. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell 2007;131:861-872]. Other embryonic-like stem cells can be generated by nuclear transfer to oocytes, fusion with embryonic stem cells or nuclear transfer into zygotes if the recipient cells are arrested in mitosis.

Adult tissue stem cells can be isolated using various methods known in the art such as those disclosed by Alison, M.R. [J Pathol. 2003 200(5): 547-50], Cai, J. et al., [Blood Cells Mol Dis. 2003 31(1): 18-27], Collins, A.T. et al., [J Cell Sci. 2001; 114(Pt 21): 3865-72], Potten, C. S. and Morris, R. J. [Epithelial stem cells *in vivo*. 1988. J. Cell

Sci. Suppl. 10, 45-62], Dominici, M et al., [J. Biol. Regul. Homeost. Agents. 2001, 15: 28-37], Caplan and Haynesworth [U.S. Pat. No. 5,486,359] Jones E.A. et al., [Arthritis Rheum. 2002, 46(12): 3349-60]. Fetal stem cells can be isolated using various methods known in the art such as those disclosed by Eventov-Friedman S, et al., PLoS Med. 2006, 3: e215; Eventov-Friedman S, et al., Proc Natl Acad Sci U S A. 2005, 102: 2928-33; Dekel B, et al., 2003, Nat Med. 9: 53-60; and Dekel B, et al., 2002, J. Am. Soc. Nephrol. 13: 977-90. Hematopoietic stem cells can be isolated using various methods known in the arts such as those disclosed by "Handbook of Stem Cells" edit by Robert Lanze, Elsevier Academic Press, 2004, Chapter 54, pp609-614, isolation and characterization of hematopoietic stem cells, by Gerald J Spangrude and William B Stayton.

Generally, isolation of adult tissue stem cells is based on the discrete location (or niche) of each cell type included in the adult tissue, *i.e.*, the stem cells, the transit amplifying cells and the terminally differentiated cells [Potten, C. S. and Morris, R. J. (1988). Epithelial stem cells *in vivo*. J. Cell Sci. Suppl. 10, 45-62]. Thus, an adult tissue such as, for example, prostate tissue is digested with Collagenase and subjected to repeated unit gravity centrifugation to separate the epithelial structures of the prostate (e.g., organoids, acini and ducts) from the stromal cells. Organoids are then disaggregated into single cell suspensions by incubation with Trypsin/EDTA (Life Technologies, Paisley, UK) and the basal, CD44-positive, stem cells are isolated from the luminal, CD57-positive, terminally differentiated secretory cells, using anti-human CD44 antibody (clone G44-26; Pharmingen, Becton Dickinson, Oxford, UK) labeling and incubation with MACS (Miltenyi Biotec Ltd, Surrey, UK) goat anti-mouse IgG microbeads. The cell suspension is then applied to a MACS column and the basal cells are eluted and re-suspended in WAJC 404 complete medium [Robinson, E.J. et al. (1998). Basal cells are progenitors of luminal cells in primary cultures of differentiating human prostatic epithelium Prostate 37, 149-160].

Since basal stem cells can adhere to basement membrane proteins more rapidly than other basal cells [Jones, P.H. et al. (1995). Stem cell patterning and fate in human epidermis. Cell 60, 83-93; Shinohara, T., et al. (1999). β1- and α6-integrin are surface markers on mouse spermatogonial stem cells. Proc. Natl. Acad. Sci. USA 96, 5504-5509] the CD44 positive basal cells are plated onto tissue culture dishes coated with

either type I collagen (52 μg/ml), type IV collagen (88 μg/ml) or laminin 1 (100 μg/ml; Biocoat®, Becton Dickinson) previously blocked with 0.3 % bovine serum albumin (fraction V, Sigma-Aldrich, Poole, UK) in Dulbecco's phosphate buffered saline (PBS; Oxoid Ltd, Basingstoke, UK). Following 5 minutes, the tissue culture dishes are washed with PBS and adherent cells, containing the prostate tissue basal stem cells are harvested with trypsin-EDTA.

It will be appreciated that undifferentiated stem cells are of a distinct morphology, which is clearly distinguishable from differentiated cells of embryo or adult origin by the skilled in the art. Typically, undifferentiated stem cells have high nuclear/cytoplasmic ratios, prominent nucleoli and compact colony formation with poorly discernable cell junctions. Additional features of undifferentiated stem cells are further described hereinunder.

Methods of detecting the expression level of RNA

The expression level of the RNA in the cells of some embodiments of the invention can be determined using methods known in the arts.

Northern Blot analysis: This method involves the detection of a particular RNA in a mixture of RNAs. An RNA sample is denatured by treatment with an agent (e.g., formaldehyde) that prevents hydrogen bonding between base pairs, ensuring that all the RNA molecules have an unfolded, linear conformation. The individual RNA molecules are then separated according to size by gel electrophoresis and transferred to a nitrocellulose or a nylon-based membrane to which the denatured RNAs adhere. The membrane is then exposed to labeled DNA probes. Probes may be labeled using radio-isotopes or enzyme linked nucleotides. Detection may be using autoradiography, colorimetric reaction or chemiluminescence. This method allows both quantitation of an amount of particular RNA molecules and determination of its identity by a relative position on the membrane which is indicative of a migration distance in the gel during electrophoresis.

RT-PCR analysis: This method uses PCR amplification of relatively rare RNAs molecules. First, RNA molecules are purified from the cells and converted into complementary DNA (cDNA) using a reverse transcriptase enzyme (such as an MMLV-RT) and primers such as, oligo dT, random hexamers or gene specific primers. Then by applying gene specific primers and Taq DNA polymerase, a PCR amplification

reaction is carried out in a PCR machine. Those of skills in the art are capable of selecting the length and sequence of the gene specific primers and the PCR conditions (*i.e.*, annealing temperatures, number of cycles and the like) which are suitable for detecting specific RNA molecules. It will be appreciated that a semi-quantitative RT-PCR reaction can be employed by adjusting the number of PCR cycles and comparing the amplification product to known controls.

RNA in situ hybridization stain: In this method DNA or RNA probes are attached to the RNA molecules present in the cells. Generally, the cells are first fixed to microscopic slides to preserve the cellular structure and to prevent the RNA molecules from being degraded and then are subjected to hybridization buffer containing the labeled probe. The hybridization buffer includes reagents such as formamide and salts (e.g., sodium chloride and sodium citrate) which enable specific hybridization of the DNA or RNA probes with their target mRNA molecules *in situ* while avoiding non-specific binding of probe. Those of skills in the art are capable of adjusting the hybridization conditions (*i.e.*, temperature, concentration of salts and formamide and the like) to specific probes and types of cells. Following hybridization, any unbound probe is washed off and the bound probe is detected using known methods. For example, if a radio-labeled probe is used, then the slide is subjected to a photographic emulsion which reveals signals generated using radio-labeled probes; if the probe was labeled with an enzyme then the enzyme-specific substrate is added for the formation of a colorimetric reaction; if the probe is labeled using a fluorescent label, then the bound probe is revealed using a fluorescent microscope; if the probe is labeled using a tag (e.g., digoxigenin, biotin, and the like) then the bound probe can be detected following interaction with a tag-specific antibody which can be detected using known methods.

In situ RT-PCR stain: This method is described in Nuovo GJ, et al. [Intracellular localization of polymerase chain reaction (PCR)-amplified hepatitis C cDNA. Am J Surg Pathol. 1993, 17: 683-90] and Komminoth P, et al. [Evaluation of methods for hepatitis C virus detection in archival liver biopsies. Comparison of histology, immunohistochemistry, *in situ* hybridization, reverse transcriptase polymerase chain reaction (RT-PCR) and *in situ* RT-PCR. Pathol Res Pract. 1994, 190: 1017-25]. Briefly, the RT-PCR reaction is performed on fixed cells by incorporating labeled nucleotides to the PCR reaction. The reaction is carried on using a specific *in*

situ RT-PCR apparatus such as the laser-capture microdissection PixCell I LCM system available from Arcturus Engineering (Mountainview, CA).

DNA microarrays/DNA chips:

The expression of thousands of genes may be analyzed simultaneously using DNA microarrays, allowing analysis of the complete transcriptional program of an organism during specific developmental processes or physiological responses. DNA microarrays consist of thousands of individual gene sequences attached to closely packed areas on the surface of a support such as a glass microscope slide. Various methods have been developed for preparing DNA microarrays. In one method, an approximately 1 kilobase segment of the coding region of each gene for analysis is individually PCR amplified. A robotic apparatus is employed to apply each amplified DNA sample to closely spaced zones on the surface of a glass microscope slide, which is subsequently processed by thermal and chemical treatment to bind the DNA sequences to the surface of the support and denature them. Typically, such arrays are about 2 x 2 cm and contain about individual nucleic acids 6000 spots. In a variant of the technique, multiple DNA oligonucleotides, usually 20 nucleotides in length, are synthesized from an initial nucleotide that is covalently bound to the surface of a support, such that tens of thousands of identical oligonucleotides are synthesized in a small square zone on the surface of the support. Multiple oligonucleotide sequences from a single gene are synthesized in neighboring regions of the slide for analysis of expression of that gene. Hence, thousands of genes can be represented on one glass slide. Such arrays of synthetic oligonucleotides may be referred to in the art as "DNA chips", as opposed to "DNA microarrays", as described above [Lodish et al. (eds.). Chapter 7.8: DNA Microarrays: Analyzing Genome-Wide Expression. In: Molecular Cell Biology, 4th ed., W. H. Freeman, New York. (2000)].

Oligonucleotide microarray – In this method oligonucleotide probes capable of specifically hybridizing with the polynucleotides of some embodiments of the invention are attached to a solid surface (e.g., a glass wafer). Each oligonucleotide probe is of approximately 20-25 nucleic acids in length. To detect the expression pattern of the polynucleotides of some embodiments of the invention in a specific cell sample (e.g., blood cells), RNA is extracted from the cell sample using methods known in the art (using e.g., a TRIZOL solution, Gibco BRL, USA). Hybridization can take place using

either labeled oligonucleotide probes (e.g., 5'-biotinylated probes) or labeled fragments of complementary DNA (cDNA) or RNA (cRNA). Briefly, double stranded cDNA is prepared from the RNA using reverse transcriptase (RT) (e.g., Superscript II RT), DNA ligase and DNA polymerase I, all according to manufacturer's instructions (Invitrogen Life Technologies, Frederick, MD, USA). To prepare labeled cRNA, the double stranded cDNA is subjected to an *in vitro* transcription reaction in the presence of biotinylated nucleotides using e.g., the BioArray High Yield RNA Transcript Labeling Kit (Enzo, Diagnostics, Affymetix Santa Clara CA). For efficient hybridization the labeled cRNA can be fragmented by incubating the RNA in 40 mM Tris Acetate (pH 8.1), 100 mM potassium acetate and 30 mM magnesium acetate for 35 minutes at 94 °C. Following hybridization, the microarray is washed and the hybridization signal is scanned using a confocal laser fluorescence scanner which measures fluorescence intensity emitted by the labeled cRNA bound to the probe arrays.

For example, in the Affymetrix microarray (Affymetrix®, Santa Clara, CA) each gene on the array is represented by a series of different oligonucleotide probes, of which, each probe pair consists of a perfect match oligonucleotide and a mismatch oligonucleotide. While the perfect match probe has a sequence exactly complimentary to the particular gene, thus enabling the measurement of the level of expression of the particular gene, the mismatch probe differs from the perfect match probe by a single base substitution at the center base position. The hybridization signal is scanned using the Agilent scanner, and the Microarray Suite software subtracts the non-specific signal resulting from the mismatch probe from the signal resulting from the perfect match probe.

Methods of detecting expression and/or activity of proteins

Expression and/or activity level of proteins expressed in the cells of the cultures of some embodiments of the invention can be determined using methods known in the arts.

Enzyme linked immunosorbent assay (ELISA): This method involves fixation of a sample (e.g., fixed cells or a proteinaceous solution) containing a protein substrate to a surface such as a well of a microtiter plate. A substrate specific antibody coupled to an enzyme is applied and allowed to bind to the substrate. Presence of the antibody is then detected and quantitated by a colorimetric reaction employing the enzyme

coupled to the antibody. Enzymes commonly employed in this method include horseradish peroxidase and alkaline phosphatase. If well calibrated and within the linear range of response, the amount of substrate present in the sample is proportional to the amount of color produced. A substrate standard is generally employed to improve quantitative accuracy.

Western blot: This method involves separation of a substrate from other protein by means of an acrylamide gel followed by transfer of the substrate to a membrane (e.g., nylon or PVDF). Presence of the substrate is then detected by antibodies specific to the substrate, which are in turn detected by antibody binding reagents. Antibody binding reagents may be, for example, protein A, or other antibodies. Antibody binding reagents may be radiolabeled or enzyme linked as described hereinabove. Detection may be by autoradiography, colorimetric reaction or chemiluminescence. This method allows both quantitation of an amount of substrate and determination of its identity by a relative position on the membrane which is indicative of a migration distance in the acrylamide gel during electrophoresis.

Radio-immunoassay (RIA): In one version, this method involves precipitation of the desired protein (*i.e.*, the substrate) with a specific antibody and radiolabeled antibody binding protein (e.g., protein A labeled with I^{125}) immobilized on a precipitable carrier such as agarose beads. The number of counts in the precipitated pellet is proportional to the amount of substrate.

In an alternate version of the RIA, a labeled substrate and an unlabelled antibody binding protein are employed. A sample containing an unknown amount of substrate is added in varying amounts. The decrease in precipitated counts from the labeled substrate is proportional to the amount of substrate in the added sample.

Fluorescence activated cell sorting (FACS): This method involves detection of a substrate *in situ* in cells by substrate specific antibodies. The substrate specific antibodies are linked to fluorophores. Detection is by means of a cell sorting machine which reads the wavelength of light emitted from each cell as it passes through a light beam. This method may employ two or more antibodies simultaneously.

Immunohistochemical analysis: This method involves detection of a substrate *in situ* in fixed cells by substrate specific antibodies. The substrate specific antibodies may be enzyme linked or linked to fluorophores. Detection is by microscopy and

subjective or automatic evaluation. If enzyme linked antibodies are employed, a colorimetric reaction may be required. It will be appreciated that immunohistochemistry is often followed by counterstaining of the cell nuclei using for example Hematoxyline or Giemsa stain.

In situ activity assay: According to this method, a chromogenic substrate is applied on the cells containing an active enzyme and the enzyme catalyzes a reaction in which the substrate is decomposed to produce a chromogenic product visible by a light or a fluorescent microscope.

In vitro activity assays: In these methods the activity of a particular enzyme is measured in a protein mixture extracted from the cells. The activity can be measured in a spectrophotometer well using colorimetric methods or can be measured in a non-denaturing acrylamide gel (*i.e.*, activity gel). Following electrophoresis the gel is soaked in a solution containing a substrate and colorimetric reagents. The resulting stained band corresponds to the enzymatic activity of the protein of interest. If well calibrated and within the linear range of response, the amount of enzyme present in the sample is proportional to the amount of color produced. An enzyme standard is generally employed to improve quantitative accuracy.

As used herein the abbreviation “vs” means versus.

As used herein the term “about” refers to ± 10 %.

The terms "comprises", "comprising", "includes", "including", “having” and their conjugates mean "including but not limited to".

The term “consisting of” means “including and limited to”.

The term "consisting essentially of" means that the composition, method or structure may include additional ingredients, steps and/or parts, but only if the additional ingredients, steps and/or parts do not materially alter the basic and novel characteristics of the claimed composition, method or structure.

As used herein, the singular form "a", "an" and "the" include plural references unless the context clearly dictates otherwise. For example, the term "a compound" or "at least one compound" may include a plurality of compounds, including mixtures thereof.

Throughout this application, various embodiments of this invention may be presented in a range format. It should be understood that the description in range format

is merely for convenience and brevity and should not be construed as an inflexible limitation on the scope of the invention. Accordingly, the description of a range should be considered to have specifically disclosed all the possible subranges as well as individual numerical values within that range. For example, description of a range such as from 1 to 6 should be considered to have specifically disclosed subranges such as from 1 to 3, from 1 to 4, from 1 to 5, from 2 to 4, from 2 to 6, from 3 to 6 etc., as well as individual numbers within that range, for example, 1, 2, 3, 4, 5, and 6. This applies regardless of the breadth of the range.

Whenever a numerical range is indicated herein, it is meant to include any cited numeral (fractional or integral) within the indicated range. The phrases "ranging/ranges between" a first indicate number and a second indicate number and "ranging/ranges from" a first indicate number "to" a second indicate number are used herein interchangeably and are meant to include the first and second indicated numbers and all the fractional and integral numerals therebetween.

As used herein the term "method" refers to manners, means, techniques and procedures for accomplishing a given task including, but not limited to, those manners, means, techniques and procedures either known to, or readily developed from known manners, means, techniques and procedures by practitioners of the chemical, pharmacological, biological, biochemical and medical arts.

As used herein, the term "treating" includes abrogating, substantially inhibiting, slowing or reversing the progression of a condition, substantially ameliorating clinical or aesthetical symptoms of a condition or substantially preventing the appearance of clinical or aesthetical symptoms of a condition.

When reference is made to particular sequence listings, such reference is to be understood to also encompass sequences that substantially correspond to its complementary sequence as including minor sequence variations, resulting from, e.g., sequencing errors, cloning errors, or other alterations resulting in base substitution, base deletion or base addition, provided that the frequency of such variations is less than 1 in 50 nucleotides, alternatively, less than 1 in 100 nucleotides, alternatively, less than 1 in 200 nucleotides, alternatively, less than 1 in 500 nucleotides, alternatively, less than 1 in 1000 nucleotides, alternatively, less than 1 in 5,000 nucleotides, alternatively, less than 1 in 10,000 nucleotides.

It is appreciated that certain features of the invention, which are, for clarity, described in the context of separate embodiments, may also be provided in combination in a single embodiment. Conversely, various features of the invention, which are, for brevity, described in the context of a single embodiment, may also be provided separately or in any suitable subcombination or as suitable in any other described embodiment of the invention. Certain features described in the context of various embodiments are not to be considered essential features of those embodiments, unless the embodiment is inoperative without those elements.

EXAMPLES

Reference is now made to the following examples, which together with the above descriptions illustrate some embodiments of the invention in a non limiting fashion.

Generally, the nomenclature used herein and the laboratory procedures utilized in the present invention include molecular, biochemical, microbiological and recombinant DNA techniques. Such techniques are thoroughly explained in the literature. See, for example, "Molecular Cloning: A laboratory Manual" Sambrook et al., (1989); "Current Protocols in Molecular Biology" Volumes I-III Ausubel, R. M., ed. (1994); Ausubel et al., "Current Protocols in Molecular Biology", John Wiley and Sons, Baltimore, Maryland (1989); Perbal, "A Practical Guide to Molecular Cloning", John Wiley & Sons, New York (1988); Watson et al., "Recombinant DNA", Scientific American Books, New York; Birren et al. (eds) "Genome Analysis: A Laboratory Manual Series", Vols. 1-4, Cold Spring Harbor Laboratory Press, New York (1998); methodologies as set forth in U.S. Pat. Nos. 4,666,828; 4,683,202; 4,801,531; 5,192,659 and 5,272,057; "Cell Biology: A Laboratory Handbook", Volumes I-III Cellis, J. E., ed. (1994); "Current Protocols in Immunology" Volumes I-III Coligan J. E., ed. (1994); Stites et al. (eds), "Basic and Clinical Immunology" (8th Edition), Appleton & Lange, Norwalk, CT (1994); Mishell and Shiigi (eds), "Selected Methods in Cellular Immunology", W. H. Freeman and Co., New York (1980); available immunoassays are extensively described in the patent and scientific literature, see, for example, U.S. Pat. Nos. 3,791,932; 3,839,153; 3,850,752; 3,850,578; 3,853,987; 3,867,517; 3,879,262; 3,901,654; 3,935,074; 3,984,533; 3,996,345; 4,034,074; 4,098,876; 4,879,219; 5,011,771 and 5,281,521; "Oligonucleotide Synthesis" Gait, M. J., ed. (1984); "Nucleic

Acid Hybridization" Hames, B. D., and Higgins S. J., eds. (1985); "Transcription and Translation" Hames, B. D., and Higgins S. J., Eds. (1984); "Animal Cell Culture" Freshney, R. I., ed. (1986); "Immobilized Cells and Enzymes" IRL Press, (1986); "A Practical Guide to Molecular Cloning" Perbal, B., (1984) and "Methods in Enzymology" Vol. 1-317, Academic Press; "PCR Protocols: A Guide To Methods And Applications", Academic Press, San Diego, CA (1990); Marshak et al., "Strategies for Protein Purification and Characterization - A Laboratory Course Manual" CSHL Press (1996); all of which are incorporated by reference as if fully set forth herein. Other general references are provided throughout this document. The procedures therein are believed to be well known in the art and are provided for the convenience of the reader. All the information contained therein is incorporated herein by reference.

GENERAL MATERIALS AND EXPERIMENTAL METHODS

Culturing undifferentiated hESCs - The human ES cell (hESC) line H9 (WA-09, XX, Wicell)-derived BAC transgenic *HES5::eGFP* line (Placantonakis, D ET AL., 2008) was cultured on mitotically inactivated mouse embryonic fibroblasts (MEFs) (Globalstem). Undifferentiated hESCs were maintained in medium containing DMEM/F12, 20% KSR (knockout serum replacement), 1 mM Glutamine, 1% Penicillin/Streptomycin, non-essential amino acids, beta-mercaptoethanol and Fibroblast growth factor 2 (FGF2) (10 ng/ml). Medium was replaced daily and cells were passaged weekly by treating cells with Dispase (6 U/ml, Worthington) followed by mechanical trituration. Undifferentiated ES cells were purified with pluripotency markers Alexa 647-conjugated Tra-1-60 and PE-conjugated SSEA-3 (BD Pharmingen).

Neural induction and rosette formation and propagation - For neural induction and generation of NE cells, hESC colonies were removed from mouse embryonic fibroblasts (MEFs) by Dispase (6 U/ml, Worthington), dissociated with Accutase (Innovative Cell Technologies, Inc.), plated at sub confluent cell density [40-50x10^3 cells/cm^2, although twice higher density or alternatively small hESC clusters work well and accelerate confluence] on Matrigel™ (1:20, BD BIOSCIENCES) coated dishes, and supplemented with MEF-conditioned media and 10 μM ROCK inhibitor (Y-27632, Tocris) with daily fresh FGF2 (10 ng/ml, R&D SYSTEMS). Alternatively, neuroepithelial cells were generated either by monolayer induction – with dissociated

ES cells plated on Matrigel™ (BD biosciences), or by co-culture on MS5 stromal cells. Confluent cultures were subjected to dual SMAD inhibition neural differentiation protocol [Chambers, S. M. et al. Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. Nat Biotechnol 27, 275-280, 2009] containing Noggin (R&D, 250 ng/ml) and SB-431542 (10 μM, Tocris), and further supplemented with LDN-193189 (100 nM, Stemgent) (denoted LNSB protocol). HES5::eGFP usually appears on day 8 or 9. To generate E-RG rosettes and subsequent progenitors, NE cells were scrapped from plates on day 10-12, pre-incubated with Ca^{+2}/Mg^{+2} free HBSS (Hanks Balanced Salt Solution) followed by collagenase II (2.5 mg/ml), Collagenase IV (2.5 mg/ml) and DNAse (0.5 mg/ml) solution (all from Worthington) (37°C, 20 minutes). Cells were then dissociated and replated at high density (500,000 cells/cm^2) on moist Matrigel™ drops, and grown for additional days till rosettes appeared (E-RG stage). Neural induction and direct formation of E-RG stage rosettes could be also formed by co-culture of hESC clusters with MS5 stromal cells as previously described [Elkabetz, Y. et al. Human ES cell-derived neural rosettes reveal a functionally distinct early neural stem cell stage. Genes & Development 22, 152-165 (2008)]. Briefly, early appearing NE cells and Neural rosettes on MS5 were harvested mechanically beginning on day 8-10 of differentiation, replated on culture dishes pre- coated with 15 μg/ml polyornithine (Sigma), 1 μg/ml Laminin (BD Biosciences) and 1 μg/ml Fibronectin (BD Biosciences) (Po/Lam/FN) till Day 14, to obtain E-RG rosettes. Under both protocols, early appearing NE cells were cultured from Day 9 with N2 medium (composed of DMEM/F12 and N2 supplement containing Insulin, Apo-transferin, Sodium Selenite, Putrecine and Progesterone), and further supplemented with low SHH (30 ng/ml), FGF8 (100 ng/ml) and BDNF (5 ng/ml) (all from R&D Systems) to induce and maintain early anterior regionalization of the neural plate. Long-term culture of E-RG rosettes was performed by a weekly mechanical harvesting of rosettes and re-plating on Po/Lam/FN coated dishes with N2 medium, SHH (sonic hedgehog) and FGF8 (fibroblast growth factor 8), till Day 28. These were gradually replaced by FGF2 (20 ng/ml) and EGF (20 ng/ml) in the following two weeks of differentiation in order to maintain a proliferative (FGF and EGF responsive) NPC state on Day 28 (all cytokines from R&D Systems). At each stage cells were either replated as clusters for next passage or subjected to FACS purification by pre-

incubation with Ca^{+2}/Mg^{+2} free HBSS followed by mechanical dissociation. FACS for GFP was used to sort *HES5::GFP* cells of NE to LRG. Similarly, EGFR antibody was used to sort for LNP cells, all as a purpose to purify for the highest NPC state for each stage. NE cells were collected at day 12 of differentiation, ERG were collected at day 14, mid neurogenesis radial glial (ERG) cells were collected at day 35, late gliogenic radial glial (LRG) cells were collected at day 80, and long term NPCs (LNP) were collected at day 220. At each stage cells were either split for the next passage or subjected to FACS purification for *HES5::GFP* as described. All replating was performed on Po/Lam/FN coated dishes. For generating mature differentiated populations, HES5+ sorted NPCs were seeded at high density and subjected to mitogen withdrawal differentiation medium for 17 days which included N2 supplemented with Ascorbic Acid (AA)/BDNF (neuronal; NEdN, ERGdN, MRGdN) or 5% Fetal Bovine Serum (FBS) (Invitrogen) (glial) (LRGdA). "dA" refers to differentiated to glial cells such as astrocytes and oligodendrocytes.

Neural patterning and cortical laminar specification - Neural patterning was performed in parallel to or immediately following neural induction. For midbrain dopaminergic neuron differentiation, hESCs were neurally induced on Matrigel™ as previously described (Kriks, S. et al. Dopamine neurons derived from human ES cells efficiently engraft in animal models of Parkinson's disease. Nature 480, 547-551, 2011), and treated with SHH C25II (100 ng/ml, R&D), FGF8 (100 ng/ml) and CHIR99021 (3 μM, Stemgent). On Day 12, GFP+ and GFP- NE cells were separated by FACS, replated at very high density (400,000 cells/cm^2), followed by terminal differentiation with Neurobasal medium (Invitrogen) supplemented with BDNF (20 ng/ml), ascorbic acid (AA) (0.2 mM, Sigma), GDNF (20 ng/ml), TGFβ3 (1 ng/ml), dibutyryl cAMP (0.5 mM, Sigma), and DAPT (10 μM, Tocris) for 14 additional days. For motoneuron differentiation, hESC derived neurally induced cells either on MATRIGEL™ or MS5 were dissociated on day 12-14, and GFP+ and GFP- cells were separated by FACS and replated on Po/Lam/FN (MS5 protocol) or MATRIGEL™ drops (Matrigel™ protocol) at medium density (200,000 cells/cm^2) and treated with Retinoic Acid (RA, 1 μM, Sigma) and SHH C25II (125 ng/ml) till Day 28 as previously described (Elkabetz, Y. et al 2008). For early cortical neurons, NE cells on Day 12 were sorted for GFP+ and GFP- populations, replated and cultured with N2 supplemented with AA and BDNF.

For inhibition of Notch during terminal differentiation (Figure 3E), DAPT was added to the differentiation medium (5 μM) from day 2 of differentiation till the rest of differentiation period. For the inhibition of Notch at early neural induction (Figures 9D-F), DAPT was added at 5 μM in day 2 or day 6 and cells were harvested for analysis on day 9.

For neuronal, astroglial or oligodendroglial differentiation of late passages, E-RG rosettes were passaged through mechanical splitting till Day 80 or Day 220 with FGF2/EGF and BDNF. Either sorted GFP+ and GFP- populations (L-RG stage) or unsorted cells (LNP stage) were replated at high density and differentiated for 14 days in the presence of AA and BDNF for neuronal progeny, with 5% Fetal Bovine Serum (FBS) (Invitrogen) for astrocytic progeny, or with AA, BDNF, SHH C25II (100 ng/ml) and FGF8 for oligodendrocytic progeny according to the inventors' previous protocol (Lafaille, F. G. et al. Impaired intrinsic immunity to HSV-1 in human iPSC-derived TLR3-deficient CNS cells. Nature 491, 769-773, 2012).

Acute lineage analysis - Neuroectodermal progenitors reaching the NE stage (day 12) using the LNSB/Matrigel™ protocol were separated into HES5+ and HES5- populations, and these were replated, and either immediately fixed and analyzed for cell fate / proliferation markers by immunostaining, or maintained for another passage as separate populations till reaching the E-RG stage. Then, these NE derived HES5+ and HES5- cell populations were again separated to newly born HES5+ and HES5- cells, thus creating four distinct lineage related populations. These were either acutely analyzed or further maintained till the end of the passage and then again analyzed. All analyses were performed 2 hours after replating. For BrdU labeling, BrdU (30 μM) was added to cells for 1 hour at the second hour following replating, following which cells were subjected to fixation and analyzed.

Immunostaining and confocal imaging - Cells were fixed in 4% paraformaldehyde, 0.15% picric acid, permeabilized and blocked with PBS, 1% FBS and 0.3% Triton solution, and stained with indicated primary antibodies (see below) followed by AlexaFluor secondary antibodies (Invitrogen). Cells were imaged in phosphate buffered saline (PBS) after staining. All cell imaging was carried out in 24 well glass bottom plates (In Vitro scientific). Fluorescence images were obtained using a confocal microscope LSM710 (Carl Zeiss MicroImaging, Germany). The confocal

and time-lapse images were captured using a 10x and a 20× objectives (NA = 0.3, 0.8 respectively, Plan-Apochromat). Fluorescence emissions resulting from Ar 488 nm, 543 nm and 633 nm laser lines for EGFP, CY3 and CY5, respectively, were detected using filter sets supplied by the manufacturer. For DAPI detection the mode-locked Ti:Sapphire, fentosecond pulsed, multiphoton laser (Chameleon Ultra II, Coherent, Inc.) at a wavelength of 720 nm was used. For live imaging, cultured cells were maintained on the microscope stage in a temperature, CO_2, and humidity-controlled environmental chamber. Time-lapse green fluorescent protein (GFP) and matched Phase contrast images were acquired using Nikon Eclipse Ti-E microscope every 5 minutes for approximately 4 hours. Images and movies were generated and analyzed using the Zeiss ZEN 2011 software (Carl Zeiss, Inc.) and NIS elements (Nikon), respectively. All images were exported in TIF and their contrast and brightness were optimized in Adobe Photoshop under the same settings per each marker across all stages and as well as across HES5+ and HES5+ populations.

Neuronal output level quantification was performed by marker/DAPI ratio calculation and statistics of the entire cells in at least two (mostly three) independently taken images for a one representative experiment performed in parallel for all differentiation markers and across all stages. Note that counting is affected also by cells with positive but weak marker appearance, depending on stage examined and epitope tested (such as RELN). Additional quantitative aspects are shown in the quantitative PCR (qPCR) charts for all genes across HES5+ and HES5- and across all stages in Figures 11A-K.

Antibody list - Antibodies for CTIP2 (ab18465, 1:500), CUX1 (ab54583, 1:500), PE-conjugated anti EGFR (ab231, 1:50), LGALS1 (ab25138, 1:1000), Lin28 (ab46020, 1:1000), PHH3 (ab5176, 1:250), PLZF (ab104854, 1:100), POU3F2 (ab94977, 1:1000), SATB2 (ab51502, 1:50), SOX1 (1:1000), SOX2 (ab79351, 1:500), TBR1 (ab31940, 1:200), TBR2 (ab23345, 1:200) were from Abcam. Antibodies for BrdU (347580, 50 µl/test), KI67 (556003, 1:1000), phycoerythrin-conjugated SSEA-3 (560237, 20 µl/test), Alexa Fluor 647-conjugated TRA-1- 60-647 (560850, 5 µl/test), Alexa Fluor 647-conjugated TUJ1 (560340, 1:500) were from BD Biosciences. Antibodies for DCX (AB2253, 1:5000), O4 (MAB345, 1:25), RELN (MAB5364, 1:200), Tyrosine Hydroxylase (AB152, 1:500) were purchased from Millipore.

Antibodies for FABP7 (51010-1-AP, 1:100), S100B (15146-1-AP, 1:100) were from ProteinTech. Antibodies for AP2α (3B5 concentrated, 1:100) and PAX6 (supernatant, 1:16) were from DSHB. Antibody for NESTIN (MO15012, 1:500) was from Neuromics. Antibody for GFAP (Z0334, 1:2000) was from DAKO. Antibody for β-3-Tubulin (PRB-435P, 1:1000) was from Covance. Antibody for GLAST (ACSA-1) (130-095-822, 1:10) was from Miltenyi Biotec.

Quantitative PCR (qPCR) analysis - RNA was extracted using miRNeasy kit (Qiagene) followed by Maxima reverse transcription reaction kit (Fermentas). 1 ng (nanogram) of cDNA was subjected to qPCR using the primers depicted hereinbelow, ABsoluteTM QPCR SYBR® Green ROX Mix (ABgene) and ViiA-7 cycler (ABI). Threshold cycle values were determined in triplicates and presented as average compared to HPRT. Fold changes were calculated using the $2^{-\Delta CT}$ method. For reverse transcriptase (RT)-PCR data evaluation for Figures 3B-D, RT- PCR data was collected in triplicates, log2 transformed and normalized to HPRT. Mean normalized expression values were then normalized to 1 across all differentiation stages of HES5+ and HES5- for each gene separately to reflect the relative expression across stages. Finally, gene expression levels were normalized to 1 for each stage to also reflect the relative abundance of each gene in each stage. The resulting values were grouped into markers for deep layer neurons TBR/RELN and CTIP/FEZF2; and upper layer neurons CUX1/CUX2/SATB2 and displayed in pie charts and bars as shown in Figures 3B and 3C, respectively.

Primer set list (all for human genes) - BRN1 (POU3F3) Forward, 5'-TGGACTCAACAGCCACGAC -3' (SEQ ID NO:1) and Reverse 5'-CTTGAACTGCTTGGCGAAC-3' (SEQ ID NO:2); BRN2 (POU3F2) Forward, 5'-TGTATGGCAACGTGTTCTCG-3' (SEQ ID NO:3) and Reverse 5'-CCTCCTCCAACCACTTGTTC-3' (SEQ ID NO:4); CTIP2 Forward, 5'-TCCAGAGCAATCTCATCGTG-3' (SEQ ID NO:5) and Reverse 5'-GCATGTGCGTCTTCATGTG-3' (SEQ ID NO:6); CUX1 Forward, 5'-CAACAAGGAATTTGCTGAAGTG-3' (SEQ ID NO:7) and Reverse 5'-CTATGGTTTCGGCTTGGTTC-3' (SEQ ID NO:8); CUX2 Forward, 5'-GAGCTGAGCATCCTGAAAGC- 3' (SEQ ID NO:9) and Reverse 5'-AGGCCTCCTTTGCAATAAGC-3' (SEQ ID NO:10); EGFR Forward, 5'-

GATAGTCGCCCAAAGTTCCGT-3′ (SEQ ID NO:11) and Reverse 5′-CTGAATGACAAGGTAGCGCTG-3′ (SEQ ID NO:12); GFAP Forward, 5′-AGAGATCCGCACGCAGTATG-3′ (SEQ ID NO:13) and Reverse 5′-TCTGCAAACTTGGAGCGGTA-3′ (SEQ ID NO:14); HES5 Forward, 5′-ACCAGCCCAACTCCAAGCT-3′ (SEQ ID NO:15) and Reverse 5′-GGCTTTGCTGTGCTTCAGGTA-3′ (SEQ ID NO:16); HPRT Forward, 5′-TGACACTGGCAAAACAATGCA-3′ (SEQ ID NO:17) and Reverse 5′-GGTCCTTTTCACCAGCAAGCT-3′ (SEQ ID NO:18); PLZF Forward, 5′-CCTTTGTCTGTGATCAGTGCG-3′ (SEQ ID NO:19) and Reverse 5′-CAGTGCCAGTATGGGTCTGC-3′ (SEQ ID NO:20); RELN Forward, 5′-AATGCCGTCACCTTCTGTG-3′ (SEQ ID NO:21) and Reverse 5′-GGAGGACAGAAGCTGTTGTTG-3′ (SEQ ID NO:22); S100β Forward, 5′-GGAAATCAAAGAGCAGGAGGTT-3′ (SEQ ID NO:23) and Reverse 5′-TCCTGGAAGTCACATTCGCC-3′ (SEQ ID NO:24); SATB2 Forward, 5′-TAGCCAAAGAATGCCCTCTC-3′ (SEQ ID NO:25) and Reverse 5′-AAACTCCTGGCACTTGGTTG-3′ (SEQ ID NO:26); TBR1 Forward, 5′-GTCACCGCCTACCAGAACAC-3′ (SEQ ID NO:27) and Reverse 5′-ACAGCCGGTGTAGATCGTG-3′ (SEQ ID NO: 28); TBR2 Forward, 5′-AGCCGACAATAACATGCAGGG-3′ (SEQ ID NO:29) and Reverse 5′-TCCTGTCTCATCCAGTGGGA-3′ (SEQ ID NO:30); TH Forward, 5′-CCTCGGATGAGGAAATTGAG-3′ (SEQ ID NO:31) and Reverse 5′-TCTGCTTACACAGCCCGAAC -3′ (SEQ ID NO:32).

Fluorescent activated cell sorting (FACS) - Cell sorting was performed using ARIA flow cytometer (Beckton Dickinson). NE cells were dissociated with collagenase II (2.5 mg/ml), Collagenase IV (2.5 mg/ml) and DNAse (10 mg/ml) (all from Worthington) solution (37°C, 20 minutes). E-RG and subsequent stages were dissociated either with Accutase (37°C, 15 minutes) or Ca^{+2}/Mg^{+2} free HBSS (RT, 1 hour). All stages were FACS sorted to GFP+ and GFP- gated populations following exclusion of dead cells with DAPI. L-RG and LNP stages were also analyzed for EGFR abundance. Undifferentiated hESCs were sorted for the pluripotency markers Tra-1-60 and SSEA-4.

Microarray data processingand analysis - For all array hybridizations, GeneChip PrimeView Human Arrays were used and deposited in GEO. (GEO# TBD). Normalized log2 transformed probe level intensities were collapsed onto MGI gene symbols yielding 19,448 gene level measurements. Next, genes were filtered for a minimum log2 change of 1 or greater across between any pair of samples as well as a minimum log2 expression level of 3 or greater in at least one sample. The results yielded 6371 gene entries, which are listed in Supplementary Data 6, which is fully incorporated herein by reference in its entirety.

Stage Wise Clustering - To get a high-resolution view of the underlying dynamics and evaluate the distinct expression patterns, the present inventors performed clustering kmeans (k=100) (n=26) on the time series of the positive samples based on a set of 8 predefined patterns. The expression patterns were defined based on all possibilities of gene up-regulation between consecutive differentiation stages, e.g. up-regulated from hESCs to NE, but down in E-RG, up- regulated from hESCs to NE and not changing from NE to E-RG but down-regulated from E-RG to M-RG etc. Differential expression between two stages was defined as a minimum log2 expression change of 1 or greater. In total, the present inventors classified 496 genes, which are listed in Table 3, hereinbelow to follow one of these patterns. Subsequently, each cluster was subjected to gene set enrichment analysis. The results are shown in Figure 6A.

Notch active specific genes - Genes expressed in a stage specific fashion for HES5+ with respect to the HES5- populations and vice versa were determined by first clustering all time points using kmeans (k=100). Subsequently, hESC expression levels were subtracted from each stage. Next, HES5+ expression levels were divided by HES5- and vice versa. Next, the present inventors selected clusters that at the same time point (i) showed an average fold change exceeding 1.4 in one of the time points; and (ii) average fold changes that are lower than 1.2 in all remaining time points. Results are reported in Table 1 and 2, respectively and in Figure 4A.

Gene enrichment analysis - Stage specific obtained gene data sets were analyzed for enriched categories using Ingenuity Pathway Analysis (IPA) and selected resulting categories were plotted as heatmap (Figure 6B), or directly referred to within text.

Chromatin immunoprecipitation followed by sequencing (ChIP Seq) - For the histone ChIP experiments, the present inventors used similar approaches to Garber M. *et al.* 2012 (*Molecular cell* 47: 810-822). Specifically, around 160.000 cells were crosslinked in 1% formaldehyde for 10 minutes at 37°C, followed by quenching with 125 mM glycine for 5 minutes at 37°C, washed with PBS containing protease inhibitor (Roche, 04693159001) and flash frozen in liquid nitrogen. To lyse the cells, the present inventors used 1% SDS, 10 mM EDTA and 50 mM Tris-HCl pH 8.1 complemented with protease inhibitor. The chromatin was then fragmented with a Branson Sonifier (model S-450D) at 4°C, calibrated to a size range of 200 and 800 bp. Chromatin was mixed with antibody and incubated at 4°C overnight. Protein-A and Protein-G Dynabeads were added to chromatin/antibody mix (Invitrogen, 100-02D and100-07D, respectively) and incubated for 1-2 hours at 4°C. Samples were washed 6 times with RIPA buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA pH 8.0, 14 mM NaCl, 1% TritonX-100, 0.1% SDS, 0.1% DOC), twice with RIPA buffer containing 500 mM NaCl, twice with LiCl buffer (10 mM TE, 250 mM LiCl, 0.5% NP-40, 0.5% DOC), twice with TE (10 mM Tris-HCl pH 8.0, 1 mM EDTA), and then eluted in elution buffer (10 mM Tris-Cl pH 8.0, 5 mM EDTA, 300 mM NaCl, 0.1% SDS; pH 8.0) at 65°C. Eluate was treated with RNaseA (Roche, 11119915001) and Proteinase K (NEB, P8102S) overnight at 65°C.

For the OTX2 ChIP cells were collected and crosslinked in 1% formaldehyde for 15 minutes on ice, quenched with 125 mM glycine for 5 minutes at room temperature and pelleted. Nuclei were then isolated and chromatin was digested at 37°C with MNase enzyme until the majority of the DNA was between 50 and 800 bp. Specifically, 25 U and 35 U of MNase enzyme were used to digest NE cells and RNS/RG cells, respectively. The chromatin was then incubated with the antibodies over night at 4°C and co-immunoprecipitation of antibody-protein complexes was performed with Protein A or G beads for 1-2 hours at 4°C.

ChIP-Seq library preparation and sequencing - To extract DNA and create the Illumina libraries the present inventors used Solid-Phase Reversible Immobilization (SPRI) beads. The SPRI beads were added to the samples, mixed 15 times, incubated for 2 minutes at room temperature. Supernatant was extracted from the beads on a magnet (4 minutes). 70% ethanol was used to wash the beads and then dried for another 4

minutes. 40 µl EB buffer (10 mM Tris-HCl pH 8.0) was used to elute the DNA. The next steps of Illumina library construction include end-repair, addition of A-base, ligation of barcoded adaptors and PCR enrichment. To minimize the loss of ChIP material throughout this procedure, the present inventors used a general SPRI cleanup procedure after each reaction step reusing the same beads. PEG buffer (20% PEG and 2.5 M NaCl) was used to re- bind chIP material to SPRI following each reaction, and washing and extraction occurred as stated above. The enzymatic reactions were carried as follows: (1). DNA end-repair: Epicenter End-IT Repair kit incubated at room temperature for 45 minutes. (2). A-base addition: Klenow (3'->5' exonuclease; New England Biolabs) incubated at 37°C for 30 minutes. (3). Adaptor ligation: DNA ligase (New England Biolabs) and indexed oligo adaptors and incubated 25°C for 15 minutes, followed by 0.7X SPRI/reaction to remove non-ligated adaptors. (4). PCR enrichment: PCR mastermix (primer set, dNTP mix, Pfu Ultra Buffer (Agilent), Pfu Ultra-II Fusion (Agilent), water), for 20 cycles. The PCR amplified libraries were cleaned up using 0.7X SPRI/reaction (size selection mode) to remove excessive primers. Roughly 5 picomoles of DNA library was then applied to each lane of the flow cell and sequenced on Illumina HiSeq 2000 sequencers according to standard Illumina protocols.

For the OTX2 ChIP, DNA libraries were constructed using standard Illumina protocols for blunt-ending, polyA extension, and ligation. MyOne Silane beads (Life Technologies 37002D), were used to purify DNA fragments following each step of the library preparation. Adapter ligation was performed overnight at 16°C. Ligated DNA was then PCR amplified and gel size selected for fragments between 150 and 700 bp. Samples were sequenced using Illumina HiSeq at a target sequencing depth of 20 million uniquely aligned reads.

Strand Specific RNA-Sequencing Library Construction - RNA was extracted using the miRNeasy kit (Qiagen, 217004). Poly(A) RNA was isolated using Oligo d (T25) beads (NEB, E7490L). The Poly(A) fraction was then fragmented (Invitrogen, AM8740). Fragments smaller than 200 bps were eliminated (Zymo, R1016) and the remaining fraction was treated with FastAP Thermosensitive Alkaline Phosphatase (Thermo Scientific, EF0652) and T4 Polynucleotide Kinase (NEB, M0201L). RNA was then ligated to a RNA adaptor as reporter previously [Engreitz, J. M. *et al. Science* 341: 1237973, (2013)] using T4 RNA Ligase 1 (NEB, M0204L), which was then used to

facilitate cDNA synthesis using Affinity Script Multiple Temperature Reverse Transcriptase (Agilent, 600105). More specifically, the following adaptors reported in Engreitz, J. M. *et al. 2013* were used:

RNA sequencing - RiL-19 3' RNA adaptor:

/Phosphate/rArGrArUrCrGrGrArArGrArGrCrGrUrCrGrUrG/ddC (SEQ ID NO: 33);

RNA sequencing - AR17 RT primer: ACACGACGCTCTTCCGA (SEQ ID NO: 34);

RNA sequencing - 3Tr3 5' DNA adaptor:

/Phosphate/AGATCGGAAGAGCACACGTCTG/ddC (SEQ ID NO: 35);

RNA sequencing - PCR enrichment:

AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTC CGATCTCAAGCAGAAGACGGCATACGAGATNNNNNNNNGTGACTGGAGTTC AGACGTGTGCTCTTCCGATCT (SEQ ID NO: 36).

RNA was then degraded and the cDNA was ligated to a DNA adaptor using T4 RNA Ligase 1 as described in Engreitz, J. M. *et al. 2013*. Final library amplification was completed using NEB Next High Fidelity 2X PCT Master Mix (M054L). To clean up the final PCR and removed adapter dimers, two subsequent 1X and 8X SPRI reactions were completed to prepare the final library for sequencing.

Pooled shRNA (short hairpin RNA) screen – The present inventors selected 244 transcription factors and epigenetic modifiers that were differentially or continuously highly expressed during the *in vitro* differentiation time course in an otherwise unbiased fashion (Supplementary data 4, which is fully incorporated herein by reference in its entirety). In addition, the present inventors included GFP, RFP, lacZ and luciferase as internal controls. A sub-pool of the human 45K shRNA pool [Luo, B. *et al. Proc. Natl. Acad. Sci. USA* 105: 20380-20385, 2008] distributed by the Broad Institute Genomic Perturbations Platform and the RNAi Consortium (TRC) against these genes. For each gene, 5 distinct shRNAs were included as well as 5 scrambled and 3 empty control vectors, amounting to a total of 1230+8 shRNAs. The plasmid for shRNA expression under the control of the constitutive U6 shRNA promoter was the lentiviral vector pLKO.1. shRNA pool production and infection conditions were performed as previously described [Luo, B. *et al. 2008]*. Subsequently, the present inventors performed calibration experiments to determine to optimal combination of MOI (multiplicity of infection) and Puromyocin concentration to ensure efficient selection.

The present inventors identified MOI 0.4 and 1 μg/ml of Puromycin as optimal parameters for all stages. The, 26 million cells were infected at each stage of NE, ERG and MRG to ensure sufficient shRNA integration events to recover the complexity of the shRNA library. 24 hours post infection and prior to full expression but after integration of the lentivirus into the genome 3 million cells were collected to determine the baseline shRNA library representation. Subsequently, the cells were subjected to 5 days of Puromycin selection and then FACS sorted the resulting populations into HES5+ and HES5- compartments. Next, the representation of the shRNA library in each of the 9 populations was assessed by retrieving all shRNA integration events from genomic DNA isolated from each sample using PCR followed by next generation sequencing as previously described [Strezoska, Z. *et al. PLoS One* **7**, e42341, (2012)]. More specifically, two rounds of PCR were performed using the following primers for the primary PCR: Primary R:CTTTAGTTTGTATGTCTGTTGCTATTAT (SEQ ID NO: 37); Primary F: AATGGACTATCATATGCTTACCGTAAC (SEQ ID NO: 38). For the second, nested PCR the following primers were used:
Nested F: GGCTTTATATATCTTGTGGAAAGGA (SEQ ID NO: 39) Nested R: GGATGAATACTGCCATTTGTCTC (SEQ ID NO: 40). Next, standard Illumina sequencing library construction was performed as outlined above for 4 technical replicates for NE and MRG and 3 technical replicates for ERG, each comprising HES5+, HES5- and 24 hours control, amounting to a total of 33 libraries. These amplicon libraries were then sequenced on a HiSeq2500 with a PhiX spike in of 25%.

Individual shRNA validation for OTX2 and PAX6 - RNA was extracted using miRNeasy kit followed by Maxima reverse transcription reaction kit (Fermentas). 1 ng of cDNA was subjected to qPCR using designed primers and the ABsolute QPCR SYBR Green ROX Mix (ABgene) on a ViiA-7 cycler (ABI). Threshold cycle values were determined in triplicates and presented as average compared to HPRT. Fold changes were calculated using the 2--ΔCT method.

WGBS and RRBS library production - WGBS libraries were generated as previously described in Gifford, C. A. *et al.* 2013 (Cell 153: 1149-1163). RRBS was carried out using the multiplexed, gel free protocol described in Boyle, P. *et al. 2012* (Genome Biol 13: R92).

Data processing - For RNA-Seq data processing, reads were trimmed to 80, 60 or 30 bp depending on their per-base quality distribution in order to achieve maximum alignment rates. Reads were mapped to the human genome (hg19) using TopHat v2.0 (Trapnell, C., et al., 2009. *Bioinformatics* 25: 1105-1111) (tophat(dot) cbcb (dot)umd (dot) edu) employing the unfiltered gencode.v19.annotation.gtf annotation as the transcriptome reference. TopHat was run with default parameters except for the coverage search being turned off. Transcript expression was estimated with Cuffdiff 2 (Trapnell, C. et al. 2013). The workflow used to analyze the data is described in detail in Trapnell et al. (2012) (alternate protocol B).

WGBS libraries were aligned using BSMap 2.7 (Ref reference assembly. Subsequently, CpG methylation calls were made using custom software as previously described (Ziller, M. J. et al., 2013) excluding duplicate, low-quality reads as well as reads with more than 10% mismatches. Only CpGs with more than 5x coverage were considered for further analysis. ChIP-Seq data were aligned to the hg19/GRCh37 reference genome using MAQ35 version 0.7.1 with default parameter settings or Bowtie 2 version 2.05 (Langmead, B. & Salzberg, S. L. 2012). Reads were filtered for duplicates and extended by 200 bp at the end of the read. Visualization of read count data was performed by converting raw bam files to .tdf files using IGV tools (Thorvaldsdottir, H., et al., 2012) and normalizing to 1 million reads. Fragment length extended, duplicate and quality-filtered reads were used for subsequent analysis.

shRNA screen data analysis - For the screen data analysis, the protocol outlined by Dai et al. (Dai, Z. et al. 2014) was followed employing the R package limma (Smyth, G. K. 2005). First, the number of times each shRNA was observed in each library was extracted and counted using the shRNA sequence as barcode and the R function *processHairpinReads*. Next, the shRNA counts were normalized to the total number of reads observed harboring a shRNA to *counts per million* (cpm) and retained only those shRNAs with more than 0.5 cpm in more than 2 samples. After further QC showing excellent reproducibility (Figures 17A-D and 22F), differential shRNA count analysis was performed between the HES5+ and 24 hour control and the HES5+ and HES5- populations for each stage. To that end the dispersion for each condition was estimated and then a negative binomial generalized linear model using the R package *edger* was fitted. The likelihood ratio test was conducted for each contrast and only retain those

shRNAs as differentially enriched at a FDR≤0.05. To determine genes with significant positive or negative impact on HES5+ maintenance or cell survival, all genes that were targeted by at least two independent shRNAs which showed a significant effect (FDR≤0.05) in the same direction were determined. A mean effect score was computed in order to rank genes by computing the weighted mean of the log fold change between the two conditions weighted by the log cpm across all significant shRNAs and targeting a particular gene with an effect in the same direction. If an equal number of shRNAs showed a significant effect in positive or negative direction, the gene was classified as not significantly affected. Otherwise the effect direction was chosen based on the majority of the shRNAs. The results from the HES5+ to 24 hour control and HES5- comparison was combined into one by taking the maximum mean effect score observed in either comparison. The resulting mean effect scores are then used for visualization and analysis purposes in main text and figures and are reported in Supplementary data 3, which is fully incorporated herein by reference in its entirety. In addition, an empirical FDR was calculated by determining the fraction of shRNAs with a statistical significant effect based on the generalized linear model but were not expressed based on the RNA-Seq data for the condition where the significant effect was observed.

For the TERA validation analysis, all motifs were ranked according to their TERA scores at each stage. Next, motifs that were not associated with at least one TF that was covered in the screen design were filtered out. The fraction of top 20 motifs (by absolute TERA values) that were linked to TFs which showed a significant effect in the corresponding stage specific shRNA screen were determined. This number is reported as the percentage of motifs recovered. Only motif- knockdown results that have a straightforward interpretation were considered as hits. These include: 1. positive TERA score and positive depletion score (gene is involved HES5+ maintenance, progression or cell survival); 2. negative TERA score and negative depletion score (impedes HES5+ maintenance, progression or apoptosis); 3. negative TERA score and positive depletion score (gene is involved HES5+ maintenance, progression or cell survival but most likely acts as a repressor by causing H3K27ac or H3K4me3/1 loss). For the comparison with the expression based analysis, all significantly differentially expressed genes were ranked by their absolute fold change and determined the fraction of top 20 TFs observed among the differentially enriched shRNAs in the screen.

Differential expression analysis - Differential expression analysis was carried out using Cuffidff 2 (Trapnell, C. et al. 2013) and genes differentially expressed at a FDR $\leq$ 0.1 for each comparison and a minimal expression level of 1 FPKM in at least one of the conditions were considered. Clustering analysis was performed using the *csCluster* function in the *cummeRbund*40 package version 2.6.1 compbio (dot) mit (dot) edu/cummeRbund/) with the Jensen-Shannon distance as metric. The number of clusters for the NPC set (ESC, NE, ERG, MRG, LRG) and the differentiated populations (NEdN, ERGdN, MRGdN, LRGdA) was determined as the number of clusters between 10 and 20 with the minimum average silhouette width across all clusters. Subsequently, a pseudocount of 1 was added to all FPKM counts followed by a log2 transformation. The resulting values were used for all further expression analysis.

ChIP-Seq data analysis and normalization - For H3K27ac and H3K4me3 histone marks, the Irreproducible Discovery Rate (IDR) framework41 with a cutoff of 0.1 in combination with the MACS242 peak caller version 2.1 was used to identify peaks taking advantage of both replicates for each condition. For MACS2 peak calling, an initial p-value cutoff of 0.01 was used and the corresponding whole cell extract (WCE) control library was used as background. All IDR peak sets can be obtained from GEO under GSE62193.

For the broad histone marks H3K27me3 and H3K4me1, all 1 kb tiles of the human genome (hg19) that were significantly enriched over background in at least one of the replicates were determined. To that end a Poisson model43 with the WCE was used as background to model the fragment count distribution in each genomic. To that end a nominal *p*-value was defined for enrichment within a given region *i* in sample *k* harboring *rik* ChIP fragments compared to the WCE control sample *l* with *ril* ChIP fragments as $P(C \geq rik)$ where (Mikkelsen, T. S. *et al., 2010*):

Formula 1 (Figure 26A).

and $eil = ril / \lambda l$, λk = (region size) x (total number of ChIP fragments in sample k)/(corrected genome size), λl= (region size) x (total number of ChIP fragments in sample l)/(corrected genome size). In order to account for regions with no/minimal WCE read counts due to sampling, $eil = max(eil,1)$ were chosen. Resulting *p*-values were adjusted for multiple testing using the Benjamini- Hochberg (Benjamini, Y. & Hochberg, Y., 1995) correction and the q-value R package (Alan Dabney, J. D. S. a. w.

a. f. G. R. W. Q-value estimation for false discovery rate control. R package version 1.34.0.). Only regions significant at a q-value ≤ 0.05 and with an enrichment level over background ≥ 1.5 were considered to be enriched.

For differential enrichment analysis of histone marks between consecutive conditions, the R-package *diffBind* (Ross-Innes, C. S. et al. 2012) was used. To normalize read counts, the effective library size was used, counting only reads in peak regions (either the IDR peaks for H3K27ac, H3K4me3 or the enriched 1kb tiles for H3K27me3 or H3K4me1). The differential analysis was then conducted using the DBA_DESEQ2 method, taking full advantage of both replicates per condition with the *bTagwise* parameter set to true. Only regions differentially between consecutive conditions at a p-value of 0.05 were reported.

In addition, a union peak set for each mark was created separately by joining overlapping peaks/enriched regions in preparation for the transcription factor epigenetic remodeling activity (TERA) analysis. For H3K4me1, the enrichment over the union of all H3K27ac regions was computed since the focus was on well more sharply defined promoter and putative enhancer regions for this mark. For H3K27ac, the focus was on distal regions only (≥1kb of nearest TSS) since the present inventors were specifically interested in putative enhancer regions for this mark. For H3K4me3, the present inventors used the union of all H3K4me3 IDR based peaks regardless of distance, accounting for most promoters and CpG islands. Then the present inventors determined the enrichment level for all regions in the union set in each replicate across all marks separately. Region enrichment was computed as follows: First, the number of tag counts in each region was determined and normalized to reads per kilobase per million reads (RPKM) sequenced using the full library size of non-duplicate reads. Next, RPKM read counts were divided by the mean RPKM counts across all WCE libraries. Subsequently, the resulting enrichment levels were log2 transformed after adding a pseudo enrichment of 1. Finally, the resulting enrichment values were quantile normalized across the entire dataset for each mark separately. The resulting values were then average across replicates to obtain a region x condition normalized enrichment matrix. The resulting matrix was used as input for the TERA analysis. The present inventors tested several ChIP normalization strategies by assessing between replicate correlation and between condition discriminative power on a large dataset of 70 REMC (roadmap epigenomics

mapping consortium) H3K27ac samples and identified this strategy as best performing one.

Footprinting detection - To determine small regions depleted of histone modifications but surrounded by regions of much greater enrichment, termed footprints, the present inventors extended an approach used for the analysis of DNAse I HS data (Neph, S. *et al. 2012*). The footprints identification algorithm consisted of three main phases: In the first phase, the present inventors identify peaks using the IDR framework (see previous section) for H3K27ac and H3K4me3 and use these as baseline regions in which footprints could be detected. In the second phase, the present inventors identified footprints located within/around peak regions in the following manner:

1. For each peak, extend by 400 bp from apex in either direction;
2. Split entire resulting region into bins of size 20 bp;
3. Compute number of RPKM counts for a central sliding window across the entire region (shifting by increments of one bin) for different window sizes ranging from two bins to ten bins in increments of one;
4. For each position of the central window and for each window size, compute the following three quantities: *Cij* – RPKM count for central window at current position *i* and window size *j*, R*ij* – RPKM count for a 200 bp stretch directly to the right of the central window and *Lij* – RPKM count for a 200 bp stretch directly to the left of the central window;
5. For each resulting position *i* and window size *j* compute the depletion score: Formula 2 (Figure 26B)

With the footprint size normalization factor $f = s / b$, with s the size of the central window and b the size of the border regions;

6. Identify non-overlapping, non-adjacent footprint candidates starting from small to larger central window sizes and recording footprint candidate iff $eij > 0 \& e_{ij} < 1 \& L_{ij} > C_{ij} \& R_{ij} > C_{ij}$ followed by removing all other potential footprints (central window+borders) of larger size overlapping the current candidate;
7. Finally, all resulting candidate footprints with a footprinting score $eij \leq 0.9$ were reported.

The latter procedure was carried out for H3K27ac and H3K4me3 independently for each sample. Subsequently, the present inventors merged all footprints from individual

samples into consensus footprints set for each epigenetic mark separately, collapsing overlapping footprints by taking the union of all regions with non-zero overlap.

DMR detection - DMR detection was carried out as previously described with slight modifications (Gifford, C. A. et al. 2013). Pairwise comparisons of consecutive samples (hESC, NE, ERG, MRG, LRG, LNP) were carried out on a single CpG level using a beta-binomial model and the beta difference distribution requiring a maximum q-value below 0.05 and an absolute methylation difference greater than 0.1. q-values were computed based on beta-binomial model p-values using Benjamini-Hochberg 1995 method. Only CpGs covered by at least 5 reads in either sample were considered. Subsequently, differentially methylated CpGs within 500 bp were merged into discrete regions. Differential CpGs without neighbors were embedded into a 100 bp region surrounding each CpG. Next, differential methylation analysis was repeated on the region level using a random effects model. Only regions significant at q-value below 0.01, an absolute methylation difference above 0.2 and harboring at least 2 differentially methylated CpGs were considered differentially methylated and used for subsequent analysis. For the DNA methylation analysis in the context of the TERA framework, the present inventors restricted the analysis to DMRs consistently covered across all conditions, including those only assessed by RRBS. This left us with 7,929 regions.

Gene set enrichment analysis - Gene set enrichment analysis for genomic regions was carried out using the GREAT toolbox (McLean, C. Y. et al. 2010) and only categories with q-values ≤ 0.05 for both the hypergeometric and the binomial test as well as a minimal region enrichment level greater than 2 were considered, following the GREAT recommendations. Due to the large number of enriched gene sets, a selected subset of the results is shown in the different figures. In addition, the present inventors utilized the Allen Brain atlas (Thompson, C. L. *et al. 2014)* to determine enrichment for distinct brain structures and developmental time points. To that end the present inventors derived gene sets from the Brain atlas data in the following fashion:

The present inventors obtained *in situ* hybridization counts for the developing mouse brain at 7 distinct fetal time points and 11 different brain substructures through direct correspondence with alleninstitute.org. Specifically, the present inventors investigated the following structures and time points: Rostral secondary prosencephalone (RSP), Telencephalon (Tel), peduncular (caudal) hypothalamus (PHy),

Hypothalamus (p3), pre-thalamus (p2), pre-tectum (p1), midbrain (M), prepontine hindbrain (PPH), pontine hindbrain (PH), pontomedullary hindbrain (PMH), medullary hindbrain (MH); and embryonic (E)11.5, E13.5, E15.5, E18.5 as well postnatal P4, P14 and P28. In total, the present inventors had 14,585 measurements for 2,105 different genes across these different regions and time points. In order to define sets of genes characteristic for each combination of time point and structure, the present inventors computed the z-scores as well as the maximum observed variation for each gene across the entire matrix of structure and developmental time point combinations. Only genes that exhibited a maximum observed variation (maximum activity – minimum activity) $\geq$ 1 were considered for gene set definition. Next, all mouse genes were mapped to their human orthologs using the biomaRt database. Finally, gene sets for each region-time point combination was defined using genes that exhibited a z-score ≥ 2 in that particular combination. Since the Allen brain atlas gene sets are defined for each developmental time point and regional identity, the visualization was simplified by focusing either exclusively on structures or developmental time points. Therefore, the gene set with the maximum gene set activity was determined at each differentiation stage across all gene sets associated with distinct developmental time points for each structure separately. Similarly, the gene set with maximum activity for each developmental time point the present inventors determined taking the maximum across all structures at each stage. The gene set activity was determined as the mean log2 transformed expression level of all gene set members in for each condition.

Motif library construction and mapping to transcription factors - The present inventors combined the position weight matrices from Transfac professional database (Fogel, G. B. et al. 2005) with the PWM collection reported in Jolma et al. (Jolma, A. et al. 2013), only retaining motifs annotated for homo sapiens or mouse. To eliminate redundant motifs, the present inventors determined pairwise motif similarities for all resulting 1,886 PWMs using the TOMTOM (Gupta, S., et al., 2007) program which is part of the MEME (Bailey, T. L., et al., 2006) suite with default parameters. Next, a pseudo-distance matrix was compiled based on the resulting pairwise motif similarities. As a proxy for motif similarity, the present inventors used the log10 transformed TOMTOM q-value which was capped at 10. To convert the resulting motif similarities into a distance matrix, the present inventors inverted the scale by subtracting

the transformed q-values from 10. Then, the resulting matrix was used to perform hierarchical clustering with Euclidean distance and Ward's method. Finally, the present inventors employed the *cutree* function with a threshold of 7 to partition the resulting clustering dendrogram into discrete clusters of motifs. For each cluster, the present inventors then determined the motif with the highest complexity based on the relative entropy compared to a genome background model with the following base frequencies: A=0.2725, C=0.189, G= 0.189,T= 0.2728. Only motifs with a relative entropy greater or equal than 8 were retained for subsequent analysis. After identification of the candidate with the highest complexity for each motif cluster, the present inventors assigned all genes mapping to any motif in each corresponding cluster to the cluster representative motif. This lead to a final motif list of 557 motifs. In order to obtain a more quantitative association of each motif with its linked genes, the present inventors computed the ETFA (epigenetic transcription factor activity) scores across 70 REMC H3K27ac or H3K4me3 cell types and correlated the results with RNA-Seq expression data across 40 cell types. This analysis gave rise to a correlation matrix containing the pearson correlation coefficient of each motif with its linked genes. This matrix was used in combination with the plain gene mapping reported in primary motif sources. For Figure 16B, the present inventors uniquely map each motif to a corresponding linked gene by computing an association score as the product of the absolute pearson correlation coefficient and the average gene expression level of the corresponding gene. Then, the present inventors chose the gene with the highest association score. For motifs without an entry in the H3K27ac correlation matrix (due to the inability to determine suitable GEV parameters on the REMC dataset), the present inventors chose the gene with the highest gene expression level. In Figure 16B, only genes expressed with at least 10 FKPM in the respective condition are considered. Then, the present inventors report the top 35 genes for each condition, where TERA scores of motifs mapping the same gene were averaged.

In Figures 18A-E and 19A-B, the present inventors incorporate the results of the shRNA screen to uniquely map motifs apply the aforementioned mapping strategy only on the genes identified as hits. If it does not map to any gene hit by the screen, the present inventors use the standard assignment strategy outlined above.

Identification of putative transcription factor binding sites - In order to determine putative binding sites in a given genomic region, the present inventors used a biophysical model of transcription factor affinities to DNA (Manke, T., et al., 2008; Manke, T., et al., 2010) to determine putative binding to the footprint sets. This biophysical model requires the training of generalized extreme value (GEV) distributions of binding affinities based on a PWM matrix for each transcription factor and each set of genomic regions in order to generate a suitable background model. In order to take the distinct properties of footprints determined from different epigenetic marks, the present inventors determined the GEV parameters for footprints arising from H3K27ac, H3K4me3 and DNAme using the framework outlined by Manke et al. (Manke, T., et al., 2008; Manke, T., et al., 2010). The resulting three binding matrices were then filtered for minimal significant binding affinity at p-values below 0.05. All other entries with higher p-values were set to one. Next, the present inventors took the negative log10 of the entire matrix as a quantitative measure of binding affinity in subsequent analysis.

Inference of transcription factor activities based on epigenetic data - In order to infer transcription factor epigenetic remodeling activities (TERA), the present inventors first computed epigenetic transcription factor activities (ETFA) from the epigenetic data. To that end, the present inventors first focused on motif activity analysis and associated each motif in a second step with its corresponding transcription factor. For each epigenetic mark, the present inventors used the normalized epigenetic enrichment scores as well as DMRs with a minimal DNA methylation difference of at least 0.2 and covered consistently in all datasets. For the DNA methylation data, the present inventors inverted the scale to obtain de-methylation scores (1=fully de-methylated, 0=fully methylated) since usually the de-methylated states coincides with gene regulatory element activity. To determine the unobserved activity of a transcription factor binding motif, the present inventors took advantage of recent developments in the microarray field (Boulesteix, A. L. & Strimmer, K., 2005; Boulesteix, A. L. & Strimmer, K. 2007) and adapted this approach to epigenetic data. To that end the present inventors modeled the enrichment level *yit* of a particular epigenetic mark at genomic region *i* and time point *t* as a linear function the unknown transcription factor activities. Considering *p* predictor variables (epigenetic motif/transcription factor activities -ETFA) and *k* time

points the present inventors describe the unknown TFA X as a p x k matrix. Incorporating all regions n meeting the above listed criteria, the present inventors employ the linear model:

$$Y = A + BX + E$$

With the observed matrix of epigenetic enrichment scores Y (n x k), a constant offset matrix A (n x k), the connectivity matrix B (n x p), describing the filtered binding affinities for all transcription factor motifs to all regions and an error term matrix E. Subsequently, the present inventors followed the approach outlined by Boulesteix and Strimmer 2005 and applied partial least square (PLS) regression and specifically the SIMPLs algorithm (Dejong et al., 1993) to determine the unknown transcription factor motif activities. The idea in PLS is to employ a linear dimensionality reduction

$$T = BR$$

where the p predictors in X are mapped onto $c \leq rank(X) \leq min(p,n)$ latent components T (n x c matrix) and to compute the weight matrix R not only based on the data matrix B but explicitly taking into account the response matrix Y. The latter strategy maximizes predictive power even for a small number of latent components.

In order to determine the number of latent components for each epigenetic mark and genomic context, the present inventors performed cross validation by randomly partitioning the dataset 20 times into 2/3 training and 1/3 test set. Then the number of components was chosen such that it minimized the prediction error. The corresponding analysis methodology was implemented in the statistical programming language R adapting the implementation provided by Boulesteix and Strimmer (Boulesteix, A. L. & Strimmer, K. 2005). To assess the significance of the resulting ETFA scores, a permutation test was performed by randomly permuting the epigenetic enrichment scores for each gene regulatory element and recomputed the ETFA values on the permuted values. This process is repeated 100 times. Positive ETFA scores are considered to be insignificant and set to 0 if a greater ETFA score is observed more than once on the randomly permuted set and vice versa for negative ETFA scores.

Finally, the TERA scores were determined by computing the differential ETFA scores between consecutive conditions. These scores were determined by subtracting ETFA scores of consecutive time points from each other. Subsequently, the significance of this difference using a permutation test by randomly permuting the epigenetic

enrichment scores across all regions, re- computing the ETFA scores for each conditions and assessing the TERA score between consecutive conditions for each motif. Positive TERA scores are considered to be insignificant and set to 0 if a greater TERA score is observed more than once on the randomly permuted set and vice versa for negative TERA scores.

Co-binding analysis - Co-binding relationships were evaluated using an empirical approach with the entire set of footprints for each epigenetic mark as background. For a given factor i, the present inventors determined the footprints set Fi relevant for the current comparison (e.g. changing their epigenetic state in particular cell state transition) that were predicted to harbor a TFBS based on the binding model outlined above. Next, the present inventors computed the frequency of motif co-occurrence s^{F}_{ij} across *Fi* for all other motifs *j* in the database. To generate a proper null distribution, the present inventors randomly sampled $K = 100$ size standardized footprint sets Gk of cardinality |Fi| from the entire footprint collection for the epigenetic mark under study and computed the same test statistic s^{Gk}_{ij} on these sets. Finally, the present inventors determined an empirical *p*-value and odds ratio based on these quantities by counting the number of instances for which $s^{Gk}_{ij} \geq s^{F}_{ij}$:

Formula 3 (Figure 26C)

Only co-binding relationships significant at p-value ≤ 0.01 were retained.

Validation analysis on ENCODE data - To validate the outlined strategy *in silico* the present inventors took advantage of publically available transcription factor ChIP-Seq data in four cell lines from the ENCODE (Bernstein, B. E. et al. 2012) project as well as H3K27ac and RNA-Seq data for 70 cell types from the REMC project. The present inventors downloaded H3K27ac data as well as processed transcription factor binding data from the ENCODE project for the cell line K562 since abundant transcription factor binding data based on ChIP-Seq was available. In addition, this dataset has been successfully used in several studies to benchmark TF binding predictions (Sherwood, R. I. et al. 2014; Thurman, R. E. et al. 2012). The present inventors then applied the TERA-pipeline to the H3K27ac datasets and computed the TF-binding affinities for a set of 557 distinct motifs. With these datasets at hand, the present inventors computed the true positive rate (TPR), the false positive rate (FPR) and the positive predictive values (PPV) for all transcription factors that could be matched to

at least one motif with available binding affinities (46/117). In the event that one factor matched multiple motifs, the present inventors chose the motif with the highest AUC.

GWAS analysis - The GWAS analysis was conducted using 11,027 GWAS SNPs from the GWAS catalog. For each footprint set, the present inventors sampled K=100 randomly selected, H3K27ac footprints determined across 57 epigenome roadmap datasets processed in the same fashion as the neural dataset. Next, the present inventors determined the overlap with GWAS SNPs for control and neural H3K27ac footprint sets. Subsequently, the present inventors computed an empirical p-value for each trait/disease i in the catalog by determining the number of trait associated SNPs $sCij$ overlapping with each control region set Cj and the number overlapping with the corresponding footprint set si according to:

Formula 4 (Figure 26D)

Determination of core network - The core network was defined as those transcription factors that were differentially expressed during neural induction from ES cell to NE and not differentially expressed between consecutive stages of NE, ERG and MRG. The present inventors did not consider the LRG stage. Furthermore, the present inventors required that each factor was expressed at least 10 FPKM or more in NE, ERG and MRG and that it's mean normalized, maximum difference in expression levels between any of the stages did not exceed one standard deviation computed across the entire dataset of 9 cell types. In addition, the present inventors also considered genes that were not differentially expressed between any consecutive stages including the ESC stage but fulfilled all other criteria. This identification procedure gave rise to the candidate list of core factors. The present inventors then intersected this list with the results of the shRNA screen and retained only those factors that were significantly depleted in the HES5+ population relative to the respective HES5- or control population in at least two stages. Since the literature supported a role for PAX6 and OTX2 for which the shRNAs showed no effect due to the pooled setup or absent knockdown (**Figures 17A-D**), the present inventors included these genes as well. Finally, the present inventors merged this list will all TFs that were depleted in the shRNA screen at all 3 stages in the HES5+ population relative to the controls and were expressed at least at 10 FPKM or more in NE, ERG and MRG. This algorithm yielded a list of 22 transcription factors or epigenetic modifiers (**Figure 18A**). The present inventors then carried out co-

binding analysis in H3K27ac footprints dynamically regulated at each stage in order to obtain putative stage specific co-binding relationships. To determine significant co-binding events, the present inventors used the permutation procedure outlined above and retained all co-binding partners with an odds-ratio ≥ 1.5 that were significant at $p \leq 0.01$ that were also identified as a significant hit in the shRNA screen at the particular stage under investigation.

Transcription factor binding site priming analysis - To determine transcription factors associated with transcription factor binding site priming prior to factor activation, the present inventors determined all transcription factors at each stage that were significantly up-regulated at the consecutive NPC time point or induced in the corresponding more differentiated cell type (q-value≤ 0.1) and showed an increase in H3K4me1 or DNAme derived TERA activity at the current stage under investigation. In addition, the present inventors required that the corresponding motif did not map to any TF that was expressed more than 2 FPKM at the current stage under investigation. From this list, the present inventors picked the pro-neural genes *NEUROD4, ASCL2 and NFIX* for further investigation due to their literature support for their pro-neural functions. Finally, the present inventors required that the potential downstream target genes were significantly enriched for differentially regulated genes at the next NPC stage or in the corresponding more differentiated cell types. To that end, the present inventors determined all putative transcription factor binding sites for a particular factor in dynamically regulated H3K27ac or H3K4me1 footprints at the stage of potential priming. The present inventors then associated each of these putative binding sites with the nearest TSS and determined the number of differentially expressed genes for each factor. To assess significance, the present inventors randomly drew 100 sets of equally sized H3K27ac footprints with no motif of the factor under investigation and determined the number of differentially expressed genes for the subsequent stages. Only factors that exhibited more differentially expressed genes compared to the control sets in more than 99 % of the cases were retained.

Next, the present inventors performed co-binding analysis in H3K27ac peaks differentially regulated between the ES cell and NE stage as outlined above and display the top 10 co- binding relationships per factor with an odds-ratio ≥ 1.5 that were significant at an permutation test based $p \leq 0.01$ in Figure 19A.

EXAMPLE 1

NOTCH ACTIVATION LINKS MAJOR NEURAL LINEAGE TRANSITIONS

The present inventors used the previously established H9 (WA09) derived *HES5::eGFP* hESC reporter line [Placantonakis, D. et al. BAC transgenesis in human ES cells as a novel tool to define the human neural lineage. Stem Cells (2008)] to monitor morphology and HES5 reporter cell expression dynamics. The present inventors defined five consecutive stages during 220 days of neural differentiation and propagation (Figures 1A-B and 8A-B). Neuroectodermal cells emerged as early as day 5-8 and expressed SOX1 followed by PAX6, but not HES5 (Figure 8C). On day 12, HES5 is widely expressed and coincides PAX6 and SOX1, along with other progenitor markers such as SOX2 and NESTIN (Figures 1C-E, and 8D), possibly marking establishment of the CNS earliest NE cells following neural induction [Lowell, S., Benchoua, A., Heavey, B. & Smith, A. G. Notch promotes neural lineage entry by pluripotent embryonic stem cells. PLoS.Biol. 4, e121 (2006)]. Shortly after, on day 14, HES5 expressing cells rapidly become elongated, maintain PAX6 expression and form neural rosettes - highly polarized structures containing radially organized columnar cells [Elkabetz, Y. et al. Human ES cell-derived neural rosettes reveal a functionally distinct early neural stem cell stage. Genes & Development 22, 152-165 (2008)] - reminiscent of RG cells residing within the developing ventricular zone (VZ) [Gotz, M., Stoykova, A. & Gruss, P. Pax6 controls radial glia differentiation in the cerebral cortex. Neuron 21, 1031-1044 (1998); Gaiano, N., Nye, J. S. & Fishell, G. Radial glial identity is promoted by Notch1 signaling in the murine forebrain. Neuron 26, 395-404 (2000)] and as suggested by other *in vitro* studies [Eiraku, M. et al. Self-organized formation of polarized cortical tissues from ESCs and its active manipulation by extrinsic signals. Cell Stem Cell 3, 519-532, 2008; Gaspard, N. et al. An intrinsic mechanism of corticogenesis from embryonic stem cells. Nature 455, 351-357, 2008; Mariani, J. et al. Modeling human cortical development in vitro using induced pluripotent stem cells. Proc Natl Acad Sci U S A 109, 12770-12775, 2012]. Neural rosettes last till approximately day 35 (Figures 1B-E and 8A). Therefore, the present inventors designated day 14 rosettes as early radial glial (E-RG) cells and day 35 rosettes as mid radial glial (M-RG) cells. HES5 continues to be expressed in progenitors throughout the progression period, albeit in progressively decreasing numbers (Figure

8B). In contrast, SOX1, SOX2 and NESTIN remained highly expressed in the majority of cells throughout the entire propagation (Figures 1C-E). This indicates that the highly proliferative conditions are not sufficient to retain the initial high Notch activation level beyond the E-RG stage. More importantly, without being bound by any theory, this may reflect the transition of early NSCs into more limited progenitors, in line with *in vivo* findings [Huttner, W. B. & Kosodo, Y. Symmetric versus asymmetric cell division during neurogenesis in the developing vertebrate central nervous system. Current opinion in cell biology 17, 648-657, 2005; Temple, S. The development of neural stem cells. Nature 414, 112-117, 2001]. This observation was also accompanied by an apparent expression of DCX at the M-RG and L-RG stages, together with a gradual loss in rosette integrity (Figures 1C-E). Taken together, these findings suggest that extensive neurogenesis occurs mainly during M-RG through L-RG stages.

Based on these observations the present inventors defined two additional post rosette consecutive stages for analysis - day 80 and day 220. Neural progenitors on day 80 represent a later radial glial (L-RG) cell population exhibiting a more gliogenic bias, based on down regulation of rosette (R-NSC) markers such as PLZF and the up regulation of glial markers such as epidermal growth factor receptor (EGFR) and S100B (Figure 8E). These were still capable of generating neurons and glia, supporting existence of subsets of NSCs (Figures 3F-G) (Elkabetz, Y. et al. Human ES cell-derived neural rosettes reveal a functionally distinct early neural stem cell stage. Genes & Development 22, 152-165 (2008)). Neural progenitors could continuously propagate for many additional passages. Day 220 represents a long-term cultured neural progenitor (LNP) stage exhibiting a further substantial increase in EGFR and S100B levels (Figure 8E), while retaining multipotency (Figure 8F). These dynamic changes in Notch activation state along with morphological features during suggest that this long-term culture system provides a suitable paradigm to study NSC state and cell fate transition.

EXAMPLE 2

NOTCH ACTIVATION CONFERS CNS ROSTRO-CAUDAL PATTERNING ABILITY

To dissect the early cell fate potential of HES5+ progenitors compared to that of HES5- purified populations, the present inventors tested whether early Notch activation

is required for NE cells to respond to early developmental cues that yield regionally specified CNS neurons. The present inventors exposed neuroectodermal cells to patterning cues directing rostro-caudal regional fates prior to onset of *HES5::eGFP* expression. When neuroectodermal cells reached the NE stage (day 12), HES5+ and HES5- cells were separated, further subjected to complete differentiation along the selected regional paradigm, and were finally assessed for their ability to yield the corresponding regional specific neuronal subtypes (Figure 2A). Remarkably, early projection neurons expressing appropriate rostral to caudal regional neuronal markers such as TBR1 forebrain cortical neurons, FOXA2/TH midbrain dopaminergic neurons and HB9 spinal motoneurons, could be generated mainly from high HES5 expressing cells (Figures 2B-E). In contrast, HES5- progenitors weakly responded to patterning cues although they were capable of generating neurons (Figure 2B, bottom; Figures 9A-C). Requirement for Notch activation in the generation of early CNS neurons was also evident for additional early cortical neuronal markers such as CTIP2, NR2F1 and PCP423 (Figures 11A-K and 12A-G). Finally, the present inventors also confirmed requirement for Notch activation by inhibiting this signaling pathway using DAPT at either day 2 or day 6 of neural induction. Both HES5 and PAX6 expression levels were reduced following DAPT addition on these time points, while the neural crest / placodal marker SIX1 was upregulated (Figures 9D-F). These findings suggest that neuroectodermal cells require high Notch activation in order to acquire appropriate CNS neuronal cell identities. To further support this latter possibility, the present inventors followed HES5+ and HES5- progenitors derived from the NE stage through the E-RG stage and assessed their cell fate and proliferative capacities with respect to Notch activation. The present inventors found that consecutively sorted HES5+ populations retained PAX6 expression, while consecutively sorted HES5- cells retained AP2A expression, confirming that CNS and neural crest fates are dictated by Notch active and inactive states, respectively (Figures 10A-C). Furthermore, additional CNS markers such as SOX2 and OTX2 were enriched in HES5+ cells at the NE stage compared to HES5- cells, while the neuronal marker DCX was enriched in HES5- cells (Figure 10D). Finally, consecutively sorted HES5+ populations retained an overall stable level of BrdU incorporation, compared to consecutively sorted HES5- cells (Figure 10E). These results suggest that Notch in neuroectodermal cells is mainly important for

segregating CNS from non-CNS cell fates and in addition may confer CNS cells with a proliferative advantage.

EXAMPLE 3

NOTCH ACTIVATION ENABLES CORTICAL LAMINATION AND GLIAL FATES

Studies on cortical differentiation from PSCs have shown how continued culture of early rosettes yields sequentially appearing cortical neuronal layers by a default intrinsic mechanism [Eiraku, M. et al. 2008; Gaspard, N. et al. 2008; Mariani, J. et al. 2012]. Here, the present inventors asked whether HES5+ NE cells generated by such a default model serve as the primary progenitor cell source also for cortical lamination. The present inventors specifically asked whether prospective purification of Notch active progenitor cells throughout the progression *in vitro* correlates with potential to yield cortical neurons in a time and cortical layer dependent manner.

The present inventors found that early NE and E-RG progenitor stages gave rise mainly to neurons populating deep layers (Figure 3a) and in a Notch dependent manner (Figures 3B-D, 11A-K and 12A-G). These included deep layer-V FEZF2+ and CTIP2+ corticospinal neurons [Arlotta, P. et al. Neuronal subtype-specific genes that control corticospinal motor neuron development in vivo. Neuron 45, 207-221, 2005], early subplate and deep layer-VI TBR1+ corticothalamic neurons [Hevner, R. F. et al. Tbr1 regulates differentiation of the preplate and layer 6. Neuron 29, 353-366, 2001], early marginal zone RELN+ Cajal Retzius neurons [Zecevic, N. & Rakic, P. Development of layer I neurons in the primate cerebral cortex. J Neurosci 21, 5607-5619, 2001], and deep layer SATB2+ callosal neurons. While the latter are mainly known in their contribution to upper layers, they have been also shown to reside within deep layers to some extent [Britanova, O. et al. Satb2 is a postmitotic determinant for upper-layer neuron specification in the neocortex. Neuron 57, 378-392, 2008]. In contrast, later M-RG and L-RG stages gave rise mainly to neurons that populate superficial layers (Figure 3A) and in a Notch independent manner (Figures 3B-D, 11A-K and 12A-G). These included CUX1+ and CUX2+28 as well as SATB2+ layers II-IV callosal neurons. It is noted that while CUX1/2 protein levels were induced in neurons derived from M-RG progenitors and onwards, at the RNA level they were induced already at early stage

derived neurons, and this transcript expression depended on Notch activation (Figures 3B-D and 11A-K). This suggested that early CUX1/2 RNA expression reflected early progenitor potential, rather than immediate competence, to generate superficial layer neurons. This result is paralleled by an *in vivo* observation according to which progenitors prospectively labeled for *Cux2* appear already during early cortical development [Franco, S. J. et al. Fate-restricted neural progenitors in the mammalian cerebral cortex. Science 337, 746-749, 2012]. Taken together, these results show that early progenitor stages require Notch activation to generate early appearing neurons, while late progenitor stages yield later derived neurons regardless of Notch activation.

Without being bound by any theory, the present inventors have hypothesized that M-RG stage progenitors did not require Notch activation for generating later derived neurons because many of them correspond to HES5- subventricular zone (SVZ)-like cells that have already accumulated from earlier stages in a Notch dependent manner. To test this, the present inventors asked whether the generation of SVZ progenitors expressing TBR2 (EOMES) [Haubensak, W., Attardo, A., Denk, W. & Huttner, W. B. Neurons arise in the basal neuroepithelium of the early mammalian telencephalon: A major site of neurogenesis. Proceedings of the National Academy of Sciences of the United States of America 101, 3196-3201 (2004)] requires Notch activation. The present inventors found that TBR2 was upregulated during differentiation of NE cells in a Notch dependent manner and that this upregulation was prevented following Notch inhibition by DAPT (Figure 3E). This is in contrast to the later (M-RG) stage, where most TBR2 levels were derived from HES5- cells, and accordingly were not inhibited by DAPT. This shows that the majority of TBR2 progenitors that were apparent at the M-RG stage were already generated from early Notch active cells rather than *de novo* at the M-RG stage. For comparison, the present inventors also tested the expression of FEZF2 - a hallmark of earliest RG progenitors [Chen, B., Schaevitz, L. R. & McConnell, S. K. Fezl regulates the differentiation and axon targeting of layer 5 subcortical projection neurons in cerebral cortex. Proc Natl Acad Sci U S A 102, 17184-17189, 2005; Guo, C. et al. Fezf2 expression identifies a multipotent progenitor for neocortical projection neurons, astrocytes, and oligodendrocytes. Neuron 80, 1167-1174, 2013]. FEZF2 expression at early stages also strictly depended on Notch activation and was fully inhibited by DAPT (Figure 3E).

The inhibition of FEZF2 and TBR2 by DAPT demonstrates that generation of both early and late progenitors and their neurons is significantly affected in the absence of Notch activation.

Additional support for stage specific differential dependence on Notch is provided by TBR1+ and RELN+ neurons. These appear not only during early sub-late and marginal zone formation, respectively, but also during mid-gestation by later SVZ progenitors [Englund, C. et al. Pax6, Tbr2, and Tbr1 are expressed sequentially by radial glia, intermediate progenitor cells, and postmitotic neurons in developing neocortex. J Neurosci 25, 247-251, 2005], and may populate more caudal cortical regions [Takiguchi-Hayashi, K. et al. Generation of reelin-positive marginal zone cells from the caudomedial wall of telencephalic vesicles. Journal of Neuroscience 24, 2286-2295, 2004]. Accordingly, the present inventors found that TBR1 and RELN neurons could be both generated also at the M-RG stage and in a Notch- independent manner (Figures 3B-D, 11A-K and 12A-G).

Similar to the critical role of Notch activation during the derivation of early progenitor cells and their neuronal progeny, also the generation of astrocytes expressing GFAP at the L-RG stage required Notch activation (Figures 3F-G). This is in contrast to neurons at that stage, which could be derived also from HES5- cells (Figures 3B-D). Taken together, these results show that the distinct progenitor units spatiotemporally organized in the developing VZ and SVZ and which are responsible for cortical lamination and glial transformation *in vivo*, can be consecutively isolated from PSCs through sustained Notch signaling *in vitro*. While the main role of Notch activation is to promote the perpetuation of potent progenitors through culture, it may not be directly involved in generating cellular diversity, but rather maintain progenitors responsive to the culture conditions, which instruct these cell fate changes.

EXAMPLE 4

NOTCH ACTIVATION LINKS HALLMARKS OF CORTICAL DEVELOPMENT

The molecular hallmarks that specifically define each of the developmental stages *in vitro* with respect to Notch activation were studied by employing a global gene expression analysis (Supplementary data 6, which is fully incorporated herein by reference in its entirety), and specifically investigating transcripts differentially

expressed in HES5+ compared to HES5- progenitors at each stage (Figure 4A and Table 1 hereinbelow). Interestingly, genes upregulated in HES5+ NE cells were mainly associated with cell cycle progression and DNA replication, and included *CDC6*, *CDK1*, *CENPH* and *TOP2A* (Figure 4A). Genes specifically enriched in HES5- cells at the NE stage included the proneural genes *NEUROD4*, *NGN1/2*, *TBR2* and *DCX* (Table 2, hereinbelow). These results are compatible with NE HES5+ acting as symmetrically dividing NSCs during early nervous system development (Huttner, W. B. & Kosodo, Y. Current opinion in cell biology 17, 648-657, 2005) and further suggest that Notch confers amenability for regional neuronal specification through promoting cell cycle.

In contrast to NE cells, HES5+ at the E-RG stage (day 14) were enriched for genes such as *ARX, FEZF2 and NR2E1* with respect to HES5- cells (Figure 4A), indicating that Notch active NE cells underwent a sharp and rapid transition towards an RG cell stage with a strong dorsocaudal telencephalic character. Notch active progenitors in the more advanced rosette M-RG stage continued to highlight cerebral developmental genes such as *POU3F2* as well as genes associated with neuroblast cell division such as *ASPM* (Figure 4A), fitting with extensive neurogenesis during the M-RG stage. Transcripts overrepresented in HES5+ cells in the L-RG stage were associated with glial fate key genes such as *OLIG1* and *PDGFRA* (Figure 4A). Finally, genes overrepresented in HES5+ vs. HES5- progenitors at the LNP stage such as *ANXA2* and *LGALS1* are associated with ependymal cells and aNSCs6 (Figure 4A), suggesting that LNP cells progressed beyond RG fates towards adult like progenitor identity. These results show that PSC-derived neural progenitors progress through distinct stages that may be possibly linked via Notch activation, from NE cell proliferation through neurogenic RG cell differentiation, glial transformation and adult NSC specification. To further support these observations the relative distribution of germinal zone genes among the various stages with respect to Notch activation were studied. The heatmap in Figures 4B-C and the histograms in Figures 13C-J show a consecutive correlation of NE/E-RG, M-RG, L-RG and LNP stages *in vitro*, with VZ, early SVZ, late SVZ and subependymal zone (SEZ) *in vivo*, respectively. Furthermore, it is evident that VZ markers are enriched in Notch active cells while SVZ markers are more comparably distributed. To summarize, the molecular data presented herein further confirm that Notch activation links the establishment of NE cells and their

transition through consecutive primary RG progenitors.

EXAMPLE 5

EARLY AND MID ROSETTES DEMARCATE GERMINAL ZONE EQUIVALENTS

The present inventors next employed immunostainings and 3D construction analyses to dissect the hierarchical progression of progenitors at the cellular and cyto architectural levels with respect to Notch signaling. The abundant occupancy of PAX6 and HES5 at all rosette cells at the E-RG stage (Figures 4D-E) indeed fits the dorsal cortex molecular identity of E-RG cells (Figure 4A). In contrast, PAX6 and HES5 spatial distribution in the M-RG stage was mainly confined to lumens, as well as to regions located distally to rosette areas (Figures 4D-E). These two PAX6 and HES5 expressing progenitor cell types possibly corresponded to VZ residing apical RG progenitors and putative outer SVZ (OSVZ) localized basal RG progenitors, respectively [Hansen, D. V., et al., Nature 464, 554-561, 2010; Fietz, S. A. et al. Nat Neurosci 13, 690-699, 2010]. 3D construction analysis of E-RG and M-RG rosettes demonstrates that HES5+ cells are composed of elongated radial fibers that cross the entire Z-section in apical to basal manner, attesting for a complex rather than flat orientation obtained via non-confocal images. E-RG rosettes are packed with HES5+/PAX6+ cells across all rosette area while dividing nuclei located luminally, at multiple Z-levels (data not shown). M-RG rosettes are characterized by HES5+/PAX6+ cells and dividing nuclei both confined to luminal regions only, at multiple Z-levels, in addition to neuronal processes accumulating at lower Z-levels (data not shown). The cell division at luminal sites is also reflected by the expression pattern of the M-phase marker PHH3, which is confined to nuclei within lumens at E-RG and M-RG rosettes, while the general cell cycle marker KI67 was apparent among all progenitors regardless of HES5 expression (Figure 13A).

Further evidence ascribing the M-RG stage rosettes as a site of mid neurogenesis is also provided by the many TBR2+ INPs (Intermediate Neural Progenitors; which is equal to "IPC", i.e., Intermediate Progenitor Cells) that appeared transiently and specifically at this stage, and were located at rosette peripheries, assigning these regions in M-RG rosettes as mid neurogenesis SVZ-like areas (Figures 4D-E and 13B). This

was further corroborated by the expression of CUX1/2 and POU3F2/3. These neuronal markers begin to be expressed in VZ/SVZ progenitors during mid-gestation, and accordingly appeared at the M-RG stage, located at rosette peripheries (Figures 4D-E and 13B).

EXAMPLE 6

ROSETTE DISASSEMBLY MARKS THE BEGINNING OF GLIOGENESIS

Evidence suggesting that M-RG rosettes serve not only as a site of extensive neurogenesis, but also of transition to glial stages is provided by the expression pattern of the RG markers GLAST and FABP7. These became evident already in E-RG rosettes, coinciding with PAX6 and HES5 (Figures 5A-B), marking neurogenic RG. At the M-RG stage these markers appeared at luminal regions together with HES5 and PAX6, but were also located at rosettes peripheries where they did not co-express PAX6 and HES5. This fits the findings that CNS progenitors prospectively tagged for GLAST and FABP7 at early stages *in vivo* were found labeling most neuronal progeny, while if prospectively marked at mid-neurogenesis, they labeled glial fates (Anthony, T. E. & Heintz, N. Neural development 3: 30, 2008).

The L-RG and LNP stages were no longer capable of forming rosettes, reflecting loss of epithelial integrity due to accumulation of basal progenitors, neurons and cells with astroglial character. HES5 and PAX6 cells further decreased in numbers (Figures 1C-E and 8B), reflective of the reduction in neurogenic NSCs. Some CUX1/2 and POU3F2 progenitors still remained at the L-RG stage, marking residual neurogenesis (Figures 4D-E). Enhanced astroglial identity is supported by the further increase in GLAST and FABP7 levels (Figures 5A-B) as well as the glial markers S100B and EGFR (Figure 8E). The increase in EGFR transcript levels was also reflected by an increase in EGFR+ cells in 10% of L-RG cells, as judged by FACS analysis (Figure 5C). EGFR labeling possibly reflected a newly established subset of progenitors at the L-RG stage, compatible with EGFR labeling mainly late SVZ progenitors *in vivo* [Burrows, R. C., et al., Neuron 19, 251-267, 1997].

In contrast to L-RG cells, most LNP cells expressed EGFR as shown by FACS analysis (Figure 5C), suggesting that following long-term culture, progenitors correspond to EGFR+ transit amplifying cells. Such can be derived from aNSC

astrocytes following activation by EGF in culture [Doetsch, F., et al. Neuron 36: 1021-1034, 2002]. This observation may explain the low GFAP levels found in the cultures throughout the progression period (Figure 5D). Interestingly, many HES5 expressing cells at the LNP stage co-localized with S100B (Figure 8E), indicative of ependymal cells [Carlen, M. et al. Nat Neurosci 12: 259-267, 2009], and in line with enrichment of genes harboring ependymal character in HES5+ LNP cells (Figure 4A).

EXAMPLE 7

MOLECULAR CHARACTERIZATION OF NEURAL CELL FATE TRANSITION

The findings described herein suggest that HES5 expression during progenitor progression links the sequential transition through distinct competences. Such a mechanism can underlie the generation of heterogeneity in culture due to the fact that many HES5- cells exist throughout culture. Accordingly, factors that share expression among HES5+ cell stages may serve as transcriptional regulators for neural development, while stage specifically expressed factors may be co-opted to drive the transition through distinct competences. In order to identify such potential candidate genes playing role in inducing, maintaining, or transitioning between distinct competences, the present inventors employed an unbiased clustering analysis on all differentially expressed genes across the HES5+ and HES5- populations. This analysis yielded 26 gene clusters that were divided to 7 distinct gene expression patterns (Figure 6A and Table 3, hereinbelow). Gene clusters upregulated early from ES to NE cells and sustained among all stages are expected to play role in inducing neural fates and maintain anterior cell character throughout prolonged periods. Accordingly, this early upregulated cluster is enriched for central nervous system [p value (pV) =8E-13] right-tailed Fisher Exact Test used by IPA) and forebrain development (pV=1.0E-8) as well as neuronal cell movement (pV=1.0E-10) GO categories (Figure 6B), and contained factors such as *FOXG1*, *PAX6*, *ZIC1* and *SP8*, which have been well implicated either in neural induction, forebrain specification and cortical area patterning. Gene clusters up-regulated at the M-RG stage and sustained throughout subsequent stages are anticipated to play role in active neurogenesis but also in the initiation of gliogenic bias, in correlation with the present findings (Figures 4A-E and 5A-D). As such, this cluster was enriched for genes involved in morphology of nervous system (pV=7.9E-7) and

formation of plasma membrane projections (pV=2.2E-6), both implicated in neuronal axon maturation. These categories included for example *NFIA* and *NFIB*, which are interestingly involved in both repressing neuronal progenitor state through Notch signaling concomitantly with activating glial fates (Piper, M. et al. 2010). Other factors included are *SLITRK*, *ASCL1* and *PREX1*, which are associated with neurite outgrowth as well as neuronal maturation and migration. In addition, the present inventors have interestingly found that *EZH2* - the histone methyltransferase of PRC2, is transiently expressed through the M-RG stage (Figures 14A-E, in particular Figure 14C). This latter observation nicely correlates the finding that Ezh2 regulates the balance between self-renewal and differentiation in the mouse cerebral cortex, as its loss leads to aberrant timing of cortical development [Pereira, J. D. et al. Proc. Natl. Acad. Sci. USA 107:15957-15962, 2010]. One particularly interesting cluster is characterized by genes exhibiting a transient expression during the NE stage (specifically in HES5- cells) followed by a transient re-expression during the M-RG stage. This cluster included *TBR2*, *RSPO2*, *NEUROD1* and *TFAP2B* - genes associated with neurogenesis and basal progenitor (INP) cell fate. This observation may intriguingly imply that the establishment of a set of transcription factors (TFs) regulating SVZ generation and cortical expansion may already originate and act during early corticogenesis. Genes up-regulated at the L-RG and LNP stages are highly enriched for genes involved in neurotransmission (pV=2.4E-7) and include *GABBR2*, *GRIA4* and *GRM3*, but are also enriched for genes strongly implicated in glial fates such as *OLIG1* and *OLIG2* (pV=3.4E-7), again manifesting how late neuronal maturation events coincide with extensive gliogenesis. Another gene expressed at the LNP stage is *LGALS1* (Figures 14A-G). Interestingly, *Lgals1* was shown to be specifically enriched in prospectively isolated GFAP+/Prominin1+ aNSCs as well as ependymal cells [Beckervordersandforth, R. et al. Cell Stem Cell 7: 744-758, 2010]. Gene clusters upregulated at the NE towards E-RG stage were enriched for nervous system morphogenesis (pV=1.2E-7) and cancer associated factors (pV=6.9E-8) and included genes such as NR2E1 and LGR5. NR2E1 is mainly expressed in Notch active cells at the E-RG stage (Figures 14A-E), compatible with *Nr2e1* role in controlling proliferation of VZ progenitors during the establishment and expansion of the SVZ (Roy, K. et al. J Neurosci 24, 8333-8345, 2004). Interestingly, NR2E1 was also moderately expressed at

the later LNP stage (Figures 14A-E, in particular Figure 14B), in correlation with its expression in mouse aNSC astrocytes as well as its role also in brain tumor initiation from NSCs45. *LGR5* - another interesting E-RG specific gene (Figures 14A-E, in particular Figure 14A) - is a major stem cell regulator of adult tissue regeneration and malignancy, and was initially identified in the stem cells of the small intestine and colon (Barker, N. et al. Nature 449: 1003-1007, 2007). The present inventors also identified a cluster of genes expressed in ES cells but also transiently in NE and E-RG stages. One such candidate is *LIN28A* (Figures 14F-G). Interestingly, this RNA binding protein is known also to play a role in reprogramming to pluripotency [Yu, J. et al. Science 318: 1917-1920, 2007], suggesting additional roles for this protein during early neural development. Accompanying in this cluster is *HMGA2* (Figure 6A) - a young (but not adult) NSC marker [Nishino, J., et al., Cell 135: 227-239, 2008] and a target for Let-7, a microRNA whose maturation and function is repressed by *LIN28A*. Finally, the present inventors also identified F11R (Figure 14F). This tight junction protein shown to be involved in platelet adhesion to the activated endothelium [Kornecki, E., et al. J Biol Chem 265: 10042-10048, 1990], but was also suggested to act in the cell-to-cell adhesion of neuroepithelial cells [Famulski, J. K. & Solecki, D. J. Trends in neurosciences 36: 163-173, 2013]. Taken together, these transcriptional trends suggest that the dynamic changes occurring in progenitor cell potency during culture are linked via Notch activation through stage specifically acting factors.

Analysis and discussion

The results presented in Examples 1-7 offer a first in depth dissection of the dynamic changes that lead to heterogeneity in PSC derived neuroepithelial cells during long-term culture, and shows that they match developmental logics and timing principles of mammalian NSC ontogeny. Moreover, this study suggests that Notch activation is a critical component orchestrating this ontogeny *in vitro*, by establishing the identity of neuroepithelial cells, regulating their numbers during progression, and linking their transition through distinct developmentally specific primary progenitor cells – which together comprise the diversity of NSC types promoting neurogenesis and gliogenesis of the CNS.

The consecutive prospective isolation of Notch active progenitors along the

entire differentiation period *in vitro* enabled the present inventors to enrich cultures for primary progenitor cells that may hold proliferative advantage and broad developmental potential, while eliminating those lacking these features. This allowed the generation of distinct progenitor modules *in vitro* temporally linked via Notch activation to serve as building blocks of nervous system establishment and neocortical construction (**Figure 7**). It is conceivable that each of the distinct HES5+ populations exhibits improved homogeneity with respect to Notch activation. This allows a more meaningful evaluation of the functional, cellular and molecular properties of distinct progenitor cell types during normal and abnormal development. The combined functional analysis and gene profiling of the isolated cell types during stage transitions provide a highly valuable resource of stably expressed as well as stage specifically expressed transcriptional regulators, which may be critical for both launching the onset of early NSCs as well as driving their progression through distinct developmental potencies, through Notch activation.

The present study provides a more accurate identification of neuroepithelial cells and their properties with respect to Notch activation. The findings described herein emphasize the ability of enhanced Notch activation to ensure the maintenance of progenitors in a state that allows them to respond to developmental cues. Importantly, high Notch activation does not prevent the progression through distinct fate competences, but rather links the progression through distinct lineages in culture, thus ensuring the execution of the full developmental potential of NE cells. Mechanistically, Notch activation first dictates CNS identity during neural induction. Second, it represses proneural transcriptional activity in NE cells and by that maintains a highly undifferentiated state. Third, Notch active NE cells display augmented expression of cell cycle components, in correlation with maintenance of BrdU incorporation in later derived HES5+ cells. Without being bound by any theory, the present inventors propose that Notch activation may confer amenability to specification cues mainly by extending the time window during which NE progenitors are exposed to these cues. This model can explain the ability of HES5+ but not HES5- progenitors to undergo complete neuronal specification for various distinct regional identities. This model is supported by *in vivo* studies showing the requirement for successive cell cycles during the specification of both spinal motoneurons and cortical neurons [Ericson, J., et al., Cell

87: 661-673, 1996; Rodriguez, M., et al., Development 139: 3870-3879, 2012]. Several intriguing aspects on the molecular forces that drive NSC progression can be drawn from the present study. The findings that genes such as *SOX2*, *FOXG1*, *OTX2* and *PAX6* are expressed throughout the culture progression support a model according to which CNS identity is determined during early stages by a core of stably expressed transcriptions factors (TFs). Nonetheless, the significantly differentially expressed genes sets among stages indicate that stage specifically expressed genes are also critical for stage transition. Without being bound by any theory, the present inventors propose that the extensive remodeling capacity of NE cells through progression is provided by stably expressed TFs co-acting with consecutively and transiently appearing factors to control NSC progression through Notch activation. It is intriguing to speculate that distinct sets of Notch regulators are consecutively appearing and replacing one another in a relay mechanism to generate potency diversity, while maintaining proliferation capacity through Notch signaling. Such a model should further advance the ability to use these factors to directly induce or maintain specific modules *in vitro* - towards establishing perpetuating NSC types amenable for drug screening, disease modeling and for developing better protocols for deriving specific neuronal and glial lineages.

The progenitor module dissection approach of the present study enables new possibilities of gaining knowledge on progenitor cell dynamics during disease onset and progression. Many disease models, particularly iPS cell based, rely on the ability to generate specific neuronal types suspected to be clinically and physiologically relevant.

This approach offers a unique possibility to specifically isolate damaged or malfunctioning progenitor modules that give rise to the clinically affected neuronal or glial cell types, and to gain deep insights into pathogenic features within such defected modules such as stem cell properties, developmental potential and molecular drivers. Also, the comprehensive array data sets may help to link the expression pattern of disease causing mutated genes along the developmental stage modules with relation to Notch activation. Using the cellular system described herein for deciphering 'defective units' during pathogenesis of various nervous system diseases *in vitro* should be a great advancement to the field of disease modeling. Lissencephaly, a developmental cortical disorder, is associated with defects in 'core' genes such as *ARX*, stage specific genes such as *DCX*, and Notch active specific genes such as *VLDLR*. Similarly, Microcephaly

is associated with defects in 'core' genes such as *MCPH1* and *STIL*, stage specific genes such as *CENP*, and Notch active specific genes such as *ASPM*. The datasets presented herein may provide insights also to other nervous system disorders such as autism as well as psychiatric disorders. Altered regulation of *DISC1* associated with schizophrenia may be interesting due the fact that expression of this gene appears in culture only starting the M-RG stage. Neurodegenerative diseases associated with mutations or SNPs in genes differentially expressed in the system described herein may also shed light on the potential role of such candidates in predisposition and / or actual elderly onset. For example, the present inventors found that *SPON1* and *RRAS2*, overrepresented specifically in L-RG HES5+ cells and thus may relate to gliogenesis, contain SNPs associated with Alzheimer's disease (Pv=2.07E-4, Odds=15.26). Such findings may imply that the potential embryonic roles of these factors may be inferred also to the malfunction of such SNP-bearing genes during disease onset.

The system described herein also offers a unique possibility to look into the origin and tumorigenic properties of distinct and yet to be defined brain cancer stem cells. As many of the developmental genes have tumorigenic potential, this study may potentially advance the understanding of how Notch activation is associated with the emergence of distinct brain cancer stem cells. The association of these data sets with brain growth and tumorigenesis also reinvigorates the development of strategies to minimize heterogeneity of progenitors beyond the findings on cortical development. Such studies should also help to develop approaches to control the balance between proliferation and differentiation *in vitro*, to eliminate proliferating progenitors from their differentiated progeny, and to minimize chances of tumorigenicity – towards future implications in preclinical setups.

In addition, it will be interesting to test whether the newly described naïve PSCs [Gafni, O. et al. Derivation of novel human ground state naive pluripotent stem cells. Nature, doi:10.1038/nature12745 (2013)] can be used to generate NE cells and their progeny with employing the differentiation paradigm described herein, and whether such approach can be helpful to improve harnessing the full neurogenic and gliogenic potential of these cells.

Because the uniqueness of datasets and cellular analysis is in their proliferative nature, the present inventors envisage that the comprehensive data analyses presented

herein would serve as a powerful tool to dissect lineage transitions, to identify origin of progenitor cells, to relate them to onset and progression of brain tumors, and to address fundamental questions related to human cortical expansion.

EXAMPLE 8

EPIGENETIC CHANGES THROUGH BRAIN COMPARTMENTS AND DEVELOPMENTAL STAGES

The present inventors utilized the human ES cell line WA9 (or H9) expressing GFP under the HES5 promoter (Placantonakis, D. G. et al. 2009) to isolate defined neural progenitor populations of neuroepithelial (NE), early radial glial (ERG), mid radial glial (MRG) and late radial glial (LRG) cells based on their Notch activation state (Examples 1-7 hereinabove and Edri, R. et al. 2015), as well as long term neural progenitors (LNP) based on their EGFR expression (Figure 15A and Figure 20A). The present inventors took these defined stages to create strand-specific RNA-Seq data, chromatin immunoprecipitation followed by sequencing (ChIP-Seq) maps for H3K4me1, H3K4me3, H3K27ac, and H3K27me3 as well as DNA methylation (DNAme) data by whole genome bisulfite sequencing (WGBS) for the first four stages and reduced representation bisulfite sequencing (RRBS) for the last two (LRG and LNP) stages (Figure 15A and supplementary data 1, which is fully incorporated herein by reference in its entirety).

Global transcriptional analysis of the undifferentiated ES cells and the first four NPC stages identified 3,396 differentially expressed genes (Figures 20B-C and Supplementary data 2, which is fully incorporated herein by reference in its entirety). Pluripotency associated genes such as OCT4 and NANOG are, as expected, rapidly downregulated, and pan-neural genes are induced early and maintained throughout (Figure 20C). Using data from the mouse Allen Brain Atlas as an *in vivo* reference for genes expressed in different brain compartments and developmental stages, the present inventors observe a consecutive shift of expression signatures along the NPC differentiation trajectory (Figure 15B). NE through LRG transcripts suggest anterior neural fates, while the MRG and LRG stages show in addition some posterior identities (Figure 15B, left). Accordingly, differentiated progeny derived from these populations express deep cortical layer neuronal markers (NEdN and ERGdN) such as FEZF2 and

BCL11B and superficial layer neuronal markers (MRGdN) such as SATB2 (Figure 20D). Progression from early (NE) to late (LRG) stages was also accompanied by a transition from predominantly neurogenic to mainly gliogenic potential, although LRG cells can still generate neurons (Figure 20D). This progressive change in NPC identity aligns well with the *in vivo* order developmental events (Examples 1-7 above and Edri, R. et al. 2015).

In line with these observations, the WGBS data show changes in DNAme that can be separated into two overall patterns: the first is characterized by widespread loss and retention of the resulting hypomethylated state throughout subsequent differentiation stages (Figure 15C, top right). This pattern coincides with major cell fate decisions such as commitment from ES cells to the neural fate and the transition from ERG to MRG, the latter demarcating both peak of neurogenesis and onset of gliogenic potential (Figure 15C, right middle). The second pattern is defined by a stage-specific loss with subsequent gain at the next stage as observed during the transition from NE to ERG and also from MRG to LRG (Figure 15C, right). Conversely, regions gaining DNAme during transition from one stage to another frequently reside in a hypomethylated state in all preceding stages, indicating the possible silencing of stem cell or pan-neural gene regulatory elements (Figure 15C, left). At the histone modification level the present inventors also observe the most widespread changes during the initial neural induction (Figure 15D), although it is worth noting that the biggest gain of the repressive mark H3K27me3 occurs at the MRG stage.

EXAMPLE 9

COMPUTATIONAL ANALYSIS OF EPIGENETIC DATA IN THE VARIOUS BRAIN DEVELOPMENTAL STAGES

The coordinated epigenetic changes described in Example 8 hereinabove, are likely the result of differential transcription factor (TF) activity (Voss, T. C. & Hager, G. L. 2014; Ziller, M. J. et al. 2013; Gifford, C. A. et al. 2013).

A computational method to attribute the genome wide changes in histone modifications and DNAme at regions termed footprints (FPs) to particular TFs - Therefore the present inventors developed a computational method to attribute the genome wide changes in histone modifications and DNAme at regions termed footprints

(FPs) to particular TFs and quantified this remodeling potential (TERA: Transcription factor Epigenetic Remodeling Activity; (Figure 16A and Figures 21A-B). Interestingly, TF FPs in the NPC model were highly enriched for single nucleotide polymorphisms previously reported to be implicated in Alzheimer's disease ($p \leq 0.001$, Figure 21C) and bipolar disorders ($p \leq 0.001$) by genome wide association studies, suggesting the possibility to utilize this differentiation system to study the genetic component of complex diseases *in vitro* (Maurano, M. T. et al. 2012; Claussnitzer, M. et al. 2014). Next, in order to identify potential key regulators of onset, maintenance and transition through distinct NPC populations, the present inventors ranked all motifs and their associated TFs based on their TERA scores between consecutive time points (Supplementary data 3, which is fully incorporated herein by reference in its entirety).

Then, the present inventors retrieved the highest scoring 40 TFs for each cell state transition (Figure 15B). This analysis confirmed many well known key regulators of *in vivo* neural development and forebrain specification that are induced at the NE stage such as PAX6, OTX2, FOXG1 (Gotz, M., et al., 1998; Hanashima, C et al., 2004; Martinez-Barbera, J. P. et al. 2001) as well as various SOX proteins (Pevny, L. H. et al., 1998). Interestingly, the present inventors also found differential activity of distinct downstream components of signaling pathways such as a decrease of SMAD4 activity at the NE stage, consistent with inhibition of TGFb signaling that promotes neural induction (Chambers, S. M. et al. 2009). Another example is POU3F2 known to be involved in sub ventricular zone expansion and superficial layer neuronal specification, and TCF12, which is highly expressed in germinal zones during brain development (Uittenbogaard, M. & Chiaramello, A. 2002) (Figure 16B and Supplementary data 3, which is fully incorporated herein by reference in its entirety).

To obtain a higher-level overview of the processes and roles associated with the distinct putative regulators, the present inventors decomposed the H3K27ac data into seven distinct modules, each corresponding to a unique epigenetic dynamic, genomic region and upstream regulator set (Figure 21D, top). Gene set enrichment analysis (McLean, C. Y. et al. 2010) on the genomic regions associated with each of the distinct modules revealed that the module activated upon neural induction and sustained throughout the MRG stage is strongly associated with stem cell maintenance and differentiation related processes as well as Notch signaling (Figures 21D-E; module 2).

Further analysis of upstream regulators of this module revealed a strong association with PAX6 and FOXG1, suggesting a role for these factors in the general establishment and maintenance of the telencephalic cortical identity of the NPC states (Figure 21E).

EXAMPLE 10

IDENTIFICATION OF TRANSCRIPTION FACTORS WHICH ARE INVOLVED IN EACH DEVELOPMENTAL STAGE BY HSRNA SPECIFIC INHIBITION

Pooled shRNA screen against 244 transcription factors - To explore the relevance of predicted factors for each cellular state, the present inventors carried out a pooled shRNA screen against 244 TFs and epigenetic modifiers selected based on the RNA-Seq data (Figure 17A and Figure 22A, Supplementary data 4, which is fully incorporated herein by reference in its entirety). In total, 110 factors were recovered with a significant (Figure 17B, q-value≤0.05, mean empirical FDR=0.045) negative impact on the number of HES5+ cells in at least one differentiation stage (Supplementary data 4, which is fully incorporated herein by reference in its entirety), with high overlap between the distinct stages (Figure 17C and Figure 22B). Despite the expected high false negative rate (Sims, D. et al. 2011) the screen consistently validated more than 50% of the predicted TFs with a known motif for the top 20 motifs found at each stage (Figure 17D and Figures 22C-D), while an expression based identification yielded only ~30% recovery (Figure 22C). Among the top factors recovered from the predictions at the early stage (NE and ERG) are the RFX proteins including RFX4, which has been implicated in cortical and brain development (Blackshear, P. J. et al. 2003; Zarbalis, K. et al. 2004), FOXG1, as well as NR2F2, whose paralog NR2F1 has been shown to serve as an intrinsic factor for early regionalization of the neocortex (Zhou, C., et al., 2001; Faedo, A. et al. 2008). Gene set enrichment analysis of putative genomic targets of NR2F2 in the NE cells further expands this role suggesting involvement in telencephalon, diencephalon and posterior hindbrain development (Supplementary data 5, which is fully incorporated herein by reference in its entirety). At the MRG stage, the present inventors recover genes involved in extensive neurogenesis but also in commencing early gliogenesis such as NFIA and NFIB, which are involved in both repressing the neuronal progenitor state through Notch signaling

concomitantly with activating glial fates (Piper, M. et al. 2010), as well as REST – a major pleotropic epigenetic regulator of neural cell fate decisions (Qureshi, I. A., et al., 2010).

Transcription factors functional at all stages - Next, the present inventors selected a set of 22 core factors with evidence to be functional at all stages as assessed by RNA-Seq and the shRNA screening results (Figure 23A). In order to determine whether the subset of core factors with a DNA binding motif available (10/22) exerts the same function at each stage, the present inventors performed a co-binding analysis based on the predicted binding sites of 523 TFs in dynamically regulated H3K27ac footprints. This analysis uncovered highly stage-specific relationships that were also supported by the observed knockdown effect at each stage (Figure 18A and Figure 23B). Interestingly, most of the identified co-binding partners are either expressed in a more stage-specific fashion or are only activated in more mature neuronal or glial cell types (Figure 18B). To further validate some of these findings, the present inventors focused on OTX2 due to its high expression in all NPC populations (Figure 18B). OTX2 was enriched at more targets in NE of which around 35% overlapped with MRG bound sites (Figure 18C and Figure 23C). The shared target set is highly enriched for genes involved in stem cell maintenance and differentiation as well as various pro-neural gene sets known to act during advanced stages of forebrain and midbrain progenitor cell maturation (Figure 18D and Figure 23D). This binding pattern combined with the observation that the OTX2 target gene set reaches peak transcriptional activity in the NEdN and ERGdN populations implies a role for OTX2 in the preparation of pro-neural genes expressed at later stages (Figures 18D-E). These findings further suggest a model where a core set of TFs helps sustain NPC identity throughout the differentiation time course and at the same time participates in the progression and modulation of NPC differentiation potential through cooperation with stage-specific regulators.

To gain a better understanding of how factors that are active at distinct NPC stages contribute to their corresponding neuronal and glial cell propensities, the present inventors took advantage of the fact that many TFBSs (transcription factor binding sites) exhibit a gain of H3K27ac or H3K4me1 and loss of DNAme at the early NPC stages prior to increased expression of their associated genes in more differentiated cell types (hence referred to as epigenetic priming) (Figure 19A and Figures 24A-C). For

instance, the present inventors identified three pro-neural factors that show evidence of priming, are induced only at a later stage, and possess TFBS that are also significantly ($p \leq 0.05$ permutation test) associated with other genes differentially expressed at a later stage (Figure 19A, bold genes). Because these pro-neural genes are not expressed at the early NPC stages but at more mature cell types or later NPC stages derived from these early NPCs, the identification of such priming events highlights that the epigenetic state is useful for predicting key regulators and their downstream targets. In order to pinpoint TFs potentially involved in facilitating these priming events at the respective NPC stages, the present inventors determined significant co- binding relationships between the subset of pro-neural genes and other TFs that are concurrently expressed (Figure 19A).

To specifically investigate the hypothesis that a part of the pro-neural binding site landscape is epigenetically primed at the NPC stages, the present inventors focused on predicted NEUROD binding sites (Neuronal Differentiation) within H3K27ac footprints and defined five patterns of H3K27ac and H3K4me1 enrichments across these sites (Figure 19B). The present inventors found that genes associated with predicted NEUROD binding sites in regions gaining H3K27ac or H3K4me1 enrichment at distinct stages of NPC progression are up-regulated in more mature populations derived from the respective NPC stage (Figure 19B and Figure 24D). Consistent with the idea of a comprehensive preparation of the epigenetic landscape during lineage specification, NEUROD binding sites associated with NPC related genes that retain high levels of H3K27ac and H3K4me1 throughout the time course, are associated with various anterior and posterior cortical structures as well as early and late developmental time points (Figure 24E).

These results support a model where selected TFs at the NPC stage remodel the binding site repertoire for pro-neural factors by preparing the epigenetic landscape at their respective targets. First the general lineage landscape is established upon commitment to the neural fate, followed by the stage-specific modulation of primed pro-neural binding sites. This in turn restricts their binding space as a mechanism to ensure proper neuronal and glial differentiation capacity. In addition to these mechanistic insights, the present inventors provide a general analysis strategy to interpret differences in epigenetic landscapes based on cell fate regulatory TFs. This

strategy can be readily applied to other datasets including the extensive collection of the NIH Roadmap Epigenomics Project (Supplementary data 3, which is fully incorporated herein by reference in its entirety).

EXAMPLE 11

CLUSTER ANALYSIS DATA OF GENES DIFFERENTIALLY REGULATED DURING NEURAL PRGENITOR CELL DIFFERENTIATION

The genes differentially expressed genes across ES cells and four HES5+ NE, HES5+ ERG, HES5+ MRG and HES5+ LRG which are shown in Figure 20C were grouped into 18 clusters, and the genes included in each cluster are provided herein:

Cluster 1: ARX, CROT, PAX6, YAF2, IKZF2, CYP46A1, HDAC9, COL9A2, HHAT, GUCY1B3, MAP2, LRP2, LYRM2, CD82, CHRD, CATSPERG, TIMP3, GALNT16, KIAA0247, EYA1, NCALD, C7orf63, CDH23, LGI1, MAPK10, TBC1D9, TBC1D19, ZBTB16, LPXN, ASIC1, TSPAN11, WNT5B, NEDD9, HBEGF, PLCH1, DNAH1, OTX1, EPHA4, WLS, PROX1, PLAGL1, LTBP2, PCDHB14, CXCR4, EGR2, IKZF4, WNT1, EFNB2, BMP2, AMOT, PALLD, CBFA2T3, PNPLA7, VSTM2L, MEIS2, EMP1, ANKRD6, EPHA7, PAX3, ENPP2, IL11RA, LEF1, COL2A1, LGR5, MBNL2, CDH11, CBX8, MAPK4, PRDM16, BARHL2, RABL2A, FBLN7, ACKR3, SLIT2, SCUBE3, CRB2, TENM4, CTF1, LYPD1, ME3, CLGN, PDE1C, FBXO43, FCHO2, TSPAN18, DPYSL5, TMSB15B, C2orf81, SCUBE1, GBAP1, FRZB, ADAMTS9, FABP7, SP8, ZHX1, STOX1, FBN1, PRTG, KCNJ4, FAM107A, EMX2, DPY19L2P2, S1PR1, RP11-1055B8.7, ZNF584, BCL2, ID4, ALG14, 10-Mar, TCAP, PDIK1L, NR2F1, FOXG1, RNF152, RP11-169K16.8, C17orf96, FZD2, GAS1, NRIP1, BEGAIN, MEX3B, ZNF703, MAATS1, OAF, SOCS3, ZBTB7C, NR2F2, AHNAK2, MPPED1, FAM69C, WNT7B, HES5, SMOC1, ZNF521, WDR52, DENND1B, RP11-640M9.2, ZNF37BP, RP11-166B2.1, C1orf213, DKK1, CCL2, CDH6, FLRT3, LCN9, KLHL35, SPON2, CDC42EP5, TMEM132C, SYTL4, CHI3L1, RSPO2, TMEM163, CELF3, MSX1, PTX3, ZEB2, C17orf97, MYT1, E2F2, REEP1, TEKT2, PAK7, CA2, RIN2, CBWD5, WDR78, SLITRK5, ZNF503, CDYL2, LDLRAD4, FLRT2, FANK1;

Cluster 2: GAS7, NRXN3, FHL1, VIM, MFAP3, KIAA0556, PPP1R3F, COL3A1, GPC1, CDON, CTNNA2, NEO1, SDK2, C16orf80, FAM135A, KAT6A, BCORL1, EBF4, CPXM1, PCBP4, EFNB1, PLEKHG2, NLRP1, PTGS1, H2AFY2, PALM, 3-

Sep, ARSA, NKAIN4, E2F1, CST3, LPIN2, CD99L2, ZNF629, COTL1, METRN, SYT17, IGDCC4, STMN2, TUSC3, AMH, APLP1, MEIS3, IQCE, ANKMY2, GLI3, TLE4, ABCA2, ACTA2, KAZALD1, TSPAN14, EFNB3, KIAA1211, DTX4, ARHGEF17, WASF1, PCDHB5, H2AFY, PDGFRB, PLXNA1, STAM2, IGFBP2, IGFBP5, ODC1, RIMS3, AKT3, ST6GALNAC5, CNN3, C2orf43, C1orf198, NEK6, DENND1A, IFI27L2, PYROXD2, DUSP4, CNTFR, IFT81, MXD4, ENKD1, SOX21, PLAGL2, MCF2L, IFI6, FGFRL1, GNAZ, MGAT3, CCDC136, CDO1, LPPR3, ZSWIM6, PLXNA3, ZNF428, LRRC4B, REEP2, NES, WASF3, RFXAP, EPHB2, C1QTNF6, RNF122, FAM127A, SOX3, PHC2, ANXA1, CCDC102A, ARHGEF4, SPRY2, VEZF1, HS6ST1, FOXP4, ATAT1, FXYD6, MAPKBP1, BCAR3, SHROOM3, PIANP, VPS37B, CDH24, TSPAN3, ABHD2, CMTM3, TLDC1, SGSM2, SLC39A6, CBX4, EMP3, EPHA2, CELSR2, GPR161, VGLL4, TAGLN3, BOC, TMEM209, TMEM47, SHC3, INA, FEZ1, LPHN3, DIXDC1, DPYSL4, RNF144A, ASTN1, ZSCAN1, SORBS2, MARCKS, UBE2L6, FBRS, ZNF618, RHPN1, TMSB15A, GPR153, CACHD1, GPSM1, FAM171A2, BEND5, CAMK2N1, IER5, CCDC104, PACRGL, STK36, SPICE1, PTPN13, ABHD6, CTSB, HGSNAT, CASP7, AMOTL1, SAMD14, TUBA1A, TTYH1, RNF187, TCTN2, ZNF608, FAM110B, EFNA1, SDC2, FOS, CDH2, KIAA0232, HS6ST2, ASXL1, VAT1L, FBXL14, NBEA, INSM1, ADCY6, CNTNAP2, IGDCC3, FBXW8, SPSB4, GAL3ST3, CCDC57, VPS37D, SOX11, MAGEF1, ZNF664, MAGED1, CUEDC1, ZNF467, IDH2, RGMA, B4GALNT4, NDN, CRIP2, KIRREL, KCNH8, EFNA5, SNN, 5-Sep, H1FX, TCEAL2, KLHDC8B, TSPYL4, ARL4C, COL4A5, HES4, LDOC1L, FHIT, FAM217B, ASB13, ZNF777, CD47, SVIL, CHAMP1, TBKBP1, CCDC167, DMD, ARMCX6, ZDBF2, RING1, ZBTB12, PPP1R11, TECPR1, ZNF316, NYNRIN, RP1-152L7.5, FAM229A, ZBTB22, PCDHGC3, C4orf48, SHANK3, ZNF865, CORO1A, POU6F2, CPE, ID2, ID3, STMN1, DUSP6, CALB2, ZCCHC12, RNU4-2, S1PR3, EIF1AXP1, CALCOCO1, DKK3, CAMK2B, TRO, ATP6AP1, CRMP1, MOK, SLC22A17, MAGED2, CLIP3, CACNG7, CADM4, CHN2, MLLT6, LGALS3BP, SPARC, FGF12, HES1, GBE1, PODXL2, FNDC4, KIF21B, KLHL29, 6-Sep, CDKN2D, SAT1, MLLT1, GSTM3, NREP, KCNMB4, IGFBPL1, STRA6, FBN2, ZIC5, HERC2P2, CCDC40, LMTK3, ADAMTS10, IGLON5, PLEKHA6, HES6, IL17RD, PAM, GSN, FRMD4A, LGI2, MPP3, NYAP1, RHOH, SERINC2, GSTA4,

EFCAB12, BRSK2, PNMA1, PDE4DIP, EXOC7, TTC3, NRBP2, LSAMP, MAPT, HIST1H1C, HIST1H2BK, DDR1, C14orf132, PEG10, TUBB3, FUT8, IGSF9, DYRK1B, MEST, ING4, KHDRBS2, SRRM4, REM2, GAREM, IMMP1L, SHANK1, KLHDC9, FAM84A, CHST14, CSRNP3, RNF182, HIST3H2A, PNMAL1, EXT1, TCEAL7; **Cluster 3:** CX3CL1, DNAH9, ARHGAP6, SOAT1, RIMBP2, SLC6A16, EYA2, VASH1, CACNG4, PLXNA2, CMTM1, TF, TGFB2, CACNG1, SOX6, RFX4, PACRG, PDE10A, NME5, KAT2B, SLC1A4, PTCH2, SGIP1, PLEKHG1, NBPF14, GRK4, ZNF436, GNG11, KCNC1, MMP28, SYNE1, AVIL, STAT4, SLC40A1, KLRG1, LUM, ILDR2, UNC80, SCD5, FBXL17, TTBK1, LPAR4, SYBU, ZEB1, FAM13C, GAS2, SETBP1, ZIC1, PTPRN2, GLYATL2, FZD1, ZNF19, ZNF221, FBXL13, KLHDC8A, BBS5, PCDH19, DCHS1, MAP1A, NBPF9, LRRN2, DCLK2, MAP6, EPHX4, LY6H, SCN4B, MAF, CHRM4, SOX1, GPR1, LHFP, LRRC55, POU3F2, PROS1, SORCS2, ANKS1B, MAPK11, KLHL32, DCC, DNER, C12orf55, ZNF471, ANKRD36B, C5orf42, COLGALT2, GRK5, POU3F3, BTBD17, LY6G6C, LRRC10B, FAM71F2, DENND6B, DOK6, EGFEM1P, LDHAP4, ST18P6, LINC00537, RP11-3L10.3, LL0XNC01-240C2.1, CCDC13, ABCC6P1, AC020910.2, CTD-2054N24.2, FTLP12, SMIM17, LMO3, GLIS3, ST18, GABRA2, APCDD1, TPTE2P5, ZNF573, SP5, CACNA1G, STMN4, PMP22, MDGA1, PDE1A, NRN1, LRRC17, BAI3, LMO2, DOCK10, KIAA1644, ASCL1, RTN1, LRRC4C, KIRREL3, AKAP6, ATP2B2, AIM2, ATOH8, ZIC4, ZNF25, A2M, C12orf68, SLITRK1, UNC5C, GNG2, ZNF536, RGAG4, THSD7A, MEGF9, ST8SIA1, SATB2, PIGZ, CDKN2C, CRB1, PLCE1, PCDH10, NTRK2, LIX1L, NRGN, VCAM1, FAM198B, LMBRD2, VSTM4, TMEM100, IRX1, LRRN3, CHRNA7, IRX5, ZBTB20, GABRG3, GRID1, PCDH9, ELOVL2, PPP1R14C, ST6GALNAC4P1, AC016712.2, PNMA2, PTCHD4; **Cluster 4:** MEOX1, ITIH4, WNT8B, ARHGAP15, RBFOX1, ABCC6, CRYBA1, FOLR1, BMP5, NR2E1, RGS4, CASC1, OLFM3, NR4A3, PDE1B, TMEM255A, OPN1SW, ALDH1A2, VTCN1, BCMO1, MORN3, CLEC18B, PTH2, DMRTA2, OXSM, FEZF2, SST, TNFRSF14, TNNI1, CCDC135, LMX1A, PCDP1, HEYL, SYTL3, TMEM74, C9orf24, DACH1, CMTM5, RSPO1, TRH, SOSTDC1, NHLH1, SYNPO, SPTLC3, C4orf26, RAG2, FAM182B, DMRTA1, C8orf4, CLEC3A, ODF3B, GPBAR1, NME9, GLUD2, FAM131C, NTF3, TAL2, AC018892.9, C9orf171, ARL9, KLHL14, WDR96, MEIG1, DCAF12L2, HIST1H3A, C1orf228, ABCA4,

PPAPDC1A, CCDC160, C6orf25, CTD-2514C3.1, C21orf49, RP11-108K14.4, RP11-156F23.2, COL6A4P2, RP1-203P18.1, TSPAN19, RP11-216N14.5, SCAND3, RP11-5P18.10, CRYBB2, ALG1L12P, RP11-671M22.4, CTD-3193O13.10, RP11-353N4.6, AC104057.1, CTC-559E9.12, DMRT3, TTC29, MMRN1, HTR2C, TPPP3, LIPH, NOG, AC018816.3, NOS2, RARB, BARHL1, PAX8, SLC47A1, CORIN, LHX1, CLEC18A, OSR2, NEUROG1, RTN4RL1, FOXE3, RP11-730A19.9, PTPRH, MECOM, LHX1, NDST4, DRC1, CLSTN2, C2, CCDC60, AC010980.2; **Cluster 5:** MRC2, ATP1A2, MRVI1, KCNQ2, DPYSL2, NDRG4, PTN, B3GAT1, PDE4D, SOX9, SYT11, NOTCH2, ARHGAP29, SEMA6C, GAREML, DLC1, ARNT2, ZMAT3, CDH4, GPC6, NLGN3, COL5A2, GPR56, SYP, CENPI, PLLP, SLC4A3, NRP2, ITPKB, COL8A1, NOTCH1, GFRA1, DBI, FGFBP3, AGRN, TFPI, BRCA1, EXTL3, RUFY3, TIMP2, ATP9A, MYLK, PPP2R5B, KIFAP3, GPR137B, DLL3, RAB36, PLTP, ZNF423, NEFM, NOVA2, DKFZP761J1410, PON2, RUNDC3A, NRXN2, SERINC1, TANC1, ST3GAL5, PARD3B, C1orf61, ATP1B2, DPP6, NDFIP1, KIDINS220, GNS, NACAD, SCARB2, CPNE2, KIFC3, EVI5L, SIK1, SLC16A2, NCAM1, GPM6A, PDLIM3, KIF5A, ATP6V0D1, AGPAT3, PPAP2B, SDC3, PEA15, SPRY1, FAM219A, HTRA1, STIM1, LTBP3, KIF5C, ANTXR1, CHD3, PAQR8, SCUBE2, CDK5R1, ASB8, CTNNBIP1, ZHX2, CECR6, KREMEN1, FAM43A, PDE2A, AFAP1, XRCC2, LRP10, STMN3, PHF2, SHISA4, C1orf233, ZNF853, SOX8, ETV1, AKAP11, GPM6B, SCML1, H6PD, KIF1B, LIMCH1, HIP1K2, SNCAIP, DIP2B, MPPED2, EVI5, TRIB2, EPN2, NOTCH3, RBL1, PHLPP1, EPB41L3, SEMA5B, GNA11, OSBPL8, MAP3K1, MICALL1, NIN, CDC25B, RASSF2, KIAA0226L, CDK6, TSPAN12, PTPRZ1, SUSD1, ARHGAP21, CLCN3, CORO1C, CDK2AP1, FAM184A, FAM65B, BVES, QKI, GNB4, ARHGEF26, ABCD3, PCDH17, KCNIP2, UBL3, FKBP9, GLIPR2, CHST3, FAM199X, ATXN1, VPS13C, ARHGEF6, SORT1, ROCK2, LRP4, ARHGAP32, APC, TTYH3, ABHD17C, ENTPD1, SSFA2, 11-Sep, ETV6, KIF21A, GAS2L3, RB1, SLAIN1, NOVA1, TCF12, ZCCHC14, NPC1, TTYH2, RNF157, IFNAR1, CDC42BPA, MEIS1, CAMK2D, ANK2, PDGFC, PM20D2, TMEM181, PURB, TACC1, NACC2, EIF4EBP2, SESN3, ENDOD1, SOGA1, FREM2, SACS, C5orf28, MGAT5, DDAH1, GRIA1, ELMO1, PSD3, MMP16, ATAD2, LSS, CLPB, ZYG11B, LRRC58, SRGAP2, IFI16, WDR41, TP53INP1, FAT3, REEP3, CERCAM, SCARA3, PBK, GNG4,

RAB31, ADAM9, TSPAN5, PXDC1, COL22A1, ROBO1, CTNND2, TMEM133, MOB3A, LCLAT1, PEAK1, HEG1, SH3PXD2B, TMEM9B, DPP10, KBTBD11, WSB2, IRX3, TMTC2, WSCD1, ZBTB42, PAK2, ARSJ, F2R, BAI1, PDE4B, NAT8L, FNBP1, ZNF70, TDRD7, SLC35F1, SLC6A9, WNK3, IGF2R, GAL3ST4, VEPH1, GRM3, PJA2, C2orf72, C5orf51, PTPLB, SLC35F6, RP11-697E2.7, C1orf226; **Cluster 6:** DCT, VAX2, BCAN, NT5E, MTTP, NFIB, MEF2C, NALCN, CHL1, SPARCL1, GRIK1, DCN, ARHGAP31, MYO16, CA12, CACNG5, PAG1, C14orf105, SLC8A3, KLHL4, IL7, INPP4B, GBP1, ATP10B, SLITRK3, RHOJ, SCN2A, IFI44L, IFI44, CNGA3, DLX1, CNPY1, CSGALNACT1, SNTG1, SLC39A12, ADAM12, PCDH15, NPAS3, SCN3A, ANKFN1, ADAMTS3, ART3, SLC6A1, S100B, BRINP3, SYNPR, EN2, CSMD3, KIAA1161, AMER2, SLC16A4, GSX1, TM4SF1, ARPP21, DDX60L, PTPRT, PAX5, CPNE4, DPF3, AC008060.7, AC092675.3, TRAC, RP1-161N10.1, AL807752.1, CYP26B1, TENM1, ARAP2, SLC4A4, MLC1, SGCG, HOXA2, ELMOD1, TNR, LPPR5, GRIA2, SLC6A11, THSD1, MMD2, TLR4, IL33, MDGA2, ABCA8, SLC13A5, CLVS2, MKX, GPR158, GUCY1A2, PRDM8, KCNJ16, NCAM2, ACSS1, NFIA, EN1, GPR155, NKX6-1, GRIK3, SCRG1, TMEM71, SVEP1, ALK, KCND3, HOPX, KIAA1239, KCNA2, KCNJ10, RGS6, CSMD1, GRIN2A, OLIG1, C3orf70, LHFPL3, SLC18A3, C15orf52, C2orf80, POU3F4, GPR123, RYR3, AC138783.11, C21orf62, AC005019.2, TMEM257, UBBP3, RP11-434H6.2, RP11-598P20.3, HNRNPUP1, ZNF818P, GRIN2B; **Cluster 7:** CPA1, KRT23, UCP1, ITGB6, CSF3R, TNNT3, CGA, ELF5, FGFBP1, TRIM29, ERP27, CYP1A1, CLEC18C, SCGB3A1, S100A9, S100P, AGBL1, MYEOV, C1orf127, GLDN, FAM90A27P, MYL4, RP11-15J10.1, SLCO6A1, OR7E101P, RP11-4K3__A.3, AC012441.1, RP11-436I9.6, APOC4, RP11-551L14.7, RP11-579D7.8, TMEM176A, TFAP2B, SIX1, IL6, TMPRSS4, P2RY6, ANKRD34B, CASP4, CCDC162P, IGFL4, GCGR, GATA3, AIPL1, HPSE2, RNVU1-6, CLPSL1; **Cluster 8:** ZMYND10, ANO2, SEMA3C, PPEF1, SLC17A6, EBF3, UNCX, DYNLRB2, VWA3A, CNTN2, GRP, SLC25A48, CPLX2, SLC30A2, NSG2, AMBN, HIST3H2BB, PAX7, EFCAB1, IPCEF1, MOXD1, EPHA6, MYH7, SPEF1, MYOM1, CBLN1, RUNDC3B, PPP1R17, HGFAC, CNTN3, C20orf85, PCSK2, GDF5, FOXP2, ISLR, RFTN1, SYT4, RGS8, TBR1, GAD2, LHX9, PSD2, GPR6, RSPO3, CRH, PTPRO, TMEM132D, SH2D6, CAPSL, ABI3BP, FGF18, KCNB1, FGF17, GJD2, ZFYVE28, UBXN10, DAPL1,

CCDC141, EOMES, TLX3, SAMD3, KCNC2, SVOP, WDR16, MS4A8, ISLR2, TEKT1, FSTL5, PARM1, SCRT1, PDGFD, CDK5R2, TMEM196, LINGO2, TTC9B, RTP1, ST8SIA3, HIST3H2BA, GALNT9, PCP4, C11orf88, OPCML, SLITRK6, TMEM173, AKAP14, SHISA7, COL25A1, C1orf192, RELN, SERPINA3, HRCT1, BRINP2, SAMD5, RP1-65P5.1, RP11-12A20.4, MCIDAS, SHISA8, SMIM22, ACTL6B, P2RX6, APOL4, RSPH4A, NEUROD4, HDC, C1orf222, VAX1, CDH12, CDC20B, ZNF804A, KCNK3, MYT1L, RNVU1-14, EBF2, C17orf72, XKR7, RN7SL53P; **Cluster 9:** NFIX, SEZ6, TLE2, EEF1A2, PCYT1B, SCG3, CALB1, CPT1A, OAS3, LIFR, CSMD2, ITPR2, PMEPA1, UNC13A, LGR6, DCLK1, HEY2, CNOT6L, MEGF10, GRIK4, USP43, ZIC3, NMNAT2, ABTB2, PARP14, RAI14, YL, METTL7A, NRG3, TRIM69, SPRY4, OGDHL, RN7SL721P, CCDC80, PREX1, TFAP2A, SLC15A2, FAM181B, SLITRK2, S100A2, MAMLD1, LRRC7, TNC, LZTS1, SLC12A2, PLD1, DCX, OPHN1, GNAO1, RAPGEF4, SEZ6L, DOK5, PCSK1N, CORO2B, AP3B2, SPOCK2, SFXN3, GHR, NR3C1, WNT5A, BCL11A, PDZRN3, CD97, LPGAT1, MASP1, MCHR1, GAD1, NCAN, CMPK2, TAOK3, PCDH8, SEMA6D, TMEM117, DISP2, TGFB1I1, CACNG8, CASQ1, ARHGAP18, CHST7, CDKN2A, NR3C2, TNIK, MEGF11, ABCG1, CHRNB2, NFASC, BSN, MGAT5B, VSTM2A, LONRF2, FUT9, ADCY5, LPL, TSHZ1, NTM, ABAT, MAML2, MANEAL, NTNG2, LPAR1, TOX, PTCHD2, LRRC37A16P, NTN1, ACTN2, ADCYAP1R1, SLC1A3, TNS1, JAK2, TRIM9, ACSBG1, ITGB8, TBC1D12, LAMA4, SEMA5A, BCHE, CSPG5, CNR1, CPNE5, DOCK4, TMOD2, TULP4, CPM, EDNRB, GABBR2, PARP9, NFIC, IGSF11, TENM2, PLEKHH2, GRIA4, PRKCA, PIEZO2, NTNG1, KIF26B, RFTN2, KCNF1, RNF180, ZCCHC24, KIAA1462, CACNB2, NETO1, GBX2, GPR37, HAS2, MCC, B3GALT1, GNG7, NCKAP5, ZDHHC22, KCND2, POMK, C1QL4, SLIT1, LUZP2, TET3, SHISA6, CTD-3088G3.8, LAMA2, DCHS2, ZNF460, PLXNB3, BMPR2, CYS1, RP3-340B19.2, RPL19P20, RP11-390F4.2, FRRS1L, CICP14; **Cluster 10**: NFE2L3, ADAM11, CAPN6, NEBL, LHX5, MFNG, GPR143, PTPN5, RHOU, OLFML3, CDK18, PACSIN1, PRKCG, RASL11B, FEZF1, APOE, RAX, CCDC64, MYOF, RGS16, CXCL14, TPBG, KCNMA1, CLIC6, OTX2, ELFN2, IGF2, SPINT2, ROR2, AFAP1L2, RAB3B, MT1E, CMTM8, KRT8, NRTN, HSD11B2, IFITM1, FAM183A, MT1X, SEMA4A, C6orf141, COL4A6, AP000783.1, HECW1, SYT7, F7, DSP,

TNNT1, FBXO2, NPY, ITIH5, BMP4, SIK1, WNT3A, MEGF6, PAPPA, SCN5A, HIST1H4D, MT1F, C6orf100, MT1M, C20orf26, LEPREL1, BAMBI, TSPAN15, MXRA5, SLC17A7, CNFN, SH3GL2, ICAM2, GALNT12, KIAA1217, IL13RA1, PPP1R1A, KCNK1, CRABP2, OCIAD2, ZNF185, NR4A2, PLEKHG4B, DMTN, ALOX15, ALPL, ECEL1, RCAN2, EGFL7, CSPG4, SLIT3, RP11-923I11.7, ARHGAP44, GABRA3, SYT13, POU2F2, PTPRU, MYO3B, PPP2R2C, IL4R, TRPM3, FER1L4, SLC7A8, PICK1, CTSH, BMF, NDRG1, LSR, UNC5B, NMU, KIAA1324, SLC8A2, PPP4R4, MSX2, PLXDC2, EPHX2, SDC4, PTGIS, BCAS4, ZNF391, TMEM74B, DACH2, TNNI3, PRRG3, CASZ1, PODNL1, ZDHHC8P1, MST4, IER3, SMAD6, SIX3, FRAS1, CCDC33, SH3GL3, TMC6, FHAD1, RGS5, MPP7, SPAG17, KIT, CACNA2D3, DYNC1I1, WNT4, RBM47, GRM2, TMEM130, SEMA6B, PHYHIP, GPRC5C, MGMT, EMB, KNDC1, PCSK1, EFCAB4A, SHISA2, MAPK15, KCNQ3, FMNL1, FAM86B1, GCNT1, ZNF204P, GOLGA2B, CFB; **Cluster 11:** SLC25A5, PSMC4, KCNG1, ISOC2, WDR1, DAZAP1, FAM50A, NAALAD2, OSBPL6, EZR, HSP90AB1, CBX7, C20orf27, AHCY, GPR50, HTATSF1, TOX3, TIMM50, GPI, ISYNA1, DNAJB6, LFNG, HSPB1, PFN1, RAB34, CCND1, HSPA9, DPYSL3, AMOTL2, ACTR1B, ERRFI1, GADD45A, NID1, MYCL, SLC2A1, PRDX1, CTGF, DUSP1, EGR1, LYPLA1, ZSCAN18, BBS9, NR4A1, PLP1, AHNAK, ID1, CDC42EP1, LOXL1, SHC2, JUND, COL5A1, DKC1, RAN, SERPINF1, EIF5A, TMEM2, MSI1, TXN, SULF1, GABARAPL1, PSMB6, SERBP1, PSMA5, RAB13, HSPD1, SFRP2, PLK2, CCT6A, GPC3, GSTO1, ALDOA, UCHL1, TSC22D3, CSRP1, BTG2, CBS, SHC1, COL26A1, TKT, 8-Sep, CITED2, RPL36AL, ZNF219, HSP90B1, PPIB, CYCS, SUCLG2, FAM195A, EHBP1L1, PPP1R14B, MRPL11, TP53I11, BASP1, METRNL, TALDO1, DPM3, MYADM, SKIDA1, GAS6, ARMCX2, SUMO3, DLK1, ZFP36L1, POU3F1, EMID1, PTMA, H1F0, TOMM7, NRARP, ARC, DLL1, APRT, FAM127B, MT-RNR1, TWF2, MYH9, PGRMC1, RBP1, CLU, LAMA5, MYCN, FST, KLF4, CLSTN3, IFITM3, COL6A2, FAM57B, MARVELD1, MMP14, STC1, SPATA18, WNK2, AGPAT2, ANXA2, S100A4, KLHL13, SPATA20, TACC3, DEPDC1B, FLT4, ZIC2, YBX3, POLD1, ABCA7, IPO5, ASNS, MCM2, ENO1, MCM6, MCAM, CAMSAP3, SESN1, TRIP6, AURKA, AARS, ORC6, PHGDH, LGALS1, MCM5, FAM118A, PHF5A, MYBL2, SLC7A5, SFRP1, ANKRD27, FBL, PLD3, SLC1A5, BCAT2, PRKAG2, PTGR1, NPDC1,

ALDH2, MCM3, GMNN, GNPDA1, EEF1B2, SDC1, HDAC1, IVNS1ABP, LPHN2, TXNIP, CDC20, PRDX6, MTHFD1L, TRIM24, CKS2, OBSL1, CSE1L, SNRPB, BCL2L12, DLGAP5, PAICS, FAM64A, AJUBA, CDKN1C, SYT5, SESN2, TCEAL4, BEX1, DDB2, YARS, CDCA8, PSAT1, ATP5G2, HMGA1, ATIC, FAM60A, SLC7A1, PCSK6, GREB1L, CYR61, RHOB, CSRNP1, CCNA2, ABRACL, CDCA5, PLIN2, LATS2, GJA1, CENPH, FZD7, CCNB2, IGF2BP1, CELF5, SNX7, ATF3, PKDCC, NUAK2, RPL39L, CDC25A, HMGB2, PTTG1, SLC16A9, PLEKHF1, CRABP1, BLCAP, FBXO22, HID1, TK1, RPSA, NT5DC2, RFWD3, FAM84B, SLC30A1, PFKFB3, PGAM1, SPSB1, CTPS1, CEBPB, CKS1B, SEZ6L2, CHST2, IMPDH2, AURKB, GEMIN4, HIST1H2AC, SIAH2, SHMT2, GJC1, IRAK1, ISG15, EIF4EBP1, GREB1, IARS, MYO1C, RPL12, TPM2, INPP5F, GPC2, CTB-63M22.1, KIFC1, ITGA3, CYB561, RPS20, HEBP1, GRAMD1B, AP2S1, RCN1, PLEKHA5, KCNH2, PHF21B, RPL18, CA11, NDUFB4, POLB, RPL31, SMARCE1, CTTNBP2, FTL, C14orf166, RPL6, HEPH, NAT14, CD200, SH2D3C, ATP5D, NDUFB7, RPL3, LGMN, MAP1LC3A, MYL12A, EIF3E, RPS16, CDC37, RPL18A, EDF1, RPL19, SYNGR2, RPL34, MDK, C12orf57, CD83, RPS12, SLC12A7, RPL24, REEP6, PCSK4, RPS15, LAMTOR2, RGS2, TMEM9, MFAP2, RPS25, HDHD3, RPL21, RPL5, ALKBH7, RPL23, SNRPD2, COX6B1, PRMT1, RRAS, SH3BP4, RPL36, C16orf13, ATPIF1, COX4I1, RPL27, BEX2, CTSL, COX5B, RPL35, RPS6, RPLP1, RPS24, SRP14, TPM1, RPS2, NDUFB10, NOB1, ARRB2, ZNF787, RPS11, RPL11, RPS8, CREG1, EFNA3, RPL32, AMT, RPS3A, NDUFS6, RPL37, BOD1, RNF44, ABT1, IGSF1, RPL7, SERPING1, RPS3, HMGA2, FARP1, BCL2L11, FAM49B, SSBP3, TPRG1L, TAGLN2, C21orf59, PSMD4, CHCHD6, DEDD2, LY6E, RPL26, RPL29, KIAA1522, MRPL55, CCDC74A, H3F3A, FSTL1, RPL9, CAMKV, RPS14, NSD1, TMED3, GPX4, RCOR2, NDUFS5, ATP5I, NSMCE1, CCDC8, RPS9, C9orf16, TBCA, RPS21, RPS7, MZT2A, SLC19A1, UQCRH, ZNHIT2, RPL4, NDUFA11, ARL4D, EID2, RPLP2, RPS27, FUCA1, SSR4, TRAPPC5, LIMK2, RPL35A, C2CD4C, NDUFB1, UQCR10, NIPSNAP1, CEND1, SIVA1, IFITM2, KIAA1598, RPL14, AJAP1, ADARB1, RPS26, RPL37A, RPS4X, RPL23A, UBL5, RPL10A, SELM, MT-ND3, RPL39, LBH, TSPAN4, AP000350.4, FAM19A5, RP11-742N3.1, RPL41, PET100, RPS18, TMA7, RPS28, WBP1, TMEFF1, RPL17; **Cluster 12:** CYP51A1, CD9, KPNA6, CD44, ARID4A, TSPAN17, SLC4A8, MSMO1,

CCAR1, TNPO3, OAT, PDIA5, ELOVL1, KLF6, PDZD4, ROCK1, TFE3, RASGRP2, DNAJA2, SLC44A1, WBSCR22, PVR, FBLN1, FDFT1, HSP90AA1, STK17B, XPO1, SH3BP2, PXN, USP48, SEMA6A, RFFL, SORBS1, SLC25A1, TNRC6B, PCNX, PLS3, DNAJC3, MMP15, SQLE, GSR, CCDC61, OLFM2, GRIN2D, CAV1, TMEM248, CCDC6, PSMD11, CCDC34, SC5D, HSPA8, DDX6, HPS5, SASH1, C6orf62, HMGCS1, HMGCR, INO80D, SPTBN1, COQ10B, PDCL3, BIRC6, DHCR24, ECE1, AKAP7, CCND2, TRIM25, ZC3H13, AGO2, ARFGEF2, SOX4, RRBP1, NCLN, VGF, GNL3L, LDLR, MAP1S, OLFM1, MPP1, GSE1, RAMP2, TRIM22, LANCL2, SWAP70, HSD17B4, VAV3, MDM2, KLHL36, RC3H1, TUBB2A, ARRB1, PRCP, SLTM, PPM1B, PHLDA1, UBE2Q2, ULK3, BRD4, SNX27, SPATA5, PHIP, ADM, DGKZ, COMMD7, PLCB3, FOXO1, BICD1, RABGAP1L, PLOD2, TRAPPC8, LPCAT1, THY1, PITPNC1, AZIN1, TMEM55A, LARP1, PDIA4, PPP2R2B, HK1, LRP8, DYRK1A, EIF5B, ZNF714, SLC25A44, FAM102B, IGSF8, RNF168, CCDC112, CASP3, CDCA7L, ADCY1, ABCA1, ARHGAP42, GABRB3, CYB5A, MVD, ATCAY, STAT3, PLEKHA2, LIN52, LINGO1, HNRNPA3, INSR, SHCBP1, GAP43, DHCR7, FAM222B, ARHGAP1, BOK, PPFIA3, PFAS, RCC2, CALR, TMEM64, NQO1, DHTKD1, YIPF6, ATP6AP2, HIST2H2BE, NPIPB4, 1-Mar, INSIG1, NF2, USP7, DAPK1, SCN8A, ZNF841, HELZ, GFPT1, FAM155A, ATXN2, RAB12, ARHGAP19, PLEKHO2, TMEM158, LYN, GAN, SEMA3A, COL9A3, SCD, SGK3, DLX5, PTGDS, IRX4, WWC1, CHPF, IDH1, COL6A1, CREB3L1, CDK1, RASGRP1, SGK223, RP11-1166P10.1, ST7, MYLIP, CAPN1, GRN, EPB41L2, KIF3C, WDR47, PPP1R13B, CHRNA4, CELF4, SMAD7, SALL1, SHD, ATP1A3, CLIP2, ELAVL2, FN1, DLGAP3, FAAH, KIF14, IRF2BPL, GDAP1L1, SERPINB6, TMEM35, KIF1A, C19orf66, ITM2C, CENPE, RGS3, SLC20A1, STK11IP, LRIG1, SLC18B1, JAK1, SERPINI1, CADPS, KIF2C, SLC43A2, CXXC5, PRNP, ETV4, FAM20C, BCOR, PTCH1, LRRC16B, FBLL1, S100A16, ANXA6, L1CAM, COL11A2, TAPBP, SEPP1, CREBBP, PROM1, SAMD4A, GCLM, TRIO, RRM2B, LIMA1, QSER1, TM7SF3, PPP2R5A, PRR11, NDE1, TRAF4, SYNJ2, LRRC16A, CLDND1, ITGB5, STARD7, DOCK9, LZTS3, RCOR1, GOLGA3, CLSPN, XYLB, CRKL, GINS1, EEF2K, IMPAD1, PPP2CB, PIK3R2, SIPA1L3, ZC3HAV1, WASL, AKNA, NCS1, GAB1, CBL, HIPK3, FOXM1, MAN1A1, ASCC3, GPR126, TMEM30A, ACAP2, CCDC88A, STRN, MARK1,

TROVE2, MAN1C1, CCNI, CYP20A1, PHF19, ZC3H7A, 7-Sep, SLC25A16, CDKN1A, AIF1L, TEP1, MAP1B, TOP2A, STARD13, KRAS, SBF2, PTGFRN, RNFT2, AKIRIN2, FAM129A, SERPINE2, SMC2, TLN1, CASC5, KIF11, TMEM132B, RNPEPL1, LMO4, GATAD2B, SFXN5, SH3KBP1, MSN, DOCK5, RSU1, MKI67, LIN7C, CELF1, ATM, QDPR, UHMK1, SEMA3D, LONRF1, AGPAT5, ELMSAN1, SAMD8, SKI, B4GALT5, ZYX, VMA21, CCNF, SLC25A45, LRP5, USP24, NOL9, VANGL2, SGCB, SGOL2, PRSS12, FOXK1, SGPL1, C18orf54, TMEM194A, PAFAH1B2, HOOK3, SOGA2, FASN, TRIM8, WIPF2, HCFC1, PDE3A, DAG1, ZBTB4, JUN, PLEC, SOCS4, RAP2B, HS3ST4, TRAK1, 9-Sep, ZNF445, KIF18B, TMSB4XP8, TMEM201, COL27A1, GMFB, FAM102B, STYX, MDM4, CALM1, PHACTR4, LGR4, C15orf59, FXYD7, SALL3, TMEM178B; **Cluster 13:** MATK, MLXIPL, PRICKLE3, PRSS8, DRD4, ANKRD24, WFDC2, PPP1R16B, EPHX3, CORO2A, CPZ, CCKBR, SLC35F2, PTPN6, SLC16A10, PRSS16, POMC, GRB14, VAMP8, EPCAM, A4GALT, KLK10, RHBDF2, APOC1, LIN28A, RAP1GAP2, RAB25, MTL5, MLK4, ZNF385B, MARVELD2, SLC7A7, RAB11FIP1, SUSD3, DUSP2, FAM46B, VSNL1, TERT, GYLTL1B, SPINT1, TAC3, RBFOX3, CA4, CYP2S1, UGT3A2, PCSK9, CDC42BPG, KRT19, FUT1, CPNE7, HLA-DQB1, CLDN6, SYN3, FAM83G, BEND4, CLDN4, RASGEF1A, LST1, ANKRD18B, ARHGAP8, ACTN3, SLC22A31, CTD-2616J11.4, PLEKHG6, CD74, COL17A1, PVALB, RHOV, KRT18, B3GNT2, IQCA1, KRT7, ANXA3, SYTL1, THNSL2, ANKRD1, FTCD, DMKN, ELF3, CLDN3, AQP3, RHOD, SFN, RNF212, SP6, TGM2, CD79B, VRK2, ADAM28, LLGL2, STXBP2, GSDMD, EBI3, FGF21, DOCK8, TCIRG1, STC2, TNNT2, TEK, NPPB, LRRC61, KLK8, TSPAN2, ALPK3, INHBE, CYP11A1, IRF8, PADI3, LEFTY2, PLAC8, PAMR1, CACNA2D4, ACTC1, MAMDC2, ALDH1A1, PKDCC, L2, STX3, KLK6, NAALADL1, CCDC88B, KLK7, SLC16A5, KCNS3, PRKCDBP, TNFRSF10C, CD7, AIM1L, RP11-551L14.1, FAM83H, TXNRD2, GSAP, PLEKHG4, HLA-DOA, TMEM114, CTD-2021H9.3, RP11-554D14.2, PKP2, TJP3, SIT1, MPZ, ARHGAP27, NOS3, FES, AIFM3, RP11-527L4.2; **Cluster 14:** PRDM1, CDH3, ME1, ATP2A3, FOSL2, SLC4A11, FXYD5, ARHGAP4, TGM1, TRIM14, PDE3B, IM1, COL1A1, HTATIP2, ARTN, CNN1, GDF15, SYT6, FAM71F1, MAP7, SORL1, HERC5, FGF2, CNKSR1, ACTA1, MALL, OSBPL10, USP53, N4BP3, IGFBP3,

DIAPH2, STOM, NR6A1, F11R, NPR2, FGFR4, DHRS3, FZD5, KCNK5, COL1A2, GPR176, TMEM88, BNC2, TNFRSF10D, PRIMA1, COL14A1, S100A10, TEAD4, ANKRD35, MICB, TRIM71, SHISA9, APOBEC3C, FAM86EP, ICA1, PRKCH, ANK1, HES2, CXCL2, PAPLN, SALL4, AURKC, PERP, ECHDC2, EMILIN2, THBS1, CGN, SLC5A12, RLTPR, F2RL1, WDR72, INPP5D, PTRF, NPW, SP140L, PDGFA, MYO15B, SERPINB1, CAPG, FGFR3, NEDD4, TPD52, CECR1, CARD10, PCK2, CTSZ, TFPI2, HPS1, FBXW4, PTPMT1, SH2B3, KCNQ4, FILIP1, TRIM6, ADCY7, TMEM54, ANXA11, WIPF3, SLC12A4, ADM2, ARHGAP22, CHAC1, SPOCD1, TMBIM1, USP44, DRAM1, SLC39A8, JDP2, CDH8, NLRX1, PRKCD, SCG5, VWCE, STXBP6, PLAC1, LRRC34, MYD88, NPAS4, NRIP3, FOSL1, OPLAH, NSUN7, ARID3B, CALN1, UPP1, ADAP2, ADRA2C, SIGIRR, PMEL, LIN28B, SH2D5, S100A6; **Cluster 15:** CFLAR, YBX2, JARID2, AASS, MGST1, TCEB3, RALBP1, NLRP2, UBR2, HSPA5, USP28, COL23A1, FOXN3, NOP58, MTA3, ERBB3, PFKP, PPAP2A, PDK3, TCOF1, NLE1, ST6GAL1, ADD2, REXO2, ICAM3, GPC4, PHF17, KIAA0020, SEPHS1, PTPN4, DDX18, DNMT3B, TESC, ESF1, TNRC6A, EPB41L4B, SIRT1, CECR2, DDT, PRODH, PLA2G3, HMOX1, SPTLC2, RIMS4, SAMHD1, FAM83D, ADNP2, LIPG, NKAP, PIM2, EMD, CRYM, BFAR, CEP152, GABPB1, TWISTNB, SERPINE1, KDM4C, DDX58, NPM3, KLHL2, GAR1, CRYAB, CTSC, ZNF259, GNPTAB, PHC1, TTK, SLC29A1, MAN2A1, BRIX1, RRP9, CEBPZ, MEF2D, MAD2L2, PRPF3, SLC19A2, RCAN3, ESYT2, SPP1, SENP5, HEATR1, RBM25, SLC17A5, HELLS, HSPH1, DPPA4, FAM126A, PAIP2B, USP9X, MRS2, NQO2, UPF3B, TRIP10, VASP, WDR60, TPST2, PODXL, MAP7D3, DOHH, FAM155B, GADD45G, SULT4A1, EIF2S3, ZNF331, GFPT2, DNAJB1, MYBBP1A, SLC39A11, RRAS2, CDK7, LDHA, ELP2, GLS2, GCC2, FLNB, SKIL, KIAA0368, RABEPK, CTSV, MYC, PIM1, PPP2R1B, LRPPRC, TMEM180, PPIG, CIR1, HCN4, LARP1B, ETNK1, PIF1, MFGE8, CDH13, MYO19, TP53, PMAIP1, URB1, MTF2, POU2F1, GOLGA4, TERF1, PTGES2, PPRC1, PAK1, SAP18, C12orf45, UPF2, CWF19L2, JMY, RANBP2, INO80C, RSPH10B, OTUD6B, GRAMD3, MCU, PCDH1, RBPMS, MYO1E, LDLRAP1, C1orf51, CBR1, ATAD3B, CILP2, PDXK, GPATCH4, FDXR, PDPN, FLVCR1, PQLC3, ADCK3, METAP1, EDIL3, FABP5, MID1IP1, SLC7A3, NEMF, TSC1, DDX21, C10orf118, PACSIN3, TAF1D, TRMT61A, ZNF143, RBPMS2, PLK1,

MRPL16, RBPJ, DNAJC21, PGM2, BUB1, UGP2, OTUD3, B3GNT2, IRX2, RSL1D1, SLC25A33, ATF7IP, EIF2AK3, PPP1R3B, PTPN2, PLA2G16, PUS1, BEND3, PHC1P1, AEN, FRAT2, FAM101B, CHEK2, CMSS1, TOP1MT, ATL3, WBSCR17, CD24P4, ZBTB3, VSIG10L, PALM3, PTAR1, LITAF, RRP7A, ZBTB44, ZNF34, C6orf106, ZNF398, MAFG, RPF2, ZNF121, ZNF649, WDHD1, FAM169A, LRIG2, HSPA1B, HSPA1A, MICA, CTD-2328D6.1, AL627309.1, TRNP1, MYH14, PLAU, MT2A, ZFP36, ANO4, RDH13, SHANK2, S100A11, SLC2A14, BCAM, ARHGAP23, HIST1H3F, PLXND1, TNFRSF12A, FGFR2, ORC1, POLE2, CAV2, GARS, WDR75, FHL2, CTH, STIL, ASS1, DNA2, DIAPH3, PIP4K2A, BUB1B, WEE1, CDT1, PHLDA3, COL13A1, LCAT, FADS3, ST3GAL1, PREX2, ST6GALNAC2, PIEZO1, RPS6KA1, BRIP1, DTNB, TET1, NLRC5, SLC37A1, UBXN7, PRR14L, MYO6, HTT, ARHGAP11A; **Cluster 16:** CNTN1, FAM65C, PHF15, PTPRN, TRHDE, SNCB, ITM2A, BRINP1, RUNX1T1, CADPS2, CHGB, SUSD2, CHGA, CRISPLD2, FOXF1, DHDH, TLE6, COBL, GHRHR, RARRES2, CXCL12, KCNA1, NT5DC3, KIAA1244, POLR3G, FGF1, STEAP3, SLC30A3, LRAT, TSHZ3, FABP3, TWIST1, CHST8, HIST1H4B, MT1G, BFSP1, GRPR, KDR, CGNL1, FIBCD1, GMFG, PPP1R1B, HSPA12B, DOCK2, ETS1, HRK, GPNMB, AOX1, NAAA, PNMT, PPAP2C, VASH2, CPNE9, DCDC2, SLC43A1, ADRA2A, MX1, FOXH1, FGF19, ACTG2, PITX2, RASEF, ZNF488, CLMP, LDHD, CA5B, KCNAB1, SCN9A, WNT10B, NPTX1, SYT12, HES3, FAM53A, UGT8, IDES, MGAT4C, B3GALT5, LCN12, THBS2, FAM78B, FAM179A, RYR1, UAP1L1, HIST4H4, HIST1H2AL, ZYG11A, HLA-G, RPL31P11, RPL13AP7, CKMT1A, HDGFP1, RP1-228H13.1, CKMT1B, CCDC169, RP11-603B24.2, UNGP3, GABRP, EDN2, MARVELD3, THRB, MAPK13, MIXL1, RPL9P29, CHD5, CHRNB4, REPS2, GPR160, SSC5D, AIF1, RP11-1396O13.1, RP11-74E22.4, GRID2, GNA14, RIDA, D1, TMEM151A, TRPV2, RNU5B-1, DYNLL1P1; **Cluster 17:** CASP10, IL32, CDH1, LAMC3, STYK1, GAL, NMRK2, NOS1, CYP26A1, KCNK6, PNPLA5, SLC17A9, SLC52A3, RSPO4, ACP5, FA2H, ESRP1, MYL7, SCNN1A, RASAL1, AIM1, TFCP2L1, MLPH, IRF6, FAM124B, GRM4, RAB17, SH2D3A, AP1M2, ZSCAN10, FAM83F, VRTN, DYSF, LECT1, LCP1, CUZD1, ADAMTS14, HERC6, RAB20, GRB7, CBLC, VWDE, VENTX, ACOXL, B3GNT7, CDA, LAD1, GPR114, BNIPL, CXCL5, CDCP1, HPGD, FAM160A1, ELOVL7, GLB1L3, SCNN1G, HRASLS5,

SERPINB9, PAH, UTF1, GPR64, C1orf172, ZFP42, PKP3, PDE6G, C6orf132, PLA2G2A, RUFY4, KLRG2, CARD11, POU5F1, C9orf135, MT1H, RGL3, TRBC2, POU5F1B, LTB, L1TD1, SPIB, PRSS22, ACPP, RUNX3, EPN3, LAMC2, ATP2C2, MYBPC2, OXT, VSIG1, VGLL1, ESRP2, HPN, SGCA, BSPRY, C3, RBBP8NL, GALNT6, ADAMTS17, DPEP3, TINAGL1, KISS1, C2orf54, B3GNT3, TMPRSS2, TACSTD2, IGLL3P, FAM209B, KIAA0040, C1orf210, RP11-2C24.5, MT1L, RP11-694I15.7, BTK, IZUMO2, FAM162B, RN7SL471P; **Cluster 18:** SCT, AFP, TMPRSS11E, FLT1, CLC, AMBP, NANOG, OPRD1, RARRES3, TRPC6, CD109, IL34, AIRE, CAPN13, PF4, KLKB1, UCMA, GPR182, NLRP7, KLK13, ITGAM, PYDC1, FAM71D, RP11-683L23.1, TDRD12, RP11-458F8.3, LCK, GDF3, U52111.12, C16orf54, TRDN, HBA2, HIST1H4A, MME, RNY3, RP5-886K2.1, DPPA5, NPM1P46, EEF1A1P16, EEF1A1P15, KB-1683C8.1, RP11-420L9.2, RN7SKP127, HLA-DPB2, RP11-490K7.1, TDGF1P3, IPPKP1, RP11-543B16.2, AC005077.12, PPIAP19, KRT18P1, RP11-185B14.1, RP11-382D8.5, RP4-612B15.2, CRYGFP, AC097523.1, GAPDHP65, RPL30P4, HIST2H2BB, RPL17P20, RN7SL585P, AC006445.6, RP11-392E22.3, RN7SL532P, RP11-468N14.7, RP11-384K6.4, EEF1A1P17, OLA1P3, RP11-297M9.1, RN7SL136P, RN7SL600P, NRBF2P1, AP001888.1, HNRNPA1P52, CTC-412M14.6, RP11-180P8.5, AHSP, CTA-242H14.1.

EXAMPLE 12

CLUSTER ANALYSIS DATA OF GENES DIFFERENTIALLY REGULATED IN DIFFERENTOATED CELLS

Gene differentially expressed across four more mature cell populations (corresponding to the sequentially generated cortical layers) obtained through differentiation of NE, ERG or MRG cells to neuronal like cells (NE/ERG/MRGdN) and astrocyte like cells (LRGdA) derived from the LRG stage, which are shown in Figure 20D, were grouped into 12 clusters, and the genes included in each cluster are provided herein:

Cluster 1: CD74, CAPG, DRD4, LLGL2, CLUL1, CTSZ, WFDC2, CBLN1, MAP4K1, TNNT1, PRX, CDH23, CCL2, KRT18, PERP, STC2, KCNQ4, 2-Mar, CHST8, PAIP2B, EDN2, GDF15, LIN28A, SERPINF1, DDB2, MMRN1, INHBE, ZIC5,

PPFIA4, ZNF385B, STAC, SRCRB4D, RGS10, ELF3, SYNPR, ANKRD33B, CDC42EP5, COL3A1, FAM84B, FSTL5, INPP5D, ROR2, CMTM8, MGMT, RASGRP1, RHOD, KCNIP1, NEB, CLDN6, PMEL, SAMD5, AP001065.7, ULK4P2, CTD-2550O8.7, SCML2P2, HFE, DCN, MTMR11, COL9A2, CAPN6, NDRG1, CHRNB4, OLFM3, IL13RA2, LRRC61, PPP1R1B, DUOX1, ITGB7, PRDM16, WNT9A, LIX1, HKDC1, MMP14, SCUBE1, PTH1R, COL1A2, RAB3B, REPS2, TM4SF1, ECEL1, CSPG4, TCAP, PHYHD1, RPL37AP8, SCN4B, GRM8, TH, SERHL2, SLIT3, C1orf95, C6orf99, CFB, RN7SL336P, RP11-1348G14.6, RP11-694I15.7, PDK4, YBX2, PLEKHB1, IL4R, CDK18, SDC4, BMP2, PRKCG, HOOK1, STRA6, NDST4, TPBG, EXTL1, DUSP23, ADAMTS9, RAB26, SOSTDC1, LRRC48, GAS1, SHISA2, MAPK15, OAF, BCAM, COL4A6, CDHR5, GPR143, 4-Sep, TPD52L1, SARDH, ZNF391, RASL11B, CDKN2A, SLC25A27, DDR2, KCNJ4, GRAMD1C, BACE2, EMILIN3, KLHL14, PTPRH, MYH14, ZNF185, PLEKHH2, MERTK, ARHGAP25, RAB3IL1, BTN3A2; Cluster 2: BTN3A1, ZFYVE16, EPHA3, CYBA, SYNE2, ALPK1, TRIP6, PNPLA3, PCK2, SALL4, CHRNA4, EEF2K, TRPS1, RELB, BBC3, PLEKHA4, LSR, LFNG, HSPB1, PTGR1, CUBN, WNT5B, MDM1, B3GAT2, PDGFRB, XRN1, TXNIP, SLC2A1, CCND2, KLHL29, P4HA1, PLEKHG3, A4GALT, TPST2, APOE, NOTCH2, LDHA, C9orf40, LEF1, STON2, SIK1, EPHA2, CRABP2, IL17RD, SCUBE3, CREB5, GPC3, NR6A1, LCN9, CDHR1, TMEM132D, USP43, RAB11FIP1, ZIC3, ATP2B2, FZD1, IL34, CLSTN2, RAPGEF6, HK2, TPPP3, ABCG1, COL26A1, SLC30A7, ATF3, FRZB, NUAK2, RPL39L, INTU, HSPA4L, STARD4, SP8, CRYL1, C10orf118, SMCO4, PRTG, TMPRSS5, STX3, TMIGD2, AFAP1L2, SLC50A1, SLC30A1, GPRC5C, PYGO1, NHLH1, CEBPB, FUT9, PDP2, MAF, ARID3B, VWA1, PENK, GPR3, FAM89A, PCP4, GAS6, FAM19A1, KCNQ3, ATL3, OSBP2, INSIG1, EMID1, LIN28B, ARL4C, ZC3H6, HES4, S100A4, HES5, PIM3, TPM2, ARC, CCDC160, EBF2, FTLP3, CTB-63M22.1, RP4-612B15.2, HLA-B, AC007875.3, KIAA1456, ATP1A2, GRAMD1B, CDON, SDK2, ACACB, BCORL1, IGSF9, LTBP4, RAPGEF4, COL9A3, CATSPERG, MFNG, SFRP1, POU6F2, GDF10, PPM1H, SLC29A1, CTH, FILIP1, HIF3A, PSPN, MST4, COL6A2, SLIT2, CACNA2D1, C6orf57, MAPK13, HYDIN, STC1, PKN3, FGF11, NBPF10, PHYHIP, AXIN2, AGPAT2, KIAA1919, LINGO2, ULK4P3, KCTD12, ZNF467, CHST15, UNC5C, OPCML, FLRT2, FAM43A, KCNJ11, RELN,

NTRK1, HSPA1B, NYNRIN, ASIC3, RP11-458I7.1, RPL5P34, RP11-295P9.3, SEPP1, RPL9P29, CCDC109B, MYLIP, ZIC2, PTPRU, COL11A1, FGFR3, BLVRB, TMEM101, TSPAN15, TSPO, NFATC4, BMP7, BMF, KAZALD1, GALK1, CTSC, HES1, RHOU, EBPL, IFI6, ITGA7, FBN2, COL2A1, IGFBP4, hsa-mir-1199, TTC12, HMGA2, RILPL2, WDR78, KLF10, ANKRD9, C1orf51, LY6E, MEGF6, ELFN2, SLC19A1, PIFO, CHST2, COL18A1, EFNA5, KLHDC8B, WNT7B, COL4A5, C1orf192, FCHSD1, SELM, MICA, TSPAN4, WDR54, CD9, HEBP1, CA11, EFNB1, TEKT2, CECR1, LAMA1, NXT2, EIF3E, CLEC11A, PRUNE2, PALD1, LTA4H, VEGFA, SLC12A7, REEP6, PCSK4, RPS6KA1, EPHX2, TMEM74B, C1RL, IMPA2, ACKR3, SLC18B1, GLB1L2, EMC10, KIAA1522, CCDC74A, PCDP1, ABHD6, SPINT2, SEMA6B, DDIT4, RGS14, CNTNAP2, TMEM107, C2CD4C, MAATS1, FOXO4, TPCN1, ACN9, ADA, SIRPA, C1orf228, DMD, GSTM2, TMEM159, PKP2, MYLK, PPP2R5A, ATP2B3, LRRC16A, SLCO1A2, TYRO3, MAP3K1, CDH20, PLXDC2, SORBS3, SWAP70, FRAS1, USP53, CDH12, UNC5D, SERINC5, RBPMS2, GPR37, EPS8L2, ANXA2, FMNL1, COL27A1; Cluster 3: NOS2, FMO1, PRSS8, TYR, DCT, TRPM3, TYRP1, CPZ, SCNN1A, PRSS16, CNTN3, PAPPA2, CPXM2, BARHL1, C3, GDF5, BST2, TRPM1, KRT7, PSTPIP1, DMRTA2, HAPLN1, HTR2C, CACNA2D4, ABI3BP, WNT3A, DKK2, PPARGC1B, COLEC12, SIM2, COL6A3, NEUROD6, CMTM5, CYP2S1, SOST, HIST3H3, KRT8, PRKCDBP, KRT19, AIM1L, AMBN, RPS17L, RP5-1120P11.4, CLDN4, AC004832.1, BTN3A2, TRBV20OR9-2, GMNC, TBC1D26, CES5AP1, AC004383.3, AC007386.3, AC072052.7, RP11-452D2.2, CHCHD2P6, HSPB1P1, RP11-69M1.1, ACTN3, CDC42P4, RP11-629O19.1, RP5-878I13.2, CFH, NOS1, KCNJ5, ITIH5, SIX1, ATP1A4, SLC2A5, TINAGL1, GDF7, COX6B2, LMX1A, TSHR, MEI1, TEKT1, COL8A2, AQPEP, EVX2, MAP3K19, RIPPLY3, C11orf88, DNAH2, RYR1, HLA-DOA, RP11-277B15.1, C17orf98, HMGN2P18, RSL24D1P6, AC104809.3, MCIDAS, GAPDHP52, RN7SL547P, RN7SL338P, RP11-1016B18.1, NMRK2, C19orf77, JAK3, ICAM2, TULP1, GAL3ST1, CCDC33, GPA33, C10orf11, TNFRSF14, CRYAA, FBLN2, RBM47, C2, TRPV3, SPTLC3, RBM11, MORN5, P2RX2, NCCRP1, S100A6, MSX2P1, AC004074.3, RPL24P2, INSL3, LINC00908, MLXIPL, RIMS1, FER1L4, LIM2, ACCS, BMP5, TTR, CRHR1, TTLL10, CLDN19, CDC20B, ALS2CL, RP11-5P18.10, KCNJ2; Cluster 4: GABRA3, PTGER3, DSP, SLC17A7,

CCKBR, NR2E1, POMC, NRN1, PCSK2, ADM2, RAX, PDGFRA, USP44, TBR1, C1orf222, BARHL2, LHX9, OSR1, DCDC2, CYP39A1, RSPO3, GABRA2, VWA5B1, DMKN, CTRC, EOMES, SAMD3, CLDN3, PCSK9, RSPO1, SLFN11, EGFL7, DMRTA1, FZD8, LRRC3B, FAM83H, SCN5A, ADRA2C, HIST3H2BB, C6orf141, GPRIN2, SP5, RPS26P8, RP11-296L22.7, DMRT3, ATP2A3, TP73, LHX5, EVX1, LGI1, GHR, PTPRB, APOC1, ZDHHC8P1, EMX1, ADCYAP1, RGS16, RSPO2, GJD2, PLXDC1, UBXN10, KCNC2, WDR16, IGF2, PARM1, EMX2, TRH, PRL, FOXG1, NHLH2, TMEM132C, NEUROG1, C9orf171, MYBPC1, RP3-340B19.2, PABPC1P3, RP11-2C24.5, ARX, EFCAB1, WNT8B, SPAG6, MECOM, SLC7A8, OTX1, GALNT12, LMO2, HDC, VAX1, FEZF2, FGF18, PRKCB, RHOH, GREM2, RTN4RL1, GCNT1, FAP, ADAMTS2, LHX2, EFEMP1, ACOT2, LYPD3, PACSIN1, TMEM255A, DSC2, FHAD1, THSD7B, OCIAD2, ME3, DRC1, CCDC135, C1orf115, OTX2, GFRA2, KCTD19, DCDC1, NPTX1, HLA-DQB1, LYNX1, SLC35F3, AC016757.3, MPPED1, C1orf168, CD55, AJAP1, ZNF204P, HLA-DPB1, SLCO4A1, ATP8B3; Cluster 5: ALDH3B1, LAMA3, SEMA3C, GADD45B, PLS3, HPX, CTGF, LTBP2, EGR1, ECHDC2, ANXA11, SNAI1, RRAS, HELB, ZFP36, KANK4, POSTN, VAV3, EMP1, ANXA1, HEY2, GPNMB, ITGA11, MYOF, ANXA3, HAPLN3, SYTL1, CYR61, CSRNP1, IGFBP3, PLOD2, TTC39B, CREB3L1, FAM46B, ACTG2, S100A11, CLDN1, AQP3, PKD1L2, SPDYE6, TAC3, TMEM88, EFNA1, WNT10B, SEMA3E, COL24A1, ANO5, RP11-79L9.2, P2RF, ERBB4, TMEM64, TXNRD2, SOCS3, IFITM1, THBS2, PLEKHG4, MME, RP5-886K2.1, HMGN2P15, RPL41P1, PRADC1P1, RP11-182J1.16, RN7SL473P, TNFRSF12A, SCT, CACNG5, CYP26A1, BAMBI, MXRA5, CPED1, PTGDS, COL1A1, GBP1, MSX2, COL5A1, BHLHE40, CYP1B1, RBMS3, SLC39A4, PAMR1, ACE, SGPP2, FAM160A1, C6orf132, LAMA2, TCEA3, VGLL3, AC055811.5, ST6GALNAC4P1, DNM1P51, NONOP2, RP11-598P20.3, RP11-359M6.3, AC006116.20, CLDN11, SNAI2, PLD1, TFAP2C, UNC5B, UNC93B1, CD83, OLFML3, PPP4R4, AHNAK, WNT1, DOCK6, C19orf66, GFPT2, IL13RA1, SLC39A11, IER3, ARRDC4, ATHL1, GJA1, GRAMD3, LCA5L, TGFBR2, TNFRSF11B, PLEKHF1, RILP, COL22A1, PDGFD, RARG, NTF3, HLA-E, RGL3, AC007875.2, AP001055.1, IL32, FOSL2, RARB, COBLL1, LEPREL1, LAMB1, ENG, BVES, AP1M2, CNN1, DRAM1, ARHGAP29, SLC40A1, TAGLN, ACTC1,

ZMAT3, NABP1, MSRB3, AHNAK2, OLFML2A, TRPV2, FAM114A1, RPL35P2, SLC12A8, EPPK1, PTCHD4, VRK2, ACP5, SMAD6, RGS5, CCDC81, EDNRA, DAB2, MSX1, SLFN5, TAP1, SYNPO, BNC2, EFCAB4A, CCDC162P; Cluster 6: CYP51A1, TFPI, CACNA1G, RPS20, LYPLA2, PTBP1, KPNA6, KCNG1, H6PD, RCN1, MSMO1, DCBLD2, YBX3, QSER1, POLD1, RPL18, WDR18, FOXJ2, MPPED2, PDZD4, ASNS, CNGB1, RPL31, FAM50A, MRPS34, TRAF4, DGKD, PAK3, FBLN1, N4BP2, FDFT1, MRPL22, RPS5, PPIE, EFR3B, MRPL28, FTL, C14orf166, ASAP3, RPL6, GOLGA3, AARS, ALKBH5, PHGDH, SH2D3C, SCD, MYO9B, ATP5D, CDC34, RANBP1, SMARCB1, NHP2L1, DDX17, RAB36, TMED8, SPTLC2, LGMN, PNN, NOL4, MIB1, CENPI, FNDC3A, NDRG4, HCFC1R1, NME4, METRN, SQLE, BNIP3L, OAZ1, RPS16, SLC1A5, RPS19, BCAT2, RPL18A, ISYNA1, FKBP8, ZKSCAN1, FSD1L, TGFBR1, SPOCK2, NPM3, RPL28, MLLT6, PSMB3, RPL19, DHX40, ENO3, RPL34, SC5D, CCDC86, CBL, MDK, NDUFS8, C12orf57, MCM3, RPS12, COX7A2, HMGCS1, HMGCR, AMOTL2, RPL24, ACVR2B, KLHL24, RPS15, SDC1, DHCR24, FAM20B, HDAC1, MIIP, MTR, CDC20, AKR1A1, PRDX1, CNN3, ARID1A, RPS25, FOXO3, CCNI, TJP2, TRIM67, PPP2R4, IFI27L2, NPC2, IRF2BPL, DUSP1, MOB3B, RPL21, RPL5, BBS9, PRDX4, PFDN5, NR4A1, FAM199X, LPGAT1, AGO2, CKS2, ATP8A1, RPS10, SOX9, CHCHD5, ALKBH7, RPL23, SNRPD2, COX6B1, BCL2L12, PRMT1, UXT, BCL11B, MASP1, CHTF18, AKAP9, POR, SAT1, LDLR, LRCH2, RPL36, PGLS, LSM4, TAF4, CHMP2A, C16orf13, EIF3G, COX4I1, RPL27, ACLY, VPS25, KDM6B, OSER1, KRAS, LPIN1, CRB1, SOX3, DTNA, PSAT1, ATP5G2, CD63, KCNMB4, RC3H1, GCC2, ARHGEF4, RPL35, GRHPR, IGFBPL1, RPS6, PIM1, HMGA1, MAPKBP1, RPLP1, IFI44, RPS24, ATIC, TMEM117, CDH11, RPS2, TOB1, RNF165, SH3GL1, ZFP14, MYO1F, ZNF787, RPS11, RPL11, RPS8, IGSF3, XPR1, MCL1, POGZ, SYT14, SLC39A1, KCNN3, GATAD2B, RHOB, SLC4A10, RPL32, STXBP5L, ATP10D, KLHL8, RPL37, MEGF10, BOD1, ABRACL, SLC16A2, PLIN2, RPL7A, ZEB1, LRRC4C, CELF1, RPS3, B3GAT3, FAU, TMEM219, THYN1, RABGAP1L, ASTN1, DCLRE1C, ZSCAN1, CENPH, CMIP, MSI2, FZD7, PSD3, C14orf2, UQCRB, RPL30, UBE2L6, KAT6B, LRP8, TSC22D3, C9orf91, TMSB15A, NBL1, PTMS, BTG2, CHCHD6, SSU72, CBS, CSTB, LSS, IER2, RPL8, MPP3, RPL26, RPL29, ZYG11B, VANGL2, INHBB, KRTCAP2,

UBXN7, RAD54L2, HMGB2, TLX3, RPS14, GPR85, PTTG1, BRI3, FOXK1, STOML2, SLITRK5, AMER2, SPINT1, RPL27A, MCM7, PLK1, DUSP18, GPX4, YIF1B, DAPK3, HDGFRP2, C19orf48, ICT1, TK1, ECI1, ABCA3, SLC3A2, RPSA, RBPJ, DUS1L, USP47, CDCA4, RPS9, INSR, ATF7IP, FBXL14, RPS21, RPS7, C1QTNF4, BSG, MRPL52, RPL38, HECTD4, TRMT112, ADCY5, MZT2A, STOX2, UQCRH, CBX2, MRPL11, BRSK2, FGFBP3, UBE2C, PACS1, CSRP2, TP53I11, ARL10, CCDC85B, LRRN1, SOX11, MTHFR, RPLP2, CHID1, RPS27, IMPDH2, EIF3K, AURKB, DPM3, CYC1, CDH4, C17orf96, MYADM, CITED4, FZD2, ZNF609, SSR4, TRAPPC5, MRPL41, SHMT2, YBEY, CEP97, RPL35A, SPATA13, CADM1, RUVBL2, PRR14L, FBXL7, CBX6, CNTN2, DUSP8, ARMCX2, SUMO3, SIVA1, BRI3BP, MAFF, ZNF445, METTL7A, POMK, KRT10, PDE2A, RSBN1L, ZNF70, NCR3LG1, H2AFX, RPS2P46, FAT4, TOMM7, ARID5A, LAGE3, PDGFA, ZNF460, RPL37A, MYO1C, RPL12, VPS13A, RPS4X, 5-Mar, ZNF770, RPL23A, GFPT1, NRARP, MDM4, TTC37, DLL1, RPL10A, MT-CO1, KIAA1737, SREBF2, RPL39, APRT, ARMCX6, LSM2, HLA-C, GNB2L1, NT5M, MT-RNR1, ARL2, RP11-742N3.1, RPL41, RPS18, RPS28, EIF1AXP1, C4orf48, NME1-NME2, C1orf213, RP11-603B24.2, GAN, UBBP4, RPL17, RPAP3, CLCN6, CALCOCO1, TACC3, CAPN1, DEPDC1B, MYO16, MXD1, ISOC2, ABCA7, ANKS1A, NDUFB4, PPP2R5B, ST6GALNAC2, VASH1, DAZAP1, PDCD2, LPHN1, HMMR, MCM2, CLASP1, NUP37, MCM6, BAZ2A, UBE2T, TNRC6C, CBFA2T2, KIF22, REX3, KAT6A, BAX, DOCK3, GCN1L1, RPLP0, KCNH4, TNRC6A, SLC22A17, HNRNPC, CREM, CDKN3, PFDN4, CDC25B, AHCY, LIPG, RBBP7, EEF1D, OLFM2, FBL, CLIP3, PTPRS, GRIK5, SIPA1L3, NUDT1, ABCA2, NSMCE4A, LGALS3BP, TMEM104, RPS13, PARP11, ATN1, LDHB, MDGA1, GMNN, DNPH1, EEF1B2, CLIP4, ODC1, RPL22, MAD2L2, RIMS3, RPA2, KMT2A, CPSF3, CTNNAL1, DNMT3A, BCL11A, FABP3, TRIM24, CSE1L, MRPS26, TRAP1, TIMM17B, DLGAP5, COX7C, TNFRSF19, PAICS, CDKN1C, KIF1A, PRRC2B, TMEM160, PLXNA3, RAB11FIP4, FLOT2, CTNNBL1, DSTYK, CCNB1, MYCN, RPS15A, MSI1, KLHL36, TTYH3, ANP32B, ARHGEF39, IFT172, METTL5, FAM117B, HADH, CLSTN3, ANP32A, SEC11A, NOB1, ANAPC11, BRD4, NFIC, COL6A1, IGLON5, CASQ1, RPS27A, HES6, STK11IP, RPS3A, BTF3, RPL7, SIGMAR1, SEC16A, NOTCH1, SSRP1, TENM4, SIDT2, MZT2B, TMEM163, ZIC1,

CCSAP, MARCKS, LARP1, ADAMTS3, CCT8, BUB1B, DYRK1A, SKI, CACHD1, NAE1, SPON2, RDH13, TONSL, USP24, FLVCR1, DAPL1, HIPK1, RPL9, RNF168, H2AFZ, CDC25A, NDUFAF2, MED30, UNCX, PSIP1, WNK2, PRDX3, GOLGA2, STIM1, MVD, KIFC2, HID1, SRRM2, CPLX1, HINT1, BUB1, UBE2E1, USP38, LRRN2, TRIM8, EXOSC1, SUCLG2, HCFC1, CHCHD1, DHCR7, CKS1B, RPL15, VCPIP1, LPL, RPRM, AGTRAP, ZBTB41, KLHL11, COX5A, ZHX2, DCTPP1, PPA1, GPR137C, SOX2, SNRPE, NIPSNAP1, CMSS1, HIST2H2BE, H1FX, SORCS2, NAT8L, RPS23, KIAA1598, AGRN, TUBB4B, GTF2F2, SPATS2L, PLXNB2, PCNXL3, ADH5, SPG7, WDHD1, OSTC, PJA2, C6orf25, CCHCR1, DDR1, ATXN2, ZBTB10, AP000350.4, ARHGAP23, CTD-2328D6.1, RPL36A, H2AFJ, HMBS, FLJ00418, RP11-574K11.27, MPND, KAL1, SLC25A39, PLEKHA5, TMEM161A, TRAM1, GSTP1, CPXM1, CRKL, ADRBK2, CST3, TUSC3, PLIN3, RABAC1, P2RX3, CD81, ACAP2, FNDC4, IGFBP2, OBSL1, IQSEC2, STK35, ABHD17A, NINJ1, CTSL, CASC5, TET1, LGR5, MFGE8, FKBP10, ADAM15, BOC, PAM, CHST7, GSTO1, SERPING1, ATM, FARP1, LSM11, RHPN1, IGF2BP1, CCNF, CLPB, IER5, BDH2, SPRY1, CITED2, BRWD3, IFI27L1, TMED3, HSP90B1, MFSD3, SERINC2, RP11-220D10.1, GLB1, ORMDL3, B3GNT1, ZBTB18, HIST1H2AC, SKIDA1, GLRX5, GSTT1, HIST1H1C, ZNF292, INPP5F, SYT15, TMEM256, HLA-A, RP11-848P1.7, CACNA2D2, CYB561, FAM65A, AKR7A2, PLEKHH1, PFKP, HYAL2, ATP2B1, RIF1, ZFAND6, NAT14, GALNT16, PGRMC1, NECAB2, TMEM205, TMEM147, POLD2, INTS2, B9D1, EFNB3, MAP2K6, ELP4, KHDRBS2, ERGIC1, MAPKAPK3, TMEM9, PLAGL1, ENKD1, MPST, CDO1, STK33, BEX2, TMTC1, PPP1R1A, FOXP4, CPNE2, SCRN2, ARRB2, RPL13A, PHIP, EBP, CETN2, IMMP1L, UHMK1, CFDP1, OXA1L, ELMSAN1, UBXN11, C21orf59, UBXN1, FABP7, TMUB1, HGSNAT, DPCD, RPL13, RCOR2, CETN1C, UBB, DENND5B, PAQR8, ZNF672, TBCA, VAT1L, LRRC20, NDNF, VPS37D, KIAA2018, C8orf59, MLF1, NDN, NR2F2, MORN2, RPL14, ARID2, C20orf96, MYO6, HIST1H2BK, ZNF587, ZNF521, FANK1, FAM19A5, PLEKHM1, WBP1, ARHGAP44, EXTL3, EML1, KLF6, NCK2, FSCN1, FGFR1, TPX2, SLC25A1, DICER1, MID1, ABCC1, STX10, TBC1D9, ECT2, KIAA1324, PKN1, AHDC1, TMEM8A, JUND, PAK4, TOP2A, SPATA6, RIN2, WASF3, CDCA8, GLUL, RAC1, NRM, KIF11, CENPE, REM2, EFHD2, ALDH9A1, RALB, CDCA7, MYO10, RNF44,

DENND2A, TCF7L1, KIT, CACNA2D3, WASF2, RACGAP1, WNT4, PKDCC, SRGAP2, CDCA7L, MID1IP1, DACH1, TMEM130, C18orf54, TMEM194A, TTYH1, RAC3, CDK1, VEGFB, DPY19L1, GNG7, CLVS1, JUN, GRINA, LDLRAD3, WSCD1, BAI1, 9-Sep, ADARB1, SH3BGRL2, CLIC1, C1orf226; Cluster 7: SYT13, PRDM1, TLE2, CD82, SLC17A6, SEZ6L, GPR50, CALB1, RASD1, INPP4B, PTPN6, CSMD2, SSUH2, STAT5A, ASS1, GMFG, KIAA1644, FOXN4, CBLN2, PPP1R1C, PPP2R2B, ALPL, BRINP3, HEYL, TMEM74, PTCHD1, SLC7A3, ZNF488, ISLR2, RCAN2, CHST1, BEGAIN, SLITRK6, TMEM173, DLK1, CXXC11, HIST1H4A, HIST4H4, SMOC1, GRIN3A, RN7SL262P, WI2-2334D6.1, ZMYND10, DNAH9, TFAP2B, SNCB, C20orf26, PPP1R17, KIAA1244, GPR83, CCNA1, LMX1B, NR4A2, FGF17, ZFYVE28, SLC6A5, GSX1, VSTM2A, KCNS3, CALB2, ZIC4, RTP1, ST8SIA3, CPNE4, C1orf186, CACNA1E, BRINP2, SFMBT2, ONECUT3, RPL5P18, RP11-408A13.1, RP11-331F4.4, CTC-512J14.5, CTBP2P7, RP11-460N20.7, SYT7, SLC11A1, OAS1, FAM83E, GABRG2, SLC30A3, IQSEC3, ISLR, MTUS2, CDH18, CRH, VCAM1, SHH, NSD1, INSM2, RP11-551L14.1, RGS7, GRIN2A, FAM162B, ZBTB7C, C11orf87, BEND4, PTPRT, SLC6A17, RNY3, DAZAP2P1, AC093106.5, SALL3, CAMK1G, RPH3A, RIMS4, MAP1LC3A, C7orf63, NMU, PRRX1, CLYBL, DPP6, SLC43A3, GPR6, ADHFE1, FAM84A, NEUROD1, MYO1D, OTOS, FAM43B, TCEAL2, LEMD1, ERICH2, LRRC10B, SLC18A1, PARP12, EYA4, NKX2-1, SLC13A5, FZD9; Cluster 8: ICA1, CROT, NFIX, AKAP11, VCL, TLN, DKK3, EIF2AK2, KMT2C, ATP2B4, LZTS1, TRHDE, FRYL, TMEM131, PAG1, GPC4, ACER3, TNS1, LRP2, EPB41L3, NCOA1, CCDC80, XYLB, FKBP5, LGALS1, MYL12A, DGKH, KATNAL1, HSF4, NFAT5, QPRT, IMPAD1, PLAT, ITGB8, GLIS3, PDLIM1, SFXN3, MED13, CLCN3, CPT1A, HIPK3, PPFIBP1, TENC1, CUX2, LOX, PCDHB5, CDH6, LIFR, WNT5A, FOXP1, TANC1, ERRFI1, ASH1L, SLC35D1, GADD45A, TROVE2, NID1, PRDX6, AVPI1, TSHZ3, BICC1, MED13L, TOX2, PMEPA1, C1orf61, AGO3, CDC42EP1, PODXL, ATP1B2, PVRL2, ROCK2, LRP4, MDM2, SKIL, KLF4, IFI44L, SECISBP2L, NAAA, CNOT6L, JDP2, FAM181A, TPM1, MAPK4, IFITM3, ST6GAL2, CNGA3, COL8A1, SFRP2, PLK2, SDK1, NFIB, SERPINH1, ARID5B, LPHN3, PRSS23, ANKRD50, ZNF827, TNIK, STARD9, ACOX1, SHANK1, ADCY9, RFTN2, GPR155, GPR98, HEY1, FAM219A, ABCA1, PCBD1, TUB, CASC4, PLEKHA2, ANTXR1, FOS, CYTL1, GAA, CXXC5,

ID4, LRRN3, PEAK1, HEG1, RNF213, PHLDA3, PRIMA1, ZDHHC21, NCKAP5, ADAMTSL1, PLEC, TSHZ1, RNF182, AEN, FAM181B, PAPPA, CAMK1D, UPP1, KREMEN1, PCDH9, ROBO2, IFITM2, POU3F1, HEXIM1, ISG15, PTAR1, PDLIM7, SRGAP1, SERTAD1, PGAP1, ZNF624, HHLA3, FAM49A, MAN1A2, HELZ, BMPR2, PTCHD2, C5orf51, TAPBP, ZNF585B, FRRS1L, RP11-173M1.5, AFF4, NOTCH3, SEMA3A, CELSR1, GPR137B, OPHN1, MYOM1, SMAD7, SGK3, LAPTM4B, WNT3, RAB34, PPARGC1A, SH2B3, RFX4, NEDD9, SPARC, NR3C1, WWC1, RBP1, PIKFYVE, IGFBP5, FHL2, ID3, CHST3, CHPF, USP9X, FOXP2, PDE11A, MPP1, PTGFRN, GNS, RGS8, PAX3, TFAP2A, SCARB2, PHLDA1, GAS2L3, IGF1R, NKD1, CREG1, RAB13, ITPKB, UNC80, MTFR2, ADM, WNT7A, PIEZO2, ARSE, UBN2, MEGF11, FNDC5, AGPAT3, FDXR, NEXN, SLC15A2, EDIL3, IKBIP, CERCAM, LTBP3, SOGA2, PGM2, SDC2, NRIP3, TOM1L2, CRIPAK, DDX60L, AP1S2, ZNF703, MAML2, RPS27L, ERAS, ZNF573, SLC35F1, GAL3ST4, LRP10, TEAD4, COL5A2, AC092675.3, PET100, C9orf172, APOBEC3C, RP11-266L9.6, RP11-51F16.9, GRN, ANO2, TNFRSF1A, LAPTM4A, EPHA6, IGSF9B, PLOD1, CD59, MMP2, HEPH, PLTP, CDK6, MAP3K8, MDFI, FN1, WLS, CA14, SLC16A7, SMAD9, ADCY7, SLC12A4, CDKN1A, ID1, SORT1, PRCP, FURIN, EMP3, SELENBP1, LHFPL2, GSN, RHOC, TAGLN2, SHC1, PDPN, IGSF8, PBXIP1, CTSB, BAALC, C9orf89, ZNF503, UBTD1, HTRA1, B2M, LAMB2, FIGN, NOG, CSF1, PROS1, NPIPB4, ANXA6, HTT, COLGALT2, S1PR3, RP11-274B21.2, PRR24, CTD-2021H9.3, CFLAR, PLXND1, ITGA3, PROM1, MRC2, SAMD4A, GCLM, LIMA1, FOXN3, RC3H2, PPP1R12A, MYO9A, ROCK1, MAST4, NEDD4, PTPN21, MCAM, PPP1R12B, SERTAD4, OSBPL8, MICALL1, MYH9, EP300, PCNX, DNAJC3, PIEZO1, DENND3, ZC3HAV1, PRKAG2, ACTA2, SH3PXD2A, SHOC2, GAB1, MAN1A1, TMEM30A, LAMA4, MAN2A1, KAT2B, GNB4, ITGA4, PKD2, SATB2, C1orf198, TGFBI, FAM126A, FKBP9, GLIPR2, ARFGEF2, PALLD, ZSWIM6, C1QL1, RARA, RHPN2, TRIM22, LOXL2, ARHGAP32, TMEM2, TLN1, SYTL2, SORL1, THBS1, AMIGO2, URB1, RGL1, LYST, LPP, CAMK2D, SPATA5, SETD7, BDP1, TMEM47, GABRQ, NACC2, SLC5A12, CAPN5, LATS2, LYPD1, CRIM1, WWC2, PTPN14, DDAH1, SAMD8, B4GALT5, LARP4, NFIA, LMBRD2, RHOBTB3, DLC1, TP53INP1, KIAA1958, SLC16A9, ARHGAP42, ADAMTS15, NAV2, VWCE, HOOK3, STAT3, ADAM9, RALGAPB, PFKFB3, ZNF440, AFF1,

LCLAT1, NUDT4, DPP10, SOCS4, POLR2A, HS3ST4, GPC6, WBSCR17, B3GALTL, TSC22D2, POTEF, IGF2R, S100A10, CASC10, LGR4, ZNF611, RP11-274B21.4, MICAL3, NPIPB5, RRAGD, BCAR1, BCAT1, ITGB5, ABLIM1, KLHL4, PMP22, FAM46A, EVA1A, ID2, LOXL1, CASZ1, FLNB, TNFRSF21, MGAT5, PLEKHG4B, STEAP2, FSTL1, NNMT, LDLRAD4, PXDC1, KIRREL, GAS2L1, SAMD11, TMSB4XP8, POTEI, AL627309.1, TMEM158; Cluster 9: CD44, CP, ASB2, GHRHR, CRYAB, FOSB, EMILIN2, GALNT6, ADAM12, HNMT, EVA1C, FOSL1, MGAT4C, DPYD, CFI, AL590762.6, RP11-460N11.3, ARHGDIB, FGF1, HOXB3, RHCG, ANKRD1, OR51E2, CCDC36, SSC5D, AFP, BATF3, GFAP, A2M, CHSY3, HIST1H3A, C3orf72, RP11-289I10.2, RANP1, AL807752.1, ARAP2, SP100, SYNJ2, HOXA9, PDGFB, EBI3, CAV2, CAV1, SERPINE1, SLC35F2, STEAP3, NPPB, LRAT, CPA4, ITGB4, NTS, DOCK2, PARP9, GLIPR1, FBLN5, NLRC5, PLAC8, FLI1, HSPB8, MAP3K7CL, RUNX1, SLC25A45, FAM198B, TMEM100, GREM1, HR, IL7R, MT1E, PDE3A, ARSJ, C12orf56, CLDN24, COL14A1, C15orf52, COL13A1, PDCD1LG2, ALPK2, LAYN, HMGB3P11, RP11-302K17.3, AC068279.2, RP11-255N24.4, RP11-336N8.4, RP11-745A24.2, HOXA10, CTD-2014B16.3, AC093323.1, CTC-412M14.6, CASP8, CRHR2, AOX1, IL6R, ADRA1D, C5orf46, RGS6, COL25A1, RBM20, TBC1D3B] Cluster 10: ARHGAP33, CREBBP, CCDC124, PSMB1, SPAG9, MAMLD1, STMN4, RUFY3, POU2F2, SARS, ARID4A, TMSB10, TIMP2, TRIO, AP2S1, KIAA2022, SLC4A8, PHPT1, PTPRN, KIF1B, ATP9A, DSD, HIPK2, SLC12A2, ANKRD44, HSD17B10, CRMP1, CA12, ZZEF1, C19orf10, KCNQ2, RASAL2, CACNG4, KIFAP3, DCX, CAPZB, RUNX1T1, SMARCA2, SESN1, KIF3C, ORC1, PPP1R15A, GNAO1, C20orf194, ERP29, SPAG7, ORC6, AGO1, DPYSL2, MRPS18A, ACOT7, NDUFB7, POLR2E, 3-Sep, XBP1, TNRC6B, PHF5A, MYBL2, EEF1A2, TRIB3, MANBAL, DNTTIP1, ANKRD12, PCSK1N, ZNF423, COTL1, SLC7A5, CORO2B, AP3B2, STMN2, TUBB4A, NOVA2, C19orf53, FAM32A, CCNE1, CDC37, GRIN2D, DKFZP761J1410, RAB3A, COPE, SSBP1, COBL, GARS, BCL7B, FKTN, MEGF9, EDF1, ARHGAP21, SORCS1, AATF, RUNDC3A, PFN1, CPE, FRG1, TRIM2, NRXN2, DDX6, CARS, AIP, ALDH2, TMEM14C, SERINC1, CPEB4, PRKAR2A, INO80D, LAMTOR2, KIF21B, MRPS15, HMGN3, CNR1, KLF12, GRIA2, MTHFD1L, DPPA4, SLITRK3, DDX39A, PLP1, SNRPC, PPDPF, SNRPB, DYNLRB1, RBM42, PRDX5, PIN1,

ECHS1, SHFM1, ATF4, LRRC17, VPS13C, CDKN2D, GAMT, NSUN5, TULP4, OLFM1, SESN2, ATPIF1, GSE1, KIF3A, MAP1B, PCNA, SYT11, CHL1, KIDINS220, YARS, NREP, CCDC64, CCT7, KIAA0513, EGLN1, ITM2C, COX5B, NACAD, TXN, STXBP1, PSMB7, MYC, CNPY3, TUBB2B, UQCC2, PPP3CA, KIF21A, PPFIA2, SETD1B, VPS37B, HERC2P2, ZFHX3, CMTM3, NDUFB10, EIF4A3, ADAMTS10, PSMB6, POU2F1, SNX27, VASH2, NVL, SOX13, PLEKHA6, TMEFF2, NBEAL1, IQSEC1, NDUFS6, PPIP5K2, ST18, LIN7C, ZC3H12C, NCAM1, SOGA1, GPM6A, WDR17, DLG5, ACSL1, SETBP1, EPG5, UBALD1, PRKCA, AFF2, KIF5A, EXOG, SMG1, RBPMS, CCNB2, SLC35B2, PSMG3, TPRG1L, KCNB1, DMTN, PSMD4, CELF3, U2AF1, CELF5, C16orf59, ZSWIM5, JAK1, DRAXIN, NFASC, CLASP2, CADPS, WDFY3, NDUFS4, ACSL6, KIAA1161, PGM2L1, KBTBD6, DACT1, SALL2, PSMC3, DCHS1, CRABP1, BLCAP, MAP1A, SNRPD1, TBC1D16, ATCAY, MLST8, FTH1, POLR2G, PAFAH1B2, KIF5C, DTYMK, NDUFS5, RNF181, ATP5I, ATF5, MN1, ZRSR2, ONECUT1, CTNND2, OTUD7A, SF3B5, CHD3, NSG2, RNF150, LONRF2, SMAGP, C9orf16, GRIK1, SCAND1, PPIH, GAP43, MOB3A, POP7, MYEOV2, PPP1CA, NBEA, MRP63, FKBP2, SPTBN2, GNG5, NDUFA11, ZNF25, DRAP1, RAB6A, AURKAIP1, SPHK1, CDK5R1, POLR2L, C12orf68, ASB8, PDE4DIP, GADD45GIP1, NOP10, RGMA, ARL6IP4, SOX1, ABAT, BCOR, MRPL54, UQCR10, TSPYL2, 5-Sep, NDUFA12, FAF1, NRG3, TRIM69, NDUFA13, GNG2, NF2, LRRC16B, MAPT, RPS19BP1, MRPS21, DCC, TET3, THSD4, EIF4EBP1, DNER, LAMTOR4, LDOC1L, UROS, BLOC1S2, MYT1, LCOR, XRCC6, ARHGEF12, ZNF398, MRPL21, PHF2, RPS26, UBL5, ASNA1, ZNF536, STK39, SLC9A6, MT-CO2, MT-ND2, MT-ND3, TOX, RUSC2, TSEN15, MT-ND1, L1CAM, POU3F3, CCDC167, PHACTR4, ZBTB12, GPR56, LCMT1, MRPL23, C14orf132, C19orf24, AC079250.1, TMA7, GPX1, AKAP2, PEG10, RP11-274B21.1, XKR7, TMEM178B, MRPL12, RN7SL2, LRRC37A16P, CICP14, LAS1L, NDUFAB1, SLC25A13, ST7, POLR2J, MAP3K9, MGST1, POLA2, MDH1, TOMM34, FUT8, PHLPP2, PSMA4, RFC2, ARID1B, FAM168A, NOP58, YBX1, HDAC4, POLR1A, C16orf80, WSCD2, SLC46A1, NAALAD2, EPB41L2, SRCAP, NDC80, PHLPP1, GSK3B, WBP11, TAF9, AKR1B1, AURKA, GNA11, SLC23A2, SEMA6A, CDC45, CEP170B, RTCB, DDHD1, NOP56, E2F1, SYP, SMS, MMP15, ASF1B, SHD, LIG1, RUNDC3B, CHCHD3, CLIP2,

PHYH, CUEDC2, BCCIP, EBF3, GLRX3, MTMR4, RAI1, PPP1R9B, MRPL27, FZD10, COX6A1, SASH1, FBXO5, OPRM1, SRF, TTK, IRX4, PDCD10, ATP6V1A, SLC4A3, SF3B14, PSMD14, PDE1A, SPTBN1, ST3GAL5, BIRC6, SRSF7, SLC1A4, PARK7, SMG7, ST6GALNAC5, RPF1, CENPF, NDUFA8, SLIRP, MAPK8IP1, KIAA1549, C19orf43, SOX4, SNRPB2, UROD, STYXL1, FAM64A, GNL3L, KIAA1683, RLIM, ZFYVE20, UTP3, ARHGEF11, NASP, SYT4, POMP, CHRM3, TIMM10, MPHOSPH6, PLXNC1, BZW2, NUSAP1, SLC44A5, COX17, SNRPF, TARBP2, DISP2, GCSH, ZCCHC14, COPS3, PIP4K2B, SOD1, DOPEY2, SAE1, LMTK3, PSMA5, CELSR2, SEMA6C, HAX1, ILF2, PYCR2, SNRPG, TPRKB, UBR3, LRIG1, TAGLN3, LYAR, C7orf50, CCT6A, NDUFB11, POLR2K, FAM57B, CCT5, AKAP6, MLF1IP, ACSS1, DBI, ADK, ZNF618, ALG8, FLAD1, ADAR, SHANK2, C11orf48, MAGOH, BOLA3, CCT3, PPM1L, THOC7, PPM1K, LSM6, EBF1, TMEM200A, FABP5, LMTK2, CSMD3, COX6C, DCAF13, RPP25L, ZNF367, RPL36AL, ATP5C1, WEE1, KMT2D, MPLKIP, CDK5R2, MAP6, COQ2, PARP14, DAG1, INSM1, RGMB, ZCCHC12, ARHGAP1, ZBTB8OS, TALDO1, SRSF3, SEPW1, CDC42EP4, C7orf41, RTKN2, CECR6, PSMG1, NDUFB1, TRAIP, CHEK2, LRRC55, GPR173, POU6F1, ZNF93, APOO, NDUFA6, MAPK11, SBK1, NDUFA4, LITAF, SLC35E2B, PPIA, ELAVL3, AFAP1, MYO18A, HNRNPAB, RPF2, GRM3, MT-CO3, CENPW, QTRT1, PAM16, NME1, ATP5J2, EIF6, RGAG4, TUBB3, MYCBP2, E2F2, PAX7, NUP160, HSPA5, CELF2, MON2, SNAP91, RAP1GAP, MAP2, SEMA5B, NKAIN1, ZFHX4, HECTD1, DAAM1, USP31, PON2, PTN, TSPAN13, NCS1, DTX4, GNPTAB, BACH2, DNAJC16, MYL12B, ATAD2B, CXCR4, ZC3H13, FGF13, SYNE1, ACACA, PIK3C2B, STARD13, AGO4, RNFT2, URB2, ARRB1, RASGEF1B, SRRM4, G6PC3, ABL2, ZNF462, KLHL35, ANK3, SACS, PDLIM3, ZKSCAN2, GAREML, AUTS2, SV2A, VMA21, NOL9, CAMK2N1, RPRD2, TRIM66, MAPRE2, IRGQ, MIDN, SLC43A2, GNG4, FAM161A, ZNF562, MYD88, WSB2, CTXN1, PAK2, TRAK1, CD24P4, BCAP31, SLX4, WNK3, VEPH1, SPTAN1, MBP, C1orf122, RPS6KL1, RN7SK, BTBD17, KLHL23, ZNF324B, AC007382.1, ARF5, BTBD7, DIP2B, ERC1, KLHL42, TTC28, CHMP4B, MAGED2, TOX3, PIK3R2, CHCHD2, KIAA1211, CORO1C, C12orf49, RGS2, CAMSAP2, SNX19, BAZ2B, SLC36A1, DGCR6L, RABL5, SH3BP4, VSTM2L, MICAL1, TCF12, RERE, ATP1B1, MGST3, CDC42BPA, TCF7L2, PIP4K2A, ASAP1,

KLHDC8A, H3F3A, GTF2A1, NEMF, CYB5A, ZNF91, TSPAN5, YWHAG, AKAP13, MGA, SSNA1, CTNNBIP1, FNBP1, SRGAP3, NLGN3, NTNG2, MVB12B, SFT2D2, SCRT2, C1orf233, RP11-274B21.3, ETV5, ETV1, ST3GAL1, NRXN3, LTBP1, HOMER3, KCNH2, PHF21B, TRIB2, RBFOX2, PLCB4, CACNG7, PTPRZ1, STIM2, CCDC34, FOXM1, TBC1D30, DPYSL3, NRP2, CNTFR, ACSL3, HS3ST3B1, DACH2, AIF1L, AKAP12, SH3BP5, PDZD2, SERPINE2, SSFA2, LMO4, SFXN5, SCD5, CDCA5, CACNA1B, FREM2, THY1, MMP16, SLC37A1, NTNG1, KLF15, CAMKV, PBX3, DHRS13, CHST11, TPPP, SEZ6L2, KCNA2, ZNF48, CUEDC1, LIMK2, GRID1, SPNS2, CEND1, SLC6A9, CALM1, LDB1, PLXNB3, SHISA4, TMSB4X; Cluster 11: THSD7A, DPF1, CNTN1, LRRC7, RAB27B, LMO3, CAMK2B, REEP1, ADCYAP1R1, OSBPL6, SLC1A3, MOXD1, WDR47, PHACTR3, CHGB, FXYD5, PITPNM2, NRCAM, UNC13D, SORBS1, TRIM9, CHGA, DOK5, RASSF2, PAK7, CELF4, PCYT1B, RHOV, ATP1A3, ELAVL2, SEMA5A, HBEGF, CSPG5, SGIP1, PCDH17, DUSP4, PREX1, GDAP1L1, MT2A, SOX21, TMOD2, CHAC1, RHBDF2, NCAN, UNC13A, RFTN1, SNAP25, DCLK1, TAOK3, FAIM2, SPRY2, EDNRB, RAPGEF5, HCRTR2, TUBB2A, SLCO5A1, LBX1, DUSP6, ASCL1, RTN1, TGFB1I1, ASXL3, GREB1L, CBX4, AFF3, SCN1A, 4-Mar, ZNF300, CPLX2, TENM2, PSD2, CSGALNACT1, SNX30, INA, GRIK4, NPAS3, PTPRO, ANO4, GFRA1, SPARCL1, SCN3A, ANKFN1, PCDH1, NMNAT2, SYNJ1, ELAVL4, BSN, LRFN5, ABTB2, SVOP, SCN3B, MGAT5B, GPRIN1, KCNAB1, IRX2, NEGR1, RND1, ARPP21, RALGAPA1, GALNTL6, SPSB4, RIMKLA, TMTC2, NRXN1, SLC8A1, RALYL, FAM174B, MYT1L, MT1X, C11orf96, FAM69C, SHISA7, NHS, S100A16, GPR123, STMN3, C21orf62, TRIM71, SOGA3, PLXNA4, RP11-323D18.5, RN7SL721P, HECW1, KLHL13, ARHGAP31, RIMBP2, PLXNA2, KIAA0020, ABCB1, DLL3, NALCN, SMPD3, SCG3, NEFM, APBA1, HGFAC, SLC1A2, ULBP1, CHD5, F3, ONECUT2, MND1, PDZRN3, ATXN1, SYT5, FST, BAI3, GAD2, SEMA6D, ST8SIA2, PMAIP1, CACNG8, 1-Mar, FRMPD3, SYBU, NTRK2, THRB, SLC6A1, SHROOM4, CHRNB2, KCNJ9, KIF26B, PTX3, PCDH19, GABRB3, CACNG2, SLC16A4, DCLK2, EMB, SCRT1, FGGY, ABLIM3, TMEM196, TTC9B, FIBIN, SLITRK1, CPNE7, SLC17A8, NTM, GPR19, MANEAL, C3orf70, MAGI2, NKAIN2, PCDHB16, SZT2, PPP1R14C, AC006547.14, ATP5F1P5, ANK1, LIMCH1, SNCAIP, FRY,

ACTL6B, KCNK2, NIN, PAPLN, SALL1, NPTX2, SUSD1, KIAA1549L, PTPN5, PRMT8, UHRF1BP1L, FAM65B, FGF12, DLGAP3, DNAJC6, PTCH2, PIK3R3, PROX1, CNRIP1, UTP20, RAB9B, TMX4, DOCK4, MTUS1, ATP8A2, DUSP26, UNC79, DTX1, GABBR2, PAK6, HCN4, TMEM132B, INTS7, MEIS1, ANK2, TACC1, CDH8, PRDM8, EPHB1, PTPRN2, GRIA1, C10orf12, FMN2, ELMO1, KALRN, PRICKLE2, ADCY1, HSPA12A, TET2, CXXC4, STK32A, ZEB2, MCC, ARNT2, SCUBE2, PCSK1, ETV4, PGBD5, KCNJ10, BEND3, KCNH8, KCND2, POU3F2, PDE4B, PTCH1, PCLO, SLIT1, LUZP2, FBLL1, SHISA6, GREB1, POU3F4, C15orf59, NPTXR, SMIM13, RN7SL532P, TMEM179, GPM6B, CAMK2A, SLC4A4, MTMR9, DFNA5, CPEB3, HCFC2, APC2, SLC8A2, CPNE5, TRPC4, MEIS2, APC, MAP7, FAM129A, LMO7, PLCE1, NOVA1, ZNF41, MYO1E, RCAN1, VSNL1, ADORA1, RNF180, FAT3, SMPD1, RAB31, ALCAM, PCDHB13, MYO5A, SLC35F6, EML6, FXYD7, RP11-578F21.12, PRKAR2B, SOX8, HIVEP2, MYL7, BCHE, KCNIP2, CDKN2C, PDE1B, LANCL2, BEX1, PCDH10, IGSF11, CBWD5, RAB3C, NRGN, KCNMA1, ARHGAP27, KCNF1, PPIP5K1, GRIN1, PLEKHM3, CSRNP3, NAP1L3, SLC18A3, ANKRD36B, FAM212B, ELOVL2, LPAR1, GRK5, C2orf72, UBE2QL1, LYN; Cluster 12: MEF2C, TESC, ANKRD24, DLX2, PAX8, BCAN, NT5E, GPR45, DUSP5, MTTP, SNTG1, FZD5, GRIK3, SCN9A, NPAS4, XKR4, BRINP1, NOX4, GLRA2, RHOJ, DDC, PCDH8, SCN2A, NOV, ABCA8, GLRA1, CNTN5, SLC45A2, EN2, C1QL4, DPF3, RP1-161N10.1, CYP26B1, SEZ6, NTN1, RIMS1, MLC1, SLC32A1, PPP1R16B, ACSBG1, HOXA2, GATA3, ELMOD1, PDE4D, LPPR4, ATP10B, SPP1, VSX2, LAMP5, FGD3, VGF, SLC6A11, HRK, CD36, DOCK10, SCML4, NKX6-2, GRID2, GRIA4, STK32B, LHFPL4, S100B, AKNAD1, EN1, NKX6-1, PRSS12, HPGD, PDZRN4, NETO1, GRM5, C8orf46, KCNK3, KCND3, DSCAM, IRX5, IRX3, AMER3, CHRM2, ZBTB20, SLITRK2, SPRY4, C2orf80, CTD-3088G3.8, PAX5, SPOCK3, ESRRG, DCHS2, SP9, AC080125.1, RP11-460N20.4, RP11-384K6.4, VN1R84P, GRIN2B, TNC, CAPN3, TLL2, SYNDIG1, ST8SIA5, STS, DRP2, PLCL1, GBP3, ITPR2, GRM4, ARHGEF6, ANO1, MICAL2, SPOCD1, DYSF, DLX1, DOCK5, MFSD6, GUCY1A2, ANGPT1, DGKI, IFI16, SCRG1, CACNB2, GPR64, SUSD5, CSMD1, SP140L, MUC19, HRH3, TPO, LPPR5, NR4A3, ESRRB, FIBCD1, SLCO1C1, GFOD1, GPR158, GRIK2, NELL1, GBX2, ALK, GATA2, RYR3, ISPD, AC018470.1.

Table 1
HES5+ vs. HES5- expression intensity ratios during NSC progression

Gene Symbol	*Polyp. SEQ ID NO:*	*Polyn. SEQ ID NO:*	*cluster*	*NE*	*ERG*	*MRG*	*LRG*	*LNP*
TOP2A	4213	2429	NE	2.05	-0.01	0.15	0.5	-1.4
HIST1H4C	3289-3302	1012-1027	NE	1.69	0.18	-0.01	0.08	-0.62
TRIM71	4238	2465	NE	1.66	-0.15	-0.06	-0.27	0
PPIG	3816	1814-1815	NE	1.55	-0.26	-0.17	-0.18	-0.34
MLLT4	3536-3537	1382	NE	1.35	0.33	-0.36	-0.23	0
TNC	4204	2416	NE	1.26	-2.72	0.71	0.47	-0.67
CDK1	2879-2882	384-388	NE	1.22	0.27	0.54	0.45	-0.63
OIP5	3698	1622-1623	NE	1.21	0.37	0.23	0.43	-0.38
GDF15	3210	889-890	NE	1.19	-1.38	0.38	-0.17	-0.68
MCM6	3499	1327-1328	NE	1.18	0.53	-0.05	-0.03	0.15
TP53TG1		2437-2438	NE	1.18	0.5	-0.05	0.02	0.18
FAM83D	3126	757-758	NE	1.13	0.4	0.17	0.5	-0.59
FANCI	3129-3130	762-764	NE	1.12	0.29	0	0.73	-0.16
GINS2	3216	899-900	NE	1.12	0.27	-0.14	0.16	-0.02
KDM5A	3383	1145	NE	1.11	-0.24	-0.21	-0.17	-0.41
GSTM3	3263	969-971	NE	1.1	0.18	-0.38	0.02	0
FAM64A	3123	751-752	NE	1.09	0.14	0.18	0.54	-0.45
LIMS1	3439	1232-1233	NE	1.07	-1.25	0.28	0.16	-0.86
CENPH	2899	408-409	NE	1.06	0.07	0.18	0	-0.72
KIF2C	3394	1163-1164	NE	1.05	0.35	0.23	0.42	-0.44
ATAD2	2757	188-189	NE	1.03	0.44	0.48	0.2	-0.43
DTL	3037	628-629	NE	1.03	0.23	0.71	0.39	-0.46
CDCA5	2878	382-383	NE	1.03	0.45	0.52	0.75	-0.32
ARHGEF6	2739	156-157	NE	1.03	0.48	-0.03	-0.04	-0.15
LIPA	3440-3441	1234-1235	NE	1.03	0.02	-0.21	0.24	0.2
POLE2	3802	1791	NE	1.02	0.42	0.67	0.29	-0.35
RRM2	3935-3936	2006-2007	NE	1.02	0.23	0.58	0.62	-0.66
MAD2L1	3464	1269	NE	0.99	-0.53	0.21	-0.22	-0.94
CKS1B	2919	438-440	NE	0.98	-0.47	-0.09	-0.2	-0.41
TTK	4245-4246	2478-2479	NE	0.98	0.18	0.47	0.31	-0.83
DHFR	3012	584	NE	0.97	0.09	0.67	-0.11	-0.7
S100A4	3946-3947	2025-2027	NE	0.94	-1.4	-0.43	-0.75	-0.04
NUP37	3688	1604-1605	NE	0.94	0.09	-0.35	-0.08	0
PMAIP1	3797	1782-1783	NE	0.93	-1.29	0.26	0.15	-0.42
CENPN	2900-2902	410-413	NE	0.93	0.32	-0.17	0.32	-0.19
RNASEH2A	3907	1960-1961	NE	0.92	0.16	0.45	0.13	-0.2
BST2	2799	252	NE	0.91	-0.26	0.16	0.05	0.25
MCM10	3495-3496	1319-1322	NE	0.91	0.58	0.27	0.53	-0.35
MAF	3465-3466	1270-1272	NE	0.9	0.2	0.08	0.09	-0.22
KIAA0101	3384-3385	1146-1148	NE	0.9	0.24	0.38	0.14	-0.23
C8ORF4	2821	290-291	NE	0.87	-1.85	0.44	0.5	0.03
E2F7	3045	641-642	NE	0.86	0.4	0.41	0.54	-0.61
CENPA	2896-2897	405-406	NE	0.86	0.02	0.44	0.52	-0.34
UBE2T	4264	2508-2509	NE	0.85	0.34	0.03	0.14	-0.49
RAB13	3859	1885	NE	0.84	-0.46	-0.08	-0.16	0

Gene Symbol	*Polyp. SEQ ID NO:*	*Polyn. SEQ ID NO:*	*cluster*	*NE*	*ERG*	*MRG*	*LRG*	*LNP*
TMEM126A	4181	2378-2379	NE	0.83	0.23	-0.35	0.2	0.12
MAGT1	3471	1278-1279	NE	0.83	0.08	-0.24	-0.01	-0.01
CDC6	2877	381	NE	0.82	0.28	0.2	0.55	-0.27
C6ORF211	2818	284-285	NE	0.81	0.11	-0.4	0.26	0.31
RFC5	3892-3894	1938-1942	NE	0.81	0.12	0.5	0.34	-0.4
PSMD1	3837	1849	NE	0.8	-0.38	-0.27	-0.45	-0.28
HMMR	3308-3311	1034-1039	NE	0.8	0.01	0.23	0.12	-0.54
UNG	4265-4266	2510-2512	NE	0.8	-0.12	0.18	0.05	0.25
UBE2C	4259-4262	2499-2506	NE	0.8	0.34	0.54	0.14	-0.48
GINS1	3215	897-898	NE	0.79	0.29	0.16	0.43	-0.46
AURKB	2770	206	NE	0.76	0.45	0.4	0.26	-0.43
LEPREL1	3420-3421	1204-1206	NE	0.76	-1.68	0.11	-0.13	0.34
SBNO1	3957-3958	2041-2043	NE	0.75	-0.27	0.26	0.36	-0.76
ZWINT	4366-4368	2678-2682	NE	0.74	0.38	0.18	0.47	-0.26
MKI67	3534-3535	1380-1381	NE	0.73	0.36	0.25	0.34	-0.47
CCAR1	2856	351-352	NE	0.72	-0.49	0.13	-0.26	-0.68
FKBP5	3145-3148	788-791	NE	0.72	-0.12	0.13	-0.33	-0.33
PVRL3	3854	1877-1878	NE	0.7	-0.4	-0.21	-0.18	-0.34
CCNB1	2860	358-359	NE	0.7	-0.02	0.41	0.19	-0.35
NOP58	3652	1553-1554	NE	0.68	-0.16	0.07	-0.12	-0.36
COL4A1	2940	478	NE	0.66	-0.52	0.19	-0.46	-0.1
GGH	3214	896	NE	0.64	-0.06	0.31	-0.11	-0.35
LSM6	3455	1257-1258	NE	0.63	-0.02	0.23	-0.22	-0.67
EID1	3059	659-660	NE	0.59	-0.04	-0.21	-0.26	-0.23
GPX8	3246	949	NE	0.59	-0.71	0.22	-0.36	-0.33
STC2	4116	2273	NE	0.59	-1.64	-0.4	-0.22	-0.14
CD276	2871-2872	374-376	NE	0.57	-0.16	0.18	-0.32	0.15
HS2ST1	3328-3329	1066-1068	NE	0.55	-0.42	0.03	-0.06	-0.55
EIF5B	3063	664-665	NE	0.54	-0.65	0.07	-0.13	-0.5
HDGF	3271-3273	983-985	NE	0.52	-0.32	0.14	-0.15	-0.31
NOL7	3651	1551-1552	NE	0.5	-0.49	0.21	-0.01	-0.18
NR2E1	3661	1565	ERG	0.2	2.87	0.77	1.28	0.13
HES5	3281	999	ERG	0.73	2.5	0.29	0.91	-0.5
ARX	2753	180-181	ERG	-0.02	2.32	0.45	1.14	0.14
C1ORF61	2809	268-269	ERG	-0.24	2.23	0.4	0.58	0.43
FRZB	3170	825	ERG	0.67	2.08	0.42	-0.1	0.08
GRM3	3254	959	ERG	0.75	2.06	0.56	-1.11	-0.5
EPHA3	3088-3089	700-703	ERG	-0.46	2.05	0.22	1.04	-0.89
NAV3	3573	1439-1440	ERG	0.12	2.03	0	-0.33	0.04
EGR2	3054-3057	654-657	ERG	0.11	2	0.68	-0.48	-0.07
RGMA	3899-3904	1950-1956	ERG	0.04	2	-0.04	0.71	0.16
NRXN3	3681-3683	1592-1596	ERG	0.01	1.98	-0.09	-0.61	-1.16
FAM107A	3116-3117	739-741	ERG	0.52	1.91	0.15	0.75	0.55
FABP7	3115	738	ERG	0.41	1.84	0.13	0.14	0.03
EGR3	3058	658	ERG	0.15	1.83	0.53	-0.67	-0.45
ZNF385B	4348-4350	2648-2651	ERG	0.75	1.83	0.34	-0.08	-0.38
TTYH1	4247-4248	2480-2482	ERG	1.01	1.81	0.43	0.26	0.26
SNCAIP	4035	2157-2158	ERG	0.11	1.78	-0.32	0.89	0.13

Gene Symbol	*Polyp. SEQ ID NO:*	*Polyn. SEQ ID NO:*	*cluster*	*NE*	*ERG*	*MRG*	*LRG*	*LNP*
NRARP	3676	1585	ERG	0.94	1.72	-0.13	0.34	0.26
PLP1	3790-3792	1772-1775	ERG	0.45	1.68	-0.12	0.42	0.08
LIX1	3445	1242-1243	ERG	0.54	1.62	-0.03	0.25	-0.34
LFNG	3422-3425	1207-1210	ERG	0.33	1.58	0.03	0.07	-0.8
HES4	3279-3280	996-998	ERG	0.53	1.54	0.38	0.19	0.69
CD82	2873-2874	377-378	ERG	0.9	1.53	0.32	0.78	0.19
HS6ST1	3330	1069-1070	ERG	0.24	1.48	0.48	-0.09	-0.42
PTPRZ1	3852	1875	ERG	0.46	1.47	0.27	0.39	0.09
CACHD1	2824	295-296	ERG	0.33	1.47	0.4	0.08	-0.22
DACH1	2981-2983	537-541	ERG	-0.04	1.46	0.16	0.57	-0.14
FEZF2	3139	779-780	ERG	0.24	1.45	0.85	-0.22	-0.24
DTX4	3038	630-631	ERG	0.27	1.44	0.25	0.62	0.28
FUT9	3182	844-845	ERG	0.19	1.4	0.55	0.81	0.32
WNT5B	4303-4304	2571-2574	ERG	0.42	1.37	0.02	0.18	-0.01
ENPP2	3076-3078	683-686	ERG	0.73	1.36	0.32	1.1	-0.5
POU3F3	3806	1796-1797	ERG	0.88	1.34	0.19	-0.27	0.04
EMX2	3073-3074	680-681	ERG	0.32	1.32	0.16	0.93	-0.08
MECOM	3503-3508	1334-1341	ERG	0.29	1.32	0.01	0.04	-0.08
XYLT1	4308	2580-2581	ERG	0.2	1.31	0.72	0.07	-0.79
ARMCX2	2746-2747	170-173	ERG	0.64	1.3	0.39	-0.31	-0.05
FOS	3155	799-800	ERG	0.16	1.28	0.16	0.06	0.34
PPAP2B	3815	1813	ERG	0.31	1.28	0.1	0.45	0.23
NOS2	3653	1555	ERG	0.33	1.26	0.29	-0.02	0.01
LRP2	3451	1252	ERG	0.45	1.25	0.49	0.64	-0.12
SOX9	4065	2203	ERG	0.52	1.24	0.37	0.32	-0.18
NLGN3	3648-3650	1546-1550	ERG	0.35	1.22	0.51	0.62	-0.41
TMEM2	4191-4192	2395-2397	ERG	0.27	1.18	0.35	-0.31	-0.05
CXCR7	2973	525-526	ERG	0.22	1.16	0.63	-0.05	-1.1
EPHA7	3090	704	ERG	0.7	1.15	0.1	-0.26	-0.14
SMOC1	4031-4032	2151-2153	ERG	0.14	1.15	0.31	-0.07	0.13
TBC1D9	4138	2310-2311	ERG	0.28	1.14	0.47	0.3	-0.44
FAT4	3132	767-768	ERG	-0.06	1.13	0.33	0.54	-1.51
SCUBE3	3967	2056-2057	ERG	0.58	1.12	0.01	0.45	-0.19
FUT8	3179-3181	839-843	ERG	0.1	1.11	0.41	0.15	-0.02
CSPG5	2957	504-505	ERG	-0.2	1.11	0.6	0.45	-0.25
DLL1	3017	592-593	ERG	-0.2	1.11	0.61	-0.23	-0.55
BOC	2798	250-251	ERG	0.28	1.1	0.01	-0.16	0.11
ID4	3337	1078	ERG	0.36	1.07	0.28	0.14	-0.46
EGR1	3053	653	ERG	0.03	1.06	-0.11	0.15	0.3
ALPL	2702-2703	97-98	ERG	0.23	1.06	0.32	0.3	-0.6
RFX4	3897-3898	1946-1949	ERG	0.3	1.06	0.58	0.42	-0.3
GALNT12	3203	879-880	ERG	0.41	1.06	-0.06	-0.12	0.08
CBX2	2854-2855	347-350	ERG	0.2	1.03	0.23	0.14	-0.23
FHOD3	3143	785-786	ERG	-0.2	1.02	0.43	0.24	0
SORBS2	4045-4052	2173-2181	ERG	-0.33	1.01	0.21	-0.16	0.28
GUCY1B3	3266	976	ERG	0.22	1	0.3	0.28	-0.03
MBIP	3493-3494	1316-1318	ERG	0.29	0.99	0.03	0.22	0
FBXO16	3136	773-774	ERG	0.48	0.99	0.67	0.29	-0.62
SHISA2	3987	2088	ERG	-0.07	0.98	0.29	-0.15	0.41

Gene Symbol	*Polyp. SEQ ID NO:*	*Polyn. SEQ ID NO:*	*cluster*	*NE*	*ERG*	*MRG*	*LRG*	*LNP*
DAB1	2980	535-536	ERG	-0.03	0.97	0.18	0.37	-0.25
GLI3	3224	915	ERG	0.22	0.94	0.36	0.03	0.16
FZD3	3189	854-855	ERG	0	0.93	0.45	-0.03	-0.09
SEMA5B	3973	2066	ERG	-0.03	0.92	0.02	0.3	0.37
LGALS3BP	3428	1215-1216	ERG	0.46	0.9	0.39	-0.13	0.19
SFRP1	3981	2078	ERG	0.31	0.9	-0.21	0.17	-0.42
C1QL1	2811	272-273	ERG	-0.24	0.89	-0.63	0.18	0.13
RING1	3906	1959	ERG	0	0.87	-0.01	-0.01	0.22
GPRC5B	3243	944-945	ERG	-0.08	0.87	-0.05	0.15	0.27
ZNF710	4360	2668-2669	ERG	0.07	0.86	0.18	0.08	0.06
WSCD1	4305	2575-2576	ERG	0.2	0.86	0.37	0.25	-0.22
VPS37B	4289	2553-2554	ERG	0.26	0.84	0.26	0.02	-0.36
ZIC2	4333	2623-2624	ERG	0.04	0.84	0.19	0.17	0.2
SDK2	3969	2060	ERG	0.4	0.83	0.27	-0.07	0.02
DOCK11	3031	618-619	ERG	0.12	0.81	-0.02	0	-0.58
GAS1	3208	887	ERG	0.16	0.81	0.16	-0.14	-0.01
ZNF436	4353-4354	2656-2658	ERG	0.37	0.8	0.37	-0.51	-0.04
TMSB15A	4201-4202	2410-2413	ERG	0.14	0.76	-0.03	0.18	-0.18
IER2	3339	1080-1081	ERG	-0.09	0.74	-0.19	-0.06	0.2
FEZ1	3137-3138	775-778	ERG	0.21	0.73	0.08	0.17	0.14
CELF2	2892-2895	400-404	ERG	0.28	0.73	0	0.12	-0.25
SFT2D3	3982	2079-2080	ERG	0.17	0.7	-0.28	0.07	0.12
NCALD	3574-3581	1441-1449	ERG	-0.15	0.7	-0.11	-0.02	-0.02
AKAP7	2698-2700	90-95	ERG	0.13	0.68	-0.33	-0.12	-0.11
MYADM	3551-3555	1404-1409	ERG	-0.12	0.67	-0.19	0.13	0.03
NEDD4L	3593-3601	1471-1480	ERG	-0.35	0.65	-0.16	0.14	-0.04
PHC2	3768-3769	1735-1737	ERG	0.27	0.65	0.29	-0.21	0.15
PI4KAP2	#N/A	1754-1755	ERG	-0.42	0.62	-0.57	-0.24	0.19
STARD3	4100-4102	2250-2253	ERG	-0.11	0.58	-0.21	0.13	0.11
CAMK1D	2843-2844	327-330	ERG	-0.02	0.57	-0.35	0.01	-0.1
FZD10	3188	852-853	MRG	-0.38	0.64	1.72	-0.26	-0.53
ZEB2	4325-4326	2611-2613	MRG	-0.25	-0.93	1.68	-0.98	-1.5
EN2	3075	682	MRG	-0.11	0.05	1.68	-1.38	0.17
ST20	4087-4088	2232-2233	MRG	-0.4	-1.13	1.44	0.4	-0.42
CDKN2C	2888-2889	394-396	MRG	0.04	0.64	1.4	0.15	-1.28
RAB10	3858	1883-1884	MRG	-0.32	-1.5	1.34	-0.36	-1.61
WASF1	4291-4294	2556-2559	MRG	-0.49	-0.04	1.31	-0.15	-0.62
ZBED4	4314	2590-2591	MRG	-0.61	-1.29	1.3	-0.46	-1.47
EZH2	3113-3114	735-737	MRG	-0.23	-0.33	1.29	-0.17	-0.75
PPA2	3811-3814	1805-1812	MRG	0.21	-0.33	1.28	0.13	-0.27
H1F0	3270	981-982	MRG	-0.75	-0.26	1.26	0.18	0.24
CCNJ	2866-2868	367-370	MRG	-0.43	-0.51	1.24	-0.5	-0.74
ITGB8	3371	1128	MRG	-0.77	-0.05	1.24	-0.01	-0.5
SH3BGRL3	3985	2084-2085	MRG	-0.79	-0.05	1.21	-0.36	0.02
IRX2	3365-3366	1119-1121	MRG	-0.84	0.63	1.2	-0.45	-0.5
KIF23	3392-3393	1159-1162	MRG	0.58	-0.25	1.19	0.03	-1.01
PEG10	3759-3760	1722-1724	MRG	-1.05	-0.64	1.18	-0.69	-0.53

Gene Symbol	*Polyp. SEQ ID NO:*	*Polyn. SEQ ID NO:*	*cluster*	*NE*	*ERG*	*MRG*	*LRG*	*LNP*
SMC3	4029	2147-2148	MRG	-0.28	0.02	1.16	-0.05	-0.59
NUSAP1	3691-3692	1609-1612	MRG	0.66	0	1.12	0	-1.13
APLP1	2734-2735	148-150	MRG	-0.83	0.35	1.12	-0.17	-0.07
ADAMTS3	2695	85-86	MRG	0.22	-0.34	1.12	0.09	-0.18
RACGAP1	3867-3869	1898-1903	MRG	-0.08	-0.14	1.1	0.52	-0.93
LIMCH1	3434-3438	1226-1231	MRG	0.13	-0.19	1.1	0.22	0.02
ETNK1	3092-3093	706-708	MRG	-0.83	-0.83	1.09	-0.1	-1.01
RNF13	3908-3909	1962-1965	MRG	-0.5	-0.86	1.08	-0.19	-0.79
ARID1B	2740-2741	158-161	MRG	-1.21	-0.84	1.08	-0.24	-1.34
TRIM28	4235	2459-2460	MRG	-0.08	0.42	1.07	-0.48	0.09
CNOT8	2928	457-458	MRG	-0.36	-0.79	1.05	-0.27	-1.1
CRNDE	4369	4370-4374	MRG	-1.22	-1	1.05	-0.39	-0.6
TWSG1	4251	2486-2487	MRG	-0.01	-0.65	1.04	-0.69	-1
NT5DC2	3684-3685	1597-1599	MRG	-0.36	0.24	1.04	-0.54	0.58
NAA50	3571	1435-1436	MRG	0.43	-0.62	1.03	0.14	-0.78
NUF2	3686-3687	1600-1603	MRG	0.71	-0.22	1.02	0.16	-1.06
ABCE1	2687-2688	75-76	MRG	0.16	-0.52	1.01	0.41	-0.21
PLTP	3794-3795	1777-1780	MRG	-0.34	-0.06	1.01	-0.42	0.24
FBRSL1	3134	770	MRG	-1	-0.16	1	0.43	-0.16
DCAF16	2991	550-551	MRG	-0.33	-0.58	1	-0.32	-0.88
OGT	3696-3697	1618-1621	MRG	-0.28	-0.5	0.99	-0.42	-0.47
ZFYVE16	4330-4331	2619-2621	MRG	0.14	-0.19	0.96	-0.32	-0.52
FOXM1	3161-3163	810-815	MRG	0.41	0.19	0.96	0.33	-0.44
PM20D2	3796	1781	MRG	0.08	-0.73	0.96	0.44	-0.67
POU3F2	3805	1794-1795	MRG	-0.06	0.45	0.96	-0.4	-0.08
MCM4	3497-3498	1323-1326	MRG	0.48	-0.1	0.95	0.13	-0.75
HERPUD2	3277	992-993	MRG	-0.05	-0.64	0.94	0.02	-0.58
VRK1	4290	2555	MRG	0.33	-0.25	0.94	0	-0.91
TRIM41	4236-4237	2461-2464	MRG	-0.29	0.18	0.94	-0.59	0.51
SATB1	3955-3956	2039-2040	MRG	-1.07	-0.33	0.93	-0.09	-0.93
HOMER1	3323	1058	MRG	-0.56	-0.8	0.93	0.24	-0.53
CCNG1	2863-2864	363-365	MRG	0.34	-0.55	0.93	-0.29	-0.63
ATF2	2758	190	MRG	-0.42	-0.14	0.92	-0.02	-0.41
AP1AR	2726-2727	136-138	MRG	-0.63	-0.35	0.92	0.06	-0.43
GABPA	3195	866	MRG	-0.24	-0.37	0.92	-0.16	-0.55
STXBP3	4122	2283-2284	MRG	0	-0.84	0.91	-0.21	-0.76
SMC5	4030	2149-2150	MRG	0.21	-0.35	0.91	0.07	-0.42
CDKN1B	2884	390	MRG	-0.03	-0.19	0.91	-0.17	-0.75
NUPL1	3689-3690	1606-1608	MRG	0.02	-0.25	0.91	-0.24	-0.7
UBA1	4254-4255	2491-2493	MRG	-0.52	0.36	0.9	-0.92	0.54
CYTH2	2976-2977	529-531	MRG	-1.03	0.32	0.9	-0.4	0.43
FXYD6	3183-3187	846-851	MRG	-0.44	0.47	0.9	-0.52	-0.02
ISYNA1	3369-3370	1125-1127	MRG	0.02	0.34	0.89	-0.16	0.07
DOCK1	3029	615	MRG	0.35	-0.26	0.88	0.09	-0.29
41527	#N/A	#N/A	MRG	0.45	-0.23	0.87	0	-0.38
LPHN2	3450	1250-1251	MRG	-0.16	0.05	0.87	-1.28	-0.1
IDI1	3338	1079	MRG	-0.19	-0.12	0.87	0.11	-0.63
PXMP2	3855	1879-1880	MRG	0.19	-0.47	0.86	0.32	-0.57
U2AF2	4252-4253	2488-2490	MRG	-0.35	-0.74	0.86	-0.57	-0.72

Gene Symbol	*Polyp. SEQ ID NO:*	*Polyn. SEQ ID NO:*	*cluster*	*NE*	*ERG*	*MRG*	*LRG*	*LNP*
ARHGAP12	2738	154-155	MRG	-0.68	-0.16	0.86	-0.15	-0.35
KLHL24	3402	1178-1179	MRG	-1.15	-0.53	0.85	-0.28	-0.53
CKAP2	2915-2916	432-434	MRG	0.36	-0.47	0.85	-0.07	-0.96
ZNF238	4344-4345	2641-2644	MRG	-0.39	0	0.85	-0.28	-0.61
PARP6	3720	1657-1658	MRG	-0.83	0.12	0.85	-0.14	0.27
NHSL1	3644-3645	1539-1541	MRG	0.14	0.27	0.84	-0.11	-0.02
PBRM1	3736	1684-1685	MRG	-0.34	-0.06	0.84	-0.38	-0.33
BAZ1A	2776-2777	215-218	MRG	-0.35	-0.45	0.84	0.18	-0.37
MAP4K5	3480-3481	1294-1297	MRG	-0.4	-0.3	0.84	-0.12	-0.21
TSPAN12	4243	2474-2475	MRG	0.4	-0.31	0.84	-0.15	-0.06
SH3GLB1	3986	2086-2087	MRG	-0.41	-0.65	0.84	0.12	0.02
ASPM	2756	186-187	MRG	0.43	0.22	0.83	0.46	-0.57
ANKLE2	2710	108-109	MRG	-0.38	-0.47	0.83	-0.04	-0.54
SPG20	4068-4071	2207-2211	MRG	0.44	-0.58	0.83	-0.21	-0.4
MAP4K4	3477-3479	1288-1293	MRG	-1.58	-0.33	0.83	0.15	-0.21
CASC4	2847-2848	335-338	MRG	-0.04	-0.44	0.83	-0.2	-0.44
FUBP3	3178	838	MRG	0.04	-0.44	0.82	-0.1	-0.82
ARSB	2751-2752	177-179	MRG	-0.1	-0.9	0.81	0.07	-0.94
BTAF1	2800	253	MRG	-0.23	-0.53	0.81	-0.07	-0.68
SIKE1	3988-3989	2089-2091	MRG	-0.08	-0.55	0.81	0.15	-0.36
VEZT	4282	2540-2541	MRG	-0.18	-0.07	0.81	-0.24	-0.52
PBX3	3738-3739	1687-1693	MRG	-1.4	-0.51	0.81	-0.09	-0.66
CBL	2851	342-343	MRG	-0.3	-0.04	0.81	-0.01	-0.34
EIF2AK4	3061	662	MRG	0	-0.55	0.81	-0.34	-0.53
API5	2731-2733	142-147	MRG	-0.12	-0.15	0.81	-0.36	-0.22
MTSS1	3550	1402-1403	MRG	-1.33	0.29	0.8	-0.32	-0.23
NET1	3610-3611	1492-1494	MRG	0.17	-0.29	0.8	0.05	-0.79
CHD6	2907	419-420	MRG	-0.19	-0.1	0.8	-0.1	-0.45
ZNF117	4338	2630-2631	MRG	-0.67	0.24	0.8	-0.3	-0.12
PNMA1	3799	1785-1786	MRG	-0.33	-0.36	0.8	0.06	-0.55
PTPN13	3845-3848	1862-1869	MRG	0.17	-0.77	0.8	-0.12	-0.61
MTIF2	3548-3549	1400-1401	MRG	-0.08	-0.51	0.79	-0.1	-0.29
SSFA2	4083-4084	2226-2228	MRG	-0.12	-0.75	0.79	0.21	-0.76
KIAA1279	3388	1152-1153	MRG	-0.49	-0.31	0.79	0.11	-0.34
STRN3	4120-4121	2280-2282	MRG	-0.08	-0.5	0.79	-0.09	-0.4
WIPF1	4301-4302	2569-2570	MRG	0.51	-0.47	0.78	0.19	-0.67
MEIS2	3519-3525	1356-1368	MRG	-0.11	-0.58	0.78	-0.29	-0.66
ZC3H4	4321	2603-2604	MRG	-0.17	-0.01	0.78	0.09	-0.42
DYNC1I2	3039	632	MRG	-0.36	-0.09	0.78	-0.67	0.17
RTN4	3940-3944	2014-2023	MRG	-0.92	-0.95	0.78	-0.61	-0.71
TAF2	4134	2305	MRG	0	-0.18	0.78	-0.19	-0.38
RASA1	3875-3876	1909-1911	MRG	-0.26	-0.47	0.77	0.02	-0.26
OSBPL8	3706-3707	1636-1638	MRG	-0.47	-0.33	0.77	-0.05	-0.62
SKA2	3992-3993	2096-2098	MRG	-0.05	0.26	0.77	-0.11	-0.64
IGF1R	3344	1089	MRG	-0.51	-0.95	0.77	-0.59	-0.52
RNF6	3916-3918	1976-1981	MRG	-0.3	-0.41	0.76	0.4	-0.37
SGTB	3984	2082-2083	MRG	-0.71	-0.91	0.76	0	-0.65
TMEM131	4183	2382-2383	MRG	-0.22	-0.04	0.76	-0.09	-1.13

Gene Symbol	*Polyp. SEQ ID NO:*	*Polyn. SEQ ID NO:*	*cluster*	*NE*	*ERG*	*MRG*	*LRG*	*LNP*
HIATL1	3284	1003-1004	MRG	0.22	-0.5	0.76	0.19	-0.35
TGIF1	4162-4169	2346-2360	MRG	0.38	-0.32	0.76	0.13	-0.12
TMEM170B	4188	2391	MRG	-1.16	-0.32	0.76	-0.21	-0.53
PSAT1	3835-3836	1845-1848	MRG	0.05	-0.6	0.76	0.26	-0.47
ACBD5	2689-2690	77-79	MRG	-0.44	-0.42	0.75	0.38	-0.31
HECTD2	3275-3276	988-991	MRG	0.04	-0.02	0.75	-0.26	-0.37
ASF1A	2755	184-185	MRG	0.24	-0.1	0.75	-0.16	-0.31
LAMB1	3408	1188	MRG	-0.59	-1.56	0.75	-0.93	-1.03
GLS	3227	919-920	MRG	-0.2	-0.62	0.75	-0.57	-0.35
DDX39	3009	578-579	MRG	0.06	0.03	0.75	0.14	-0.18
DGCR2	3011	582-583	MRG	-0.5	-0.61	0.75	-0.06	-1.1
EIF1AX	3060	661	MRG	-0.2	-0.11	0.75	0.12	-0.45
SALL1	3950-3951	2032-2033	MRG	0.1	-0.28	0.75	0.03	-0.66
GOLPH3	3233	931-932	MRG	-0.28	-0.41	0.75	-0.23	-0.72
PTBP2	3841	1856-1857	MRG	-0.23	0.49	0.74	-0.56	-0.49
GRIP1	3253	957-958	MRG	-0.37	-0.02	0.74	-0.02	-0.48
PNPLA8	3801	1789-1790	MRG	-0.79	-0.7	0.74	-0.12	-0.58
VASH2	4274-4276	2529-2532	MRG	-1.94	-0.5	0.74	-0.31	-0.86
SUDS3	4123	2285-2286	MRG	-0.07	-0.29	0.73	0.25	-0.61
PFDN4	3761	1725	MRG	-0.32	-0.44	0.73	-0.32	-0.58
BAZ2A	2778	219-220	MRG	-0.73	-0.3	0.73	-0.25	-0.14
PRKDC	3831-3832	1838-1840	MRG	0.24	-0.3	0.72	-0.32	-0.77
GLYR1	3230	925-926	MRG	-0.3	-0.42	0.72	0.1	-0.65
DAZAP2	2984-2989	542-548	MRG	0.03	-0.08	0.72	-0.3	-0.58
PCMTD1	3748	1709-1710	MRG	-0.88	-0.17	0.72	-0.27	-0.23
SENP6	3974-3975	2067-2069	MRG	-0.19	-0.39	0.72	-0.16	-0.35
CLINT1	2921	443-444	MRG	-0.13	-0.45	0.72	-0.09	-0.19
RECQL	3889-3890	1933-1935	MRG	0.05	-0.34	0.71	0.07	-0.41
CNTNAP2	2930	461-462	MRG	-0.3	-0.15	0.71	0.14	-0.93
CTBP1	2958-2959	506-507	MRG	-0.41	0.21	0.71	-0.03	-0.06
C10ORF18	2801	254-255	MRG	-0.14	-0.59	0.71	0.02	-0.26
CDON	2890	397-398	MRG	-0.3	0.37	0.71	0.17	-0.08
B4GALT6	2771	207	MRG	0.05	-0.07	0.71	0.22	-0.2
CSNK1G3	2954-2956	501-503	MRG	-0.25	-0.54	0.71	-0.03	-0.47
STAT5B	4108	2263-2264	MRG	-0.06	-0.01	0.7	-0.15	-0.47
TMEM60	4198	2405-2406	MRG	0.17	-0.22	0.7	0.3	-0.69
HNRNPH2	3320-3322	1053-1057	MRG	-0.21	-0.64	0.7	-0.31	-0.5
TACC2	4130-4133	2297-2304	MRG	-1.44	-0.26	0.7	0.31	-0.35
CCNG2	2865	366	MRG	-0.07	-0.25	0.7	-0.33	-0.58
FSCN1	3171	826	MRG	-0.65	0.01	0.7	-0.67	0.16
CCNA2	2859	357	MRG	0.31	-0.25	0.7	0.39	-0.74
C21ORF45	2812	274-275	MRG	0.25	-0.34	0.7	-0.07	-0.87
PLRG1	3793	1776	MRG	0.42	-0.31	0.7	0.03	-0.36
ZFHX3	4327-4328	2614-2616	MRG	-0.56	0.27	0.69	-1.31	-0.52
UBE2A	4257-4258	2496-2498	MRG	-0.29	-0.34	0.69	-0.1	-0.49
DMTF1	3021-3023	599-606	MRG	-0.12	-0.01	0.69	-0.01	-0.2
TRA2A	4229	2448-2449	MRG	-0.19	-0.06	0.69	-0.18	-0.23
MYO5A	3566-3567	1428-1429	MRG	-0.93	-0.74	0.69	0.23	-1.11

Gene Symbol	*Polyp. SEQ ID NO:*	*Polyn. SEQ ID NO:*	*cluster*	*NE*	*ERG*	*MRG*	*LRG*	*LNP*
FAM96A	3127-3128	759-761	MRG	-0.16	-0.55	0.69	-0.13	-0.83
IFT80	3340	1082-1083	MRG	0.02	0.02	0.69	0.17	-0.09
VPS26A	4287-4288	2550-2552	MRG	-0.01	-0.56	0.69	-0.07	-0.59
MRPL50	3542	1391-1392	MRG	0.15	-0.4	0.69	-0.27	-0.45
ACYP1	2694	84	MRG	0.07	-0.3	0.69	0.15	-0.59
WDR11	4299	2565-2566	MRG	-0.45	-0.49	0.68	0.03	-0.43
PLDN	3789	1770-1771	MRG	-0.05	-0.2	0.68	0.14	-0.64
RPRD1A	3928	1997-1998	MRG	-0.19	-0.14	0.67	-0.15	-0.32
MEAF6	3502	1332-1333	MRG	-0.23	-0.15	0.67	0.08	-0.37
CKAP5	2917-2918	435-437	MRG	0.07	-0.07	0.67	0.17	-0.11
YTHDC2	4312	2587-2588	MRG	-0.24	-0.34	0.67	-0.09	-0.35
GABARAPL2	3193	862-863	MRG	-0.49	-0.36	0.67	0.16	-0.42
IPO5	3358	1111	MRG	0.21	-0.2	0.67	0.18	-0.58
PGAP1	3762	1726-1727	MRG	-0.06	0.32	0.67	-0.12	-0.5
C14ORF147	2804	259-260	MRG	0.2	-0.18	0.67	-0.01	-0.55
CD200	2869-2870	371-373	MRG	-0.42	-0.3	0.67	-0.14	-0.34
MST4	3543-3545	1393-1396	MRG	-0.04	-0.42	0.67	-0.06	-0.41
PPT1	3824-3825	1827-1828	MRG	0.36	-0.28	0.67	0.12	-0.03
ANKRD50	2712-2713	112-114	MRG	-0.84	-0.18	0.66	-0.66	-0.56
HPS3	3327	1064-1065	MRG	0.09	-0.39	0.66	-0.15	-0.42
CCNC	2861-2862	360-362	MRG	-0.08	-0.4	0.66	0.02	-0.31
THRAP3	4173	2365-2366	MRG	-0.08	-0.15	0.66	0.1	-0.47
TWF1	4250	2485	MRG	0.11	-0.65	0.66	0	-0.18
CYP51A1	2974-2975	527-528	MRG	-0.45	-0.63	0.66	-0.03	-0.45
PSPC1	3840	1853-1855	MRG	0.03	0.08	0.65	0.07	-0.28
WDR75	4300	2567-2568	MRG	-0.4	-0.35	0.65	0.02	-0.24
CAST	2849-2850	339-341	MRG	-0.18	-0.09	0.65	-0.05	-0.29
SEPW1	3976	2070	MRG	-0.05	0.08	0.65	-1.05	0.08
C21ORF59	2813	276-277	MRG	-0.67	-0.66	0.65	0.02	-0.7
PIK3C2A	3780	1756	MRG	-0.12	-0.38	0.65	0.05	-0.31
GNG5	3232	929-930	MRG	0.16	-0.45	0.65	-0.2	-0.68
MED4	3511	1345-1346	MRG	-0.32	-0.29	0.65	-0.36	-0.45
GIPC1	3217-3221	901-910	MRG	-0.02	-0.56	0.64	0.03	-0.33
STK39	4118	2276-2277	MRG	-0.65	-0.37	0.64	0.15	-0.13
KIAA1715	3389	1154-1155	MRG	-0.82	-0.46	0.64	0.03	-0.25
PHF6	3770-3772	1738-1742	MRG	0.08	-0.38	0.64	0.15	-0.81
PPTC7	3826	1829-1830	MRG	0.03	-0.41	0.64	0.21	-0.27
SOCS4	4036-4037	2159-2162	MRG	-0.28	-0.24	0.64	-0.08	-0.42
PPM1B	3818-3821	1818-1823	MRG	-0.34	-0.48	0.64	-0.26	-0.45
UQCRB	4267	2513-2516	MRG	-0.35	-0.11	0.63	-0.47	-0.45
C10ORF84	2802-2803	256-258	MRG	-0.26	-0.29	0.63	-0.07	-0.26
SLAIN1	3994-3995	2099-2101	MRG	-0.05	-0.03	0.63	0.12	-0.49
RAB6A	3862-3865	1890-1895	MRG	-0.82	-0.5	0.63	0.01	-0.27
SOS2	4053	2182-2183	MRG	-0.16	-0.14	0.63	0.07	-0.34
KLF10	3396-3397	1167-1169	MRG	-0.33	0.06	0.63	-0.4	-0.6
RNF4	3914	1971	MRG	0.03	-0.06	0.63	0.14	-0.18
C3ORF63	2814-2815	278-280	MRG	0	-0.2	0.63	-0.03	-0.4

Gene Symbol	*Polyp. SEQ ID NO:*	*Polyn. SEQ ID NO:*	*cluster*	*NE*	*ERG*	*MRG*	*LRG*	*LNP*
INSIG1	3354-3356	1104-1109	MRG	-0.68	-0.16	0.62	-0.32	-0.68
CPSF1	2943	482-483	MRG	-0.18	0.36	0.62	-0.08	0.13
DNAJC4	3027	611-612	MRG	-0.79	-0.02	0.62	-0.03	-0.19
ATP2B4	2768-2769	204-205	MRG	0.08	-0.23	0.62	-0.26	-0.74
PPP2R1A	3823	1825-1826	MRG	-0.26	0.24	0.62	-0.78	0.07
TRIM22	4233	2456-2457	MRG	0.32	-0.34	0.62	-0.02	-0.02
SDC3	3968	2058-2059	MRG	-0.8	-0.02	0.61	-0.4	-0.16
TSNAX	4242	2472-2473	MRG	-0.28	-0.46	0.61	0.09	-0.58
PPIL4	3817	1816-1817	MRG	-0.03	-0.41	0.61	-0.11	-0.47
ZDHHC2	4324	2609-2610	MRG	0.12	-0.22	0.61	0.15	-0.41
ZBTB44	4318	2597-2598	MRG	-0.19	-0.23	0.61	-0.1	-0.27
ANO6	2714-2716	115-117	MRG	0.01	-1.09	0.61	-0.21	-0.81
PPP2CB	3822	1824	MRG	-0.16	-0.51	0.61	0.12	-0.46
UBA2	4256	2494-2495	MRG	-0.06	-0.45	0.6	-0.21	-0.76
BBS2	2779	221-222	MRG	-0.06	0.13	0.6	0.17	-0.33
ZNF423	4352	2654-2655	MRG	-0.42	-0.38	0.6	0.09	-0.38
RNF5	3915	1972-1975	MRG	-0.18	-0.08	0.6	-0.18	-0.35
C1ORF31	2808	267	MRG	0.25	-0.41	0.6	0.07	-0.3
IFT81	3341-3343	1084-1088	MRG	0.09	-0.29	0.6	-0.15	-0.74
CPSF6	2944	484-485	MRG	-0.05	-0.03	0.6	-0.09	-0.37
KLHL9	3406	1185-1186	MRG	-0.17	-0.05	0.6	-0.03	-0.15
FAM164A	3120	746-747	MRG	-1.12	-0.13	0.6	0.03	-0.54
TTC35	4244	2476-2477	MRG	-0.01	-0.22	0.6	0.23	-0.44
CCDC90B	2858	355-356	MRG	-0.1	-0.91	0.6	-0.14	-0.67
TM9SF4	4176	2370-2371	MRG	-0.15	0.06	0.6	0	-0.14
SEC24B	3970-3971	2061-2063	MRG	-0.01	-0.53	0.59	0.12	-0.21
SMARCC2	4026-4028	2143-2146	MRG	-0.38	0.18	0.59	-0.14	-0.41
CAP2	2846	333-334	MRG	0.15	-0.59	0.59	-0.14	-0.21
SAR1A	3953-3954	2036-2038	MRG	-0.01	-0.43	0.59	-0.09	-0.5
THY1	4174	2367-2368	MRG	0	0.05	0.59	-0.96	-0.46
RBPMS2	3886	1928-1929	MRG	0.09	-0.42	0.59	-0.01	-0.36
EIF3A	3062	663	MRG	-0.03	-0.57	0.59	-0.18	-0.53
DZIP1	3040-3041	633-636	MRG	0.17	-0.45	0.58	0.07	-0.36
ARL6IP1	2744	166-167	MRG	0.2	-0.17	0.58	0.01	-0.67
SACM1L	3949	2030-2031	MRG	-0.2	-0.2	0.58	-0.1	-0.34
PAPD4	3713-3715	1647-1650	MRG	-0.43	-0.34	0.58	0.03	-0.32
SCG2	3959	2044	MRG	-0.02	-0.7	0.58	-1.09	-0.88
TCF3	4146-4147	2325-2326	MRG	0	0.17	0.58	0.18	-0.12
EFHA1	3049	647-648	MRG	-0.37	-0.27	0.58	-0.04	-0.42
HNRNPA2B1	3313-3314	1042-1044	MRG	-0.29	-0.37	0.57	-0.47	-0.95
EWSR1	3106-3110	724-730	MRG	-0.14	-0.03	0.57	-0.34	-0.42
STAG2	4096-4099	2245-2249	MRG	0.01	0.1	0.57	-0.12	-0.34
YEATS4	4309	2582-2583	MRG	0.27	-0.17	0.57	-0.19	-0.63
PAQR3	3717	1652	MRG	-0.82	-0.78	0.57	0.07	-0.46
GAR1	3206-3207	883-886	MRG	0.03	-0.53	0.57	0.02	-1.11
FTH1	3177	835-837	MRG	-0.33	-0.48	0.57	0.04	-0.51
C19ORF43	2805	261-262	MRG	-0.38	0.3	0.57	0.03	-0.11
TMEM14C	4184-4185	2384-2386	MRG	0.04	-0.42	0.56	-0.05	-0.27

Gene Symbol	*Polyp. SEQ ID NO:*	*Polyn. SEQ ID NO:*	*cluster*	*NE*	*ERG*	*MRG*	*LRG*	*LNP*
CCDC104	2857	353-354	MRG	-0.36	-0.32	0.56	-0.02	-0.42
PSMD12	3838-3839	1850-1852	MRG	-0.22	-0.2	0.56	-0.18	-0.46
DCTD	3001-3002	566-567	MRG	0.06	-0.24	0.56	0.04	-0.41
SSR1	4085	2229	MRG	0.13	-0.6	0.56	0.04	-0.57
HMGCS1	3306-3307	1032-1033	MRG	-0.44	-0.55	0.56	-0.33	-1.4
HMGB3	3305	1030-1031	MRG	0.3	-0.33	0.56	-0.2	-0.39
KIF3A	3395	1165-1166	MRG	-0.75	-0.03	0.56	-0.09	-0.31
TMEM128	4182	2380-2381	MRG	0.02	-0.45	0.56	-0.15	-0.4
PATZ1	3721-3724	1659-1666	MRG	-0.17	0.13	0.55	0.13	-0.2
RBL2	3877	1912-1913	MRG	-0.29	-0.22	0.55	0.13	-0.54
ARFGAP3	2736-2737	151-153	MRG	0.01	-0.32	0.55	-0.08	-0.45
DNAJB5	3024-3026	607-610	MRG	-1.11	-0.24	0.55	-0.03	-0.23
TMED7	4177-4179	2372-2375	MRG	0.02	-0.41	0.54	-0.4	-0.66
G3BP2	3190-3192	856-861	MRG	-0.23	-0.68	0.54	-0.46	-0.77
BMPR1A	2797	249	MRG	-0.03	-0.05	0.54	-0.28	-0.5
FMR1	3151	794	MRG	-0.42	-0.54	0.54	-0.23	-0.66
TPST2	4227-4228	2446-2447	MRG	0.03	-0.62	0.54	0.02	-0.39
TMSB4X	4203	2414-2415	MRG	-0.06	-0.79	0.54	-0.27	-0.74
RP2	3927	1995-1996	MRG	-0.08	-0.58	0.54	0.28	-0.6
CEP170	2903-2905	414-417	MRG	-0.89	-0.29	0.54	-0.21	-0.54
KLHL23	3401	1176-1177	MRG	-0.16	-0.37	0.54	0.09	-0.45
RNF7	3919-3920	1982-1985	MRG	0.27	-0.47	0.54	-0.07	-0.23
HNRNPH1	3319	1051-1052	MRG	-0.06	-0.61	0.54	-0.29	-0.52
MARCKS	3484	1302	MRG	-0.52	-0.17	0.54	-0.12	-0.58
HNRNPD	3315-3318	1045-1050	MRG	-0.07	-0.12	0.54	-0.14	-0.45
TOB1	4212	2427-2428	MRG	0.02	-0.78	0.53	-0.49	-0.61
UTP11L	4273	2527-2528	MRG	0.06	-0.57	0.53	-0.26	-0.64
RFK	3895	1943-1944	MRG	0.01	-0.55	0.53	0.17	-0.64
DHX36	3015-3016	589-591	MRG	-0.35	-0.14	0.53	-0.01	-0.45
LCOR	3411-3413	1192-1195	MRG	-0.94	0.2	0.53	0.07	-0.21
WBP5	4295-4298	2560-2564	MRG	0.1	-0.31	0.53	-0.18	-0.5
PHLDB2	3774-3778	1745-1751	MRG	-0.69	-0.66	0.53	-0.46	-1.11
USP33	4270-4272	2521-2526	MRG	-0.43	-0.12	0.53	-0.06	-0.53
EFNB2	3051	651	MRG	-0.08	0.09	0.52	0.08	-0.17
C6ORF62	2819	286-287	MRG	-0.39	-0.64	0.52	0	-0.5
MEX3B	3527	1371-1372	MRG	-0.54	0.12	0.52	-0.21	-0.4
ABCD3	2685-2686	73-74	MRG	-0.05	-0.62	0.51	0.23	-0.76
ATG3	2767	202-203	MRG	-0.08	-0.43	0.51	0.1	-0.42
ARID4B	2742-2743	162-165	MRG	-0.4	-0.02	0.51	-0.19	-0.56
C7ORF11	2820	288-289	MRG	-0.01	-0.57	0.51	-0.26	-0.75
EPB41	3081-3086	690-697	MRG	-1.12	0.17	0.51	0.08	-0.28
TCF12	4141-4145	2316-2324	MRG	-0.29	-0.82	0.51	-0.29	-0.68
CDK8	2883	389	MRG	0.04	0.03	0.51	0.09	-0.01
CMIP	2923-2924	447-450	MRG	-1.1	-0.18	0.51	0.07	-0.34
ATG12	2766	201	MRG	0.22	-0.35	0.5	0	-0.16
CETN3	2906	418	MRG	0.03	-0.35	0.5	-0.41	-0.65
ZNF217	4341	2635-2636	MRG	-0.12	-0.2	0.5	-0.11	-0.42
TMEM55A	4197	2403-2404	MRG	-0.86	-0.21	0.5	-0.31	-0.15
UBE2N	4263	2507	MRG	0	-0.3	0.5	-0.15	-0.45

Gene Symbol	*Polyp. SEQ ID NO:*	*Polyn. SEQ ID NO:*	*cluster*	*NE*	*ERG*	*MRG*	*LRG*	*LNP*
PMP2	3798	1784	LRG	-0.02	-0.03	-0.24	3.48	-0.7
GABBR2	3194	864-865	LRG	-1.97	0.19	0.08	2.42	1.03
BCAN	2784-2785	229-232	LRG	-0.32	0.18	-0.07	2	-0.5
LUZP2	3461	1265	LRG	0	-0.55	0.14	1.96	0.01
SALL3	3952	2034-2035	LRG	0.08	0.88	0.35	1.81	0.42
SYNM	4124-4125	2287-2290	LRG	0.95	-0.95	-0.1	1.74	0.47
DCT	2999-3000	564-565	LRG	-0.12	0.01	0.98	1.63	-1.63
OLIG1	3701	1628-1629	LRG	-0.16	0.13	-0.26	1.57	-0.45
SPON1	4076	2217-2218	LRG	0.23	0.56	0.23	1.56	-0.69
PDGFRA	3750	1712-1713	LRG	-0.17	-0.89	-0.3	1.53	0.14
COL22A1	2938	475-476	LRG	0.34	0.61	0.16	1.43	0.6
KIAA1239	3387	1151	LRG	-0.18	-0.02	0.29	1.42	-0.05
PCDH10	3740-3741	1694-1697	LRG	0.15	0.01	-0.29	1.39	0.38
LPAR4	3449	1248-1249	LRG	0.79	0.92	0.16	1.37	-0.04
VAV3	4277-4278	2533-2535	LRG	-1.33	-0.37	0.39	1.36	-0.62
CADM2	2836-2838	316-319	LRG	0.18	0	0.29	1.36	-0.61
SOX6	4060-4063	2195-2200	LRG	0.27	0.98	-0.23	1.35	-0.19
SLC6A1	4003	2114	LRG	-0.14	-0.11	0.33	1.35	-0.72
DPP6	3032-3034	620-623	LRG	-1.43	0.57	-0.02	1.35	-0.04
FGFR3	3140-3142	781-784	LRG	0.15	0.97	0.88	1.33	0.01
PDE3B	3749	1711	LRG	-0.05	-0.17	-0.01	1.31	-0.13
MOXD1	3541	1389-1390	LRG	0.15	-1.39	-0.1	1.29	0.27
TNFRSF19	4206-4207	2418-2421	LRG	-0.86	-0.27	0.35	1.29	0.13
PYGL	3856-3857	1881-1882	LRG	0.29	-0.08	0.3	1.29	0.19
GPC6	3238	937-938	LRG	0.74	-0.52	-0.21	1.28	-0.24
COL11A1	2931-2933	463-467	LRG	0.37	-0.34	0.28	1.27	0.37
TRIM9	4239-4240	2466-2469	LRG	-0.17	-0.69	0.43	1.26	-0.41
GABRB3	3201-3202	876-878	LRG	-2.49	-0.83	0.02	1.25	0.29
TFPI	4160-4161	2343-2345	LRG	0.59	-0.67	0.49	1.23	0
CREB5	2950-2953	494-500	LRG	-0.42	0.37	0.3	1.15	-0.73
RAB3GAP 2	3861	1888-1889	LRG	-0.22	0.17	-0.23	1.14	0.23
NCAN	3582	1450	LRG	-2.01	0.27	0.43	1.14	-0.01
EFHD1	3050	649-650	LRG	0.56	0.47	0.19	1.11	-0.4
SLITRK2	4013-4020	2127-2135	LRG	0.39	0.17	0.28	1.1	-0.08
PAX6	3733-3735	1681-1683	LRG	0.34	-0.1	0.2	1.1	-0.2
SLC1A4	3999	2107	LRG	0.54	-0.64	0.02	1.09	0.39
GPR155	3241-3242	941-943	LRG	-1.2	-0.06	0.54	1.08	0.14
GPD2	3239-3240	939-940	LRG	0.37	0.44	0.49	1.08	0.3
CHST11	2911	425-426	LRG	-0.11	0.2	0.58	1.06	-0.34
PAQR8	3718	1653-1654	LRG	0.03	0.06	-0.16	1.05	-0.01
MT2A	3547	1398-1399	LRG	-0.15	-0.27	-0.17	1.04	-0.2
GPC3	3234-3237	933-936	LRG	0.25	0.36	-0.22	1.04	0.01
TMEM51	4193-4196	2398-2402	LRG	0.64	0.56	0.13	1.02	0.17
CHST3	2912	427	LRG	0.21	0.19	0.26	1.02	0.27
PAG1	3710	1643-1644	LRG	-0.45	-0.15	0.55	1.02	0.12
MYO5C	3568	1430-1431	LRG	0.51	0.29	0.02	1.02	-0.22
CACNB2	2825-2833	297-312	LRG	0.2	0.5	-0.12	1.02	0.01
NDRG2	3583-3590	1451-1466	LRG	0.14	0.64	0.01	1.01	0.29

Gene Symbol	*Polyp. SEQ ID NO:*	*Polyn. SEQ ID NO:*	*cluster*	*NE*	*ERG*	*MRG*	*LRG*	*LNP*
ST3GAL5	4089-4090	2234-2235	LRG	-0.14	0.33	-0.04	1.01	-0.35
TPD52L1	4223-4226	2442-2445	LRG	-0.41	-1.08	-0.18	1	0.33
TRIB1	4230	2450-2451	LRG	0.52	0.55	-0.03	1	0.2
PRKCA	3830	1837	LRG	-0.23	-0.54	0.3	1	0.2
BCKDHB	2787-2788	234-236	LRG	0.32	0.14	-0.25	0.99	-0.03
GLT25D2	3228	921-922	LRG	0.57	-0.4	0.04	0.99	-0.21
LITAF	3442-3444	1236-1241	LRG	-1.56	-0.85	-0.11	0.99	0.08
PLCB1	3787-3788	1766-1769	LRG	-0.2	0.26	-0.02	0.98	-0.12
TIMP3	4175	2369	LRG	0.43	0.08	0.14	0.96	0.14
ZBTB46	4319	2599-2600	LRG	-0.58	-0.24	0.47	0.95	-0.03
OPCML	3704-3705	1634-1635	LRG	-1.37	0.07	-0.17	0.92	0.55
CTDSPL	2960-2961	508-510	LRG	-0.41	-0.1	0.25	0.89	0.28
MDGA2	3500-3501	1329-1331	LRG	-0.32	-0.05	0.04	0.86	0.03
MEGF10	3517	1353-1354	LRG	0.44	-0.24	0.57	0.86	-0.05
EYA2	3111-3112	731-734	LRG	-2.03	-0.35	0.33	0.86	-0.24
KANK1	3378-3379	1136-1139	LRG	-0.18	-0.17	0.28	0.85	-0.22
RAB31	3860	1886-1887	LRG	-1.58	-1.4	-0.29	0.85	0.33
TRIL	4232	2454-2455	LRG	-0.59	-0.62	-0.16	0.83	-0.12
FAM171B	3121	748-749	LRG	0.04	0.49	0.24	0.82	0.23
ALCAM	2701	96	LRG	-1.47	-0.87	-0.05	0.82	-0.12
RAB6B	3866	1896-1897	LRG	-1.38	0.38	-0.25	0.8	0.19
PGM2L1	3763	1728-1729	LRG	-1.19	-0.09	0.15	0.8	0.01
LARGE	3409-3410	1189-1191	LRG	-0.29	0.05	0.14	0.8	-0.25
HPCAL1	3325-3326	1061-1063	LRG	-1.03	-0.14	-0.32	0.8	0.3
HTRA1	3334	1075	LRG	0.18	0.51	-0.08	0.79	0.11
TRPS1	4241	2470-2471	LRG	-0.06	0.17	-0.2	0.79	0
TRIB3	4231	2452-2453	LRG	-0.18	-0.89	0.06	0.77	0.45
IGF2BP2	3347-3348	1093-1095	LRG	-0.28	0.03	0.29	0.77	-0.18
PITPNC1	3781-3782	1757-1760	LRG	-0.32	-0.06	0.34	0.77	-0.09
CMTM4	2925-2926	451-454	LRG	0.12	-0.17	0.36	0.76	0.16
IAH1	3335	1076	LRG	0.19	-0.25	0.04	0.75	0.01
DHTKD1	3013	585-586	LRG	-0.01	0.14	0.02	0.74	0.13
SNAP29	4034	2155-2156	LRG	-0.11	-0.22	0.12	0.74	-0.26
CTNNBIP 1	2967-2968	516-518	LRG	0.13	0.39	0.1	0.72	0.19
NQO2	3658	1561	LRG	0.33	-0.16	-0.09	0.72	0.01
MAP4	3474-3476	1284-1287	LRG	0.18	-0.06	0.11	0.71	0.43
CBR1	2853	346	LRG	0.13	0.08	-0.28	0.71	-0.23
LTBP1	3456-3460	1259-1264	LRG	0.28	-0.66	0.13	0.71	0.12
C5ORF32	2816	281-282	LRG	0.37	-0.1	-0.26	0.7	0.24
MARK1	3485	1303-1304	LRG	-0.38	-0.13	0.08	0.69	0.19
AASS	2684	71-72	LRG	0.28	0.11	0.09	0.68	-0.09
CISD1	2914	430-431	LRG	-0.14	-0.02	0.01	0.68	0.1
DSC2	3035-3036	624-627	LRG	-0.58	-0.41	0.09	0.68	0.03
SLC25A33	4002	2112-2113	LRG	0.39	-0.05	-0.09	0.68	-0.05
RIMS3	3905	1957-1958	LRG	-1.51	-0.09	-0.08	0.66	-0.27
ZIC3	4334	2625	LRG	-1.6	-0.33	0.37	0.66	-0.36
EGF	3052	652	LRG	0.26	-0.12	0.4	0.66	0.06
SRGAP2	4081-4082	2223-2225	LRG	-0.07	0.03	-0.1	0.65	0.22

Gene Symbol	*Polyp. SEQ ID NO:*	*Polyn. SEQ ID NO:*	*cluster*	*NE*	*ERG*	*MRG*	*LRG*	*LNP*
RANGAP1	3870	1904	LRG	-0.11	0.05	-0.19	0.63	0.01
SCRG1	3965	2052-2053	LRG	0.05	0.06	0.14	0.63	0.07
PRCP	3827-3828	1831-1834	LRG	0.12	-0.13	-0.2	0.63	0.13
CA12	2822-2823	292-294	LRG	0.05	0.17	-0.1	0.62	-0.11
HEATR5A	3274	986-987	LRG	0.06	-0.32	0	0.62	-0.08
ZNF503	4355	2659-2660	LRG	-0.1	-0.92	0.28	0.6	0.25
GYG2	3267-3268	977-978	LRG	0.07	0.13	-0.11	0.6	0.21
ANAPC1	2704	99-100	LRG	0.24	-0.24	0.24	0.6	-0.01
C19ORF63	2806-2807	263-266	LRG	-0.34	-0.31	0.04	0.59	-0.02
ASAP1	2754	182-183	LRG	-0.4	-0.21	0.11	0.59	-0.19
C1ORF96	2810	270-271	LRG	-1.5	-0.18	-0.01	0.58	-0.27
DHX33	3014	587-588	LRG	0.08	-0.07	-0.24	0.57	0.24
FASTKD1	3131	765-766	LRG	-0.04	0.16	0.06	0.57	0.11
STAU2	4109-4115	2265-2272	LRG	-0.46	-0.41	0.25	0.56	0.3
MAML2	3473	1282-1283	LRG	-1.65	-0.81	-0.14	0.56	0.13
RRAS2	3929-3930	1999-2001	LRG	0.2	-0.08	0.06	0.56	0.02
GLTP	3229	923-924	LRG	0.15	0.2	0.01	0.55	-0.14
VPS13B	4283-4286	2542-2549	LRG	0.09	0.05	0.21	0.54	-0.02
GPT2	3244-3245	946-948	LRG	0.17	-0.22	0.08	0.53	0.11
NKAIN4	3647	1544-1545	LRG	-0.02	0.05	-0.22	0.51	-0.04
ZC3HAV1	4322-4323	2605-2608	LRG	0.13	0.28	-0.04	0.67	0.11
ANXA2P2	#N/A	132-133	LNP	-0.64	-2.14	-0.06	-2.2	2.71
ANXA2	2721-2724	124-131	LNP	-0.64	-2.03	-0.33	-2.27	2.65
FRAS1	3166-3167	819-821	LNP	-0.26	0.45	-0.06	-0.83	2.63
SPOCK1	4075	2216	LNP	0.51	1.24	0.32	0.81	1.9
PCDHB15	3746	1705-1706	LNP	-1.17	0.35	0.42	0.64	1.84
SLC10A4	3996	2102-2103	LNP	0.35	0.11	0.36	0.01	1.74
TPBG	4221-4222	2439-2441	LNP	0.25	0.36	0.3	-0.87	1.65
C5ORF39	2817	283	LNP	0.06	0.3	-0.49	0.55	1.46
MMP14	3539	1385-1386	LNP	-0.28	-1.24	-0.52	0.49	1.45
TNFRSF10D	4205	2417	LNP	0.75	-0.89	0.41	0.14	1.39
S100A6	3948	2028-2029	LNP	0	-2.2	-0.37	-0.49	1.3
RNF182	3910-3913	1966-1970	LNP	-1.4	0.24	-0.14	-0.27	1.28
LGALS1	3426	1211	LNP	0.66	-0.98	-0.45	0	1.26
ISL1	3368	1124	LNP	0.05	-0.51	-0.19	0.21	1.25
SPINK5	4072-4074	2212-2215	LNP	-0.82	0.83	0.2	0.16	1.25
DOCK10	3030	616-617	LNP	-1.38	-0.24	0.09	0.47	1.18
LECT1	3414-3415	1196-1198	LNP	0.7	0.08	0.72	0.44	1.13
LYPD1	3462-3463	1266-1268	LNP	0.66	-1.95	0.35	-0.19	1.11
ARMCX1	2745	168-169	LNP	-0.33	0.49	0.29	-0.03	1.09
NAP1L2	3572	1437-1438	LNP	-0.56	-0.28	0.17	-0.06	1.08
COL4A6	2941-2942	479-481	LNP	0.36	0.75	0.16	0.62	1.07
GSN	3255-3262	960-968	LNP	-0.31	-0.41	-0.23	-0.5	1.01
PLAG1	3784-3786	1763-1765	LNP	0.08	0.05	0.74	0	1
MMD	3538	1383-1384	LNP	0.09	0.24	0.15	0.26	0.99
PTGR1	3842-3844	1858-1861	LNP	0.66	0.08	0.32	0.26	0.96
PDP1	3751-3754	1714-1718	LNP	-0.42	-1.56	0.16	0.14	0.95
COL18A1	2934-2936	468-473	LNP	-0.46	-0.18	-0.14	-0.24	0.95

Gene Symbol	*Polyp. SEQ ID NO:*	*Polyn. SEQ ID NO:*	*cluster*	*NE*	*ERG*	*MRG*	*LRG*	*LNP*
ZIC4	4335-4337	2626-2629	LNP	0.69	0.04	0.15	-0.19	0.95
BASP1	2775	213-214	LNP	-0.52	0.31	0.41	0.1	0.93
AHNAK	2696-2697	87-89	LNP	0.14	-0.78	-0.16	-0.08	0.87
REC8	3887-3888	1930-1932	LNP	-0.06	0.36	0.32	0.29	0.84
KLHDC8B	3400	1174-1175	LNP	-0.17	0.22	0.3	0.15	0.82
FRMD6	3168-3169	822-824	LNP	0.09	-0.91	0.18	-0.52	0.7
MYL9	3558-3559	1412-1415	LNP	0.37	-0.94	-0.81	-0.35	0.69
RBMS1	3880-3881	1917-1919	LNP	-0.13	-1.06	-0.4	-1.04	0.66
TNFRSF21	4208	2422-2423	LNP	-0.21	-0.82	-0.02	0.07	0.65
FAM38A	3122	750	LNP	-0.13	-1.17	-0.35	-0.38	0.56

Table 1: HES5+ (HES5 expressing cells) vs. HES5- (HES5 not expressing cells) expression intensity ratios during NSC progression. Provided are the log 2 scores of the ratios between the two population of cells (HES5+ versus HES5-). Genes whose expression is overrepresented in HES5+ cells compared to HES5- cells at the different stages are shown. Results are arranged according to the gene clusters at each developmental stage from genes highly expressed in HES5+ cells to those less expressed in HES5+ cells. The intensity level of each gene in ES cells was subtracted from all values prior to ratio calculation. Representative polynucleotides (Polyn.) and polypeptide (Polyp.) sequences for the indicated genes are provided.

Table 2

HES5- vs. HES5+ expression intensity ratios during NSC progression

Gene Name	**Cluster**	***Polyp. SEQ ID NO:***	***Polyn. SEQ ID NO:***	**NE (HES5-) vs (HES5+) score**	**ERG (HES5-) vs (HES5+) score**	**MRG (HES5-) vs (HES5+) score**	**LRG (HES5-) vs (HES5+) score**	**LNP (HES5-) vs (HES5+) score**
LHX1	NE	3430	1219-1220	5.8135	1.8275	-0.0616	0.0189	-0.14668
CNTN2	NE	2929	459-460	5.6375	1.4368	-0.1379	0.6039	0.138649
ST18	NE	4086	2230-2231	5.5601	1.2563	-0.7274	0.2071	-0.2604
EBF3	NE	3048	646	5.4952	0.90561	0.4081	0.2194	-0.13904
NFASC	NE	3614-3619	1498-1504	5.4652	0.5897	-0.6463	0.542	0.1019
FSTL5	NE	3174-3176	831-834	5.432	0.9048	-0.3377	-0.7043	0.1242
ONECUT2	NE	3703	1632-1633	5.3885	0.2146	-0.3585	0.0845	0.194
SLC17A6	NE	3997	2104-2105	5.3052	0.509	-0.5857	0.3641	0.04147
EBF1	NE	3047	644-645	5.2666	1.1739	-0.447	0.9108	-0.0797
SLIT1	NE	4012	2126	4.975	1.34902	-0.4829	0.5256	-0.073
SYT4	NE	4129	2295-2296	4.967	0.54587	1.2702	1.2208	-0.1282

Gene Name	Cluster	*Polyp. SEQ ID NO:*	*Polyn. SEQ ID NO:*	NE (HES5-) vs (HES5+) score	ERG (HES5-) vs (HES5+) score	MRG (HES5-) vs (HES5+) score	LRG (HES5-) vs (HES5+) score	LNP (HES5-) vs (HES5+) score
NEFM	NE	3603-3604	1483-1485	4.9055	2.2333	-0.182	2.1336	-0.1433
NEUROD1	NE	3612	1495	4.7667	0.8952	-0.9178	0	0.2146
PARM1	NE	3719	1655-1656	4.6809	0.4767	0.0729	-0.8023	1.0749
CHN2	NE	2909-2910	423-424	4.6105	0.5617	-0.5152	0.8935	0.530562
DNER	NE	3028	613-614	4.5936	0.691	-0.703	-0.129	0.205
HMP19	NE	3312	1040-1041	4.5456	0.8298	-0.4776	0.9251	-0.256636
TFAP2B	NE	4157	2339	4.5269	2.5872	-0.2459	1.1232	0.2224
DCX	NE	3003-3006	568-574	4.4703	-0.231	-0.5015	0.2863	0.972
KLHL35	NE	3403	1180	4.2551	0.0492	0.1523	-0.0322	0.0492
PAPPA	NE	3716	1651	4.162	2.5698	-0.2682	0.10733	-0.23581
OLFM1	NE	3699-3700	1624-1627	4.0785	0.6821	-0.1767	-0.5756	0.1824
NHLH1	NE	3641	1534-1535	4.0581	0.1168	-0.1595	0.3518	0.4972
RTN1	NE	3938-3939	2010-2013	3.8472	-0.4485	0.1738	-0.245	-1.2731
GAP43	NE	3204-3205	881-882	3.7821	2.27942	-0.5219	-0.5752	0.1127
GFRA1	NE	3211-3213	891-895	3.7036	1.8269	-0.9674	0.031	0.6612
CHL1	NE	2908	421-422	3.6646	0.3618	-0.544	-0.0669	0.1306
FNDC5	NE	3152-3154	795-798	3.6318	0.3431	-0.3787	-0.4577	-0.0818
SCN3A	NE	3962-3964	2048-2051	3.62889	-0.021	-0.3752	0.9483	1.17
NPTX2	NE	3657	1560	3.5953	0.6951	-0.37091	1.4171	-0.5266
EOMES	NE	3079	687-688	3.5952	0.8243	0.3486	0.35682	0.378
CADPS	NE	2839-2841	320-324	3.5658	0.1174	-0.6736	-0.5961	-0.5865
NHLH2	NE	3642-3643	1536-1538	3.5261	1.2844	0.1873	0.0317	0.073
TMEM163	NE	4187	2389-2390	3.5043	-0.2491	-0.4101	0.163	-0.2184
STMN3	NE	4119	2278-2279	3.4853	0.2112	-0.2191	-0.4416	0.0455

Gene Name	Cluster	*Polyp. SEQ ID NO:*	*Polyn. SEQ ID NO:*	NE (HES5-) vs (HES5+) score	ERG (HES5-) vs (HES5+) score	MRG (HES5-) vs (HES5+) score	LRG (HES5-) vs (HES5+) score	LNP (HES5-) vs (HES5+) score
LRRN3	NE	3452-3454	1253-1256	3.4672	2.1439	-0.5248	-0.4061	-0.1709
NEFL	NE	3602	1481-1482	3.3021	0.7617	-0.0325	0.9243	0.263
ROBO2	NE	3921-3922	1986-1987	3.2613	0.0933	-1.027	0.0685	0.7086
INA	NE	3353	1102-1103	3.256	0.6429	-0.4096	0.2427	0.264
PHLDA1	NE	3773	1743-1744	3.1599	2.2102	0.03841	-0.1558	0.0438
GRIA1	NE	3248-3249	952-953	3.1538	1.5794	-0.9874	-0.0593	-0.0719
GRIA2	NE	3250-3252	954-956	3.0962	0.2147	-0.2607	-0.0969	0.1527
DCLK1	NE	2992	552	3.0869	-0.4287	-0.6781	-0.1876	-0.0462
CRABP1	NE	2945	486	3.07309	-0.1754	-0.8178	0.2218	-0.1841
OLIG2	NE	3702	1630-1631	3.0648	-0.40358	0.05765	-1.6016	-0.8114
SCG3	NE	3960-3961	2045-2047	3.0594	0.0743	-0.8797	-0.8527	0.1681
TMEM158	NE	4186	2387-2388	2.9842	0.808	-0.3393	-0.3487	0.274
FBXL16	NE	3135	771-772	2.9367	0.7521	-0.9621	-0.0437	0.1049
FAM123A	NE	3118-3119	742-745	2.7682	0.48464	-0.3491	-0.7135	0.6321
SYP	NE	4128	2294	2.72982	0.0404	-0.0221	-0.029	0.0362
KIF21B	NE	3391	1157-1158	2.6734	-0.13298	-0.3799	0.1031	0.3907
PCDH9	NE	3744-3745	1701-1704	2.641	0.4482	0.3322	0.1216	0.1173
CDKN1C	NE	2885-2887	391-393	2.6135	0.7614	-0.396	0.4872	0.9298
IGFBPL1	NE	3349	1096	2.56732	0.18504	-0.5459	0.6174	1.7607
RSPO3	NE	3937	2008-2009	2.5065	1.0753	0.2531	0.2123	0.1943
GABRB3	NE	3201-3202	876-878	2.4931	0.8282	-0.0165	-1.2497	-0.2895
TAGLN3	NE	4135-4137	2306-2309	2.3959	-0.2427	-0.3897	0.0477	-0.0469
KCNH2	NE	3380-3382	1140-1144	2.2063	-0.2456	0.0933	-0.02682	0.3444
EPB41L3	NE	3087	698-699	2.12175	-0.2889	-0.33	-0.55	0.2377

Gene Name	Cluster	*Polyp. SEQ ID NO:*	*Polyn. SEQ ID NO:*	NE (HES5-) vs (HES5+) score	ERG (HES5-) vs (HES5+) score	MRG (HES5-) vs (HES5+) score	LRG (HES5-) vs (HES5+) score	LNP (HES5-) vs (HES5+) score
EYA2	NE	3111-3112	731-734	2.0311	0.3503	-0.33118	-0.8595	0.2416
TMOD2	NE	4199-4200	2407-2409	2.0209	-0.3455	-0.4043	-0.4689	0.7347
NCAN	NE	3582	1450	2.0081	-0.2711	-0.4259	-1.1379	0.0122
GABBR2	NE	3194	864-865	1.96646	-0.1865	-0.0788	-2.4245	-1.0319
D4S234E	NE	2978-2979	532-534	1.9639	-0.8658	-0.044	0.5762	0.1084
PI15	NE	3779	1752-1753	1.9268	0.87692	0.12517	-0.4014	-0.0565
ANK2	NE	2705-2707	101-104	1.9126	0.7869	-0.6777	-0.8437	0.1061
SLC1A2	NE	3998	2106	1.8732	-0.6327	-0.4227	-0.0056	0.0009
NRCAM	NE	3677-3678	1586-1588	1.8582	-0.1457	-0.9091	0.3443	0.0719
CBLB	NE	2852	344-345	1.6652	0.22119	-0.1036	0.2755	0.0497
CAMK2N1	NE	2845	331-332	1.6143	-0.3898	-0.4733	0.4673	0.3984
ZIC3	NE	4334	2625	1.5991	0.3292	-0.3749	-0.6636	0.3592
PTPRN2	NE	3849-3851	1870-1874	1.5641	-0.6982	-0.4356	0.1863	0.1616
SORBS1	NE	4038-4044	2163-2172	1.552509	0.3051	0.085	-0.2424	-0.3264
NXPH1	NE	3693	1613-1614	1.54302	0.41304	-0.4469	0.9047	0.1682
BAALC	NE	2772-2773	208-210	1.4852	-0.184	0.1304	0.3051	0.0423
CLASP2	NE	2920	441-442	1.4709	0.11842	0.0433	-0.44565	-0.0016
DPP6	NE	3032-3034	620-623	1.4278	-0.56697	0.0216	-1.3459	0.0372
MAP6	NE	3482-3483	1298-1301	1.4164	-1.0266	-0.178	-0.2896	0.0489
FIGNL2	NE	3144	787	1.4021	0.4434	0.0074	-0.0563	0.1233
KIAA0802	NE	3386	1149-1150	1.2965	0.1156	-0.06708	-0.3461	-0.0269
KIF1A	NE	3390	1156	1.2544	0.14328	0.1678	-0.53255	-0.288
TMEM170B	NE	4188	2391	1.1584	0.3155	-0.7605	0.2069	0.5281
SLC22A23	NE	4000-4001	2108-2111	1.1468	-0.3646	0.2083	0.0627	-0.1698

Gene Name	Cluster	*Polyp. SEQ ID NO:*	*Polyn. SEQ ID NO:*	NE (HES5-) vs (HES5+) score	ERG (HES5-) vs (HES5+) score	MRG (HES5-) vs (HES5+) score	LRG (HES5-) vs (HES5+) score	LNP (HES5-) vs (HES5+) score
TMEM178	NE	4189-4190	2392-2394	1.1378	-0.3558	-0.2492	-0.3128	0.1337
CTNND2	NE	2969	519	1.1059	-0.8586	-0.759	-0.152	0.2469
CADM1	NE	2834-2835	313-315	1.0745	0.416	-0.2601	0.073	0.2928
LGALS3	NE	3427	1212-1214	1.0667	-0.0439	0.2515	-0.2314	-0.4109
TCEAL7	NE	4140	2314-2315	1.0629	-0.2261	-0.1816	-0.1126	0.7301
CD9	NE	2875	379	0.9335	-0.7716	0.2032	-1.0329	0.09014
GLB1L2	NE	3222	911-912	0.9036	-0.7014	-0.0807	-0.06902	0.3665
NFIA	NE	3626-3629	1516-1520	0.8777	0.0255	0.095	-0.6388	0.1731
YPEL3	NE	4310-4311	2584-2586	0.8681	-0.2503	-0.0673	0.0896	-0.15553
TNFRSF19	NE	4206-4207	2418-2421	0.8594	0.2733	-0.3454	-1.2914	-0.1321
SPINK5	NE	4072-4074	2212-2215	0.8238	-0.8276	-0.2033	-0.16463	-1.2485
PNMA2	NE	3800	1787-1788	0.8223	0.2398	-0.4278	-0.0582	0.4584
IRX5	NE	3367	1122-1123	0.75519	0.4756	-0.304	0.0965	-0.3108
SIN3B	NE	3990	2092-2093	0.6838	-0.1671	-0.3826	-0.34452	-0.27595
STK38	NE	4117	2274-2275	0.6181	-0.0217	-0.0158	-0.0806	-0.07283
NR3C1	NE	3664-3670	1570-1576	0.6164	-0.201	0.156347	-0.835	-0.0207
SOX8	NE	4064	2201-2202	0.6128	0.0941	-0.1613	-0.0501	-0.1152
SLC6A8	NE	4007-4009	2118-2121	0.5933	-0.3232	-0.2501	-0.0799	0.01185
SYNPR	NE	4126-4127	2291-2293	0.58652	0	-0.1157	-0.2807	-0.563
SGK223	NE	3983	2081	0.54028	0.2185	-0.1014	-0.1997	-0.1089
BASP1	NE	2775	213-214	0.5188	-0.3106	-0.4113	-0.1026	-0.9285
APC	NE	2728-2730	139-141	0.5098	-0.4177	-0.3206	-0.7791	0.1049
GREM1	ERG	3247	950-951	0.1823	3.709	-1.4281	-0.1756	-0.30164
COL3A1	ERG	2939	477	2.345	3.6453	0.1318	1.05404	0.4239

Gene Name	Cluster	*Polyp. SEQ ID NO:*	*Polyn. SEQ ID NO:*	NE (HES5-) vs (HES5+) score	ERG (HES5-) vs (HES5+) score	MRG (HES5-) vs (HES5+) score	LRG (HES5-) vs (HES5+) score	LNP (HES5-) vs (HES5+) score
PCDH8	ERG	3742-3743	1698-1700	0.5348	3.3585	-1.6013	0.9937	-1.42624
SEMA3C	ERG	3972	2064-2065	2.317	3.1944	-1.0751	1.6029	-0.19453
BMP4	ERG	2795-2796	246-248	1.5686	2.976	-0.3235	0.2193	0.0986
NID2	ERG	3646	1542-1543	-0.4011	2.957	-0.089	0.1567	0.6132
TNC	ERG	4204	2416	-1.2593	2.7215	-0.7102	-0.4651	0.6681
COL1A2	ERG	2937	474	0.1418	2.5245	0.0421	0.0169	1.1218
ANKRD1	ERG	2711	110-111	0.3982	2.4561	-0.10987	-0.218768	0.0884
ANXA1	ERG	2717	118	0.4038	2.408	-0.2234	-0.4331	0.5578
TMEFF2	ERG	4180	2376-2377	1.0094	2.0606	-0.2996	0.5763	-0.0072
PDZRN3	ERG	3755	1719-1720	1.42208	1.834	-1.1358	0	0.5287
ANXA3	ERG	2725	134-135	-0.0125	1.801	-0.185	0.4877	-0.2323
KRT8	ERG	3407	1187	-0.0954	1.7326	0.3095	0.3271	0.5396
LEPREL1	ERG	3420-3421	1204-1206	-0.7574	1.6786	-0.1093	0.1315	-0.3397
NOX4	ERG	3654-3656	1556-1559	-0.09831	1.5884	0.04407	-0.2657	-0.1353
LAMB1	ERG	3408	1188	0.5897	1.5638	-0.7512	0.933	1.0289
FLNC	ERG	3149-3150	792-793	0.0742	1.486	0.2195	-0.1665	-0.4711
FST	ERG	3172-3173	827-830	-0.427	1.4764	-0.399	0.39983	0.6635
IMMP2L	ERG	3352	1100-1101	-0.173	1.4155	-0.1224	-0.43338	0.3043
S100A4	ERG	3946-3947	2025-2027	-0.9445	1.4036	0.4282	0.7475	0.0422
GDF15	ERG	3210	889-890	-1.1888	1.3756	-0.3829	0.1729	0.6778
PHACTR2	ERG	3764-3767	1730-1734	-0.2164	1.3524	-0.2903	-0.5159	-0.1993
METTL7A	ERG	3526	1369-1370	0.1343	1.3365	0.00952	-0.3431	0.4039
MAMDC2	ERG	3472	1280-1281	-0.6966	1.261	-0.6412	0.0634	0.5807
DDIT4	ERG	3008	576-577	-0.18479	1.0765	-0.21834	-0.0734	-0.0246
BCHE	ERG	2786	233	0.2464	1.0578	-0.9311	0.2318	0.5528

Gene Name	Cluster	*Polyp. SEQ ID NO:*	*Polyn. SEQ ID NO:*	NE (HES5-) vs (HES5+) score	ERG (HES5-) vs (HES5+) score	MRG (HES5-) vs (HES5+) score	LRG (HES5-) vs (HES5+) score	LNP (HES5-) vs (HES5+) score
OCIAD2	ERG	3694-3695	1615-1617	0.0784	1.0078	-0.3357	0.3066	0.4547
TNFRSF10D	ERG	4205	2417	-0.7479	0.8947	-0.4108	-0.1375	-1.3873
BBS9	ERG	2780-2783	223-228	-0.4758	0.86141	-0.4031	-0.372	-0.1048
ELOVL2	ERG	3072	678-679	-0.0124	0.8101	-0.664	-0.1011	0.1843
TUBA1C	ERG	4249	2483-2484	-0.2299	0.8007	-0.3942	-0.5482	-0.9278
CHST7	ERG	2913	428-429	0.0268	0.7718	0.0976	-0.0757	-0.4617
RBM47	ERG	3878-3879	1914-1916	-0.448	0.7712	0.40737	-0.3882	0.0426
TFPI	ERG	4160-4161	2343-2345	-0.5901	0.6703	-0.4895	-1.2285	0.0048
NEBL	ERG	3591-3592	1467-1470	-0.4069	0.538	0.1999	-0.5354	-0.2756
LHFP	ERG	3429	1217-1218	-0.3794	0.5294	-0.5056	-0.3867	0.1616
ACSS1	MRG	2691	80-81	-0.270467	-0.45911	0.5024	-0.4398	-0.0211
THBS1	LRG	4170	2361	-0.35176	0.75918	0.0237	2.1356	-0.0732
KLHL4	LRG	3404-3405	1181-1184	0.7061	0.5011	-0.42555	1.7673	1.3683
A2M	LRG	2683	70	0.47855	1.0222	-0.0532	1.7296	-0.5758
EN2	LRG	3075	682	0.11218	-0.052472	-1.6755	1.3766	-0.17283
SLC6A6	LRG	4004-4006	2115-2117	0.1108	-0.3772	0.0016	1.3488	0.1808
ACTA2	LRG	2692-2693	82-83	-0.6445	0.875	-0.1667	1.2355	0.8073
ST6GAL1	LRG	4091-4093	2236-2240	0.3897	0.1542	0.0576	1.2307	0.089
SLC7A8	LRG	4010-4011	2122-2125	-0.0729	0.4361	-0.3991	1.2015	0.0084
GRM3	LRG	3254	959	-0.7458	-2.0569	-0.5614	1.1071	0.498762
FAM65B	LRG	3124-3125	753-756	0.1798	-0.3971	-0.6371	1.0444	0.5189
CALB1	LRG	2842	325-326	0.3992	-0.0362	0.0183	0.9471	0.0915
MYLK	LRG	3560-3565	1416-1427	-0.3245	-0.2773	-0.0767	0.9439	0.5814
TNNT1	LRG	4209-4211	2424-2426	0.046	-0.483	0.4134	0.897	-0.2137
PTX3	LRG	3853	1876	-0.6687	-0.9317	-0.7792	0.8822	0.4459

Gene Name	Cluster	*Polyp. SEQ ID NO:*	*Polyn. SEQ ID NO:*	NE (HES5-) vs (HES5+) score	ERG (HES5-) vs (HES5+) score	MRG (HES5-) vs (HES5+) score	LRG (HES5-) vs (HES5+) score	LNP (HES5-) vs (HES5+) score
MFAP2	LRG	3528-3531	1373-1377	-0.421	-0.199	0.0027	0.8038	-0.4364
HMGA2	LRG	3303-3304	1028-1029	-0.2672	-0.0044	-0.0879	0.7119	-0.7995
FBN2	LNP	3133	769	-0.13009	0.4253	0.15333	0.7867	3.7195
NELL2	LNP	3605-3609	1486-1491	-0.4964	-1.2767	-0.4918	-1.1752	3.2656
KAL1	LNP	3377	1135	-0.5812	0.2855	-0.3048	-0.1062	2.7506
PCDHB5	LNP	3747	1707-1708	1.3941	-0.0852	-0.227	0.3274	1.9019
ST8SIA4	LNP	4094-4095	2241-2244	0.2505	0.934	-1.6846	0.9771	1.5033
DCN	LNP	2993-2998	553-563	-0.3094	-0.4793	-0.3578	-0.3001	1.432
SLC6A1	LNP	4003	2114	-0.2028	0.6132	0.0788	0.56847	1.37993
CADM2	LNP	2836-2838	316-319	0.5464	0.21	0.1286	0.8689	1.2335
BCL11A	LNP	2789-2791	237-242	0.2219	-0.1042	-0.1375	0.0432	1.122
DDB2	LNP	3007	575	-0.0478	-0.3001	0.0455	-0.217	1.0964
ANXA11	LNP	2718-2720	119-123	0.2913	0	-0.0041	-0.3842	1.0744
PAK1	LNP	3711-3712	1645-1646	-0.4161	0.3404	0.2821	0.3439	1.0339
ID3	LNP	3336	1077	0.5448	-0.4909	-0.1503	0.2195	0.8936
IGF2BP1	LNP	3345-3346	1090-1092	0.537	-0.0391	-0.1412	0.2159	0.8008
ANK3	LNP	2708-2709	105-107	0.422	-0.3703	-0.3011	-1.1461	0.727
ZEB2	LNP	4325-4326	2611-2613	0.13877	0.1069	-0.3251	-1.3468	0.7184
CREB5	LNP	2950-2953	494-500	-0.18241	0	-0.2906	-1.3562	0.6147

Table 2: HES5- vs. HES5+ expression intensity ratios during NSC progression. Genes whose expression is overrepresented in HES5- cells compared to HES5+ cells at the different stages are shown. The intensity level of each gene in ES cells was subtracted from all values prior to ratio calculation. Representative polynucleotides (Polyn.) and polypeptide (Polyp.) sequences for the indicated genes are provided.

Table 3
Global gene expression cluster analysis for stage specifically expressed genes

ID	*Probe SEQ ID NO:*	*symbols*	*cluster*	*Pattern*	*NE day12 HES5 +*	*NE day12 HES5-*	*ERG day14 HES5 +*	*ERG day14 HES5-*	*MRG day35 HES5 +*	*MRG day35 HES5-*	*LRG day80 HES5 +*	*LRG day80 HES5-*	*LNP day220 HES5 +*	*LNP day220 HES5-*
11722904 a_at	4375	TBC1D 9	1	Early Up	5.463	5.18	6.482	5.341	4.787	4.314	3.647	3.346	2.248	2.683
11729326_at	4376	ZFHX4	1	Early Up	6.518	6.501	5.689	4.939	6.272	5.869	4.09	3.55	3.531	3.538
11740477_a_at	4377	SP8	1	Early Up	4.658	4.1	4.599	4.044	4.334	4.047	3.97	3.11	0.306	4.26
11731596_a_at	4378	ZNF503	1	Early Up	5.204	5.3	4.919	5.842	6.251	5.975	3.464	2.861	2.379	2.125
11728653_at	4379	IGFBP5	1	Early Up	6.933	6.759	5.364	5.968	7.709	7.489	5.721	7.224	4.45	4.311
11735570_a_at	4380	SORBS 2	1	Early Up	4.093	4.422	3.821	2.814	3.54	3.329	2.846	3.008	2.376	2.097
11722466_s_at	4381	PAX6	1	Early Up	6.93	6.59	7.241	7.342	6.808	6.609	5.65	4.555	5.398	5.598
11718760_a_at	4382	WLS	1	Early Up	4.872	4.199	5.105	4.973	5.93	5.852	3.641	4.537	3.191	2.965
11732790_a_at	4383	PRDM1 6	1	Early Up	6.255	6.027	5.967	4.165	6.35	6.311	3.876	1.957	4.736	4.656
11723986_s_at	4384	NR2F2	1	Early Up	6.369	6.615	6.118	6.506	5.724	5.882	5.356	5.179	4.588	4.588
11739518_a_at	4385	FLRT3	1	Early Up	5.519	5.423	6.761	7.022	6.095	6.437	5.918	6.298	4.28	4.392
11716097_a_at	4386	RAP1A	2	Early Up	8.266	4.036	7.318	2.412	6.714	0.797	7.05	2.45	8.519	2.484
11758030_s_at	4387	FOXG1	2	Early Up	7.45	7.25	7.197	6.848	7.147	6.19	6.372	5.314	6.89	7.624
11757309_a_at	4388	COL2A 1	2	Early Up	6.618	6.507	6.419	5.998	6.526	6.052	4.534	3.208	7.261	6.036
11731211_at	4389	ARX	2	Early Up	6.904	6.921	6.427	4.104	5.949	5.498	6.061	4.92	7.417	7.273
11728189_a_at	4390	CXCR4	3	Early Up	6.306	5.777	4.792	4.084	5.374	4.457	6.252	6.236	5.532	6.398
11724337_s_at	4391	NRCA M	3	Early Up	2.73	4.588	3.882	3.737	3.789	2.88	5.518	5.862	5.142	5.214
11718157_s_at	4392	PTX3	3	Early Up	3.563	2.894	5.31	4.378	3.154	2.375	6.363	7.245	5.274	5.72

ID	*Probe*	*symbols*	*cluster*	*Pattern*	*NE*	*NE*	*ERG*	*ERG*	*MRG*	*MRG*	*LRG*	*LRG*	*LNP*	*LNP*
11719241_x_at	4393	CAMK2N1	3	Early Up	4.35	5.964	3.88	3.49	3.382	2.909	5.115	5.582	3.025	3.423
11729975_a_at	4394	PTPRN2	3	Early Up	3.039	4.603	4.405	3.707	4.002	3.566	4.621	4.807	5.377	5.539
11720059_a_at	4395	TTYH1	3	Early Up	4.654	3.643	4.433	2.62	4.512	4.08	5.274	5.017	5.265	5.001
11725274_s_at	4396	RGS20	3	Early Up	4.111	4.164	4.191	4.337	3.418	2.99	6.5	6.45	6.439	6.07
11719586_a_at	4397	TAGLN3	3	Early Up	3.353	5.749	2.828	2.585	4.725	4.335	4.895	4.942	4.184	4.137
11726870_at	4398	SETBP1	3	Early Up	3.472	3.581	3.35	2.83	4.195	3.836	4.737	4.593	4.362	4.605
11721011_at	4399	SMOC1	3	Early Up	5.317	5.178	4.251	3.104	4.659	4.351	5.384	5.454	5.951	5.817
11727582_a_at	4400	EPHA3	3	Early Up	4.433	4.888	5.317	3.269	5.408	5.183	5.227	4.192	5.963	6.851
11722352_s_at	4401	ACTA2	3	Early Up	5.163	4.518	5.172	6.047	3.417	3.25	5.992	7.228	3.856	4.663
11730066_s_at	4402	MAPK10	3	Early Up	4.535	4.732	4.186	3.673	4.327	4.213	5.585	5.72	4.572	4.618
11720601_x_at	4403	SYT11	3	Early Up	3.854	4.23	3.203	2.643	3.269	3.27	4.722	4.366	4.682	4.829
11736718_a_at	4404	NCALD	3	Early Up	5.35	5.499	4.382	3.68	3.76	3.87	5.656	5.677	4.53	4.553
11716395_a_at	4405	GPR56	4	Early Up	6.226	6.794	5.033	3.986	6.633	6.225	8.343	7.339	8.114	8.225
11733694_a_at	4406	C1ORF61	4	Early Up	2.415	2.657	3.308	1.079	6.246	5.85	8.063	7.487	8.375	7.946
11720447_s_at	4407	SOX9	4	Early Up	4.537	4.015	2.92	1.685	5.048	4.679	5.639	5.322	6.165	6.35
11733832_a_at	4408	MAP2	4	Early Up	6.923	7.982	6.78	5.853	8.025	7.688	8.564	8.322	8.57	8.893
11723961_s_at	4409	PTN	4	Early Up	0.755	0.549	0.847	0.798	4.258	3.981	6.307	6.165	6.209	6.186
11721248_s_at	4410	GNG2	4	Early Up	1.309	2.848	1.48	1.205	4.223	3.957	4.835	4.799	4.65	4.658
11721545_x_at	4411	ZEB1	4	Early Up	5.289	5.404	5.787	5.202	5.912	5.672	7.504	7.371	7.66	7.875
11730201_s_at	4412	GPM6A	4	Early Up	3.896	6.602	4.589	3.945	6.131	5.919	7.839	8.077	7.334	7.732
11722081_a_at	4413	RUFY3	4	Early Up	1.118	3.482	1.377	1.316	3.985	3.888	5.14	4.885	5.242	5.361
11719659_x_at	4414	ATP1A2	4	Early Up	2.238	2.018	1.685	1.285	4.951	4.929	5.155	4.899	5.176	5.16
11723616_a_at	4415	PLTP	5	Early Up	-0.309	0.032	-0.477	-0.417	2.815	1.81	1.731	2.154	2.57	2.333

ID	*Probe*	*symbols*	*cluster*	*Pattern*	*NE*	*NE*	*ERG*	*ERG*	*MRG*	*MRG*	*LRG*	*LRG*	*LNP*	*LNP*
11724950_a_at	4416	CPEB2	5	Early Up	0.191	2.942	0.334	1.842	3.278	2.515	3.575	3.378	3.718	4.3
11728085_s_at	4417	BMPR1B	5	Early Up	1.032	0.243	1.73	1.529	2.984	2.308	4.415	4.558	4.026	4.334
11744743_a_at	4418	TSHZ1	5	Early Up	0.017	1.93	1.217	0.959	4.175	3.532	4.354	4.218	4.45	4.733
11739276_at	4419	SDC3	5	Early Up	0.878	1.673	0.992	1.011	3.213	2.6	3.488	3.886	3.831	3.989
11725245_a_at	4420	SDK1	5	Early Up	0.545	0.748	0.897	0.404	3.03	2.439	2.78	2.175	3.442	3.087
11730603_a_at	4421	FAM13C	5	Early Up	2.448	2.062	2.099	2.639	4.135	3.555	4.24	4.761	2.345	3.181
11736387_a_at	4422	ABAT	5	Early Up	0.583	1.128	0.749	0.166	4.437	3.941	4.137	4.127	4.232	3.927
11747981_a_at	4423	APCDD1	5	Early Up	0.178	0.76	1.049	2.621	5.155	4.726	5.018	4.536	3.947	4.15
11720874_at	4424	CORO2B	5	Early Up	0.353	2.025	0.313	0.196	3.172	2.761	4.311	4.337	4.688	4.835
11743686_a_at	4425	ZNF436	5	Early Up	1.633	1.264	1.923	1.121	4.249	3.882	4.712	5.222	4.047	4.084
11721736_s_at	4426	TNFRSF19	5	Early Up	2.373	3.232	3.237	3.51	4.38	4.035	4.772	3.481	4.827	4.695
11725183_at	4427	KLF12	5	Early Up	0.593	1.561	0.929	0.633	2.369	2.027	2.777	2.677	2.769	3.147
11731331_at	4428	ZIC1	5	Early Up	3.503	3.158	1.688	1.509	5.972	5.678	5.325	5.441	5.653	5.634
11739813_a_at	4429	FZD1	5	Early Up	1.367	2.035	2.165	2.597	5.216	4.954	3.376	3.639	3.324	3.199
11737722_a_at	4430	GLIS3	5	Early Up	1.388	1.602	3.552	3.846	4.482	4.257	4.329	3.859	3.986	4.533
11715168_s_at	4431	POU3F3	5	Early Up	1.92	1.036	1.61	0.266	4.12	3.931	4.354	4.628	4.265	4.22
11725618_a_at	4432	GAD1	5	Early Up	0.995	3.331	1.028	2.219	4.838	4.651	2.994	3.107	4.797	5.547
11738076_a_at	4433	ZIC4	5	Early Up	2.558	1.868	1.159	1.117	7.041	6.894	4.852	5.042	5.403	4.457
11742867_at	4434	CDH6	5	Early Up	0.315	0.719	4.291	5.67	4.742	4.647	3.968	4.494	3.196	2.836
11739856_s_at	4435	ODZ2	5	Early Up	1.051	1.746	1.941	2.516	5.483	5.412	5.786	5.971	4.35	4.017
11737931_x_at	4436	STMN4	5	Early Up	1.088	5.574	1.207	1.091	4.506	4.44	4.688	3.921	4.469	4.436
11754980_a_at	4437	SMARCD3	5	Early Up	1.937	2.125	2.228	1.687	4.189	4.143	3.992	3.519	4.488	4.332
11717215_a_at	4438	ATP9A	5	Early Up	1.999	2.249	1.496	0.617	3.121	3.123	3.483	3.536	2.952	3.264

ID	*Probe*	*symbols*	*cluster*	*Pattern*	*NE*	*NE*	*ERG*	*ERG*	*MRG*	*MRG*	*LRG*	*LRG*	*LNP*	*LNP*
11718926_a_at	4439	ARID5B	5	Early Up	1.889	2.417	3.987	4.24	5.076	5.092	5.592	5.881	4.625	4.937
11720133_a_at	4440	PROS1	5	Early Up	1.732	1.948	2.613	3.378	4.275	4.333	4.347	4.379	3.097	3.27
11736745_a_at	4441	ZBTB20	5	Early Up	3.43	3.69	3.203	3.072	4.217	4.289	5.514	6.042	4.846	5.167
11719601_a_at	4442	MOXD1	5	Early Up	0.054	-0.101	0.136	1.527	5.423	5.528	3	1.706	5.953	5.688
11727170_at	4443	TRPS1	5	Early Up	0.054	0.118	0.812	0.642	2.039	2.243	3.866	3.077	3.981	3.984
11741008_a_at	4444	FGF12	5	Early Up	-0.367	0.907	-0.975	-1.659	3.886	4.112	2.089	2.103	2.348	1.822
11722503_at	4445	ITGB8	6	Early Up	1.347	2.116	2.354	2.407	3.983	2.742	4.63	4.644	4.975	5.479
11729986_a_at	4446	ROBO2	6	Early Up	1.344	4.606	2.626	2.72	4.543	3.516	5.852	5.921	4.973	5.682
11719883_a_at	4447	APOLD1	6	Early Up	2.237	1.748	2.632	2.494	2.877	2.161	5.144	5.053	6.717	6.653
11717440_a_at	4448	EPS8	6	Early Up	1.44	0.627	1.959	2.42	2.716	2.007	3.763	3.209	3.816	4.043
11744575_a_at	4449	ELOVL2	6	Early Up	2.656	2.643	2.818	3.628	2.764	2.1	6.572	6.471	6.351	6.535
11743434_a_at	4450	CHST11	6	Early Up	1.834	1.943	1.67	1.466	1.587	1.007	4.804	3.747	5.953	6.296
11724690_a_at	4451	RFX4	6	Early Up	2.62	2.316	3.853	2.796	4.55	3.973	6.715	6.295	6.351	6.655
11726389_a_at	4452	NLGN3	6	Early Up	2.136	1.785	1.911	0.692	3.411	2.901	4.142	3.518	5.155	5.569
11743338_at	4453	MEGF9	6	Early Up	1.693	2.138	2.21	2.325	3.454	3.08	4.903	4.49	5.669	5.287
11721773_at	4454	WSCD1	6	Early Up	1.564	1.369	1.217	0.361	2.336	1.963	4.097	3.843	3.9	4.124
11727656_s_at	4455	EPHB1	6	Early Up	1.497	1.847	1.624	0.948	2.661	2.406	3.22	3.598	4.411	4.453
11732793_a_at	4456	EYA4	6	Early Up	2.299	2.354	3.916	2.859	3.971	3.733	5.877	4.992	7.157	7.087
11733796_s_at	4457	RCAN1	6	Early Up	0.404	1.137	0.749	1.506	1.581	1.374	2.7	2.543	4.963	4.454
11740921_at	4458	NKAIN3	6	Early Up	1.27	0.867	1.213	0.348	3.395	3.201	5.143	4.57	5.288	5.08
11722350_a_at	4459	DCLK2	6	Early Up	1.634	2.144	1.81	1.203	3.128	2.943	5.398	5.12	4.492	4.542
11739232_x_at	4460	RASSF2	6	Early Up	0.971	1.021	1.276	0.67	1.84	1.662	4.329	3.926	4.582	4.795
11721911_at	4461	SOX8	6	Early Up	-0.475	0.138	0.592	0.686	0.793	0.632	3.607	3.557	3.845	3.73
11716967_a_at	4462	NRP2	6	Early Up	0.379	1.56	0.045	1.638	1.783	1.702	3.825	4.233	3.303	2.98
11735340_a_at	4463	TANC1	6	Early Up	-0.117	-0.017	0.231	0.979	2.205	2.158	2.8	2.937	3.078	3.629

ID	Probe	symbols	cluster	Pattern	NE	NE	ERG	ERG	MRG	MRG	LRG	LRG	LNP	LNP
11725326_a_at	4464	CAMK2D	6	Early Up	2.365	2.159	2.365	2.569	2.266	2.221	3.929	3.442	4.679	4.587
11731800_at	4465	C3ORF70	6	Early Up	1.559	2.126	0.702	-0.274	1.151	1.114	5.225	4.14	6.128	5.832
11750740_a_at	4466	ST3GAL5	6	Early Up	1.22	1.356	1.13	0.795	2.289	2.327	3.468	2.459	3.715	4.064
11740258_a_at	4467	MAML2	6	Early Up	-1.32	0.332	-0.216	0.592	1.272	1.416	3.001	2.441	3.467	3.342
11721430_a_at	4468	SYBU	6	Early Up	-0.235	3.407	0.357	0.954	3.085	3.237	4.69	4.29	5.652	4.648
11726903_at	4469	TRIL	6	Early Up	1.271	1.862	0.796	1.414	3.938	4.096	4.776	3.946	4.824	4.939
11717733_s_at	4470	ZCCHC24	6	Early Up	0.67	0.386	0.234	0.377	0.719	0.891	3.508	3.373	4.275	4.344
11715585_a_at	4471	PMP22	6	Early Up	2.942	2.813	2.125	2.445	4.017	4.269	6.396	6.236	7.684	6.956
11725937_a_at	4472	LGALS3	6	Early Up	-0.488	0.579	-0.763	-0.806	0.658	0.909	2.324	2.093	5.118	4.707
11719518_at	4473	SEMA5A	6	Early Up	1.038	1.319	2.612	2.166	2.638	2.916	5.085	4.537	6.007	5.942
11741951_a_at	4474	PCDH9	6	Early Up	1.888	4.529	3.873	4.321	4.812	5.145	6.343	6.465	6.147	6.264
11718578_a_at	4475	GSTM3	6	Early Up	1.734	0.636	2.944	2.764	4.021	4.406	4.396	4.372	4.752	4.754
11718541_a_at	4476	MTSS1	7	Early Up	1.164	2.489	1.467	1.181	4.054	3.25	4.814	5.138	5.024	5.253
11718021_s_at	4477	SSFA2	7	Early Up	2.741	2.863	2.225	2.976	3.836	3.046	4.551	4.341	4.122	4.88
11730233_s_at	4478	CTNND2	7	Early Up	2.703	3.809	3.15	2.291	4.135	3.376	4.904	4.752	4.429	4.676
11718876_at	4479	DNER	7	Early Up	1.044	5.638	1.662	2.353	6.059	5.356	7.559	7.43	7.991	8.196
11726505_at	4480	FUT9	7	Early Up	1.058	0.865	1.662	0.262	3.578	3.029	4.506	3.692	4.912	4.595
11718197_at	4481	LHFP	7	Early Up	2.78	2.401	3.466	3.995	4.654	4.148	4.801	4.414	5.288	5.449
11736196_a_at	4482	DCX	7	Early Up	2.161	6.632	3.573	3.343	6.366	5.865	5.365	5.651	5.002	5.974
11716433_s_at	4483	SCD5	7	Early Up	4.639	4.808	5.076	4.644	6.738	6.249	7.038	6.582	7.201	7.673
11723926_x_at	4484	PNMA2	7	Early Up	2.393	3.215	3.044	3.284	4.324	3.896	5.19	5.132	4.922	5.38

ID	*Probe*	*symbols*	*cluster*	*Pattern*	*NE*	*NE*	*ERG*	*ERG*	*MRG*	*MRG*	*LRG*	*LRG*	*LNP*	*LNP*
11736629_at	4485	BTBD17	7	Early Up	3.021	3.168	2.811	1.917	5.652	5.258	5.962	5.397	5.258	5.579
11755815_a_at	4486	CDH4	7	Early Up	2.026	2.171	1.802	0.792	3.468	3.114	4.35	3.788	4.675	4.668
11734905_a_at	4487	NOVA1	7	Early Up	1.825	2.593	2.296	1.58	3.329	3.048	4.105	4.031	4.219	4.201
11724338_a_at	4488	SLC35F1	7	Early Up	1.643	1.293	1.917	1.405	3.442	3.174	5.093	4.591	5.7	5.594
11718008_a_at	4489	CPNE2	7	Early Up	1.417	1.309	0.813	0.24	3.505	3.383	3.631	4.131	4.537	4.243
11733086_a_at	4490	MAF	7	Early Up	2.614	1.711	1.773	1.577	4.455	4.378	4.309	4.214	4.996	5.214
11724166_a_at	4491	GLT25D2	7	Early Up	3.481	2.916	2.692	3.09	4.288	4.248	5.358	4.366	6.922	7.13
11734075_a_at	4492	FBXL17	7	Early Up	1.158	1.447	1.835	1.971	3.156	3.147	3.651	3.457	3.89	3.889
11727914_s_at	4493	NCAM1	7	Early Up	3.891	5.942	4.307	3.322	5.157	5.153	6.496	6.194	6.435	6.615
11716405_a_at	4494	HTRA1	7	Early Up	1.274	1.091	0.932	0.42	3.082	3.164	3.723	2.934	4.069	3.963
11720815_a_at	4495	BAALC	7	Early Up	2.787	4.272	2.008	1.824	4.591	4.722	7.503	7.808	7.405	7.447
11734891_at	4496	SOBP	7	Early Up	1.679	2.11	1.752	1.016	2.792	2.971	3.629	3.132	4.875	4.827
11726178_a_at	4497	SOX6	7	Early Up	4.58	4.307	5.087	4.109	5.335	5.56	6.325	4.971	6.862	7.049
11724367_a_at	4498	SEMA6D	7	Mid Up	-0.109	3.64	0.561	0.308	2.161	2.417	4.457	3.426	5.781	5.48
11727447_a_at	4499	ST6GALNAC5	7	Mid Up	4.918	4.26	3.209	2.686	5.32	5.621	6.638	6.586	6.193	5.977
11726673_s_at	4500	ASCL1	8	Mid Up	0.555	1.285	0.98	1.115	4.09	3.354	6.455	5.69	6.629	7.245
11718972_at	4501	LPL	8	Mid Up	-0.708	-1.185	-0.369	-1.092	3.101	2.522	4.64	4.75	4.121	3.937
11720030_a_at	4502	SCG2	8	Mid Up	-0.694	-0.678	-1.266	-0.567	3.688	3.109	3.767	4.855	2.706	3.587
11717820_a_at	4503	SLC4A4	8	Mid Up	-0.171	-0.58	-0.275	-0.491	0.574	0	6.846	6.663	6.633	6.451
11724375_at	4504	PAG1	8	Mid Up	0.219	0.664	1.25	1.401	4.902	4.348	7.021	6	7.208	7.086
11729506_a_at	4505	NPAS3	8	Mid Up	0.683	0.978	1.595	1.614	4.374	3.947	6.327	5.962	6.336	6.323
11724135_a_at	4506	BAI3	8	Mid Up	0.539	1.01	0.805	0.714	3.709	3.417	5.272	4.956	5.477	5.286
11730070_at	4507	KIAA1239	8	Mid Up	0.022	0.206	0.022	0.044	3.421	3.13	5.85	4.432	6.974	7.025

ID	*Probe*	*symbols*	*cluster*	*Pattern*	*NE*	*NE*	*ERG*	*ERG*	*MRG*	*MRG*	*LRG*	*LRG*	*LNP*	*LNP*
11721263_at	4508	GRIA2	8	Mid Up	0.518	3.614	0.355	0.57	3.289	3.028	7.196	7.099	7.68	7.833
11729015_at	4509	SCN1A	8	Mid Up	0.573	0.254	0.179	0.037	5.694	5.515	6.895	6.971	6.696	6.563
11729469_at	4510	FAM181B	8	Mid Up	-0.286	-0.218	-0.179	-0.049	5	4.856	5.047	4.057	5.089	4.545
11718605_a_at	4511	SCRG1	8	Mid Up	0.082	0.029	1.475	1.413	1.427	1.286	6.107	5.479	6.827	6.753
11740182_a_at	4512	RHOJ	8	Mid Up	0.29	-0.053	-0.053	0.821	4.538	4.481	5.594	5.66	5.926	6.056
11725023_a_at	4513	SPARCL1	8	Mid Up	-0.053	-0.381	-0.381	0.275	3.427	3.4	6.157	6.727	5.724	5.951
11728570_at	4514	ZFPM2	8	Mid Up	0.037	0.203	0.183	0.691	3.941	3.919	5.733	5.733	5.053	5.564
11729750_a_at	4515	LPAR1	8	Mid Up	0.602	0.95	0.297	1.018	2.966	2.977	5.922	5.75	6.189	5.93
11720851_s_at	4516	NFIA	8	Mid Up	-0.217	0.661	-0.147	-0.121	6.808	6.903	7.386	6.747	7.555	7.728
11726291_a_at	4517	LRRC17	8	Mid Up	0.625	0.766	1.659	1.21	4.772	5.023	5.715	6.15	5.507	5.811
11725664_a_at	4518	EDNRB	9	Mid Up	0.764	0.076	2.309	1.689	4.659	3.023	5.928	6.008	4.856	5.228
11726251_a_at	4519	SLC1A3	9	Mid Up	0.251	0.077	2.16	1.615	2.138	1.395	7.251	6.474	7.802	8.214
11716062_a_at	4520	TNC	9	Mid Up	2.583	1.323	2.417	5.139	5.254	4.544	8.447	7.982	7.063	7.732
11722981_a_at	4521	CSPG5	9	Mid Up	-1.444	-1.24	0.6	-0.506	0.628	0.032	5.085	4.639	5.2	5.455
11735203_a_at	4522	MEGF10	9	Mid Up	-3.123	-3.561	-3.864	-3.619	1.494	0.928	3.145	2.286	3.739	3.786
11726434_at	4523	CHL1	9	Mid Up	0.165	3.83	1.042	1.404	6.573	6.029	6.548	6.481	5.277	5.407
11724886_a_at	4524	LRRN3	9	Mid Up	0.632	4.099	1.184	3.328	5.093	4.568	6.233	5.827	7.385	7.214
11722685_s_at	4525	LIFR	9	Mid Up	-0.789	-0.522	-0.813	-0.279	2.838	2.315	4.609	5.026	3.69	4.097
11724892_a_at	4526	NCAN	9	Mid Up	2.287	4.296	0.867	0.596	4.441	4.015	6.978	5.84	7.733	7.745
11716848_a_at	4527	PREX1	9	Mid Up	-0.082	0.135	0.871	1.585	4.005	3.594	4.939	4.624	4.855	4.996
11739921_a_at	4528	FAM123A	9	Mid Up	0.515	3.283	-0.054	0.431	5.124	4.775	6.989	6.275	6.528	7.16
11716932_a_at	4529	CCDC80	9	Mid Up	-0.219	-0.141	0.483	1.349	4.34	4.016	5.197	5.227	5.309	5.427
11720608_a_at	4530	S100B	9	Mid Up	-0.297	0.181	-0.149	0.084	3.818	3.896	8.264	8.347	8.224	8.283

ID	*Probe*	*symbols*	*cluster*	*Pattern*	*NE*	*NE*	*ERG*	*ERG*	*MRG*	*MRG*	*LRG*	*LRG*	*LNP*	*LNP*
11720946_at	4531	CA12	9	Mid Up	-0.33	-0.379	-0.443	-0.611	6.952	7.052	7.176	6.556	6.376	6.488
11740786_a_at	4532	SYNM	9	Mid Up	0.904	-0.045	0.643	1.595	3.256	3.359	5.414	3.677	6.365	5.897
11731648_a_at	4533	NTM	9	Mid Up	-1.317	-1.06	-0.743	-1.005	2.517	2.767	3.677	3.215	3.85	3.717
11717219_s_at	4534	NFIB	9	Mid Up	-2.474	-2.388	-2.776	-2.615	1.863	2.277	3.77	3.37	4.568	4.465
11741485_a_at	4535	SLITRK 3	9	Mid Up	-0.256	0.731	-0.284	-0.714	2.991	3.5	4.177	4.247	6.086	6.11
11743051_s_at	4536	CDKN2 C	10	Mid Up	0.253	0.216	1.078	`	2.5	1.095	4.496	4.35	5.078	6.359
11746149_x_at	4537	BCHE	10	Late Up	-0.246	0	-0.218	0.839	2.217	1.286	4.071	4.303	4.597	5.149
11736575_x_at	4538	GPR155	10	Late Up	0.059	1.261	0.432	0.494	1.247	0.702	4.692	3.61	6.305	6.162
11722282_a_at	4539	LAMA4	10	Late Up	1.046	0.683	0.321	0.412	0.943	0.462	5.253	5.269	5.705	5.094
11731086_a_at	4540	NXPH1	10	Late Up	0.098	1.641	-0.032	0.382	0.308	-0.139	2.109	3.014	6.149	6.317
11730475_a_at	4541	GRIA4	10	Late Up	0	0.338	0.176	-0.051	0.652	0.22	4.703	3.703	6.361	6.143
11735282_at	4542	SLC10A 4	10	Late Up	0.179	-0.167	-0.412	-0.517	-0.01	-0.374	1.627	1.62	5.464	3.721
11733376_s_at	4543	DLX1	10	Late Up	-0.725	-0.142	-0.25	0.255	0.126	-0.221	3.56	3.015	6.591	6.743
11723862_at	4544	SLC6A1	10	Late Up	0.042	0.181	-0.257	-0.15	0.175	-0.15	4.177	2.831	5.485	6.204
11739379_a_at	4545	FAM19 8B	10	Late Up	0.386	0.249	1.881	5.579	0.695	0.392	4.914	2.596	6.157	6.525
11719174_a_at	4546	NT5E	10	Late Up	-0.192	0.021	-0.021	0.223	-0.042	-0.324	4.689	4.137	6.657	6.182
11733121_s_at	4547	DAZL	10	Late Up	-2.025	-1.621	-1.951	-2	-1.707	-1.983	-0.992	-1.435	3.514	3.561
11734197_at	4548	LUZP2	10	Late Up	-2.258	-2.253	-2.516	-1.964	-2.64	-2.776	1.323	-0.637	3.381	3.369
11731198_at	4549	LHFPL3	10	Late Up	0.051	0.029	-0.091	0.058	0.298	0.173	6.319	6.484	7.297	7.918
11733505_at	4550	C2ORF 80	10	Late Up	-0.069	-0.127	0.02	-0.083	0.314	0.195	5.727	6.219	5.283	5.676
11745172_a_at	4551	SYNPR	10	Late Up	0.034	0.621	-0.14	-0.14	1.391	1.276	2.098	1.817	4.841	4.278
11726427_at	4552	GABBR 2	10	Late Up	-0.02	1.947	-0.384	-0.57	-0.463	-0.541	2.774	0.35	5.484	4.453
11727401_x_at	4553	HOPX	10	Late Up	0.036	0.424	0.234	0.276	0.306	0.252	6.912	5.545	7.455	7.005

ID	*Probe*	*symbols*	*cluster*	*Pattern*	*NE*	*NE*	*ERG*	*ERG*	*MRG*	*MRG*	*LRG*	*LRG*	*LNP*	*LNP*
11715363_a_at	4554	A2M	10	Late Up	-0.454	0.025	-0.164	0.858	1.658	1.605	3.238	4.968	5.598	5.022
11722619_a_at	4555	AQP4	10	Late Up	-0.44	-0.147	-0.361	-0.404	-0.23	-0.246	5.223	4.843	8.462	8.399
11735180_a_at	4556	DLX2	10	Late Up	-0.104	0.58	0.216	1.528	0.465	0.457	3.566	3.302	7.357	7.327
11736090_a_at	4557	OLIG2	10	Late Up	0.373	3.438	0.455	0.051	0.065	0.122	3.362	1.76	5.414	4.602
11719950_a_at	4558	BCAN	10	Late Up	-0.371	-0.049	-0.264	-0.442	-0.29	-0.22	6.865	4.864	7.647	8.143
11740875_at	4559	KCNA2	10	Late Up	0.044	1.854	0.381	0.065	0.127	0.3	5.414	5.287	5.936	5.606
11720338_at	4560	FIBIN	10	Late Up	0.129	0.189	-0.048	0.458	0.159	0.346	3.969	2.976	6.627	5.41
11725757_at	4561	ISL1	10	Late Up	-0.385	-0.438	-0.27	0.244	-0.329	-0.138	-0.237	-0.449	5.419	4.168
11728197_at	4562	PMP2	10	Late Up	-0.034	-0.014	-0.172	-0.142	-0.258	-0.014	4.864	1.388	7.411	8.106
11721241_at	4563	OLIG1	10	Late Up	0.038	0.197	0.163	0.029	0.084	0.341	5.216	3.645	7.617	8.064
11756560_a_at	4564	NDP	10	Late Up	0.413	0.055	0.808	0.314	0.431	0.844	5.802	5.055	6.545	6.22
11723211_at	4565	KCNJ10	10	Late Up	0.209	0.042	-0.119	-0.308	0.445	0.868	4.683	4.14	5.59	5.011
11725782_at	4566	C5ORF 39	10	Late Up	0.082	0.019	0.012	-0.292	-0.067	0.425	2.261	1.706	4.304	2.848
11724102_s_at	4567	PDZRN 3	11	Late Up	-1.455	-0.033	-0.96	0.874	1.441	0.305	3.793	3.793	3.467	3.996
11731357_a_at	4568	ADAM TS3	11	Late Up	-0.625	-0.848	-0.243	0.095	2.722	1.607	3.418	3.33	3.941	4.122
11721269_a_at	4569	SCG3	11	Late Up	-4.6	-1.54	-4.6	-4.526	-3.192	-4.072	-0.272	-1.125	1.324	1.492
11731615_s_at	4570	CNR1	11	Late Up	-0.068	3.206	0.112	-0.447	2.814	2.08	4.804	4.254	6.358	5.921
11733310_a_at	4571	PCYT1 B	11	Late Up	-2.175	-2.033	-2.657	-2.544	0.088	-0.55	1.032	1.157	2.144	1.693
11725458_a_at	4572	KLHDC 8A	11	Late Up	0.93	0.956	0.717	0.186	2.055	1.451	2.859	2.428	4.459	4.202
11757631_s_at	4573	VGLL3	11	Late Up	-0.708	1.277	-1.25	0.686	0.69	0.141	2.685	2.846	3.377	4.666
11724534_a_at	4574	DOK5	11	Late Up	-0.91	-0.893	-1.139	-0.852	1.089	0.608	2.638	1.953	3.435	3.374
11741056_a_at	4575	ADORA 2B	11	Late Up	-1.014	-0.896	-1.058	-0.898	-0.991	-1.451	2.508	1.112	3.148	2.746
11729175_at	4576	KLHL4	11	Late Up	-1.096	-0.389	-1.66	-1.159	0.384	-0.042	1.331	3.099	2.312	3.68

ID	*Probe*	*symbols*	*cluster*	*Pattern*	*NE*	*NE*	*ERG*	*ERG*	*MRG*	*MRG*	*LRG*	*LRG*	*LNP*	*LNP*
11720617_at	4577	TRIM9	11	Late Up	0.083	0.257	0.816	1.504	1.662	1.237	4.249	2.99	6.606	7.014
11718492_at	4578	SLC1A2	11	Late Up	-0.193	1.68	1.278	0.645	2.258	1.836	2.805	2.799	6.408	6.409
11739769_a_at	4579	TMOD2	11	Late Up	0.853	2.874	1.346	1	1.346	0.941	3.79	3.321	5.272	6.007
11740942_at	4580	SALL3	11	Late Up	-2.782	-2.858	-1.26	-2.138	-1.016	-1.365	0.777	-1.035	1.938	1.52
11731561_a_at	4581	APC	11	Late Up	-0.108	0.402	0.315	-0.103	1.116	0.795	3.474	2.695	3.941	4.046
11733452_a_at	4582	IRX5	11	Late Up	-0.086	0.67	-0.187	0.289	0.74	0.436	4.191	4.287	3.107	2.796
11734325_a_at	4583	CADM2	11	Late Up	0.026	-0.157	0.82	0.82	1.242	0.951	4.133	2.777	4.408	5.022
11720451_at	4584	PLEKHH2	11	Late Up	0.701	0.037	1.308	0.72	1.915	1.631	3.876	3.648	4.096	3.775
11730521_a_at	4585	SLITRK2	11	Late Up	3.231	2.837	0.102	-0.072	0.709	0.429	6.293	5.191	6.315	6.399
11722791_s_at	4586	MATN2	11	Late Up	-1.433	-1.873	-1.647	-1.436	-1.227	-1.469	2.587	3.808	2.151	2.811
11736676_a_at	4587	DPP10	11	Late Up	0.224	0.426	0.14	-0.606	0.12	-0.103	2.42	3.146	3.657	3.785
11729445_at	4588	FAM150B	11	Late Up	-1.172	-1.087	-1.566	-1.294	-1.027	-1.234	2.078	2.482	2.947	4.086
11716804_at	4589	PLEKHA2	11	Late Up	-1.793	-1.7	-1.875	-1.929	-2.112	-2.179	1.303	0.814	2.733	2.753
11735207_a_at	4590	MDGA2	11	Late Up	0.242	0.565	0.005	0.053	2.404	2.364	4.719	3.854	4.918	4.89
11724002_a_at	4591	DPP6	11	Late Up	0.429	1.857	-0.099	-0.666	0.958	0.98	3.043	1.697	4.338	4.376
11731711_a_at	4592	CACNB2	11	Late Up	0.01	-0.185	0.384	-0.115	-0.058	0.064	4.357	3.341	5.143	5.134
11735194_a_at	4593	NR3C1	11	Late Up	-1.825	-1.208	-1.326	-1.527	-0.006	0.15	2.668	1.833	3.771	3.75
11721982_a_at	4594	PID1	11	Late Up	-0.507	0.713	-0.669	-0.479	-1.332	-1.12	2.077	1.694	4.051	3.808
11731431_at	4595	PCDH10	11	Late Up	0.057	-0.091	-0.042	-0.055	1.108	1.401	4.56	3.169	5.178	4.8
11743850_at	4596	PDGFRA	11	Late Up	-0.403	-0.236	-0.399	0.495	0.905	1.202	4.608	3.078	7.274	7.137
11731578_s_at	4597	ACSS1	11	Late Up	0.261	-0.01	-0.04	-0.499	-0.191	0.311	3.813	3.374	4.107	4.086
11747942_a_at	4598	ZEB2	12	Late Up	3.571	3.821	3.97	4.904	3.566	1.882	4.075	5.052	4.166	5.67

ID	*Probe*	*symbols*	*cluster*	*Pattern*	*NE*	*NE*	*ERG*	*ERG*	*MRG*	*MRG*	*LRG*	*LRG*	*LNP*	*LNP*
11736341_a_at	4599	EGR3	12	Late Up	3.709	3.561	3.497	1.662	-1.167	-1.697	1.435	2.104	1.94	2.387
11739154_at	4600	FAM129A	12	Late Up	-3.617	-3.58	-1.99	0.221	-3.456	-3.983	-0.954	-2.227	0.966	1.646
11715947_a_at	4601	KCTD12	12	Late Up	0.204	0.72	-0.466	-0.486	-0.2	-0.724	0.668	0.37	3.258	3.687
11730820_at	4602	PCDHB10	12	Late Up	1.749	2.654	0.007	-0.009	1.258	0.739	2.686	1.849	3.746	3.94
11723032_a_at	4603	PLP2	12	Late Up	0.294	0.263	-0.945	-0.856	-0.39	-0.806	-0.214	-0.884	2.862	1.386
11739822_a_at	4604	CPNE8	12	Late Up	1.391	1.374	0.651	0.996	0.182	-0.216	1.646	1.985	3.392	3.027
11719899_at	4605	C11ORF96	12	Late Up	2.441	3.425	1.618	1.309	0.369	0.024	2.762	2.16	4.046	4.499
11722544_a_at	4606	EYA2	12	Late Up	0.56	2.591	1.286	1.636	0.386	0.055	3.785	2.925	4.148	4.389
11728940_a_at	4607	KCNIP1	12	Late Up	-0.162	0.351	-0.005	-0.447	0.964	0.65	1.725	0.776	4.015	3.461
11724501_at	4608	SPON1	12	Late Up	1.883	1.656	2.592	2.032	-0.226	-0.454	3.456	1.897	3.994	4.688
11743877_a_at	4609	PPP1R1C	12	Late Up	-0.031	0.298	-1.412	-1.048	-2.131	-2.275	-2.131	-2.372	3.073	2.405
11723863_a_at	4610	DOCK10	12	Late Up	-0.375	1.007	1.131	1.376	0.007	-0.079	2.394	1.922	4.104	2.919
11716384_at	4611	CCL2	12	Late Up	0.962	0.634	3.58	6.019	2.26	2.226	1.265	2.559	2.737	3.146
11730472_a_at	4612	PDE3B	12	Late Up	-0.549	-0.497	0.309	0.477	-0.997	-0.991	-0.64	-1.954	1.675	1.803
11719363_s_at	4613	PRRX1	12	Late Up	-0.237	-0.315	-0.167	0.525	2.034	2.192	0.727	0.53	4.643	4.334
11716525_a_at	4614	SPON2	12	Late Up	2.431	4.411	-0.043	0.952	0.504	0.674	0.612	1.247	1.432	1.119
11756820_a_at	4615	IFIT1	12	Late Up	5.175	3.921	1.826	0.582	0.957	1.144	2.015	1.389	2.184	2.445
11725256_at	4616	FZD10	13	Late Up	-0.64	-0.264	0.544	-0.091	2.71	0.993	4.4	4.657	-1.045	-0.513
11735538_at	4617	EN2	13	Late Up	-0.175	-0.062	0.005	-0.048	2.905	1.23	4.941	6.317	0.014	-0.159
11732647_a_at	4618	PCDH8	13	Late Up	-2.173	-1.638	-0.47	2.889	0.472	-1.13	3.585	4.578	0.097	-1.329
11729446_a_at	4619	DCT	13	Late Up	0.7	0.816	0.714	0.7	1.981	0.996	5.035	3.409	1.337	2.972
11730172_a_at	4620	GRM3	13	Late Up	1.985	1.239	2.497	0.44	1.285	0.724	4.239	5.346	-0.504	-0.006
11732841_a_at	4621	SCN3A	13	Late Up	0.087	3.716	-0.212	-0.233	2.861	2.486	4.692	5.64	0.547	1.716

ID	*Probe*	*symbols*	*cluster*	*Pattern*	*NE*	*NE*	*ERG*	*ERG*	*MRG*	*MRG*	*LRG*	*LRG*	*LNP*	*LNP*
11737405_at	4622	ST8SIA 4	13	Late Up	0.198	1.592	1.074	0.989	4.204	3.977	5.182	5.509	0.565	2.467
11725275_at	4623	IRX1	13	Late Up	0.662	1.159	0.339	0.523	0.313	0.086	3.857	3.765	-0.571	0
11723852_a_at	4624	BCL11 A	13	Late Up	-5.24	-4.949	-5.073	-5.073	-0.087	-0.091	-1.586	-1.97	-4.079	-3.004
11725979_a_at	4625	CALB1	13	Late Up	-3.004	-2.604	-2.836	-2.872	2.578	2.596	1.556	2.503	-2.289	-2.197
11725585_at	4626	PI15	13	Late Up	0.333	2.26	-0.031	0.846	-0.11	0.015	4.427	4.025	1.714	1.658
11730353_a_at	4627	CYP26 B1	13	Late Up	-0.02	0.493	0.093	-0.31	-0.169	-0.027	4.635	5.067	0.055	0.198
11736001_s_at	4628	DCN	13	Late Up	-0.055	-0.185	-0.199	0.227	-0.171	-0.018	5.265	6.052	0.292	4.011
11720877_at	4629	ABTB2	13	Late Up	-1.047	-0.854	-0.678	-0.411	0.444	1.141	3.17	3.422	0.919	1.037
11724937_a_at	4630	DLL1	14	Early Up and Transient	5.851	6.051	5.699	4.592	4.544	3.93	3.93	4.163	3.288	3.837
11724094_a_at	4631	COL4A 5	14	Early Up and Transient	3.403	3.214	2.656	1.982	3.931	3.435	1.145	0.863	1.135	0.891
11724604_a_at	4632	HS6ST2	14	Early Up and Transient	4.575	4.382	3.716	2.924	2.656	2.251	1.992	1.406	-1.117	-0.884
11723374_at	4633	GLI3	14	Early Up and Transient	4.212	3.995	3.739	2.803	2.593	2.235	1.862	1.836	1.443	1.285
11720823_at	4634	JAG1	14	Early Up and Transient	3.345	3.052	2.506	2.709	0.183	-0.091	0.962	0.808	0.779	1.25
11721894_at	4635	GAS1	14	Early Up and Transient	4.135	3.974	4.174	3.365	4.659	4.498	2.526	2.664	2.638	2.644
11730056_at	4636	MSI1	14	Early Up and Transient	2.806	2.849	2.519	1.799	2.735	2.586	1.328	1.168	1.062	1.005
11727552_at	4637	FZD2	14	Early Up and Transient	4.217	3.577	3.642	3.21	2.99	2.876	1.509	1.638	1.167	1.25
11725409_s_at	4638	NAV1	14	Early Up and Transient	4.484	4.841	4.217	3.07	3.044	3.001	2.432	1.794	3.296	3.09
11733083_s_at	4639	PALLD	14	Early Up and	3.572	3.061	3.669	3.454	2.565	2.583	1.691	1.872	-0.161	0.676

ID	*Probe*	*symbols*	*cluster*	*Pattern*	*NE*	*NE*	*ERG*	*ERG*	*MRG*	*MRG*	*LRG*	*LRG*	*LNP*	*LNP*
				Transient										
11744107_x_at	4640	DLK1	14	Early Up and Transient	8.317	8.346	8.216	8.397	4.566	4.707	4.225	4.573	3.128	5.067
11732331_s_at	4641	ARL4C	14	Early Up and Transient	4.703	4.407	4.322	4.257	3.4	3.758	2.152	1.389	2.64	2.171
11734150_x_at	4642	PLAGL1	15	Early Up and Transient	6.868	6.632	6.221	5.585	6.462	5.683	3.228	3.749	-0.38	-0.295
11734956_a_at	4643	CNTNAP2	15	Early Up and Transient	2.464	2.762	2.426	2.575	4.089	3.376	-0.692	-0.832	-4.391	-3.46
11720326_a_at	4644	FRZB	15	Early Up and Transient	5.478	4.807	5.297	3.212	3.772	3.356	-0.42	-0.322	-1.275	-1.352
11723440_a_at	4645	EPHA4	15	Early Up and Transient	5.025	4.901	4.604	4.141	5.029	4.666	2.521	2.61	-0.352	0.487
11719118_a_at	4646	TPBG	15	Early Up and Transient	3.071	2.816	3.047	2.686	4.325	4.028	-1.847	-0.976	-0.846	-2.495
11728010_at	4647	PRTG	15	Early Up and Transient	7.785	7.829	7.775	7.764	4.545	4.336	1.141	0.722	1.417	1.258
11749204_a_at	4648	LGI1	15	Early Up and Transient	7.755	7.336	7.013	8.058	7.413	7.205	4.252	3.488	1.321	2.152
11717825_a_at	4649	STMN2	15	Early Up and Transient	3.575	8.006	1.714	4.225	7.448	7.27	3.294	3.778	-0.605	-0.71
11725910_a_at	4650	COL4A6	15	Early Up and Transient	5.363	5.003	4.177	3.425	4.865	4.704	-0.363	-0.979	-1.188	-2.26
11727088_a_at	4651	DACH1	15	Early Up and Transient	7.082	7.123	6.672	5.212	7	6.84	3.159	2.588	0.097	0.24
11723388_a_at	4652	EMX2	15	Early Up and Transient	5.661	5.345	4.772	3.456	6.732	6.573	1.706	0.775	-0.409	-0.333
11727862_at	4653	LIX1	15	Early Up and Transient	6.559	6.015	6.779	5.161	3.935	3.963	1.808	1.558	-0.058	0.284
11717416_a_at	4654	RHOU	15	Early Up and	4.093	4.022	3.137	3.216	5.06	5.159	0.874	0.706	0.525	0.564

ID	*Probe*	*symbols*	*cluster*	*Pattern*	*NE*	*NE*	*ERG*	*ERG*	*MRG*	*MRG*	*LRG*	*LRG*	*LNP*	*LNP*
				Transient										
11720433_s_at	4655	FLRT2	15	Early Up and Transient	1.674	2.645	2.471	3.595	4.074	4.216	0.813	1.569	-2.3	-1.948
11729604_s_at	4656	GPC3	15	Early Up and Transient	4.984	4.732	4.287	3.923	2.482	2.704	0.759	-0.282	-3.221	-3.235
11721399_a_at	4657	SLIT2	15	Early Up and Transient	5.195	4.397	4.252	4.685	5.985	6.548	1.457	2.709	-0.712	-0.591
11722766_a_at	4658	MBNL2	16	Early Up and Transient	4.54	4.598	4.254	5.299	4.756	3.921	2.95	2.91	3.515	3.683
11724650_a_at	4659	MEIS2	16	Early Up and Transient	5.373	5.487	4.705	5.287	3.317	2.535	2.22	2.514	2.464	3.128
11745255_a_at	4660	FBXO16	16	Early Up and Transient	3.383	2.9	1.636	0.646	2.143	1.476	0.33	0.037	-0.399	0.221
11733115_a_at	4661	CA2	16	Early Up and Transient	2.515	1.711	2.192	2.21	1.988	1.325	0.349	-0.124	-0.968	-1.056
11737404_at	4662	GREB1	16	Early Up and Transient	1.316	1.77	2.669	1.933	1.32	0.696	0.073	0.247	-0.708	-0.217
11731237_at	4663	ZDBF2	16	Early Up and Transient	1.458	2.128	1.09	1.505	1.646	1.201	0.291	0.352	-0.168	-0.145
11728428_a_at	4664	CCDC136	16	Early Up and Transient	2.426	2.386	1.376	1.038	1.694	1.279	0.283	0.494	0.05	-0.461
11739166_a_at	4665	LYPD1	16	Early Up and Transient	2.695	2.036	3.688	5.633	1.739	1.384	3.407	3.6	2.036	0.931
11719632_a_at	4666	ABCC4	16	Early Up and Transient	1.608	1.03	1.65	1.031	-0.283	-0.629	-0.627	-0.712	-1.095	-0.711
11744717_a_at	4667	C18ORF8	16	Early Up and Transient	0.467	3.828	1.006	1.749	1.077	0.749	0.649	0.479	0.581	0.573
11726581_a_at	4668	GUCY1B3	16	Early Up and Transient	4.528	4.308	3.521	2.52	2.628	2.332	3	2.72	2.518	2.55
11728973_at	4669	HES5	16	Early Up and	6.247	5.514	4.911	2.408	3.425	3.134	3.363	2.45	1.715	2.211

ID	Probe	symbols	cluster	Pattern	NE	NE	ERG	ERG	ERG	MRG	MRG	LRG	LRG	LNP	LNP
				Transient											
11728497_s_at	4670	SVIL	16	Early Up and Transient	2.937	2.896	2.8	2.238	1.678	1.423	0.685	0.45	1.34	1.439	
11717516_s_at	4671	EFHD1	16	Early Up and Transient	4.345	3.781	4.299	3.83	3.43	3.238	1.967	0.86	0.326	0.727	
11717619_at	4672	EPHA2	16	Early Up and Transient	3.038	2.566	2.846	3.584	1.565	1.386	0.703	1.012	0.794	1.042	
11729959_a_at	4673	ACCN2	16	Early Up and Transient	2.83	4.417	2.818	2.641	1.729	1.556	2.499	2.372	2.143	2.234	
11733703_x_at	4674	C10ORF140	16	Early Up and Transient	3.89	3.621	3.321	2.707	0.773	0.632	0.612	0.499	-0.61	0.041	
11727277_a_at	4675	TMEM169	16	Early Up and Transient	5.002	4.807	4.32	3.64	2.866	2.732	1.969	1.389	1.321	1.395	
11732915_a_at	4676	KCNT2	16	Early Up and Transient	2.238	2.995	2.741	2.566	1.925	1.81	2.05	1.665	0.862	0.704	
11715765_a_at	4677	DUSP1	16	Early Up and Transient	1.961	2.424	1.573	1.355	0.181	0.074	-0.37	0.262	0.293	0.218	
11716771_s_at	4678	SIK1	16	Early Up and Transient	2.691	2.653	1.959	1.52	2.145	2.047	-0.695	0.099	1.078	1.198	
11741550_a_at	4679	OSBPL6	16	Early Up and Transient	2.885	2.27	2.23	1.273	-0.004	-0.045	0.98	-0.004	-1.763	-0.708	
11737710_a_at	4680	LFNG	16	Early Up and Transient	3.903	3.577	3.105	1.527	1.064	1.03	1.483	1.416	0.268	1.064	
11719441_a_at	4681	WNT5B	16	Early Up and Transient	6.043	5.628	4.758	3.39	4.624	4.602	1.598	1.422	2.083	2.09	
11753903_a_at	4682	TP53TG1	16	Early Up and Transient	3.403	2.228	2.929	2.431	2.001	2.049	0.357	0.337	0.525	0.343	
11726012_a_at	4683	JUB	16	Early Up and Transient	2.919	2.211	2.644	2.926	0.85	0.926	-0.774	-0.246	-0.394	-0.91	
11725181_a_at	4684	ARHGE	16	Early Up and	3.858	4.057	1.631	0.536	1.547	1.623	0.568	-0.316	1.101	0.819	

ID	*Probe*	*symbols*	*cluster*	*Pattern*	*NE*	*NE*	*ERG*	*ERG*	*MRG*	*MRG*	*LRG*	*LRG*	*LNP*	*LNP*
		F4		Transient										
11726878_s_at	4685	C1ORF 213	16	Early Up and Transient	5.245	4.942	4.22	3.562	2.498	2.653	1.919	1.69	2.06	1.852
11733085_a_at	4686	ENO3	16	Early Up and Transient	3.332	3.168	2.244	2.327	1.77	1.93	-0.796	-0.901	-1.145	-1.2
11735332_a_at	4687	EXD3	16	Early Up and Transient	3.004	2.783	2.416	2.15	1.885	2.071	1.816	1.547	1.212	0.927
11720507_at	4688	FBN1	16	Early Up and Transient	3.748	3.061	3.315	2.752	3.28	3.527	0.564	0.925	0.824	0.752
11720743_x_at	4689	CROT	16	Early Up and Transient	4.506	4.016	3.178	2.591	1.853	2.241	0.957	0.458	0.858	0.78
11717065_a_at	4690	GREM1	17	Early Up and Transient	0.168	0.351	2.957	6.666	1.763	0.335	-0.183	-0.359	-0.04	-0.342
11720911_a_at	4691	CAPN6	17	Early Up and Transient	3.033	2.988	3.729	3.855	0.579	-0.254	-2.151	-2.348	-1.962	-2.121
11733365_a_at	4692	NR2E1	17	Early Up and Transient	1.237	1.039	5.275	2.402	2.096	1.322	1.362	0.082	0.636	0.504
11727122_at	4693	LRP2	17	Early Up and Transient	6.632	6.185	6.147	4.896	1.628	1.137	2.223	1.585	-0.166	-0.046
11758013_s_at	4694	C8ORF 4	17	Early Up and Transient	3.821	2.953	2.723	4.57	0.869	0.43	0.661	0.164	0.377	0.351
11716663_a_at	4695	GDF15	17	Early Up and Transient	2.33	1.141	1.93	3.305	-0.555	-0.938	-0.573	-0.4	-0.876	-0.198
11737871_a_at	4696	TMEFF 2	17	Early Up and Transient	-0.415	0.594	2.215	4.275	2.243	1.944	0.72	1.297	-0.323	-0.33
11724403_at	4697	NOS2	17	Early Up and Transient	6.968	6.641	6.754	5.493	0.941	0.647	0.155	0.173	0.121	0.111
11727014_at	4698	PAPPA	17	Early Up and Transient	0.719	4.881	4.359	6.929	0.766	0.497	-0.08	0.027	-0.045	-0.281
11717910_at	4699	CXCL1	17	Early Up and	2.446	4.989	1.365	2.179	0.784	0.528	-0.465	-0.394	-0.349	-0.327

ID	*Probe*	*symbols*	*cluster*	*Pattern*	*NE*	*NE*	*ERG*	*ERG*	*MRG*	*MRG*	*LRG*	*LRG*	*LNP*	*LNP*
		4		Transient										
11731599_a_at	4700	FOSL2	17	Early Up and Transient	5	4.719	3.283	2.96	0.225	-0.023	-0.278	-0.09	0.268	0.271
11715435_s_at	4701	TIMP3	17	Early Up and Transient	5.768	5.341	4.737	4.655	2.815	2.671	1.011	0.049	0.512	0.372
11731445_at	4702	ANKRD 1	17	Early Up and Transient	1.032	1.43	4.597	7.053	-0.066	-0.176	-0.007	-0.226	-0.201	-0.112
11743417_a_at	4703	LEPRE L1	17	Early Up and Transient	-0.215	-0.973	0.379	2.057	-1.065	-1.174	-3.452	-3.321	-3.113	-3.452
11728431_a_at	4704	SCUBE 3	17	Early Up and Transient	4.186	3.604	3.589	2.467	1.63	1.622	0.831	0.379	0.225	0.416
11731680_a_at	4705	NOX4	17	Early Up and Transient	-0.079	-0.177	2.744	4.332	0.008	0.052	-0.329	-0.594	-0.251	-0.386
11737166_at	4706	ANKRD 34B	17	Early Up and Transient	5.191	3.729	1.828	0.221	-0.38	-0.29	-0.372	-0.38	-0.232	-0.334
11728872_a_at	4707	RBM24	17	Early Up and Transient	2.979	2.76	3.09	4.041	0.686	0.814	0.451	0.754	-0.223	-0.088
11721135_s_at	4708	IGFBP3	17	Early Up and Transient	5.104	4.65	3.006	2.609	-0.654	-0.496	-0.448	-0.697	-1.136	-0.832
11742292_a_at	4709	DRD4	17	Early Up and Transient	4.657	3.845	3.272	2.33	-0.409	-0.231	-0.724	-1.014	-1.014	-1.117
11720633_a_at	4710	FLNC	17	Early Up and Transient	2.277	2.351	2.24	3.726	-1.07	-0.851	-1.078	-1.244	-1.32	-1.791
11736262_at	4711	INHBE	17	Early Up and Transient	-1.946	-1.884	0.788	2.413	-2.948	-2.726	-2.804	-3.018	-2.938	-3.249
11723666_a_at	4712	SEMA3 C	18	Unique	1.225	3.542	1.132	4.326	5.619	4.544	0.508	2.111	-0.019	-0.214
11724556_at	4713	NEURO D1	18	Unique	0.952	5.719	0.537	1.433	3.524	2.606	0.509	0.509	-0.215	0
11726937_at	4714	ST18	18	Unique	1.158	6.718	0.585	1.841	2.057	1.329	2.118	2.325	0.483	0.222

ID	*Probe*	*symbols*	*cluster*	*Pattern*	*NE*	*NE*	*ERG*	*ERG*	*MRG*	*MRG*	*LRG*	*LRG*	*LNP*	*LNP*
11725779_a_at	4715	CADPS	18	Unique	-0.251	3.315	-0.928	-0.81	2.703	2.03	-0.114	-0.71	-1.348	-1.935
11722322_at	4716	NFASC	18	Unique	-0.14	5.326	0.24	0.829	4.223	3.577	0.57	1.112	-0.343	-0.241
11731529_at	4717	SLC17A6	18	Unique	0.642	5.948	-0.124	0.385	4.827	4.242	0.345	0.71	-0.087	-0.046
11721936_a_at	4718	CHN2	18	Unique	0.347	4.958	0.679	1.241	3.464	2.949	0.21	1.104	-0.536	-0.006
11733723_a_at	4719	HMP19	18	Unique	-0.122	4.424	-0.505	0.325	4.435	3.958	1.012	1.937	-0.01	-0.266
11733554_a_at	4720	PAX7	18	Unique	0.201	0.124	0.12	0.306	4.803	4.351	0.755	0.752	0.028	0
11726800_at	4721	EBF1	18	Unique	1.231	6.498	1.611	2.785	4.343	3.896	2.987	3.898	0.395	0.316
11740091_s_at	4722	TMEM163	18	Unique	1.973	5.478	0.998	0.748	3.553	3.142	0.9	1.064	2.004	1.785
11722333_a_at	4723	FNDC5	18	Unique	2.417	6.048	1.183	1.526	1.687	1.308	1.107	0.649	0.747	0.665
11728922_at	4724	ONECUT2	18	Unique	1.108	6.497	0.798	1.013	2.246	1.887	0.403	0.488	0.546	0.74
11743826_at	4725	FSTL5	18	Unique	-0.197	5.235	-0.135	0.77	3.741	3.403	0.983	0.279	-0.535	-0.411
11733305_a_at	4726	TFAP2B	18	Unique	1.276	5.803	-0.127	2.46	0.692	0.447	0.808	1.931	0.22	0.442
11758493_s_at	4727	NHLH1	18	Unique	0.477	4.535	0.338	0.455	1.982	1.822	0.411	0.763	-0.216	0.281
11726789_a_at	4728	CNTN2	18	Unique	0.877	6.515	1.05	2.487	4.742	4.604	1.708	2.312	-0.006	0.133
11736682_at	4729	LHX1	18	Unique	1.242	7.055	0.559	2.386	3.239	3.177	1.142	1.161	0.159	0.012
11722232_a_at	4730	PCP4	18	Unique	0.62	0.281	-0.184	-0.131	7.115	7.185	1.184	1.283	2.041	2.746
11731002_a_at	4731	RSPO2	18	Unique	0.456	0.111	1.283	0.465	6.34	6.436	0.197	0.169	-0.353	-0.563
11736674_a_at	4732	KLHL35	18	Unique	1.127	5.382	0.411	0.46	0.728	0.88	0.529	0.497	0.411	0.46
11736412_s_at	4733	NHLH2	18	Unique	2.333	5.859	-0.769	0.516	1.345	1.533	-1.084	-1.052	-1.17	-1.097
11729375_s_at	4734	TAC3	18	Unique	0.563	5.03	0.375	-0.108	-1.047	-0.857	-0.949	-1.04	-0.83	-0.997
11725865_a_at	4735	EBF3	18	Unique	0.778	6.274	0.092	0.997	2.972	3.38	0.595	0.815	0.047	-0.092
11754285_a_at	4736	TTR	18	Unique	-0.056	0	-0.145	-0.287	5.712	6.45	0.017	0.008	0	-0.318
11738993_at	4737	SYT4	18	Unique	-0.29	4.677	-0.509	0.037	2.427	3.697	-0.252	0.968	-0.649	-0.777
11731121_s_at	4738	VASH2	19	Early Down	-4.534	-2.598	-4.085	-3.588	-2.807	-3.546	-4.698	-4.391	-5.193	-4.328

ID	*Probe*	*symbols*	*cluster*	*Pattern*	*NE*	*NE*	*ERG*	*ERG*	*MRG*	*MRG*	*LRG*	*LRG*	*LNP*	*LNP*
11716887_a_at	4739	EPCAM	19	Early Down	-3.938	-2.793	-6.82	-5.389	-6.307	-6.91	-7.763	-7.599	-7.574	-7.715
11716016_at	4740	MAL2	19	Early Down	-2.955	-2.323	-3.953	-2.458	-3.803	-4.259	-6.458	-5.231	-6.292	-6.277
11726385_a_at	4741	MT1X	19	Early Down	-2.059	-1.831	-2.99	-3.596	-4.123	-4.469	-4.136	-4.293	-3.97	-3.97
11743112_at	4742	S100A10	19	Early Down	-5.741	-3.4	-4.782	-1.594	-5.256	-5.557	-5.527	-4.304	-5.345	-5.562
11739396_a_at	4743	UGT8	19	Early Down	-3.373	-3.712	-3.584	-3.563	-3.645	-3.936	-3.694	-3.524	-4.345	-4.275
11743108_a_at	4744	ACTA1	19	Early Down	-2.686	-2.702	-1.882	-2.414	-2.793	-3.073	-3.513	-2.499	-3.563	-3.288
11745993_a_at	4745	MAP7	19	Early Down	-2.987	-2.308	-3.901	-4.398	-3.655	-3.842	-4.029	-3.643	-4.929	-4.956
11725039_at	4746	FGF2	19	Early Down	-3.295	-2.954	-3.187	-2.446	-4.319	-4.405	-3.858	-4.299	-4.589	-4.118
11725518_at	4747	FRAT2	19	Early Down	-3.856	-3.312	-3.856	-3.55	-4.375	-4.421	-4.443	-4.329	-4.654	-4.51
11754862_a_at	4748	RARRES2	19	Early Down	-4.742	-4.575	-5.175	-4.875	-4.432	-4.463	-4.877	-4.895	-4.268	-4.816
11721420_a_at	4749	RAP1GAP2	19	Early Down	-2.183	-0.199	-2.876	-2.876	-3.126	-3.103	-4.676	-4.074	-4.642	-4.634
11726904_a_at	4750	COBL	19	Early Down	-2.352	-1.707	-3.642	-3.297	-2.968	-2.903	-2.766	-3.356	-1.731	-2.352
11756509_a_at	4751	BAMBI	19	Early Down	-3.469	-3.392	-3.075	-1.743	-0.779	-0.677	-4.443	-4.484	-4.547	-4.527
11725584_at	4752	HAS3	19	Early Down	-4.481	-4.048	-3.865	-4.113	-3.61	-3.489	-4.471	-4.219	-4.889	-4.804
11729329_at	4753	EPB41L4B	19	Early Down	-3.175	-3.36	-3.222	-3.398	-4.117	-3.958	-4.057	-4.007	-3.597	-4.038
11719230_at	4754	PCDH1	19	Early Down	-2.382	-2.657	-2.683	-2.72	-3.165	-2.998	-2.743	-2.652	-2.775	-2.826
11736656_a_at	4755	GPR176	19	Early Down	-3.456	-2.891	-3.989	-3.989	-4.489	-4.279	-4.636	-4.335	-4.455	-4.457
11724735_a_at	4756	PDPN	19	Early Down	-3.487	-3.451	-2.643	-2.231	-3.932	-3.679	-3.305	-2.827	-4.211	-3.998
11715584_a_at	4757	CAV1	19	Early Down	-3.383	-3.901	-3.177	-2.181	-3.972	-3.565	-2.953	-2.882	-3.229	-3.622
11717385_a_at	4758	MT1G	19	Early Down	-4.218	-3.151	-4.576	-4.218	-5.105	-4.466	-4.766	-4.446	-5.278	-5.158
11721990_at	4759	LECT1	20	Early Down	-5.695	-6.39	-6.411	-6.494	-6.378	-7.094	-5.812	-6.248	-3.041	-4.169
11732657_a_at	4760	CYP26A1	20	Early Down	-5.62	-4.042	-6.013	-4.946	-5.502	-5.722	-5.975	-5.685	-5.84	-6.029
11740007_at	4761	POLR3G	20	Early Down	-5.416	-5.381	-5.239	-5.166	-5.554	-5.751	-5.409	-5.388	-5.208	-5.308

ID	*Probe*	*symbols*	*cluster*	*Pattern*	*NE*	*NE*	*ERG*	*ERG*	*MRG*	*MRG*	*LRG*	*LRG*	*LNP*	*LNP*
11734156_at	4762	PRDM14	20	Early Down	-5.879	-5.831	-5.643	-5.843	-5.861	-6.007	-5.879	-5.861	-6.041	-5.861
11722300_a_at	4763	ETS1	20	Early Down	-6.116	-4.36	-5.019	-1.837	-6.079	-6.154	-6.303	-5.019	-6.586	-6.518
11725749_a_at	4764	GAL	20	Early Down	-7.126	-6.817	-6.276	-4.136	-7.291	-7.353	-7.353	-7.477	-7.385	-7.41
11723543_a_at	4765	ESRP1	20	Early Down	-5.389	-3.058	-6.029	-6.297	-6.297	-6.335	-6.164	-6.283	-6.226	-6.453
11731707_at	4766	VRTN	20	Early Down	-5.065	-5.246	-5.274	-5.267	-5.481	-5.501	-5.563	-5.563	-5.589	-5.655
11736314_x_at	4767	POU5F1	20	Early Down	-7.76	-7.746	-7.714	-7.86	-8.126	-8.115	-7.916	-7.937	-7.885	-8.132
11755164_a_at	4768	LEFTY1	20	Early Down	-7.315	-7.162	-7.479	-7.225	-7.315	-7.26	-7.151	-7.465	-7.303	-7.353
11723363_at	4769	L1TD1	20	Early Down	-4.487	-4.237	-7.538	-6.977	-8.113	-8.026	-8.091	-8.121	-8.143	-7.797
11723418_a_at	4770	VSNL1	20	Early Down	-3.63	-3.934	-4.848	-4.848	-4.848	-4.755	-4.958	-4.99	-5.372	-5.447
11732833_a_at	4771	C9ORF135	20	Early Down	-5.654	-5.537	-5.405	-5.537	-5.235	-5.137	-5.544	-5.712	-5.537	-5.571
11725436_a_at	4772	SCGB3A2	20	Early Down	-6.764	-6.603	-7.151	-7.204	-6.585	-6.416	-7.151	-7.225	-7.115	-7.287
11743168_at	4773	IDO1	20	Early Down	-6.563	-6.542	-6.379	-6.211	-6.613	-6.417	-6.672	-6.301	-6.522	-6.337
11716905_a_at	4774	CDH1	20	Early Down	-4.991	-1.698	-7.273	-7.039	-7.224	-6.998	-7.051	-6.998	-4.782	-5.688
11729262_a_at	4775	SERPINB9	20	Early Down	-3.659	-3.06	-4.809	-2.472	-5.65	-5.32	-5.633	-5.957	-5.587	-5.76
11761896_x_at	4776	POU5F1P3	20	Early Down	-6.976	-6.988	-7.158	-6.91	-7.424	-7.07	-7.117	-7.266	-7.13	-7.193
11732576_s_at	4777	NANOG	20	Early Down	-6.867	-6.798	-6.683	-6.931	-6.503	-6.053	-6.917	-6.999	-6.924	-7.102
11717012_s_at	4778	VAMP8	20	Early Down	-3.338	-1.816	-4.46	-4.199	-4.879	-4.298	-5.09	-5.235	-4.856	-4.796
11727146_at	4779	GPR160	20	Early Down	-4.512	-4.512	-5.051	-4.787	-3.864	-3.082	-4.89	-4.995	-4.777	-5.039
11728305_s_at	4780	GULP1	21	Mid Down	0.072	-0.078	-0.067	0.055	-1.433	-1.947	-1.799	-1.56	-3.428	-2.108
11719868_a_at	4781	MYCN	21	Mid Down	-1.198	-1.016	0.402	0.546	-2.499	-2.905	-2.815	-3.16	-3.638	-3.441
11721297_at	4782	C6ORF115	21	Mid Down	-0.339	-0.372	-0.181	-0.217	-1.498	-1.882	-2.652	-2.483	-2.803	-2.547

ID	*Probe*	*symbols*	*cluster*	*Pattern*	*NE*	*NE*	*ERG*	*ERG*	*MRG*	*MRG*	*LRG*	*LRG*	*LNP*	*LNP*
11747361_x_at	4783	HMGA1	21	Mid Down	-0.714	-0.76	-0.45	-0.46	-3.021	-3.348	-3.781	-4.104	-3.819	-3.596
11715934_s_at	4784	TAGLN	21	Mid Down	-0.242	-0.35	0.153	1.221	-1.298	-1.609	-1.81	-0.439	-2.721	-2.311
11744002_s_at	4785	MTHFD 2	21	Mid Down	-1.069	-1.403	0.414	1.042	-2.156	-2.461	-2.368	-2.378	-3.74	-3.256
11742745_a_at	4786	PLIN2	21	Mid Down	0.149	-0.538	0.073	0.174	-1.857	-2.109	-1.811	-2.04	-1.805	-1.714
11718283_a_at	4787	EIF4EB P1	21	Mid Down	-1.263	-1.393	-0.515	0.053	-2.46	-2.684	-2.407	-2.519	-2.009	-2.243
11715755_a_at	4788	ATIC	21	Mid Down	-0.667	-1.096	-0.885	-0.73	-2.619	-2.635	-2.861	-3.267	-2.716	-2.781
11758564_s_at	4789	NUDT1 5	21	Mid Down	-0.697	-0.827	-0.55	-0.873	-1.585	-1.585	-2.309	-2.581	-2.685	-2.288
11720757_a_at	4790	IVNS1A BP	21	Mid Down	0.15	-0.293	0.852	2.253	-0.889	-0.854	-1.547	-1.101	-1.774	-1.797
11723804_a_at	4791	AKAP1 2	21	Mid Down	-0.259	-0.742	0.465	1.283	-1.503	-1.205	-1.533	-1.74	-1.925	-2.078
11724470_s_at	4792	MGST1	21	Mid Down	-4.563	-3.768	-1.377	-1.018	-5.312	-4.932	-3.976	-3.815	-6.775	-7.344
11736570_a_at	4793	SLC39A 14	21	Mid Down	-0.963	-1.341	-0.673	-0.317	-3.101	-2.451	-2.924	-3.071	-3.404	-3.483
11723047_at	4794	CRABP 1	22	Mid Down	-0.087	2.986	-0.223	-0.398	-1.218	-2.035	-2.364	-2.142	-3.787	-3.971
11734417_a_at	4795	MAMD C2	22	Mid Down	1.996	1.299	1.622	2.883	-1.6	-2.241	-1.657	-1.593	-2.745	-2.164
11726113_a_at	4796	FAM46 B	22	Mid Down	-1.949	-2.117	-2.472	-2.318	-3.894	-4.526	-5.931	-6.122	-5.638	-6.044
11743129_at	4797	SORL1	22	Mid Down	-1.09	-0.938	-1.566	-1.428	-4.3	-4.849	-5.208	-5.3	-1.569	-1.486
11720489_a_at	4798	ACTC1	22	Mid Down	-2.197	-1.974	-0.014	1.573	-5.984	-6.523	-6.284	-6.046	-6.281	-6.65
11742698_at	4799	SCG5	22	Mid Down	1.048	1.291	0.008	-0.161	-2.197	-2.649	-5.041	-4.634	-4.697	-4.799
11723309_at	4800	TNFRS F10D	22	Mid Down	2.665	1.917	1.904	2.799	-1.076	-1.487	-1.608	-1.745	0.619	-0.768
11728695_at	4801	ARG2	22	Mid Down	0.29	0.092	0.502	0.983	-1.391	-1.76	-3.197	-3.793	-4.022	-3.298
11723397_a_at	4802	ZNF385	22	Mid Down	4.95	4.205	2.484	0.652	-0.099	-0.439	-1.379	-1.303	-1.595	-1.212

ID	*Probe*	*symbols*	*cluster*	*Pattern*	*NE*	*NE*	*ERG*	*ERG*	*MRG*	*MRG*	*LRG*	*LRG*	*LNP*	*LNP*
		B												
11717017_a_at	4803	OCIAD2	22	Mid Down	-0.744	-0.666	-0.83	0.177	-2.357	-2.692	-5.17	-4.864	-5.036	-4.581
11734013_a_at	4804	BMP4	22	Mid Down	-0.535	1.034	-0.655	2.321	-1.106	-1.429	-3.602	-3.383	-3.285	-3.186
11729641_a_at	4805	TPD52	22	Mid Down	-1.292	0.519	-1.999	-1.752	-5.005	-5.318	-5.081	-4.272	-5.352	-5.851
11722339_a_at	4806	DIAPH2	22	Mid Down	-1.578	-1.943	-2.725	-2.106	-5.539	-5.842	-5.78	-5.907	-5.636	-5.653
11724283_a_at	4807	ANXA3	22	Mid Down	0.478	0.466	-0.883	0.917	-1.91	-2.095	-4.924	-4.436	-4.8	-5.032
11716034_a_at	4808	BST2	22	Mid Down	3.241	2.331	0.551	0.807	-3.079	-3.24	-3.849	-3.899	-3.518	-3.77
11723882_x_at	4809	CECR1	22	Mid Down	0.686	0.925	-0.111	-0.552	-2.169	-2.245	-3.683	-3.539	-4.728	-4.296
11719533_a_at	4810	STX3	22	Mid Down	0.714	0.316	-1.055	-0.391	-3.465	-3.517	-5.267	-5.14	-5.284	-5.412
11720460_x_at	4811	F11R	22	Mid Down	-1.348	-1.13	-2.425	-1.998	-4.671	-4.713	-7.066	-7.365	-7.278	-7.779
11725430_at	4812	LIN28A	22	Mid Down	-0.444	-0.325	-0.226	0.203	-8.176	-8.202	-8.65	-8.551	-8.572	-8.538
11731500_a_at	4813	KRT19	22	Mid Down	1.113	0.777	-1.331	-1.692	-4.912	-4.89	-5.248	-5.568	-5.258	-5.146
11721090_at	4814	THBS1	22	Mid Down	-0.063	-0.414	-0.026	0.733	-3.464	-3.441	-3.809	-1.673	-3.501	-3.574
11715353_s_at	4815	COL1A2	22	Mid Down	1.993	2.135	-1.297	1.227	-4.15	-4.108	-4.08	-4.063	-4.308	-3.186
11721424_at	4816	FZD5	22	Mid Down	-1.374	-1.585	0.085	0.218	-5.413	-5.352	-4.55	-3.628	-5.848	-5.803
11734427_at	4817	TRIM71	22	Mid Down	-1.346	-3.004	-2.128	-1.978	-5.066	-5.005	-5.413	-5.142	-5.964	-5.96
11720469_a_at	4818	TSTD1	22	Mid Down	-1.715	-1.374	-2.195	-2.376	-5.176	-5.013	-6.633	-6.451	-6.376	-5.994
11732465_a_at	4819	USP44	22	Mid Down	-0.921	-1.783	-1.864	-2.108	-4.02	-3.856	-6.723	-7.071	-6.555	-7.202
11730880_a_at	4820	NEBL	22	Mid Down	3.101	2.695	1.785	2.323	-1.362	-1.162	-2.169	-2.704	-3.447	-3.723
11757418_a_at	4821	SILV	22	Mid Down	2.838	3.669	1.531	1.629	-2.311	-2.01	-3.087	-2.213	-3.095	-3.14
11743904_s_at	4822	PDLIM1	22	Mid Down	-1.221	-1.615	-2.695	-2.22	-5.141	-4.834	-4.061	-4.834	-4.713	-3.841
11752634_x_at	4823	KRT8	22	Mid Down	1.278	1.183	-0.276	1.457	-2.478	-2.168	-4.778	-4.45	-4.809	-4.27
11718064_a_at	4824	S100A4	22	Mid Down	2	1.056	0.173	1.577	-1.028	-0.6	-1.929	-1.182	-3.142	-3.1
11719333_a_at	4825	DMKN	22	Mid Down	-2.259	-2.556	-3.709	-3.441	-5.797	-5.329	-5.902	-6.083	-5.896	-5.794
11715440_a_at	4826	CTGF	23	Late Down	1.452	1.084	1.683	1.904	0.47	0.294	-1.264	-0.228	-7.416	-6.242
11727101_at	4827	IGF2BP	23	Late Down	0.357	0.579	0.781	0.677	-0.413	-0.551	-2.064	-2.021	-7.044	-5.922

ID	*Probe*	*symbols*	*cluster*	*Pattern*	*NE*	*NE*	*ERG*	*ERG*	*MRG*	*MRG*	*LRG*	*LRG*	*LNP*	*LNP*
		1												
11716195_a_at	4828	ID1	23	Late Down	-4.27	-3.496	-2.086	-1.378	-0.947	-0.913	-4.509	-5.144	-7.888	-6.871
11716666_a_at	4829	ID3	23	Late Down	-2.827	-2.28	0.318	0.528	0.525	0.653	-1.109	-0.24	-5.298	-4.064
11736127_s_at	4830	DPPA4	24	Late Down	-1.081	-1.324	-1.693	-1.426	-2.772	-3.582	-4.335	-4.634	-7.179	-5.172
11720115_a_at	4831	MST4	24	Late Down	0.015	0.057	0.239	0.66	-0.661	-1.329	-4.003	-3.945	-6.545	-6.134
11725523_at	4832	DSG2	24	Late Down	-1.004	-0.681	-1.385	-1	-1.492	-1.954	-5.03	-6.384	-6.241	-6.945
11739709_a_at	4833	PHACTR2	24	Late Down	0.359	0.143	-0.2	1.152	-0.356	-0.646	-2.937	-3.453	-3.829	-4.029
11728140_at	4834	SHISA2	24	Late Down	1.838	1.906	2.767	1.792	1.99	1.701	-4.255	-4.1	-4.436	-4.843
11742851_x_at	4835	SPINT2	24	Late Down	-0.788	-0.327	-2.384	-2.006	-2.258	-2.497	-7.079	-7.055	-7.233	-7.197
11739501_a_at	4836	DSP	24	Late Down	1.344	1.53	0.166	0.641	0.135	-0.075	-6.083	-6.431	-6.349	-6.577
11720103_a_at	4837	PAK1	24	Late Down	-1.346	-0.801	-1.368	-1.859	-1.222	-1.372	-2.807	-2.588	-6.307	-5.413
11721204_a_at	4838	ANK3	24	Late Down	0.121	0.658	-0.221	-0.26	0.341	0.2	-1.765	-1.549	-4.832	-4.032
11720292_at	4839	LIN28B	24	Late Down	0.448	0.647	-0.089	0.063	-2.615	-2.721	-6.559	-6.563	-7.396	-7.403
11725530_a_at	4840	DSC2	24	Late Down	-0.291	0.292	0.082	0.49	-1.556	-1.645	-5.037	-5.714	-5.99	-6.021
11721991_a_at	4841	SLC6A6	24	Late Down	-0.601	-0.491	-0.511	-0.888	-0.728	-0.726	-3.003	-1.655	-5.572	-5.391
11756810_a_at	4842	MFAP2	24	Late Down	0.838	0.417	-0.202	-0.401	1.201	1.204	-3.133	-2.329	-3.502	-3.938
11728065_a_at	4843	SALL4	24	Late Down	1.09	0.569	-0.078	-0.537	-3.566	-3.518	-4.857	-4.764	-5.695	-5.281
11717719_a_at	4844	ST6GAL1	24	Late Down	-1.139	-0.749	-1.027	-0.872	-1.129	-1.072	-2.867	-1.636	-6.13	-6.041
11716281_a_at	4845	COL18A1	24	Late Down	-1.112	-0.65	-0.559	-0.378	0.032	0.176	-5.152	-4.911	-4.828	-5.782
11720062_s_at	4846	IER3	24	Late Down	0.607	0.092	1.011	2.462	2.442	2.707	-3.249	-1.978	-3.529	-3.727
11743968_a_at	4847	CYBA	24	Late Down	-0.894	-1.008	-1.115	-1.15	-1.419	-1.15	-4.275	-4.798	-5.151	-5.744
11754941_x_at	4848	KRT18	24	Late Down	0.785	0.706	-0.951	1.029	-3.145	-2.87	-6.943	-6.718	-6.915	-6.446
11741103_a_at	4849	ANXA11	24	Late Down	0.178	-0.238	-1.253	-0.913	-2.638	-2.356	-4.107	-3.763	-5.72	-4.686
11732727_a_at	4850	SLITRK4	24	Late Down	2.283	2.321	1.433	1.639	0.932	1.271	-0.748	-1.074	-3.304	-3.511

ID	*Probe*	*symbols*	*cluster*	*Pattern*	*NE*	*NE*	*ERG*	*ERG*	*MRG*	*MRG*	*LRG*	*LRG*	*LNP*	*LNP*
11736522_a_at	4851	TNNT1	24	Late Down	1.04	1.086	-1.571	-2.054	-1.216	-0.803	-4.805	-3.908	-4.925	-5.139
11756670_x_at	4852	TUBA1C	25	Late Down	-0.729	-0.959	-0.473	0.327	-2.399	-2.793	-4.384	-4.932	-3.103	-4.031
11730461_at	4853	OTX2	25	Late Down	0.93	0.619	0.944	0.904	0.853	0.664	-3.644	-3.351	-4.367	-4.373
11716745_s_at	4854	FAM60A	25	Late Down	-0.007	-0.297	-0.24	-0.266	-1.123	-1.246	-2.939	-2.895	-3.931	-3.936
11732032_a_at	4855	HMGA2	25	Late Down	0.575	0.308	0.931	0.927	0.977	0.89	-3.874	-3.162	-2.791	-3.591
11747282_x_at	4856	ANXA2P2	25	Late Down	-1.191	-0.554	-1.783	0.361	-1.318	-1.262	-4.831	-2.628	-2.41	-5.124
11720606_a_at	4857	SFRP2	25	Late Down	1.63	0.942	1.892	1.97	-0.711	-0.597	-2.093	-1.983	-3.714	-3.771
11738551_a_at	4858	ANXA2	25	Late Down	-1.08	-0.439	-1.712	0.317	-1.334	-1.006	-4.769	-2.501	-2.319	-4.971
11715670_a_at	4859	IFITM1	25	Late Down	-0.696	-1.287	-2.871	-3.351	-2.521	-2.01	-6.473	-4.119	-7.043	-6.454
11726396_x_at	4860	KAL1	26	Late Down	1.134	0.638	-0.315	-1.592	0.659	0.167	1.536	0.361	-4.428	-1.162
11754354_s_at	4861	PROM1	26	Late Down	-0.241	-0.791	-0.534	-0.828	-0.605	-1.027	0.107	0.948	-3.525	-2.1
11729118_a_at	4862	CDKN1C	26	Late Down	1.39	4.003	0.616	1.378	2.438	2.042	0.908	1.395	-2.284	-1.354
11717672_s_at	4863	ENPP2	26	Late Down	4.363	3.631	3.51	2.146	1.234	0.911	3.09	1.993	-1.359	-0.86
11725443_at	4864	TCEAL7	26	Late Down	1.617	2.68	2.083	1.857	2.099	1.917	1.752	1.64	-2.536	-1.806
11721837_a_at	4865	MYLK	26	Late Down	1.169	0.845	1.807	1.53	0.921	0.845	3.264	4.208	-1.262	-0.681
11725928_s_at	4866	NEFL	26	Late Down	1.005	4.307	0.649	1.411	2.431	2.398	3.891	4.815	-2.944	-2.681
11727563_at	4867	PCDHB5	26	Late Down	1.541	1.493	0.932	0.632	0.899	0.945	0.747	0.53	-4.278	-3.182
11719964_a_at	4868	UBE2E2	26	Late Down	0.12	0.202	0.116	-0.009	-0.18	-0.112	-0.676	-0.703	-3.784	-2.359
11718848_a_at	4869	BEX2	26	Late Down	-0.247	0.781	-0.043	0.172	0.779	0.885	-1.224	-1.455	-4.801	-3.648

Table 3: Global gene expression cluster analysis for stage specifically expressed genes. Selected trends of global gene expression clustering comprising 496 genes. Up = Upregulated; Down = Downregulated.

Table 4

*Sco*ring of the transcriptional activity of the factors and regulators affecting the epigenetic landscape during neural development

NE gene	*Polyn. SEQ ID NO:*	*Polyp. SEQ ID NO:*	*NE vs hESC*	*ERG gene*	*Polyn. SEQ ID NO:*	*Polyp. SEQ ID NO:*	*ERG vs NE*	*MRG gene*	*Polyn. SEQ ID NO:*	*Polyp. SEQ ID NO:*	*MRG vs ERG*
RFX4	3897-3898	1946-1949	10.23	**NFIA**	3626-3629	1516-1520	8.58	**RFX4**	3897-3898	1946-1949	22.43
PAX6	3733-3735	1681-1683	8.12	**MZF1**	3569-3570	1432-1434	7.02	**NFIA**	3626-3629	1516-1520	11.56
TCF4	4148-4149	2327-2328	5.71	**SMAD4**	4025	2141-2142	5.91	**NFATC1**	3620-3624	1505-1514	9.43
SOX9	4065	2203	5.57	**CTF1**	2962-2963	511-512	5.74	**CTF1**	2962-2963	511-512	8.07
LEF1	3416-3419	1199-1203	5.38	**SMAD3**	4021-4024	2136-2140	5.47	**NEUROD1**	3612	1495	6.47
POU3F2	3805	1794-1795	5.25	**RFX4**	3897-3898	1946-1949	5.38	**RFX1**	3896	1945	6.40
SOX8	4064	2201-2202	5.20	**MAF**	3465-3466	1270-1272	5.26	**TGIF1**	4162-4169	2346-2360	6.09
SOX21	4055	2186-2187	4.92	**TCF12**	4141-4145	2316-2324	5.17	**ETS2**	3094	709-710	5.63
TEAD1	4151	2331-2332	4.82	**NFYC**	3636-3640	1528-1533	5.11	**TFAP4**	4158	2340	5.50
NFATC1	3620-3624	1505-1514	4.80	**ZNF263**	4346	2645-2646	5.01	**TCF12**	4141-4145	2316-2324	5.48
SOX5	4057-4059	2189-2194	4.67	**MECP2**	3509-3510	1342-1344	4.69	**TFAP4**	4158	2340	5.43
TGIF1	4162-4169	2346-2360	4.59	**ZFP42**	4329	2617-2618	4.57	**SMAD3**	4021-4024	2136-2140	5.36
MEIS1	3518	1355	4.44	**ZIC1**	4332	2622	4.53	**NFIA**	3626-3629	1516-1520	5.23
TCF4	4148-4149	2327-2328	4.41	**YY1**	4313	2589	4.52	**SOX9**	4065	2203	5.22
MEIS2	3519-3525	1356-1368	4.31	**ETS2**	3094	709-710	4.49	**RFX1**	3896	1945	4.98
OTX2	3708-3709	1639-1642	4.24	**NR2C2**	3660	1564	4.45	**TCF4**	4148-4149	2327-2328	4.81
TEF	4153-4154	2334-2335	4.22	**SREBF2**	4079	2221	4.42	**MZF1**	3569-3570	1432-1434	4.38
ZBTB16	4315-4316	2592-2594	4.19	**SREBF1**	4077-4078	2219-2220	4.12	**STAT1**	4103-4104	2254-2257	4.37
SOX9	4065	2203	4.13	**MEIS2**	3519-3525	1356-1368	4.07	**MECP2**	3509-3510	1342-1344	4.34
MSX1	3546	1397	4.08	**NR4A1**	3671-3672	1577-1579	4.04	**MEIS2**	3519-3525	1356-1368	4.28
RFX1	3896	1945	3.98	**REST**	3891	1936-1937	3.95	**ZBTB33**	4317	2595-2596	4.22

NE gene	*Polyn. SEQ ID NO:*	*Polyp. SEQ ID NO:*	*NE vs hESC*	*ERG gene*	*Polyn. SEQ ID NO:*	*Polyp. SEQ ID NO:*	*ERG vs NE*	*MRG gene*	*Polyn. SEQ ID NO:*	*Polyp. SEQ ID NO:*	*MRG vs ERG*
SOX8	4064	2201-2202	3.96	**SF1**	3977-3980	2071-2077	3.84	**NFYA**	3634-3635	1525-1527	3.97
NR4A2	3673	1580-1581	3.95	**ZBTB7A**	4320	2601-2602	3.83	**ELF1**	3064-3065	666-668	3.89
MEIS2	3519-3525	1356-1368	3.92	**STAT3**	4105-4107	2258-2262	3.78	**MYBL2**	3556	1410	3.81
SOX15	4054	2184-2185	3.87	**E2F1**	3042	637-638	3.75	**LEF1**	3416-3419	1199-1203	3.77
STAT5B	4108	2263-2264	3.81	**NR1H2**	3659	1562-1563	3.74	**NFYC**	3636-3640	1528-1533	3.76
SOX8	4064	2201-2202	3.70	**NFKB1**	3630-3631	1521-1522	3.71	**MAFF**	3467-3470	1273-1277	3.28
SATB1	3955-3956	2039-2040	3.66	**NR2F6**	3663	1568-1569	3.66	**ZNF263**	4346	2645-2646	3.27
RBPJ	3882-3885	1920-1927	3.63	**GLIS3**	3225-3226	916-918	3.61	**YY1**	4313	2589	3.24
FOXK1	3160	809	3.57	**MAZ**	3491-3492	1314-1315	3.52	**POU3F3**	3806	1796-1797	3.20
MYBL2	3556	1410	3.53	**STAT1**	4103-4104	2254-2257	3.51	**TGIF1**	4162-4169	2346-2360	3.19
DMRT3	3020	597-598	3.51	**TGIF1**	4162-4169	2346-2360	3.49	**STAT3**	4105-4107	2258-2262	3.12
NFIA	3626-3629	1516-1520	3.48	**SOX9**	4065	2203	3.45	**SMAD4**	4025	2141-2142	3.07
CUX1	2970-2972	520-524	3.46	**HES1**	3278	994-995	3.32	**NR6A1**	3674-3675	1582-1584	3.04
TFAP4	4158	2340	3.45	**THRA**	4171-4172	2362-2364	3.29	**TGIF1**	4162-4169	2346-2360	3.03
MSX1	3546	1397	3.45	**GLIS3**	3225-3226	916-918	3.25	**MEIS1**	3518	1355	3.02
CDC5L	2876	380	3.43	**MEIS1**	3518	1355	3.20	**ZNF628**	4357	2663-2664	3.02
RFX1	3896	1945	3.41	**ESRRA**	3091	705	3.13	**ZFP42**	4329	2617-2618	2.98
FOXJ2	3158	805-806	3.40	**ZBTB7A**	4320	2601-2602	3.13	**FOXK1**	3160	809	2.97
POU6F1	3808	1799-1801	3.38	**NFYA**	3634-3635	1525-1527	3.11	**PRDM4**	3829	1835-1836	2.91
TEF	4153-4154	2334-2335	3.36	**MECP2**	3509-3510	1342-1344	3.08	**STAT1**	4103-4104	2254-2257	2.68
RBPJ	3882-3885	1920-1927	3.35	**PKNOX 2**	3783	1761-1762	3.07	**MAF**	3465-3466	1270-1272	2.62
PKNOX 2	3783	1761-1762	3.30	**EP300**	3080	689	3.07	**SCRT2**	3966	2054-2055	2.61
BCL6	2792-2794	243-245	3.29	**TFAP2B**	4157	2339	3.06	**CREB1**	2946-2947	487-489	2.60
PRRX1	3833-3834	1841-1844	3.25	**NFIA**	3626-3629	1516-1520	3.02	**GZF1**	3269	979-980	2.57
STAT1	4103-4104	2254-2257	3.24	**NR2F1**	3662	1566-1567	2.98	**CREB1**	2946-2947	487-489	2.56
POU3F1	3804	1793	3.14	**MZF1**	3569-3570	1432-1434	2.91	**VAX1**	4279-4280	2536-2538	2.55

NE gene	*Polyn. SEQ ID NO:*	*Polyp. SEQ ID NO:*	*NE vs hESC*	*ERG gene*	*Polyn. SEQ ID NO:*	*Polyp. SEQ ID NO:*	*ERG vs NE*	*MRG gene*	*Polyn. SEQ ID NO:*	*Polyp. SEQ ID NO:*	*MRG vs ERG*
FOXB1	3156	801-802	3.12	**ESRRA**	3091	705	2.84	**MECP2**	3509-3510	1342-1344	2.55
CTNNB1	2964-2966	513-515	3.01	**TFCP2**	4159	2341-2342	2.83	**NHLH1**	3641	1534-1535	2.53
PBX1	3737	1686	3.01	**NR2F1**	3662	1566-1567	2.80	**ETV1**	3095-3101	711-718	2.48
ZNF143	4339	2632	2.98	**ESRRA**	3091	705	2.76	**SOX9**	4065	2203	2.47
NFATC1	3620-3624	1505-1514	2.97	**TERF1**	4155-4156	2336-2338	2.74	**PEBP1**	3756-3758	1721	2.44
SOX21	4055	2186-2187	2.79	**KLF3**	3399	1172-1173	2.73	**SMAD4**	4025	2141-2142	2.41
TCF7L1	4150	2329-2330	2.69	**XBP1**	4306-4307	2577-2579	2.70	**XBP1**	4306-4307	2577-2579	2.35
ARX	2753	180-181	2.65	**RORA**	3923-3926	1988-1994	2.67	**USF1**	4268-4269	2517-2520	2.33
SOX15	4054	2184-2185	2.65	**PBX1**	3737	1686	2.65	**POU3F2**	3805	1794-1795	2.33
RXRA	3945	2024	2.59	**MYC**	3557	1411	2.63	**CREB3**	2948	490-491	2.31
TFAP4	4158	2340	2.55	**SNAI2**	4033	2154	2.56	**EP300**	3080	689	2.30
CUX1	2970-2972	520-524	2.55	**TEAD2**	4152	2333	2.52	**PBX1**	3737	1686	2.17
OTX2	3708-3709	1639-1642	2.50	**CENPB**	2898	407	2.52	**STAT3**	4105-4107	2258-2262	2.15
NR2E1	3661	1565	2.49	**PEBP1**	3756-3758	1721	2.51	**TFCP2**	4159	2341-2342	2.15
CUX1	2970-2972	520-524	2.46	**HINFP**	3287-3288	1008-1011	2.51	**POU2F2**	3803	1792	2.13
ZNF232	4343	2639-2640	2.43	**SREBF1**	4077-4078	2219-2220	2.49	**IRF3**	3360	1113	2.11
NR2F1	3662	1566-1567	2.39	**YY1**	4313	2589	2.49	**FOXB1**	3156	801-802	2.10
SOX4	4056	2188	2.27	**E2F1**	3042	637-638	2.49	**MSX1**	3546	1397	2.09
MEIS1	3518	1355	2.25	**HSF2**	3332-3333	1073-1074	2.48	**POU3F3**	3806	1796-1797	2.06
PBX1	3737	1686	2.22	**CNOT3**	2927	455-456	2.46	**ELK1**	3070-3071	675-677	2.05
CUX1	2970-2972	520-524	2.20	**MEIS2**	3519-3525	1356-1368	2.44	**DBP**	2990	549	2.04
NEUROD1	3612	1495	2.19	**MEF2A**	3512-3516	1347-1352	2.44	**CUX1**	2970-2972	520-524	2.04
MSX1	3546	1397	2.19	**RXRA**	3945	2024	2.44	**MEF2A**	3512-3516	1347-1352	1.98
ZNF652	4358-4359	2665-2667	2.19	**CREB3L1**	2949	492-493	2.41	**POU6F2**	3809-3810	1802-1804	1.98
MEF2A	3512-3516	1347-1352	2.10	**MYBL2**	3556	1410	2.38	**ARNTL**	2748-2750	174-176	1.97

NE gene	*Polyn. SEQ ID NO:*	*Polyp. SEQ ID NO:*	*NE vs hESC*	*ERG gene*	*Polyn. SEQ ID NO:*	*Polyp. SEQ ID NO:*	*ERG vs NE*	*MRG gene*	*Polyn. SEQ ID NO:*	*Polyp. SEQ ID NO:*	*MRG vs ERG*
OLIG1	3701	1628-1629	2.09	**ZNF524**	4356	2661-2662	2.37	**ZSCAN16**	4365	2676-2677	1.96
POU6F2	3809-3810	1802-1804	2.07	**TFAP4**	4158	2340	2.36	**MEF2A**	3512-3516	1347-1352	1.93
IKZF2	3350-3351	1097-1099	2.07	**RFX1**	3896	1945	2.35	**ETV4**	3103-3104	721-722	1.92
MECOM	3503-3508	1334-1341	2.05	**NFYA**	3634-3635	1525-1527	2.34	**OLIG1**	3701	1628-1629	1.87
STAT1	4103-4104	2254-2257	2.00	**ZBTB33**	4317	2595-2596	2.34	**HOMEZ**	3324	1059-1060	1.85
ESRRA	3091	705	1.99	**RREB1**	3931-3934	2002-2005	2.33	**DLX1**	3018-3019	594-596	1.84
IRF7	3361-3363	1114-1116	1.99	**NR6A1**	3674-3675	1582-1584	2.29	**PRRX1**	3833-3834	1841-1844	1.83
STAT1	4103-4104	2254-2257	1.96	**HES5**	3281	999	2.28	**MSX1**	3546	1397	1.81
MYBL2	3556	1410	1.96	**TFAP2B**	4157	2339	2.28	**MYC**	3557	1411	1.81
BCL6	2792-2794	243-245	1.94	**HIF1A**	3285-3286	1005-1007	2.25	**FOXO4**	3164-3165	816-818	1.81
ELK1	3070-3071	675-677	1.94	**INSM1**	3357	1110	2.22	**MEF2A**	3512-3516	1347-1352	1.80
ATF2	2758	190	1.83	**ZNF524**	4356	2661-2662	2.21	**MZF1**	3569-3570	1432-1434	1.80
SMAD3	4021-4024	2136-2140	1.81	**ELF1**	3064-3065	666-668	2.20	**ATF2**	2758	190	1.79
ATF4	2762-2763	194-196	1.78	**SMAD4**	4025	2141-2142	2.18	**GMEB2**	3231	927-928	1.79
DLX1	3018-3019	594-596	1.75	**STAT3**	4105-4107	2258-2262	2.17	**NFYA**	3634-3635	1525-1527	1.79
MEF2A	3512-3516	1347-1352	1.74	**TFAP2B**	4157	2339	2.14	**ESRRA**	3091	705	1.79
DBP	2990	549	1.72	**USF1**	4268-4269	2517-2520	2.13	**SOX9**	4065	2203	1.76
MAF	3465-3466	1270-1272	1.70	**GLI3**	3224	915	2.13	**PBX1**	3737	1686	1.75
MEF2A	3512-3516	1347-1352	1.65	**GLI2**	3223	913-914	2.12	**POU3F2**	3805	1794-1795	1.73
TEAD2	4152	2333	1.63	**ETV2**	3102	719-720	2.04	**MECOM**	3503-3508	1334-1341	1.67
SMAD3	4021-4024	2136-2140	1.56	**E2F1**	3042	637-638	2.03	**SMAD3**	4021-4024	2136-2140	1.63
SOX15	4054	2184-2185	1.56	**CREB3**	2948	490-491	2.02	**MAZ**	3491-3492	1314-1315	1.60
POU6F1	3808	1799-1801	1.53	**NFKB1**	3630-3631	1521-1522	2.00	**ELK1**	3070-3071	675-677	1.59
BARHL2	2774	211-212	1.53	**MYC**	3557	1411	1.99	**BCL6**	2792-2794	243-245	1.58

NE gene	*Polyn. SEQ ID NO:*	*Polyp. SEQ ID NO:*	*NE vs hESC*	*ERG gene*	*Polyn. SEQ ID NO:*	*Polyp. SEQ ID NO:*	*ERG vs NE*	*MRG gene*	*Polyn. SEQ ID NO:*	*Polyp. SEQ ID NO:*	*MRG vs ERG*
FOXG1	3157	803-804	1.46	**SRF**	4080	2222	1.98	**ELF3**	3068-3069	673-674	1.57
LHX9	3432-3433	1223-1225	1.42	**TFCP2**	4159	2341-2342	1.97	**PAX3**	3725-3732	1667-1680	1.54
MECOM	3503-3508	1334-1341	1.38	**ATF3**	2759-2761	191-193	1.97	**ELF1**	3064-3065	666-668	1.54
ARNTL	2748-2750	174-176	1.35	**ELK1**	3070-3071	675-677	1.96	**SF1**	3977-3980	2071-2077	1.53
MYC	3557	1411	1.34	**TP53**	4214-4220	2430-2436	1.92	**BCL6**	2792-2794	243-245	1.50
ZNF75A	4362	2671-2672	1.34	**E2F4**	3044	640	1.92	**EMX2**	3073-3074	680-681	1.50
NFIA	3626-3629	1516-1520	1.33	**ELF2**	3066-3067	669-672	1.91	**STAT1**	4103-4104	2254-2257	1.45
VAX1	4279-4280	2536-2538	1.33	**DEAF1**	3010	580-581	1.91	**E2F1**	3042	637-638	1.44
GBX2	3209	888	1.32	**RXRA**	3945	2024	1.89	**ELF2**	3066-3067	669-672	1.43
HOMEZ	3324	1059-1060	1.29	**ZNF423**	4352	2654-2655	1.87	**CLOCK**	2922	445-446	1.38
FOXO4	3164-3165	816-818	1.29	**FOXG1**	3157	803-804	1.85	**ELK1**	3070-3071	675-677	1.37
FOXO4	3164-3165	816-818	1.28	**ZNF628**	4357	2663-2664	1.85	**STAT3**	4105-4107	2258-2262	1.36
FOXB1	3156	801-802	1.27	**NEUROD1**	3612	1495	1.84	**ATF5**	2764	197-198	1.34
ZSCAN16	4365	2676-2677	1.25	**ARNTL**	2748-2750	174-176	1.82	**THRA**	4171-4172	2362-2364	1.34
ELK1	3070-3071	675-677	1.24	**STAT1**	4103-4104	2254-2257	1.81	**SOX9**	4065	2203	1.32
ATF2	2758	190	1.24	**PAX3**	3725-3732	1667-1680	1.81	**SRF**	4080	2222	1.31
SOX9	4065	2203	1.23	**NR6A1**	3674-3675	1582-1584	1.80	**TCF4**	4148-4149	2327-2328	1.29
CREB1	2946-2947	487-489	1.22	**ESRRA**	3091	705	1.79	**ATF3**	2759-2761	191-193	1.29
USF1	4268-4269	2517-2520	1.19	**TCF4**	4148-4149	2327-2328	1.76	**CTNNB1**	2964-2966	513-515	1.28
ESRRA	3091	705	1.16	**YY1**	4313	2589	1.76	**USF1**	4268-4269	2517-2520	1.25
ZNF282	4347	2647	1.12	**PBX1**	3737	1686	1.75	**FOXJ3**	3159	807-808	1.25
NEUROG2	3613	1496-1497	1.10	**VAX1**	4279-4280	2536-2538	1.73	**USF1**	4268-4269	2517-2520	1.24
NFYA	3634-3635	1525-1527	1.07	**NRF1**	3679-3680	1589-1591	1.70	**ZNF282**	4347	2647	1.23

NE gene	*Polyn. SEQ ID NO:*	*Polyp. SEQ ID NO:*	*NE vs hESC*	*ERG gene*	*Polyn. SEQ ID NO:*	*Polyp. SEQ ID NO:*	*ERG vs NE*	*MRG gene*	*Polyn. SEQ ID NO:*	*Polyp. SEQ ID NO:*	*MRG vs ERG*
NR4A1	3671-3672	1577-1579	1.07	**ETV1**	3095-3101	711-718	1.69	**NEUROG2**	3613	1496-1497	1.23
CTF1	2962-2963	511-512	1.05	**FOXB1**	3156	801-802	1.68	**ESRRA**	3091	705	1.22
ELK1	3070-3071	675-677	1.05	**ZNF85**	4364	2675	1.66	**REST**	3891	1936-1937	1.21
POU3F2	3805	1794-1795	1.04	**CREB1**	2946-2947	487-489	1.66	**E2F3**	3043	639	1.20
ELK1	3070-3071	675-677	1.03	**PRDM4**	3829	1835-1836	1.66	**ZBTB7A**	4320	2601-2602	1.19
HSF1	3331	1071-1072	1.03	**MYBL2**	3556	1410	1.64	**MYBL2**	3556	1410	1.18
E2F3	3043	639	1.01	**NHLH1**	3641	1534-1535	1.64	**HSF2**	3332-3333	1073-1074	1.18
CUX1	2970-2972	520-524	1.00	**CREB1**	2946-2947	487-489	1.63	**MAX**	3486-3490	1305-1313	1.18
CREB1	2946-2947	487-489	0.99	**KLF13**	3398	1170-1171	1.63	**ZNF143**	4339	2632	1.16
ELF2	3066-3067	669-672	0.98	**TRIM26**	4234	2458	1.61	**MYBL2**	3556	1410	1.14
MYBL2	3556	1410	0.97	**ZNF148**	4340	2633-2634	1.61	**SRF**	4080	2222	1.14
HMGA2	3303-3304	1028-1029	0.97	**E4F1**	3046	643	1.59	**FOXO4**	3164-3165	816-818	1.13
SRF	4080	2222	0.95	**USF1**	4268-4269	2517-2520	1.58	**NR4A2**	3673	1580-1581	1.10
ZNF410	4351	2652-2653	0.95	**JDP2**	3372-3375	1129-1133	1.57	**CUX1**	2970-2972	520-524	1.09
JDP2	3372-3375	1129-1133	0.95	**OLIG1**	3701	1628-1629	1.56	**E2F4**	3044	640	1.05
NR2F1	3662	1566-1567	0.92	**GZF1**	3269	979-980	1.56	**MSX1**	3546	1397	1.01
PAX3	3725-3732	1667-1680	0.89	**ATF2**	2758	190	1.54	**EOMES**	3079	687-688	1.01
NRF1	3679-3680	1589-1591	0.86	**CREB1**	2946-2947	487-489	1.52	**MAF**	3465-3466	1270-1272	0.99
SMAD4	4025	2141-2142	0.85	**FOXO4**	3164-3165	816-818	1.52	**MNT**	3540	1387-1388	0.98
ZNF85	4364	2675	0.85	**SMAD3**	4021-4024	2136-2140	1.48	**POU6F1**	3808	1799-1801	0.97
ZNF628	4357	2663-2664	0.82	**CLOCK**	2922	445-446	1.48	**NFATC1**	3620-3624	1505-1514	0.97
NFATC1	3620-3624	1505-1514	0.82	**ZNF282**	4347	2647	1.47	**STAT3**	4105-4107	2258-2262	0.97
CREB1	2946-2947	487-489	0.79	**ZSCAN16**	4365	2676-2677	1.46	**CREB3**	2948	490-491	0.95
E4F1	3046	643	0.77	**CUX1**	2970-2972	520-524	1.45	**SOX9**	4065	2203	0.93
NR2F6	3663	1568-1569	0.77	**NFKB1**	3630-3631	1521-1522	1.45	**ZNF85**	4364	2675	0.90

NE gene	*Polyn. SEQ ID NO:*	*Polyp. SEQ ID NO:*	*NE vs hESC*	*ERG gene*	*Polyn. SEQ ID NO:*	*Polyp. SEQ ID NO:*	*ERG vs NE*	*MRG gene*	*Polyn. SEQ ID NO:*	*Polyp. SEQ ID NO:*	*MRG vs ERG*
NHLH1	3641	1534-1535	0.76	**RFX1**	3896	1945	1.43	**CREB3L1**	2949	492-493	0.85
IRF2	3359	1112	0.75	**ETV5**	3105	723	1.43	**TCF3**	4146-4147	2325-2326	0.81
PBX1	3737	1686	0.73	**MEF2A**	3512-3516	1347-1352	1.41	**ELK1**	3070-3071	675-677	0.80
FOXJ3	3159	807-808	0.70	**GTF2IRD1**	3264-3265	972-975	1.39	**IRF2**	3359	1112	0.77
RORA	3923-3926	1988-1994	0.70	**TFAP2B**	4157	2339	1.36	**YY1**	4313	2589	0.77
IRF7	3361-3363	1114-1116	0.66	**NHLH1**	3641	1534-1535	1.32	**SOX15**	4054	2184-2185	0.75
NR6A1	3674-3675	1582-1584	0.65	**ELF1**	3064-3065	666-668	1.30	**PAX6**	3733-3735	1681-1683	0.72
LHX2	3431	1221-1222	0.64	**PAX3**	3725-3732	1667-1680	1.28	**E4F1**	3046	643	0.69
PAX3	3725-3732	1667-1680	0.64	**MNT**	3540	1387-1388	1.19	**MEF2A**	3512-3516	1347-1352	0.68
NR2E1	3661	1565	0.63	**ZNF740**	4361	2670	1.17	**ATF2**	2758	190	0.67
POU3F1	3804	1793	0.62	**CREB1**	2946-2947	487-489	1.16	**NFE2L1**	3625	1515	0.67
ZFP42	4329	2617-2618	0.61	**ESRRA**	3091	705	1.14	**NFATC1**	3620-3624	1505-1514	0.65
E2F4	3044	640	0.57	**EOMES**	3079	687-688	1.12	**ATF2**	2758	190	0.63
ETV5	3105	723	0.56	**EGR2**	3054-3057	654-657	1.12	**YY1**	4313	2589	0.63
ELF3	3068-3069	673-674	0.54	**NR4A2**	3673	1580-1581	1.11	**SRF**	4080	2222	0.60
USF1	4268-4269	2517-2520	0.54	**BCL6**	2792-2794	243-245	1.11	**ARX**	2753	180-181	0.58
ATF6	2765	199-200	0.53	**ZNF143**	4339	2632	1.10	**ETV2**	3102	719-720	0.56
TFAP2B	4157	2339	0.52	**MAFF**	3467-3470	1273-1277	1.10	**HINFP**	3287-3288	1008-1011	0.55
CUX1	2970-2972	520-524	0.51	**CUX1**	2970-2972	520-524	1.07	**MAZ**	3491-3492	1314-1315	0.52
IRF3	3360	1113	0.50	**NR2F1**	3662	1566-1567	1.07	**NR4A1**	3671-3672	1577-1579	0.51
RXRA	3945	2024	0.48	**TFAP2B**	4157	2339	1.07	**INSM1**	3357	1110	0.51
PEBP1	3756-3758	1721	0.47	**NEUROG2**	3613	1496-1497	1.05	**ZNF652**	4358-4359	2665-2667	0.48
LHX2	3431	1221-1222	0.47	**TCF3**	4146-4147	2325-2326	1.04	**USF1**	4268-4269	2517-2520	0.48
GZF1	3269	979-980	0.46	**ATF3**	2759-2761	191-193	1.03	**NFKB1**	3630-3631	1521-1522	0.46
MEF2A	3512-3516	1347-1352	0.45	**PLAG1**	3784-3786	1763-1765	1.00	**GBX2**	3209	888	0.45

NE gene	*Polyn. SEQ ID NO:*	*Polyp. SEQ ID NO:*	*NE vs hESC*	*ERG gene*	*Polyn. SEQ ID NO:*	*Polyp. SEQ ID NO:*	*ERG vs NE*	*MRG gene*	*Polyn. SEQ ID NO:*	*Polyp. SEQ ID NO:*	*MRG vs ERG*
MEF2A	3512-3516	1347-1352	0.45	**ESRRA**	3091	705	1.00	**POU3F4**	3807	1798	0.44
IRF9	3364	1117-1118	0.44	**E2F1**	3042	637-638	0.99	**IRF9**	3364	1117-1118	0.44
MGA	3532-3533	1378-1379	0.42	**KLF13**	3398	1170-1171	0.99	**POU6F1**	3808	1799-1801	0.43
VBP1	4281	2539	0.40	**MSX1**	3546	1397	0.96	**LHX2**	3431	1221-1222	0.43
GMEB2	3231	927-928	0.39	**SCRT2**	3966	2054-2055	0.96	**NR4A2**	3673	1580-1581	0.42
YY1	4313	2589	0.38	**GABPB1**	3196-3200	867-875	0.96	**POU6F2**	3809-3810	1802-1804	0.41
ELF1	3064-3065	666-668	0.36	**SOX9**	4065	2203	0.95	**FOXJ2**	3158	805-806	0.39
POU3F3	3806	1796-1797	0.34	**NFKB2**	3632-3633	1523-1524	0.95	**CEBPG**	2891	399	0.38
GTF2IRD1	3264-3265	972-975	0.34	**EOMES**	3079	687-688	0.95	**VBP1**	4281	2539	0.36
IRF3	3360	1113	0.33	**SIX3**	3991	2094-2095	0.93	**TERF1**	4155-4156	2336-2338	0.36
SRF	4080	2222	0.32	**MECOM**	3503-3508	1334-1341	0.92	**ESRRA**	3091	705	0.35
XBP1	4306-4307	2577-2579	0.30	**BCL6**	2792-2794	243-245	0.92	**PAX6**	3733-3735	1681-1683	0.33
ESRRA	3091	705	0.30	**POU3F3**	3806	1796-1797	0.91	**ZIC3**	4334	2625	0.33
HEY1	3282-3283	1000-1002	0.27	**HINFP**	3287-3288	1008-1011	0.91	**IRF2**	3359	1112	0.32
NFKB2	3632-3633	1523-1524	0.23	**NFATC1**	3620-3624	1505-1514	0.90	**MAF**	3465-3466	1270-1272	0.32
IRF2	3359	1112	0.22	**TBR1**	4139	2312-2313	0.89	**SOX21**	4055	2186-2187	0.31
EOMES	3079	687-688	0.21	**MYBL2**	3556	1410	0.88	**CREB1**	2946-2947	487-489	0.30
FOXB1	3156	801-802	0.20	**USF1**	4268-4269	2517-2520	0.86	**CREB3**	2948	490-491	0.27
NR2F1	3662	1566-1567	0.19	**SP1**	4066-4067	2204-2206	0.86	**IRF7**	3361-3363	1114-1116	0.26
NR4A2	3673	1580-1581	0.17	**ATF2**	2758	190	0.83	**POU3F3**	3806	1796-1797	0.26
STAT3	4105-4107	2258-2262	0.16	**LHX2**	3431	1221-1222	0.83	**NFKB2**	3632-3633	1523-1524	0.25
SP1	4066-4067	2204-2206	0.16	**MYBL2**	3556	1410	0.82	**ATF4**	2762-2763	194-196	0.25
RARA	3871-3874	1905-1908	0.16	**POU3F3**	3806	1796-1797	0.79	**MYBL2**	3556	1410	0.24
CREB1	2946-2947	487-489	0.14	**HEY1**	3282-3283	1000-1002	0.79	**SREBF2**	4079	2221	0.24
NR2F1	3662	1566-1567	0.14	**MEF2A**	3512-3516	1347-1352	0.78	**SOX9**	4065	2203	0.21
FOXJ3	3159	807-808	0.13	**MGA**	3532-3533	1378-1379	0.77	**YY1**	4313	2589	0.21

NE gene	*Polyn. SEQ ID NO:*	*Polyp. SEQ ID NO:*	*NE vs hESC*	*ERG gene*	*Polyn. SEQ ID NO:*	*Polyp. SEQ ID NO:*	*ERG vs NE*	*MRG gene*	*Polyn. SEQ ID NO:*	*Polyp. SEQ ID NO:*	*MRG vs ERG*
HSF1	3331	1071-1072	0.13	**SOX9**	4065	2203	0.74	**RBPJ**	3882-3885	1920-1927	0.17
MYBL2	3556	1410	0.11	**E2F7**	3045	641-642	0.74	**FOXJ3**	3159	807-808	0.17
SRF	4080	2222	0.11	**IRF3**	3360	1113	0.70	**HSF1**	3331	1071-1072	0.16
ETV2	3102	719-720	0.11	**SRF**	4080	2222	0.69	**HMGA2**	3303-3304	1028-1029	0.14
MECP2	3509-3510	1342-1344	0.09	**POU3F2**	3805	1794-1795	0.68	**CUX1**	2970-2972	520-524	0.11
E2F1	3042	637-638	0.09	**SREBF2**	4079	2221	0.68	**POU3F2**	3805	1794-1795	0.05
FOXJ2	3158	805-806	0.09	**NFKB1**	3630-3631	1521-1522	0.64	**EOMES**	3079	687-688	0.05
JUN	3376	1134	0.07	**MAF**	3465-3466	1270-1272	0.64	**ZNF423**	4352	2654-2655	0.04
SCRT2	3966	2054-2055	0.06	**ZNF784**	4363	2673-2674	0.64	**ESRRA**	3091	705	0.02
DLX1	3018-3019	594-596	0.05	**NR1H2**	3659	1562-1563	0.63	**KLF13**	3398	1170-1171	0.02
E2F1	3042	637-638	0.04	**MSX1**	3546	1397	0.60	**NR6A1**	3674-3675	1582-1584	0.00
E2F1	3042	637-638	0.02	**RORA**	3923-3926	1988-1994	0.60				
ATF2	2758	190	0.01	**LMO2**	3446-3448	1244-1247	0.60				
				HINFP	3287-3288	1008-1011	0.60				
				ZNF23	4342	2637-2638	0.60				
				FOXO4	3164-3165	816-818	0.58				
				NR4A2	3673	1580-1581	0.58				
				POU6F2	3809-3810	1802-1804	0.58				
				ID4	3337	1078	0.58				
				RARA	3871-3874	1905-1908	0.57				
				KLF13	3398	1170-1171	0.56				
				LEF1	3416-3419	1199-1203	0.55				
				PATZ1	3721-3724	1659-1666	0.53				
				ELK1	3070-3071	675-677	0.52				
				DLX1	3018-3019	594-596	0.50				
				E2F4	3044	640	0.47				
				E2F4	3044	640	0.47				
				ELK1	3070-3071	675-677	0.42				

NE gene	*Polyn. SEQ ID NO:*	*Polyp. SEQ ID NO:*	*NE vs hESC*	*ERG gene*	*Polyn. SEQ ID NO:*	*Polyp. SEQ ID NO:*	*ERG vs NE*	*MRG gene*	*Polyn. SEQ ID NO:*	*Polyp. SEQ ID NO:*	*MRG vs ERG*
				MAZ	3491-3492	1314-1315	0.41				
				ATF6	2765	199-200	0.39				
				TFAP4	4158	2340	0.37				
				RARA	3871-3874	1905-1908	0.37				
				PAX6	3733-3735	1681-1683	0.33				
				EGR2	3054-3057	654-657	0.33				
				HOMEZ	3324	1059-1060	0.30				
				ELK1	3070-3071	675-677	0.28				
				E2F1	3042	637-638	0.28				
				GMEB2	3231	927-928	0.25				
				ATF4	2762-2763	194-196	0.24				
				ATF2	2758	190	0.22				
				MAX	3486-3490	1305-1313	0.22				
				MGA	3532-3533	1378-1379	0.21				
				NFKB2	3632-3633	1523-1524	0.19				
				ESRRA	3091	705	0.16				
				E2F4	3044	640	0.16				
				STAT3	4105-4107	2258-2262	0.16				
				TCF3	4146-4147	2325-2326	0.15				
				NR2E1	3661	1565	0.12				
				ELF3	3068-3069	673-674	0.11				
				SOX4	4056	2188	0.06				
				USF1	4268-4269	2517-2520	0.04				
				ZNF652	4358-4359	2665-2667	0.01				

Table 4: Provided are the scoring results of the transcriptional epigenetic landscape activity of the factors in three developmental stages: (i) The differentiation of human pluripotent embryonic stem cells (hESCs) to HES5+ neuroepithelial cells (marked as "NE_vs_hESC") and the factors are termed "NE_gene"; (ii) The differentiation of HES5+ neuroepithelial cells to HES5+ early radial glial cells (marked as "ERG_vs_NE") and the factors are termed "ERG_gene"; (iii) The differentiation of HES5+ early radial glial cells to HES5+ mid radial glial cells (marked as "MRG vs ERG") and the factors are termed "MRG gene".

Table 5
Sequence information of additional genes

Gene Symbol	alias	Polypeptide SEQ ID NO:	Polynucleotide SEQ ID NO:
ANXA2		4889 - 4892	4870 - 4873
CUX2		4893	4874
EGFR		4894 - 4897	4875 - 4878
NFIB		4898	4879
S100B		4899	4880
SOX1		4900	4881
SOX2		4901	4882
GLAST	SLC1A3	4902 - 4904	4883 - 4885
plzf	ZBTB16	4905 - 4906	4886 - 4887
nestin	nes	4907	4888

Table 5: provided are representative polynucleotide and polypeptide sequences of genes of some embodiments of the invention.

SUPPLEMENTARY DATA CONTAINED IN A CD

Supplementary data 1: Summary statistics and overview of all datasets generated for this study.

Supplementary data 2: Gene expression levels in FPKM for all RNA-Seq datasets used in this study based GENCODE annotations.

Supplementary data 3: TERA scores for all epigenetic marks individually as well as averaged for all NPC stages as well as H3K27ac and H3K4me3 based scores for most REMC cell types.

Supplementary data 4: shRNA screen raw data and evaluation.

Supplementary data 5: Gene set enrichment analysis for predicted individual TF binding in dynamically regulated H3K27ac regions at each stage.

Supplementary data 6: Global gene expression arrays for HES5+ and HES5- populations during NSC progression. Gene expression intensities for ES cells as well as HES5+ and HES5- cells during 220 days of propagation *in vitro*. Normalized and collapsed log2 transformed probe level intensities are shown.

Supplementary data 7: RNA sequencing analysis for cells in all developmental stages and their differentiated progeny in FPKM. Global Gene Expression PROGENITORS AND DIFFERENTIATED PROGENY RNA SEQ.

Although the invention has been described in conjunction with specific embodiments thereof, it is evident that many alternatives, modifications and variations will be apparent to those skilled in the art. Accordingly, it is intended to embrace all such alternatives, modifications and variations that fall within the spirit and broad scope of the appended claims.

All publications, patents and patent applications mentioned in this specification are herein incorporated in their entirety by reference into the specification, to the same extent as if each individual publication, patent or patent application was specifically and individually indicated to be incorporated herein by reference. In addition, citation or identification of any reference in this application shall not be construed as an admission that such reference is available as prior art to the present invention. To the extent that section headings are used, they should not be construed as necessarily limiting.

REFERENCES
(ADDITIONAL REFERENCES ARE CITED IN TEXT)

1. McKay, R. D. Stem cells in the central nervous system. Science 276, 66-71 (1997).
2. Taverna, E., Gotz, M. & Huttner, W. B. The Cell Biology of Neurogenesis: Toward an Understanding of the Development and Evolution of the Neocortex. Annual review of cell and developmental biology, doi:10.1146/annurev-cellbio-101011- 155801 (2014).
3. Alvarez-Buylla, A. & Temple, S. Stem cells in the developing and adult nervous system. Journal of Neurobiology 36, 110 (1998).
4. Shen, Q. et al. The timing of cortical neurogenesis is encoded within lineages of individual progenitor cells. Nat Neurosci 9, 743-751, doi:10.1038/nn1694 (2006).
5. Gotz, M. & Barde, Y. A. Radial glial cells: Defined and major intermediates between embryonic stem cells and CNS neurons. Neuron 46, 369-372 (2005).
6. Beckervordersandforth, R. et al. In vivo fate mapping and expression analysis reveals molecular hallmarks of prospectively isolated adult neural stem cells. Cell Stem Cell 7, 744-758, doi:10.1016/j.stem.2010.11.017 (2010).
7. Kriks, S. et al. Dopamine neurons derived from human ES cells efficiently engraft in animal models of Parkinson's disease. Nature 480, 547-551, doi:10.1038/nature10648 (2011).
8. Li, X. J. et al. Specification of motoneurons from human embryonic stem cells. Nat Biotechnol 23, 215-221, doi:10.1038/nbt1063 (2005).
9. Eiraku, M. et al. Self-organized formation of polarized cortical tissues from ESCs and its active manipulation by extrinsic signals. Cell Stem Cell 3, 519-532, doi:S1934-5909(08)00455-4 [pii] 10.1016/j.stem.2008.09.002 (2008).
10. Gaspard, N. et al. An intrinsic mechanism of corticogenesis from embryonic stem cells. Nature 455, 351-357, doi:nature07287 [pii] 10.1038/nature07287 (2008).
11. Mariani, J. et al. Modeling human cortical development in vitro using induced pluripotent stem cells. Proc Natl Acad Sci U S A 109, 12770-12775, doi:10.1073/pnas.1202944109 (2012).
12. Hansen, D. V., Rubenstein, J. L. & Kriegstein, A. R. Deriving excitatory neurons of the neocortex from pluripotent stem cells. Neuron 70, 645-660, doi:10.1016/j.neuron.2011.05.006 (2011).
13. Lancaster, M. A. et al. Cerebral organoids model human brain development and microcephaly. Nature 501, 373-379, doi:10.1038/nature12517 (2013).
14. van de Leemput, J. et al. CORTECON: A Temporal Transcriptome Analysis of In Vitro Human Cerebral Cortex Development from Human Embryonic Stem Cells. Neuron 83, 51-68, doi:10.1016/j.neuron.2014.05.013 (2014).
15. Elkabetz, Y. et al. Human ES cell-derived neural rosettes reveal a functionally dinstinct early neural stem cell stage. Genes & Development 22, 152-165 (2008).
16. Kageyama, R. & Ohtsuka, T. The Notch-Hes pathway in mammalian neural development. Cell Res. 9, 179-188 (1999).
17. Placantonakis, D. et al. BAC transgenesis in human ES cells as a novel tool to19/39 define the human neural lineage. Stem Cells (2008).
18. Lowell, S., Benchoua, A., Heavey, B. & Smith, A. G. Notch promotes neural lineage entry by pluripotent embryonic stem cells. PLoS.Biol. 4, e121 (2006).
19. Gotz, M., Stoykova, A. & Gruss, P. Pax6 controls radial glia differentiation in the cerebral cortex. Neuron 21, 1031-1044 (1998).

20. Gaiano, N., Nye, J. S. & Fishell, G. Radial glial identity is promoted by Notch1 signaling in the murine forebrain. Neuron 26, 395-404 (2000).
21. Huttner, W. B. & Kosodo, Y. Symmetric versus asymmetric cell division during neurogenesis in the developing vertebrate central nervous system. Current opinion in cell biology 17, 648-657, doi:10.1016/j.ceb.2005.10.005 (2005).
22. Temple, S. The development of neural stem cells. Nature 414, 112-117 (2001).
23. Elkabetz, Y. & Studer, L. Human ESC-derived neural rosettes and neural stem cell progression. Cold Spring Harb Symp Quant Biol 73, 377-387, doi:sqb.2008.73.052 [pii] 10.1101/sqb.2008.73.052 (2008).
24. Molyneaux, B. J., Arlotta, P., Menezes, J. R. & Macklis, J. D. Neuronal subtype specification in the cerebral cortex. Nature reviews. Neuroscience 8, 427-437, doi:nrn2151 [pii] 10.1038/nrn2151 (2007).
25. Arlotta, P. et al. Neuronal subtype-specific genes that control corticospinal motor neuron development in vivo. Neuron 45, 207-221, doi:S0896627304008530 [pii] 10.1016/j.neuron.2004.12.036 (2005).
26. Hevner, R. F. et al. Tbr1 regulates differentiation of the preplate and layer 6. Neuron 29, 353-366 (2001).
27. Zecevic, N. & Rakic, P. Development of layer I neurons in the primate cerebral cortex. J Neurosci 21, 5607-5619 (2001).
28. Britanova, O. et al. Satb2 is a postmitotic determinant for upper-layer neuron specification in the neocortex. Neuron 57, 378-392, doi:10.1016/j.neuron.2007.12.028 (2008).
29. Nieto, M. et al. Expression of Cux-1 and Cux-2 in the subventricular zone and upper layers II-IV of the cerebral cortex. J Comp Neurol 479, 168-180, doi:10.1002/cne.20322 (2004).
30. Franco, S. J. et al. Fate-restricted neural progenitors in the mammalian cerebral cortex. Science 337, 746-749, doi:10.1126/science.1223616 (2012).
31. Haubensak, W., Attardo, A., Denk, W. & Huttner, W. B. Neurons arise in the basal neuroepithelium of the early mammalian telencephalon: A major site of neurogenesis. Proceedings of the National Academy of Sciences of the United States of America 101, 3196-3201 (2004).
32. Chen, B., Schaevitz, L. R. & McConnell, S. K. Fezl regulates the differentiation and axon targeting of layer 5 subcortical projection neurons in cerebral cortex. Proc Natl Acad Sci U S A 102, 17184-17189, doi:0508732102 [pii] 10.1073/pnas.0508732102 (2005).
33. Guo, C. et al. Fezf2 expression identifies a multipotent progenitor for neocortical 20/39 projection neurons, astrocytes, and oligodendrocytes. Neuron 80, 1167-1174, doi:10.1016/j.neuron.2013.09.037 (2013).
34. Englund, C. et al. Pax6, Tbr2, and Tbr1 are expressed sequentially by radial glia, intermediate progenitor cells, and postmitotic neurons in developing neocortex. J Neurosci 25, 247-251, doi:25/1/247 [pii] 10.1523/JNEUROSCI.2899-04.2005 (2005).
35. Takiguchi-Hayashi, K. et al. Generation of reelin-positive marginal zone cells from the caudomedial wall of telencephalic vesicles. Journal of Neuroscience 24, 2286- 2295 (2004).
36. Hansen, D. V., Lui, J. H., Parker, P. R. & Kriegstein, A. R. Neurogenic radial glia in the outer subventricular zone of human neocortex. Nature 464, 554-561, doi:10.1038/nature08845 (2010).
37. Fietz, S. A. et al. OSVZ progenitors of human and ferret neocortex are epithelial-like and expand by integrin signaling. Nat Neurosci 13, 690-699,

doi:10.1038/nn.2553 (2010).
38. Dominguez, M. H., Ayoub, A. E. & Rakic, P. POU-III Transcription Factors (Brn1, Brn2, and Oct6) Influence Neurogenesis, Molecular Identity, and Migratory Destination of Upper-Layer Cells of the Cerebral Cortex. Cereb Cortex 23, 2632-2643, doi:10.1093/cercor/bhs252 (2013).
39. Anthony, T. E. & Heintz, N. Genetic lineage tracing defines distinct neurogenic and gliogenic stages of ventral telencephalic radial glial development. Neural development 3, 30, doi:10.1186/1749-8104-3-30 (2008).
40. Burrows, R. C., Wancio, D., Levitt, P. & Lillien, L. Response diversity and the timing of progenitor cell maturation are regulated by developmental changes in EGFR expression in the cortex. Neuron 19, 251-267 (1997).
41. Doetsch, F., Petreanu, L., Caille, I., Garcia-Verdugo, J. M. & Alvarez-Buylla, A. EGF converts transit-amplifying neurogenic precursors in the adult brain into multipotent stem cells. Neuron 36, 1021-1034 (2002).
42. Carlen, M. et al. Forebrain ependymal cells are Notch-dependent and generate neuroblasts and astrocytes after stroke. Nat Neurosci 12, 259-267, doi:10.1038/nn.2268 (2009).
43. Piper, M. et al. NFIA controls telencephalic progenitor cell differentiation through repression of the Notch effector Hes1. J Neurosci 30, 9127-9139, doi:10.1523/JNEUROSCI.6167-09.2010 (2010).
44. Pereira, J. D. et al. Ezh2, the histone methyltransferase of PRC2, regulates the balance between self-renewal and differentiation in the cerebral cortex. Proc Natl Acad Sci U S A 107, 15957-15962, doi:10.1073/pnas.1002530107 (2010).
45. Roy, K. et al. The Tlx gene regulates the timing of neurogenesis in the cortex. J Neurosci 24, 8333-8345, doi:10.1523/JNEUROSCI.1148-04.2004 (2004).
46. Liu, H. K. et al. The nuclear receptor tailless induces long-term neural stem cell expansion and brain tumor initiation. Genes Dev 24, 683-695, doi:10.1101/gad.560310 (2010).
47. Barker, N. et al. Identification of stem cells in small intestine and colon by marker gene Lgr5. Nature 449, 1003-1007, doi:10.1038/nature06196 (2007).
48. Yu, J. et al. Induced pluripotent stem cell lines derived from human somatic cells. Science 318, 1917-1920 (2007).
49. Nishino, J., Kim, I., Chada, K. & Morrison, S. J. Hmga2 promotes neural stem cell self-renewal in young but not old mice by reducing p16Ink4a and p19Arf Expression. Cell 135, 227-239, doi:10.1016/j.cell.2008.09.017 (2008).
50. Kornecki, E., Walkowiak, B., Naik, U. P. & Ehrlich, Y. H. Activation of human platelets by a stimulatory monoclonal antibody. J Biol Chem 265, 10042-10048 (1990).
51. Famulski, J. K. & Solecki, D. J. New spin on an old transition: epithelial parallels in neuronal adhesion control. Trends in neurosciences 36, 163-173, doi:10.1016/j.tins.2012.10.002 (2013).
52. Ericson, J., Morton, S., Kawakami, A., Roelink, H. & Jessell, T. M. Two critical periods of Sonic Hedgehog signaling required for the specification of motor neuron identity. Cell 87, 661-673 (1996).
53. Rodriguez, M., Choi, J., Park, S. & Sockanathan, S. Gde2 regulates cortical neuronal identity by controlling the timing of cortical progenitor differentiation. Development 139, 3870-3879, doi:10.1242/dev.081083 (2012).
54. Gafni, O. et al. Derivation of novel human ground state naive pluripotent stem cells. Nature, doi:10.1038/nature12745 (2013).

55. Chambers, S. M. et al. Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. Nat Biotechnol 27, 275-280, doi:nbt.1529 [pii] 10.1038/nbt.1529 (2009).
56. Lafaille, F. G. et al. Impaired intrinsic immunity to HSV-1 in human iPSC-derived TLR3-deficient CNS cells. Nature 491, 769-773, doi:10.1038/nature11583 (2012).
57. Imayoshi, I., Sakamoto, M., Yamaguchi, M., Mori, K. & Kageyama, R. Essential roles of Notch signaling in maintenance of neural stem cells in developing and adult brains. *J. Neurosci.* **30:** 3489-3498, doi:10.1523/JNEUROSCI.4987-09.2010 (2010).
58. Shimojo, H., Ohtsuka, T. & Kageyama, R. Dynamic expression of notch signaling genes in neural stem/progenitor cells. *Frontiers in neuroscience* **5**, 78, doi:10.3389/fnins.2011.00078 (2011).
59. Edri, R. *et al.* Analyzing human neural stem cell ontogeny by consecutive isolation of Notch active neural progenitors. Nat. Commun. 2015, Mar 23; 6: 6500.
50. Placantonakis, D. G. *et al.* BAC transgenesis in human embryonic stem cells as a novel tool to define the human neural lineage. *Stem Cells* **27**, 521-532, doi:10.1634/stemcells.2008-0884 (2009).
51. Voss, T. C. & Hager, G. L. Dynamic regulation of transcriptional states by chromatin and transcription factors. *Nat Rev Genet* **15**, 69-81, doi:10.1038/nrg3623 (2014).
52. Ziller, M. J. *et al.* Charting a dynamic DNA methylation landscape of the human genome. *Nature* **500**, 477-481, doi:10.1038/nature12433 (2013).
53. Gifford, C. A. *et al.* Transcriptional and epigenetic dynamics during specification of human embryonic stem cells. *Cell* **153**, 1149-1163, doi:10.1016/j.cell.2013.04.037 (2013).
54. Maurano, M. T. *et al.* Systematic localization of common disease- associated variation in regulatory DNA. *Science* **337**, 1190-1195, doi:10.1126/science.1222794 (2012).
55. Claussnitzer, M. *et al.* Leveraging cross-species transcription factor binding site patterns: from diabetes risk loci to disease mechanisms. *Cell* **156**, 343-358, doi:10.1016/j.cell.2013.10.058 (2014).
56. Hanashima, C., Li, S. C., Shen, L., Lai, E. & Fishell, G. Foxg1 suppresses early cortical cell fate. *Science* **303**, 56-59, doi:10.1126/science.1090674 (2004).
57. Martinez-Barbera, J. P. *et al.* Regionalisation of anterior neuroectoderm and its competence in responding to forebrain and midbrain inducing activities depend on mutual antagonism between OTX2 and GBX2. *Development* **128**, 4789-4800 (2001).
58. Pevny, L. H., Sockanathan, S., Placzek, M. & Lovell-Badge, R. A role for SOX1 in neural determination. *Development* **125**, 1967-1978 (1998).
59. Uittenbogaard, M. & Chiaramello, A. Expression of the bHLH transcription factor Tcf12 (ME1) gene is linked to the expansion of precursor cell populations during neurogenesis. *Brain Res Gene Expr Patterns* **1**, 115- 121 (2002).
60. McLean, C. Y. *et al.* GREAT improves functional interpretation of cis-regulatory regions. *Nat Biotechnol* **28**, 495-501, doi:10.1038/nbt.1630 (2010).
61. Sims, D. *et al.* High-throughput RNA interference screening using pooled shRNA libraries and next generation sequencing. *Genome Biol* **12**, R104, doi:10.1186/gb-2011-12-10-r104 (2011).

62. Blackshear, P. J. *et al.* Graded phenotypic response to partial and complete deficiency of a brain-specific transcript variant of the winged helix transcription factor RFX4. *Development* **130**, 4539-4552, doi:10.1242/dev.00661 (2003).
63. Zarbalis, K. *et al.* A focused and efficient genetic screening strategy in the mouse: identification of mutations that disrupt cortical development. *PLoS Biol* **2**, E219, doi:10.1371/journal.pbio.0020219 (2004).
64. Zhou, C., Tsai, S. Y. & Tsai, M. J. COUP-TFI: an intrinsic factor for early regionalization of the neocortex. *Genes Dev* **15**, 2054-2059, doi:10.1101/gad.913601 (2001).
65. Faedo, A. *et al.* COUP-TFI coordinates cortical patterning, neurogenesis, and laminar fate and modulates MAPK/ERK, AKT, and beta-catenin signaling. *Cereb Cortex* **18**, 2117-2131, doi:10.1093/cercor/bhm238 (2008).
66. Qureshi, I. A., Gokhan, S. & Mehler, M. F. REST and CoREST are transcriptional and epigenetic regulators of seminal neural fate decisions. *Cell Cycle* **9**, 4477-4486 (2010).
67. Garber, M. *et al.* A high-throughput chromatin immunoprecipitation approach reveals principles of dynamic gene regulation in mammals. *Molecular cell* **47**, 810-822, doi:10.1016/j.molcel.2012.07.030 (2012).
68. Engreitz, J. M. *et al.* The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. *Science* **341**, 1237973, doi:10.1126/science.1237973 (2013).
69. Luo, B. *et al.* Highly parallel identification of essential genes in cancer cells. *Proc. Natl. Acad. Sci. U.S.A.* **105:** 20380-20385, doi:10.1073/pnas.0810485105 (2008).
70. Strezoska, Z. *et al.* Optimized PCR conditions and increased shRNA fold representation improve reproducibility of pooled shRNA screens. *PLoS One* **7**, e42341, doi:10.1371/journal.pone.0042341 (2012).
71. Boyle, P. *et al.* Gel-free multiplexed reduced representation bisulfite sequencing for large-scale DNA methylation profiling. *Genome Biol* **13**, R92, doi:10.1186/gb-2012-13-10-r92 (2012).
72. Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics***25**, 1105-1111, doi:10.1093/bioinformatics/btp120 (2009).
73. Trapnell, C. *et al.* Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nature Biotechnology* **31**, 46-+, doi:Doi 10.1038/Nbt.2450 (2013).
74. Trapnell, C. *et al.* Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature protocols* **7**, 562- 578, doi:10.1038/nprot.2012.016 (2012).
75. Xi, Y. & Li, W. BSMAP: whole genome bisulfite sequence MAPping program. *BMC Bioinformatics* **10**, 232, doi:10.1186/1471-2105-10-232 (2009).
76. Li, H., Ruan, J. & Durbin, R. Mapping short DNA sequencing reads and calling variants using mapping quality scores. *Genome Res* **18**, 1851- 1858, doi:gr.078212.108 [pii] 10.1101/gr.078212.108 (2008).
77. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357-359, doi:10.1038/nmeth.1923 (2012).
78. Thorvaldsdottir, H., Robinson, J. T. & Mesirov, J. P. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform*, doi:10.1093/bib/bbs017 (2012).

79. Dai, Z. *et al.* shRNA-seq data analysis with edgeR. *F1000Res* **3**, 95, doi:10.12688/f1000research.3928.110.12688/f1000research.4204 (2014).
80. Smyth, G. K. in *Bioinformatics and computational biology solutions using R and Bioconductor Statistics for biology and health* (ed Robert Gentleman) (Springer Science+Business Media, 2005).
81. Goff, L., Trapnell, C. & Kelley, D. cummeRbund: Analysis, exploration, manipulation, and visualization of Cufflinks high-throughput sequencing data. (2012). <www.compbio.mit.edu/cummeRbund/>.
82. Li, Q. H., Brown, J. B., Huang, H. Y. & Bickel, P. J. Measuring Reproducibility of High-Throughput Experiments. *Annals of Applied Statistics* **5**, 1752-1779, doi:Doi 10.1214/11-Aoas466 (2011).
83. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137, doi:10.1186/gb-2008-9-9-r137 (2008).
84. Mikkelsen, T. S. *et al.* Comparative epigenomic analysis of murine and human adipogenesis. *Cell* **143**, 156-169, doi:10.1016/j.cell.2010.09.006 (2010).
85. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. *J Roy Stat Soc B Met* **57**, 289-300 (1995).
86. Alan Dabney, J. D. S. a. w. a. f. G. R. W. Q-value estimation for false discovery rate control. **R package version 1.34.0**.
87. Ross-Innes, C. S. *et al.* Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature* **481**, 389-393, doi:10.1038/nature10730 (2012).
88. Neph, S. *et al.* An expansive human regulatory lexicon encoded in transcription factor footprints. *Nature* **489**, 83-90, doi:10.1038/nature11212 (2012).
89. Thompson, C. L. *et al.* A high-resolution spatiotemporal atlas of gene expression of the developing mouse brain. *Neuron* **83**, 309-323, doi:10.1016/j.neuron.2014.05.033 (2014).
90. Fogel, G. B. *et al.* A statistical analysis of the TRANSFAC database. *Bio Systems* **81**, 137-154, doi:10.1016/j.biosystems.2005.03.003 (2005).
91. Jolma, A. *et al.* DNA-binding specificities of human transcription factors. *Cell* **152**, 327-339, doi:10.1016/j.cell.2012.12.009 (2013).
92. Gupta, S., Stamatoyannopoulos, J. A., Bailey, T. L. & Noble, W. S. Quantifying similarity between motifs. *Genome Biol* **8**, R24, doi:10.1186/gb-2007-8-2-r24 (2007).
93. Bailey, T. L., Williams, N., Misleh, C. & Li, W. W. MEME: discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Res* **34**, W369-373, doi:34/suppl_2/W369 [pii]10.1093/nar/gkl198 (2006).
94. Manke, T., Roider, H. G. & Vingron, M. Statistical modeling of transcription factor binding affinities predicts regulatory interactions. *PLoS Comput Biol.* **4**, e1000039, doi:10.1371/journal.pcbi.1000039 (2008).
95. Manke, T., Heinig, M. & Vingron, M. Quantifying the effect of sequence variation on regulatory interactions. *Human mutation* **31**, 477-483, doi:10.1002/humu.21209 (2010).
96. Boulesteix, A. L. & Strimmer, K. Predicting transcription factor activities from combined analysis of microarray and ChIP data: a partial least squares approach. *Theoretical biology & medical modelling* **2**, 23, doi:10.1186/1742-4682-2-23 (2005).

97. Boulesteix, A. L. & Strimmer, K. Partial least squares: a versatile tool for the analysis of high-dimensional genomic data. *Brief Bioinform* **8**, 32-44, doi:10.1093/bib/bbl016 (2007).
98. Dejong, S. Simpls - an Alternative Approach to Partial Least-Squares Regression. *Chemometr Intell Lab* **18**, 251-263, doi:Doi 10.1016/0169-7439(93)85002-X (1993).
99. Bernstein, B. E. *et al.* An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57-74, doi:10.1038/nature11247 (2012).
100. Sherwood, R. I. *et al.* Discovery of directional and nondirectional pioneer transcription factors by modeling DNase profile magnitude and shape. *Nat Biotechnol* **32**, 171-178, doi:10.1038/nbt.2798 (2014).
101. Thurman, R. E. *et al.* The accessible chromatin landscape of the human genome. *Nature* **489**, 75-82, doi:10.1038/nature11232 (2012).

WHAT IS CLAIMED IS:

1. A method of isolating neural progenitor cells, the method comprising:

(a) culturing pluripotent stem cells having been transformed to express a Notch-activated reporter under culture conditions suitable for differentiation of said pluripotent stem cells into neural progenitor cells; and

(b) successively isolating progenitor cells of interest based on activation of said Notch-activated reporter.

2. An isolated population of cells comprising at least 10 % HES5+ cells, wherein said HES5+ cells are:

(i) early radial glial (E-RG) cells;

(ii) mid radial glial (M-RG) cells;

(iii) late radial glial (L-RG) cells; or

(iv) long term neural progenitor (LNP) cells.

3. The isolated population of cells of claim 2, further comprising HES5+ neuroepithelial (NE) cells.

4. The isolated population of cells of claim 3, wherein said HES5+ NE cells are capable of differentiating into HES5+ E-RG cells and into HES5- central nervous system cells neurons.

5. The isolated population of cells of claim 4, wherein said HES5+ E-RG cells are capable of differentiating into said HES5+ M-RG cells and into HES5- neural progenitor cells.

6. The isolated population of cells of claim 2, wherein said HES5+ M-RG cells are capable of differentiating into said HES5+ L-RG cells and into HES5- intermediate progenitor cells (INPs).

7. The isolated population of cells of claim 2, wherein said HES5+ L-RG cells are capable of differentiating into said HES5+ LNP cells and into HES5- neurons and astrocytes.

8. The isolated population of cells of claim 2, wherein said HES5+ LNP cells which comprise HES5+ adult neural stem cells (aNSCs) and into HES5- neurons, oligodendrocyte and astrocytes.

9. The isolated population of cells of claim 3, or 4, wherein said HES5+ neuroepithelial cells are characterized by a higher expression level of at least one gene selected from the group consisting of: TOP2A, HIST1H4C, TRIM71, PPIG, MLLT4, TNC, CDK1, OIP5, GDF15, MCM6, TP53TG1, FAM83D, FANCI, GINS2, KDM5A, GSTM3, FAM64A, LIMS1, CENPH, KIF2C, ATAD2, DTL, CDCA5, ARHGEF6, LIPA, POLE2, RRM2, MAD2L1, CKS1B, TTK, DHFR, S100A4, NUP37, PMAIP1, CENPN, RNASEH2A, BST2, MCM10, MAF, KIAA0101, C8ORF4, E2F7, CENPA, UBE2T, RAB13, TMEM126A, MAGT1, CDC6, C6ORF211, RFC5, PSMD1, HMMR, UNG, UBE2C, GINS1, AURKB, LEPREL1, SBNO1, ZWINT, MKI67, CCAR1, FKBP5, PVRL3, CCNB1, NOP58, COL4A1, GGH, LSM6, EID1, GPX8, STC2, CD276, HS2ST1, EIF5B, HDGF, and NOL7 as compared to the expression level of said at least one gene in HES5- differentiated cells obtained by culturing undifferentiated pluripotent stem cells (PSCs) under culture conditions suitable for differentiation said PSCs into HES5+ neuroepithelial cells.

10. The isolated population of cells of claim 2, wherein said HES5+ early radial glial cells are characterized by a higher expression level of at least one gene selected from the group consisting of: NR2E1, HES5, ARX, C1ORF61, FRZB, GRM3, EPHA3, NAV3, EGR2, RGMA, NRXN3, FAM107A, FABP7, EGR3, ZNF385B, TTYH1, SNCAIP, NRARP, PLP1, LIX1, LFNG, HES4, CD82, HS6ST1, PTPRZ1, CACHD1, DACH1, FEZF2, DTX4, FUT9, WNT5B, ENPP2, POU3F3, EMX2, MECOM, XYLT1, ARMCX2, FOS, PPAP2B, NOS2, LRP2, SOX9, NLGN3, TMEM2, CXCR7, EPHA7, SMOC1, TBC1D9, FAT4, SCUBE3, FUT8, CSPG5, DLL1, BOC, ID4, EGR1, ALPL, RFX4, GALNT12, CBX2, FHOD3, SORBS2,

GUCY1B3, MBIP, FBXO16, SHISA2, DAB1, GLI3, FZD3, SEMA5B, LGALS3BP, SFRP1, C1QL1, RING1, GPRC5B, ZNF710, WSCD1, VPS37B, ZIC2, SDK2, DOCK11, GAS1, ZNF436, TMSB15A, IER2, FEZ1, CELF2, SFT2D3, NCALD, AKAP7, MYADM, NEDD4L, PHC2, PI4KAP2, STARD3, and CAMK1D as compared to the expression level of said at least one gene in HES5- differentiated cells obtained by culturing HES5+ neuroepithelial cells under culture conditions suitable for differentiation said HES5+ neuroepithelial cells into HES5+ early radial glial cells.

11. The isolated population of cells of claim 2, wherein said HES5+ mid radial glial cells are characterized by a higher expression level of at least one gene selected from the group consisting of: FZD10, ZEB2, EN2, ST20, CDKN2C, RAB10, WASF1, ZBED4, EZH2, PPA2, H1F0, CCNJ, ITGB8, SH3BGRL3, IRX2, KIF23, PEG10, SMC3, NUSAP1, APLP1, ADAMTS3, RACGAP1, LIMCH1, ETNK1, RNF13, ARID1B, TRIM28, CNOT8, CRNDE, TWSG1, NT5DC2, NAA50, NUF2, ABCE1, PLTP, FBRSL1, DCAF16, OGT, ZFYVE16, FOXM1, PM20D2, POU3F2, MCM4, HERPUD2, VRK1, TRIM41, SATB1, HOMER1, CCNG1, ATF2, AP1AR, GABPA, STXBP3, SMC5, CDKN1B, NUPL1, UBA1, CYTH2, FXYD6, ISYNA1, DOCK1, 41527, LPHN2, IDI1, PXMP2, U2AF2, ARHGAP12, KLHL24, CKAP2, ZNF238, PARP6, NHSL1, PBRM1, BAZ1A, MAP4K5, TSPAN12, SH3GLB1, ASPM, ANKLE2, SPG20, MAP4K4, CASC4, FUBP3, ARSB, BTAF1, SIKE1, VEZT, PBX3, CBL, EIF2AK4, API5, MTSS1, NET1, CHD6, ZNF117, PNMA1, PTPN13, MTIF2, SSFA2, KIAA1279, STRN3, WIPF1, MEIS2, ZC3H4, DYNC1I2, RTN4, TAF2, RASA1, OSBPL8, SKA2, IGF1R, RNF6, SGTB, TMEM131, HIATL1, TGIF1, TMEM170B, PSAT1, ACBD5, HECTD2, ASF1A, LAMB1, GLS, DDX39, DGCR2, EIF1AX, SALL1, GOLPH3, PTBP2, GRIP1, PNPLA8, VASH2, SUDS3, PFDN4, BAZ2A, PRKDC, GLYR1, DAZAP2, PCMTD1, SENP6, CLINT1, RECQL, CNTNAP2, CTBP1, C10ORF18, CDON, B4GALT6, CSNK1G3, STAT5B, TMEM60, HNRNPH2, TACC2, CCNG2, FSCN1, CCNA2, C21ORF45, PLRG1, ZFHX3, UBE2A, DMTF1, TRA2A, MYO5A, FAM96A, IFT80, VPS26A, MRPL50, ACYP1, WDR11, PLDN, RPRD1A, MEAF6, CKAP5, YTHDC2, GABARAPL2, IPO5, PGAP1, C14ORF147, CD200, MST4, PPT1, ANKRD50, HPS3, CCNC, THRAP3, TWF1, CYP51A1, PSPC1, WDR75, CAST, SEPW1, C21ORF59,

PIK3C2A, GNG5, MED4, GIPC1, STK39, KIAA1715, PHF6, PPTC7, SOCS4, PPM1B, UQCRB, C10ORF84, SLAIN1, RAB6A, SOS2, KLF10, RNF4, C3ORF63, INSIG1, CPSF1, DNAJC4, ATP2B4, PPP2R1A, TRIM22, SDC3, TSNAX, PPIL4, ZDHHC2, ZBTB44, ANO6, PPP2CB, UBA2, BBS2, ZNF423, RNF5, C1ORF31, IFT81, CPSF6, KLHL9, FAM164A, TTC35, CCDC90B, TM9SF4, SEC24B, SMARCC2, CAP2, SAR1A, THY1, RBPMS2, EIF3A, DZIP1, ARL6IP1, SACM1L, PAPD4, SCG2, TCF3, EFHA1, HNRNPA2B1, EWSR1, STAG2, YEATS4, PAQR3, GAR1, FTH1, C19ORF43, TMEM14C, CCDC104, PSMD12, DCTD, SSR1, HMGCS1, HMGB3, KIF3A, TMEM128, PATZ1, RBL2, ARFGAP3, DNAJB5, TMED7, G3BP2, BMPR1A, FMR1, TPST2, TMSB4X, RP2, CEP170, KLHL23, RNF7, HNRNPH1, MARCKS, HNRNPD, TOB1, UTP11L, RFK, DHX36, LCOR, WBP5, PHLDB2, USP33, EFNB2, C6ORF62, MEX3B, ABCD3, ATG3, ARID4B, C7ORF11, EPB41, TCF12, CDK8, CMIP, ATG12, CETN3, ZNF217, TMEM55A and UBE2N as compared to the expression level of said at least one gene in HES5-differentiated cells obtained by culturing HES5+ early radial glial cells under culture conditions suitable for differentiation said HES5+ early radial glial cells into HES5+ mid radial glial cells.

12. The isolated population of cells of claim 2, wherein said wherein said HES5+ late radial glial cells are characterized by a higher expression level of at least one gene selected from the group consisting of: PMP2, GABBR2, BCAN, LUZP2, SALL3, SYNM, DCT, OLIG1, SPON1, PDGFRA, COL22A1, KIAA1239, PCDH10, LPAR4, VAV3, CADM2, SOX6, SLC6A1, DPP6, FGFR3, PDE3B, MOXD1, TNFRSF19, PYGL, GPC6, COL11A1, TRIM9, GABRB3, TFPI, CREB5, RAB3GAP2, NCAN, EFHD1, SLITRK2, PAX6, SLC1A4, GPR155, GPD2, CHST11, PAQR8, MT2A, GPC3, TMEM51, CHST3, PAG1, MYO5C, CACNB2, NDRG2, ST3GAL5, TPD52L1, TRIB1, PRKCA, BCKDHB, GLT25D2, LITAF, PLCB1, TIMP3, ZBTB46, OPCML, CTDSPL, MDGA2, MEGF10, EYA2, KANK1, RAB31, TRIL, FAM171B, ALCAM, RAB6B, PGM2L1, LARGE, HPCAL1, HTRA1, TRPS1, TRIB3, IGF2BP2, PITPNC1, CMTM4, IAH1, DHTKD1, SNAP29, CTNNBIP1, NQO2, MAP4, CBR1, LTBP1, C5ORF32, MARK1, AASS, CISD1, DSC2, SLC25A33, RIMS3, ZIC3, EGF, SRGAP2, RANGAP1, SCRG1, PRCP, CA12, HEATR5A, ZNF503, GYG2, ANAPC1, C19ORF63, ASAP1, C1ORF96,

DHX33, FASTKD1, STAU2, MAML2, RRAS2, GLTP, VPS13B, GPT2, NKAIN4 and ZC3HAV1 as compared to the expression level of said at least one gene in HES5- differentiated cells obtained by culturing HES5+ mid radial glial cells under culture conditions suitable for differentiation said HES5+ mid radial glial cells into HES5+ late radial glial cells.

13. The isolated population of cells of claim 2, wherein said HES5+long term neural progenitor cells are characterized by a higher expression level of at least one gene selected from the group consisting of: ANXA2P2, ANXA2, FRAS1, SPOCK1, PCDHB15, SLC10A4, TPBG, C5ORF39, MMP14, TNFRSF10D, S100A6, RNF182, LGALS1, ISL1, SPINK5, DOCK10, LECT1, LYPD1, ARMCX1, NAP1L2, COL4A6, GSN, PLAG1, MMD, PTGR1, PDP1, COL18A1, ZIC4, BASP1, AHNAK, REC8, KLHDC8B, FRMD6, MYL9, RBMS1, TNFRSF21, and FAM38A as compared to the expression level of said at least one gene in HES5- differentiated cells obtained by culturing HES5+ late radial glial cells under culture conditions suitable for differentiation said HES5+ late radial glial cells into HES5+ long term neural progenitor cells.

14. The isolated population of cells of claim 3, wherein said HES5+ cells are neuroepithelial cells (NE) which constitute at least about 80 % of the isolated population of cells.

15. The isolated population of cells of claim 2, wherein said HES5+ cells are early radial glial cells (E-RG) which constitute at least about 70-80 % of the isolated population of cells.

16. The isolated population of cells of claim 2, wherein said HES5+ cells are mid radial glial cells (M-RG) which constitute at least about 30 % of the isolated population of cells.

17. The isolated population of cells of claim 2, wherein said HES5+ cells are late radial glial cells (L-RG) which constitute at least about 10-15 % of the isolated population of cells.

18. The isolated population of cells of claim 2, wherein said HES5+ cells are long term neural progenitors (LNP); which constitute at least about 7-10 % of the isolated population of cells.

19. The isolated population of cells of any one of claims 2-18, wherein said HES5+ cells are genetically modified.

20. According to an aspect of some embodiments of the invention, there is provided an isolated population of cells comprising at least 10 % HES5- cells, wherein said HES5- cells are:

(i) non-CNS cells comprising neural crest cells, placodal cells, non-neuroepithelial cells; and CNS cells which exhibit an NEUROD4+/NGN1+/NGN2+/TBR2+/DCX+ expression signature and which form neurons of layers 1 and 6 of the brain cortex;

(ii) neural progenitor cells which belong to the CNS, having a lower proliferative capacity as compared to the HES5+ ERG cells, which form layers 1 5 and 6 of the brain cortex;

(iii) intermediate progenitor cells (INPs) which belong to the CNS, and which are capable of differentiating into the neurons forming layers 4, and 2 of the brain cortex;

(iv) HES5- neurons and astrocytes, wherein said neurons form layers 2, 4 and 3 of the brain cortex; or

(v) neurons, oligodendrocyte and astrocytes, wherein said neurons comprise neurons reaching the olfactory bulb.

21. The method of claim 1 or the isolated population of cells of any one of claims 2-20, wherein the cells are human cells.

22. The method of claim 1 or the isolated population of cells of any one of claims 2-21, derived from a subject having a CNS disease or disorder.

23. The isolated population of cells of any one of claims 2-22, wherein the cells having been subjected to an ex-vivo differentiation protocol.

24. The isolated population of cells of any one of claims 2-23 for use in the treatment of a CNS disease or disorder.

25. The method of claim 1, wherein said successive isolation comprises at least two isolation steps following at least two culturing steps, wherein a first isolation of said at least two isolation steps is effected up to 12 days of a first culturing of said at least two culturing steps, and wherein a second isolation of said at least two isolation steps is effected up to 5 days of a second culturing of said at least two culturing steps.

26. The method of claim 25, wherein said successive isolation comprises at least three isolation steps following at least three culturing steps, wherein a third isolation of said at least three isolation steps is effected up to 21 days of a third culturing of said at least three culturing steps.

27. The method of claim 26, wherein said successive isolation comprises at least four isolation steps following at least four culturing steps, wherein a fourth isolation of said at least four isolation steps is effected up to 45 days of a fourth culturing of said at least four culturing steps.

28. The method of claim 27, wherein said successive isolation comprises at least five isolation steps following at least five culturing steps, wherein a fifth isolation of said at least five isolation steps is effected up to 140 days of a fifth culturing of said at least five culturing steps.

29. The method of claim 25, wherein said first isolation results in a population of cells comprising HES5+ neural epithelial cells.

30. The method of claim 25, wherein said second isolation results in a population of cells comprising HES5+ early radial glial cells.

31. The method of claim 26, wherein said third isolation results in a population of cells comprising HES5+ mid radial glial cells.

32. The method of claim 27, wherein said fourth isolation results in a population of cells comprising HES5+ late radial glial cells.

33. The method of claim 28, wherein said fifth isolation results in a population of cells comprising HES5+ long term neural progenitor cells.

34. The method of any one of claims 1 and 25-33, further comprising qualifying presence of a neural progenitor cell of interest according to at least one marker comprised in an expression signature of said neural progenitor cells, wherein:

(i) an expression signature of HES5+ neuroepithelial cells comprises HES5+/SOX1+/PAX6+/SOX2+/Nestin+/CDC6+/CDX1+/CENPH+/TOP2A+;

(ii) an expression signature of HES5+ early radial glial cells comprises HES5+/ARX+/FEZF2+/NR2E1+;

(iii) an expression signature of HES5+ mid radial glial cells comprises HES5+/POU3F2+/GLAST+/FABP7+;

(iv) an expression signature of HES5+ late radial glial cells comprise HES5+/OLIG1+/PDGFRA+/CUX1+/CUX2+/POU3F2+/S100B+/EGFR+;

(v) an expression signature of HES5+ long term neural progenitor cells comprise HES5+/ANXA2+/LGALS1+/EGFR+/ S100B+.

35. The method of any one of claims 1 and 25-33, further comprising qualifying presence of a neural progenitor cell of interest according to epigenetic analysis functional phenotype and/or morphological phenotype.

36. The method of any one of claims 1 and 25-35, wherein said stem cells are derived from a subject having a CNS disease or disorder.

37. The method of claim 36, wherein said CNS disease or disorder comprises a motor-neuron disease.

38. The method of claim 36, wherein said CNS disease or disorder is characterized by cortex damage.

39. A culture medium for neuroepithelial differentiation comprising noggin, LDN-193189and SB-431542.

40. The culture medium of claim 39, further comprising sonic hedgehog.

41. The isolated population of cells of claim 20, wherein said HES5- cells of (i) are characterized by a higher expression level of at least one gene selected from the group consisting of: LHX1, CNTN2, ST18, EBF3, NFASC, FSTL5, ONECUT2, SLC17A6, EBF1, SLIT1, SYT4, NEFM, NEUROD1, PARM1, CHN2, DNER, HMP19, TFAP2B, DCX, KLHL35, PAPPA, OLFM1, NHLH1, RTN1, GAP43, GFRA1, CHL1, FNDC5, SCN3A, NPTX2, EOMES, CADPS, NHLH2, TMEM163, STMN3, LRRN3, NEFL, ROBO2, INA, PHLDA1, GRIA1, GRIA2, DCLK1, CRABP1, OLIG2, SCG3, TMEM158, FBXL16, FAM123A, SYP, KIF21B, PCDH9, CDKN1C, IGFBPL1, RSPO3, GABRB3, TAGLN3, KCNH2, EPB41L3, EYA2, TMOD2, NCAN, GABBR2, D4S234E, PI15, ANK2, SLC1A2, NRCAM, CBLB, CAMK2N1, ZIC3, PTPRN2, SORBS1, NXPH1, BAALC, CLASP2, DPP6, MAP6, FIGNL2, KIAA0802, KIF1A, TMEM170B, SLC22A23, TMEM178, CTNND2, CADM1, LGALS3, TCEAL7, CD9, GLB1L2, NFIA, YPEL3, TNFRSF19, SPINK5, PNMA2, IRX5, SIN3B, STK38, NR3C1, SOX8, SLC6A8, SYNPR, SGK223, BASP1 and APC as compared to the expression level of said at least one gene in HES5+ neuroepithelial cells.

42. The isolated population of cells of claim 20, wherein said HES5- cells of (ii) are characterized by a higher expression level of at least one gene selected from the group consisting of: GREM1, COL3A1, PCDH8, SEMA3C, BMP4, NID2, TNC, COL1A2, ANKRD1, ANXA1, TMEFF2, PDZRN3, ANXA3, KRT8, LEPREL1, NOX4, LAMB1, FLNC, FST, IMMP2L, S100A4, GDF15, PHACTR2, METTL7A, MAMDC2, DDIT4, BCHE, OCIAD2, TNFRSF10D, BBS9, ELOVL2, TUBA1C, CHST7, RBM47, TFPI, NEBL and LHFP as compared to the expression level of said at least one gene in HES5+ early radial glial cells.

43. The isolated population of cells of claim 20, wherein said HES5- cells of (iii) are characterized by a higher expression level of the ACSS1 gene as compared to the expression level of said gene in HES5+ mid radial glial cells.

44. The isolated population of cells of claim 20, wherein said HES5- cells of (iv) are characterized by a higher expression level of at least one gene selected from the group consisting of: THBS1, KLHL4, A2M, EN2, SLC6A6, ACTA2, ST6GAL1, SLC7A8, GRM3, FAM65B, CALB1, MYLK, TNNT1, PTX3, MFAP2 and HMGA2 as compared to the expression level of said at least one gene in HES5+ late radial glial cells.

45. The isolated population of cells of claim 20, wherein said HES5- cells of (v) are characterized by a higher expression level of at least one gene selected from the group consisting of: FBN2, NELL2, KAL1, PCDHB5, ST8SIA4, DCN, SLC6A1, CADM2, BCL11A, DDB2, ANXA11, PAK1, ID3, IGF2BP1, ANK3, ZEB2 and CREB5 as compared to the expression level of said at least one gene in HES5+ long term neural progenitor cells.

46. The isolated population of cells of claim 2, wherein said HES5+ neuroepithelial cells are characterized by an increased transcriptional epigenetic activity of at least one transcription factor encoded by the gene selected from group consisting of: RFX4, PAX6, TCF4, SOX9, LEF1, POU3F2, SOX8, SOX21, TEAD1, NFATC1, SOX5, TGIF1, MEIS1, TCF4, MEIS2, OTX2, TEF, ZBTB16, MSX1, RFX1, NR4A2, MEIS2, SOX15, STAT5B, SATB1, RBPJ, FOXK1, MYBL2, DMRT3, NFIA, CUX1, TFAP4, MSX1, CDC5L, RFX1, FOXJ2, POU6F1, TEF, RBPJ, PKNOX2, BCL6, PRRX1, STAT1, POU3F1, FOXB1, CTNNB1, PBX1, ZNF143, NFATC1, SOX21, TCF7L1, ARX, SOX15, RXRA, TFAP4, CUX1, OTX2, NR2E1, CUX1, ZNF232, NR2F1, SOX4, MEIS1, PBX1, CUX1, NEUROD1, MSX1, ZNF652, MEF2A, OLIG1, POU6F2, IKZF2, MECOM, STAT1, ESRRA, IRF7, STAT1, MYBL2, BCL6, ELK1, ATF2, SMAD3, ATF4, DLX1, MEF2A, DBP, MAF, MEF2A, TEAD2, SMAD3, SOX15, POU6F1, BARHL2, FOXG1, LHX9, MECOM, ARNTL, MYC, ZNF75A, NFIA, VAX1, GBX2, HOMEZ, FOXO4, FOXO4, FOXB1, ZSCAN16, ELK1, ATF2, CREB1, USF1, ESRRA, ZNF282,

NEUROG2, NFYA, NR4A1, CTF1, ELK1, POU3F2, ELK1, HSF1, E2F3, CUX1, CREB1, ELF2, MYBL2, HMGA2, SRF, ZNF410, JDP2, NR2F1, PAX3, NRF1, SMAD4, ZNF85, ZNF628, NFATC1, CREB1, E4F1, NR2F6, NHLH1, IRF2, PBX1, FOXJ3, RORA, IRF7, NR6A1, LHX2, PAX3, NR2E1, POU3F1, ZFP42, E2F4, ETV5, ELF3, USF1, ATF6, TFAP2B, CUX1, IRF3, RXRA, PEBP1, LHX2, GZF1, MEF2A, MEF2A, IRF9, MGA, VBP1, GMEB2, YY1, ELF1, POU3F3, GTF2IRD1, IRF3, SRF, XBP1, ESRRA, HEY1, NFKB2, IRF2, EOMES, FOXB1, NR2F1, NR4A2, STAT3, SP1, RARA, CREB1, NR2F1, FOXJ3, HSF1, MYBL2, SRF, ETV2, MECP2, E2F1, FOXJ2, JUN, SCRT2, DLX1, E2F1, E2F1 and ATF2as compared to the transcriptional epigenetic activity of said at least one transcription factor in human ESC cells.

47. The isolated population of cells of claim 2, wherein said HES5+ early radial glial cells are characterized by an increased transcriptional epigenetic activity of at least one transcription factor encoded by the gene selected from group consisting of: NFIA, MZF1, SMAD4, CTF1, SMAD3, RFX4, MAF, TCF12, NFYC, ZNF263, MECP2, ZFP42, ZIC1, YY1, ETS2, NR2C2, SREBF2, SREBF1, MEIS2, NR4A1, REST, SF1, ZBTB7A, STAT3, E2F1, NR1H2, NFKB1, NR2F6, GLIS3, MAZ, STAT1, TGIF1, SOX9, HES1, THRA, GLIS3, MEIS1, ESRRA, ZBTB7A, NFYA, MECP2, PKNOX2, EP300, TFAP2B, NR2F1, MZF1, ESRRA, TFCP2, NR2F1, ESRRA, TERF1, KLF3, XBP1, RORA, PBX1, MYC, SNAI2, TEAD2, CENPB, PEBP1, HINFP, SREBF1, YY1, E2F1, HSF2, CNOT3, MEIS2, MEF2A, RXRA, CREB3L1, MYBL2, ZNF524, TFAP4, RFX1, NFYA, ZBTB33, RREB1, NR6A1, HES5, TFAP2B, HIF1A, INSM1, ZNF524, ELF1, SMAD4, STAT3, TFAP2B, USF1, GLI3, GLI2, ETV2, E2F1, CREB3, NFKB1, MYC, SRF, TFCP2, ATF3, ELK1, TP53, E2F4, ELF2, DEAF1, RXRA, ZNF423, FOXG1, ZNF628, NEUROD1, ARNTL, STAT1, PAX3, NR6A1, ESRRA, TCF4, YY1, PBX1, VAX1, NRF1, ETV1, FOXB1, ZNF85, CREB1, PRDM4, MYBL2, NHLH1, CREB1, KLF13, TRIM26, ZNF148, E4F1, USF1, JDP2, OLIG1, GZF1, ATF2, CREB1, FOXO4, SMAD3, CLOCK, ZNF282, ZSCAN16, CUX1, NFKB1, RFX1, ETV5, MEF2A, GTF2IRD1, TFAP2B, NHLH1, ELF1, PAX3, MNT, ZNF740, CREB1, ESRRA, EOMES, EGR2, NR4A2, BCL6, ZNF143, MAFF, CUX1, NR2F1, TFAP2B, NEUROG2, TCF3, ATF3, PLAG1, ESRRA, E2F1, KLF13, MSX1, SCRT2,

GABPB1, SOX9, NFKB2, EOMES, SIX3, MECOM, BCL6, POU3F3, HINFP, NFATC1, TBR1, MYBL2, USF1, SP1, ATF2, LHX2, MYBL2, POU3F3, HEY1, MEF2A, MGA, SOX9, E2F7, IRF3, SRF, POU3F2, SREBF2, NFKB1, MAF, ZNF784, NR1H2, MSX1, RORA, LMO2, HINFP, ZNF23, FOXO4, NR4A2, POU6F2, ID4, RARA, KLF13, LEF1, PATZ1, ELK1, DLX1, E2F4, E2F4, ELK1, MAZ, ATF6, TFAP4, RARA, PAX6, EGR2, HOMEZ, ELK1, E2F1, GMEB2, ATF4, ATF2, MAX, MGA, NFKB2, ESRRA, E2F4, STAT3, TCF3, NR2E1, ELF3, SOX4, USF1 and ZNF652 as compared to the transcriptional epigenetic activity of said at least one transcription factor in HES5+ neuroepithelial cells.

48. The isolated population of cells of claim 2, wherein said HES5+ mid radial glial cells are characterized by an increased transcriptional epigenetic activity of at least one transcription factor encoded by the gene selected from group consisting of: RFX4, NFIA, NFATC1, CTF1, NEUROD1, RFX1, TGIF1, ETS2, TFAP4, TCF12, TFAP4, SMAD3, NFIA, SOX9, RFX1, TCF4, MZF1, STAT1, MECP2, MEIS2, ZBTB33, NFYA, ELF1, MYBL2, LEF1, NFYC, MAFF, ZNF263, YY1, POU3F3, TGIF1, STAT3, SMAD4, NR6A1, TGIF1, MEIS1, ZNF628, ZFP42, FOXK1, PRDM4, STAT1, MAF, SCRT2, CREB1, GZF1, CREB1, VAX1, MECP2, NHLH1, ETV1, SOX9, PEBP1, SMAD4, XBP1, USF1, POU3F2, CREB3, EP300, PBX1, STAT3, TFCP2, POU2F2, IRF3, FOXB1, MSX1, POU3F3, ELK1, DBP, CUX1, MEF2A, POU6F2, ARNTL, ZSCAN16, MEF2A, ETV4, OLIG1, HOMEZ, DLX1, PRRX1, MSX1, MYC, FOXO4, MEF2A, MZF1, ATF2, GMEB2, NFYA, ESRRA, SOX9, PBX1, POU3F2, MECOM, SMAD3, MAZ, ELK1, BCL6, ELF3, PAX3, ELF1, SF1, BCL6, EMX2, STAT1, E2F1, ELF2, CLOCK, ELK1, STAT3, ATF5, THRA, SOX9, SRF, TCF4, ATF3, CTNNB1, USF1, FOXJ3, USF1, ZNF282, NEUROG2, ESRRA, REST, E2F3, ZBTB7A, MYBL2, HSF2, MAX, ZNF143, MYBL2, SRF, FOXO4, NR4A2, CUX1, E2F4, MSX1, EOMES, MAF, MNT, POU6F1, NFATC1, STAT3, CREB3, SOX9, ZNF85, CREB3L1, TCF3, ELK1, IRF2, YY1, SOX15, PAX6, E4F1, MEF2A, ATF2, NFE2L1, NFATC1, ATF2, YY1, SRF, ARX, ETV2, HINFP, MAZ, NR4A1, INSM1, ZNF652, USF1, NFKB1, GBX2, POU3F4, IRF9, POU6F1, LHX2, NR4A2, POU6F2, FOXJ2, CEBPG, VBP1, TERF1, ESRRA, PAX6, ZIC3, IRF2, MAF, SOX21, CREB1, CREB3, IRF7, POU3F3, NFKB2, ATF4, MYBL2, SREBF2, SOX9, YY1, RBPJ, FOXJ3, HSF1,

HMGA2, CUX1, POU3F2, EOMES, ZNF423, ESRRA and KLF13 as compared to the transcriptional epigenetic activity of said at least one transcription factor in HES5+ early radial glial cells.

FIG. 1A

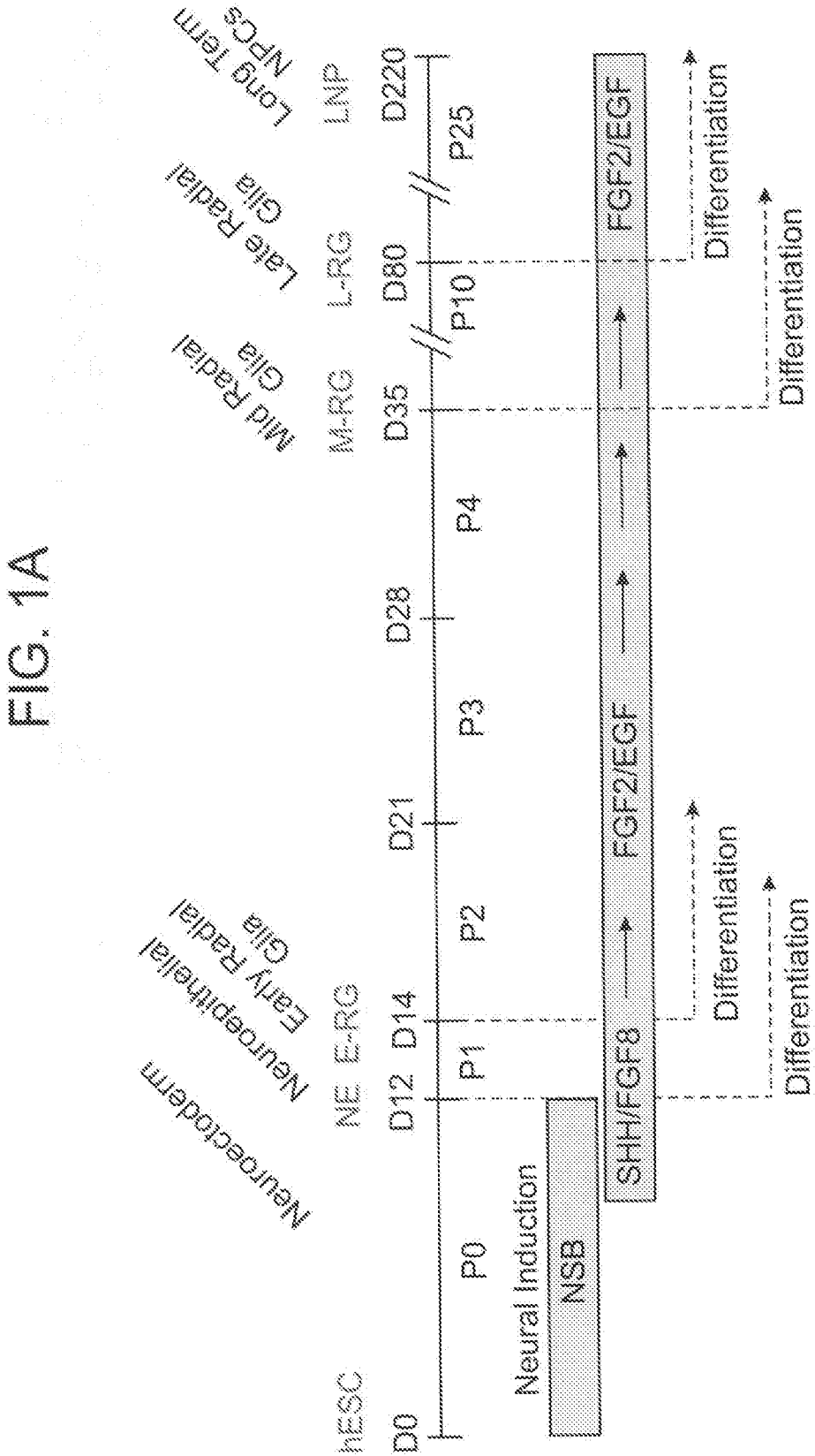

hESC
NE
E-RG
M-RG
L-RG
LNP
D0
D12
D14
D21
D28
D35
D80
D220
P0
P1
P2
P3
P4
P10
P25
Neural Induction
NSB
SHH/FGF8
FGF2/EGF
FGF2/EGF
Differentiation
Differentiation
Differentiation
Neuroectoderm
Neuroepithelial
Early Radial
Glia
Mid Radial
Glia
Late Radial
Glia
Long Term
NPCs

FIG. 1B

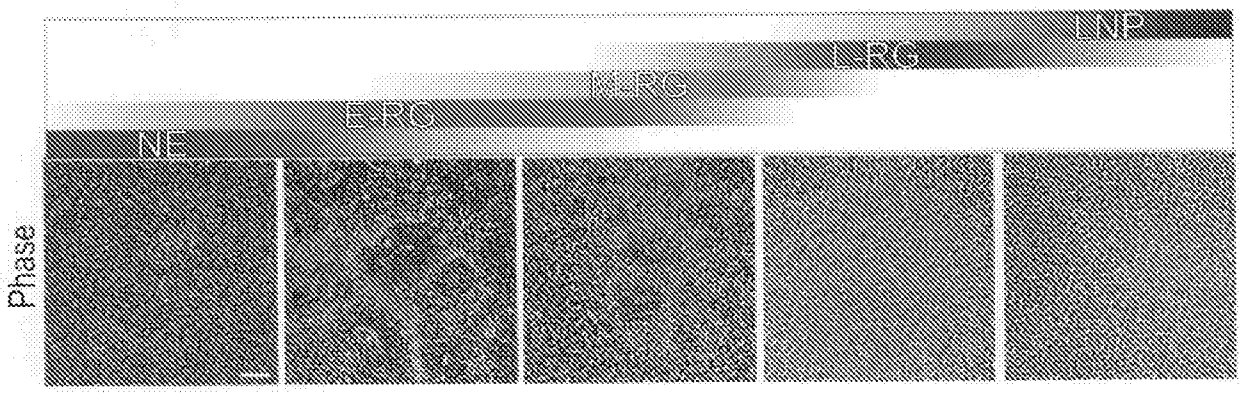


FIG. 1C

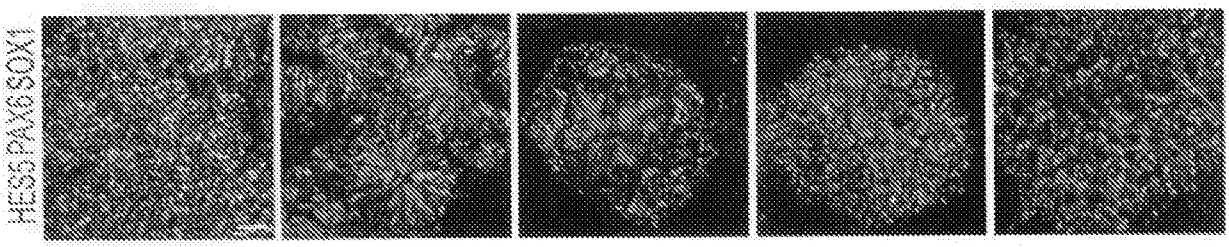


FIG. 1D

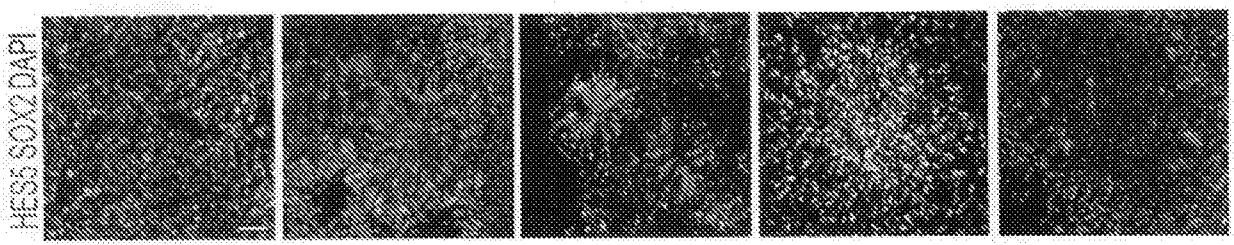


FIG. 1E

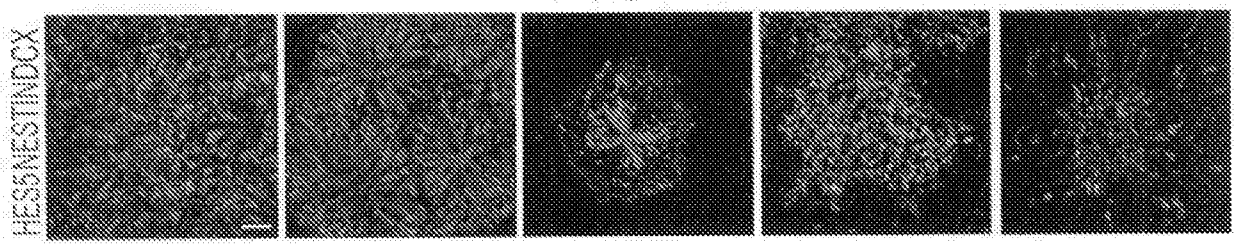


FIG. 2A

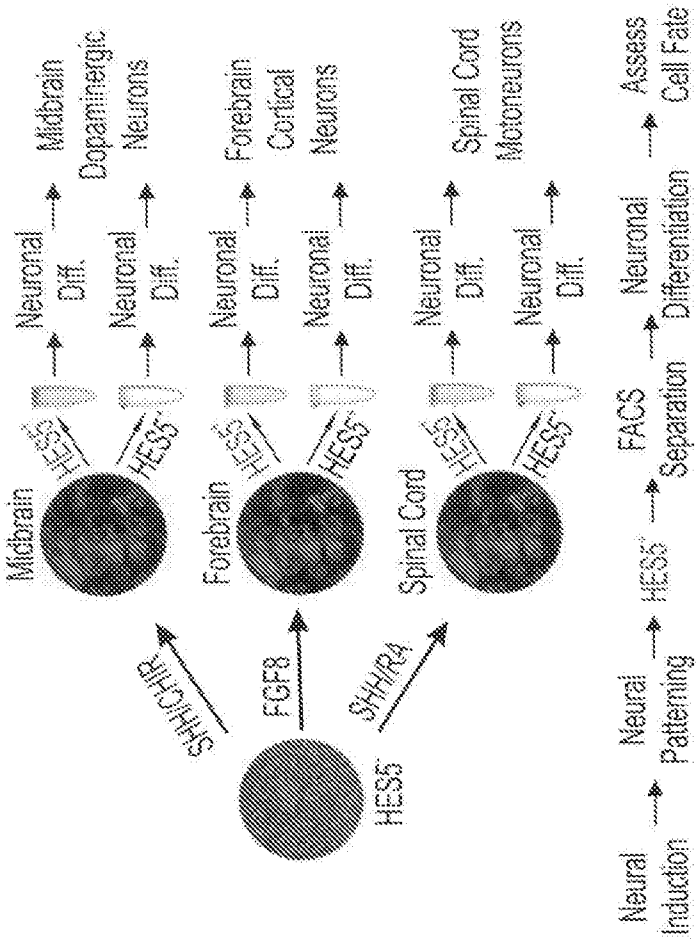

Neural Induction
Neural Patterning
HES5
HES5
FACS Separation
Neuronal Differentiation
Assess Cell Fate
SHH/C25II
FGF8
SHH/R4
Midbrain
Forebrain
Spinal Cord
HES5
HES5
Neuronal Diff.
Midbrain Dopaminergic Neurons
Forebrain Cortical Neurons
Spinal Cord Motoneurons

FIG. 2B

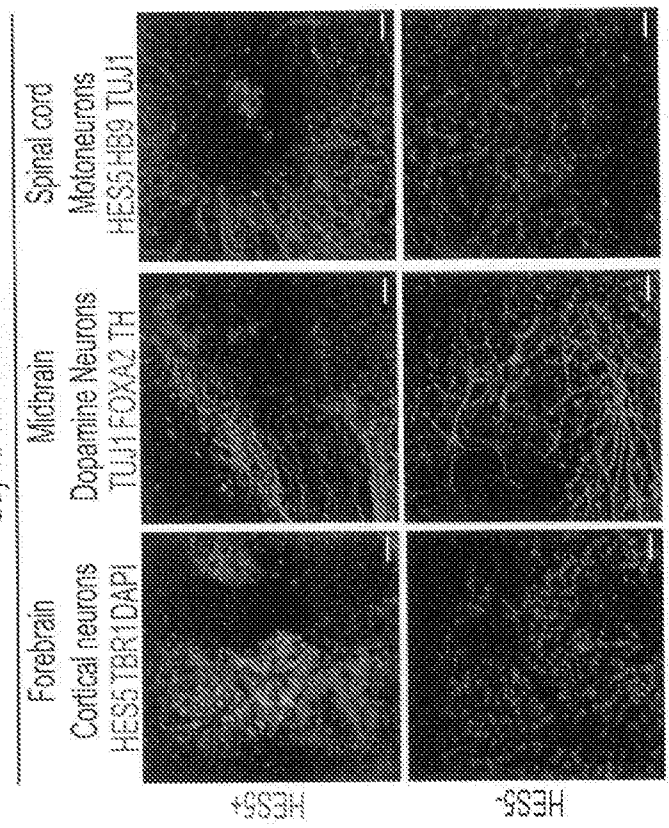

Day 12 derived neurons
Forebrain
Cortical neurons
HES5 TBR1 DAPI
Midbrain
Dopamine Neurons
TUJ1 FOXA2 TH
Spinal cord
Motoneurons
HES5 HB9 TUJ1
HES5
HES5

FIG. 2C

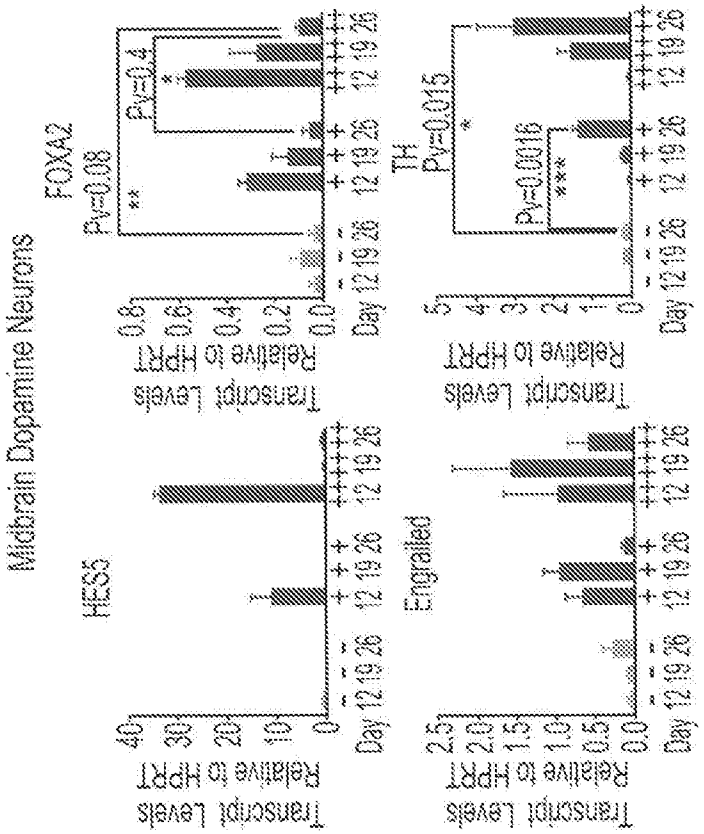


FIG. 2D

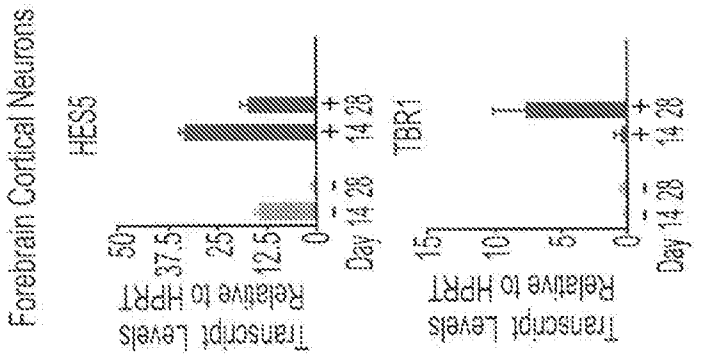


FIG. 2E

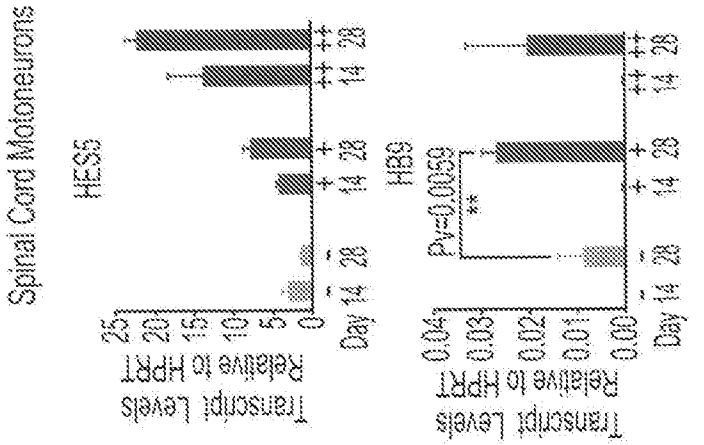


FIG. 3A

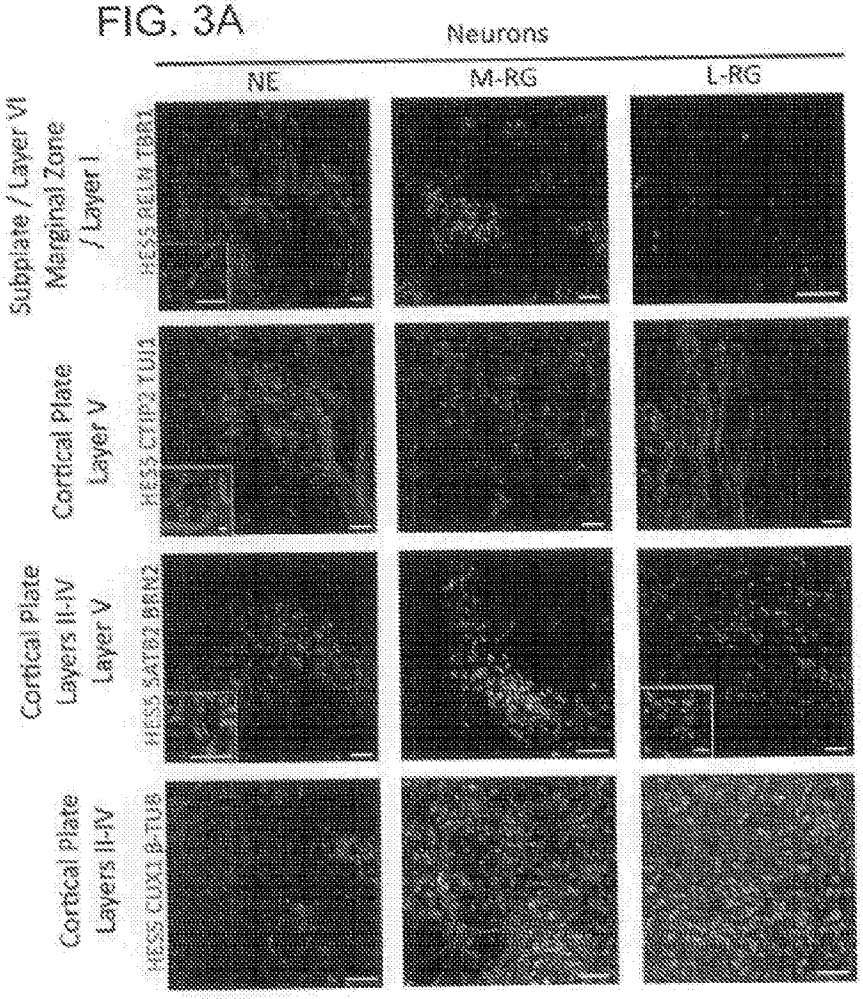


FIG. 3E

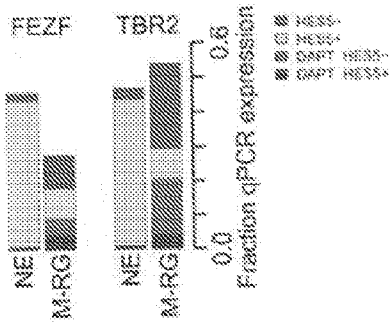
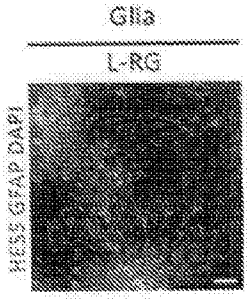


FIG. 3F



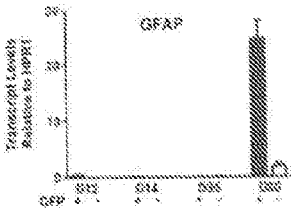


FIG. 3G

FIG. 3B

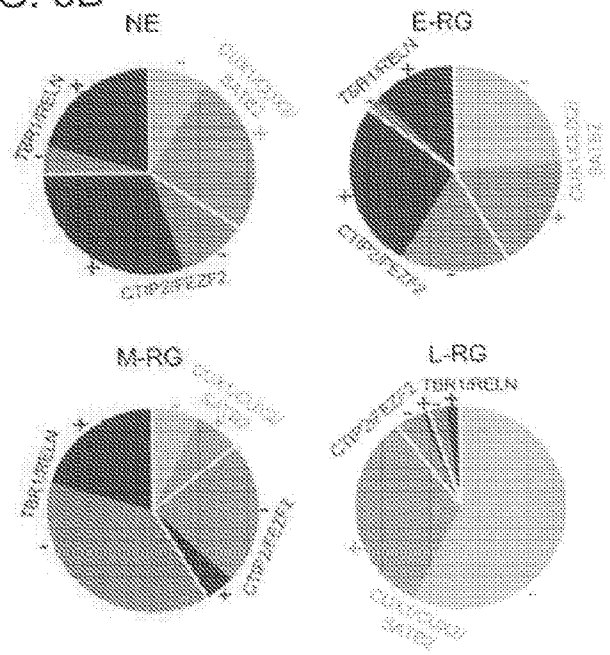


FIG. 3C

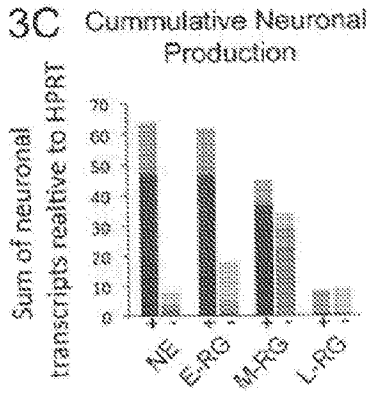


FIG. 3D

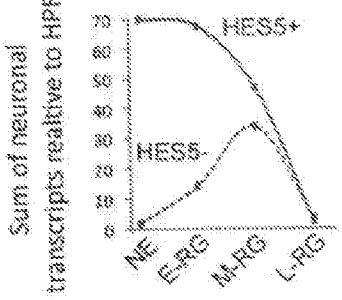


FIG. 4A

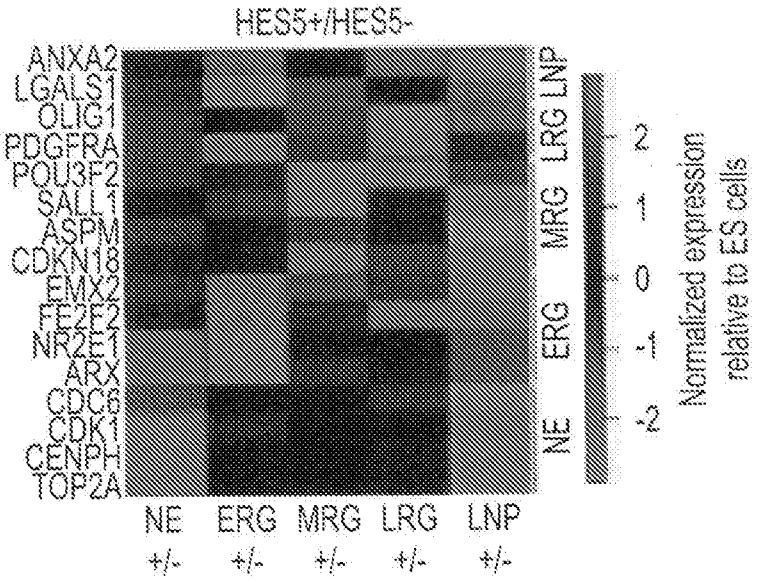


FIG. 4B

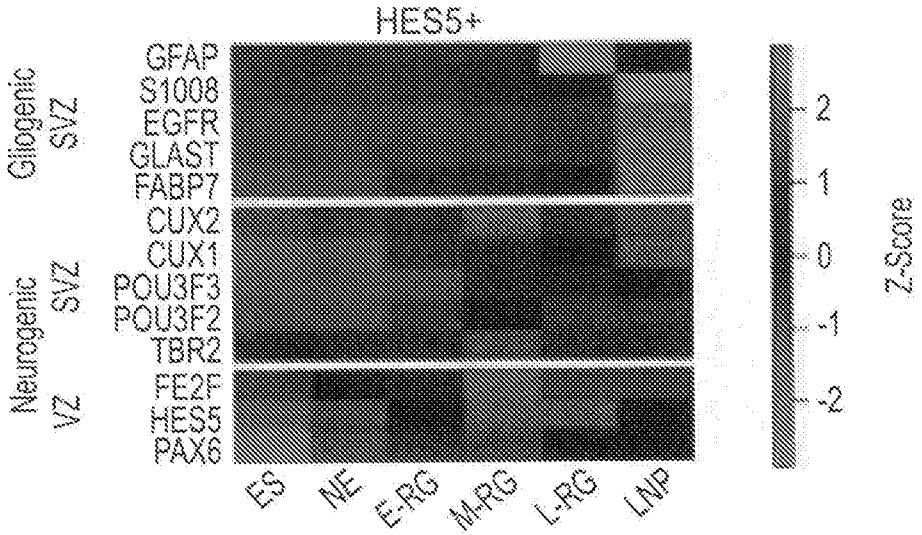


FIG. 4C

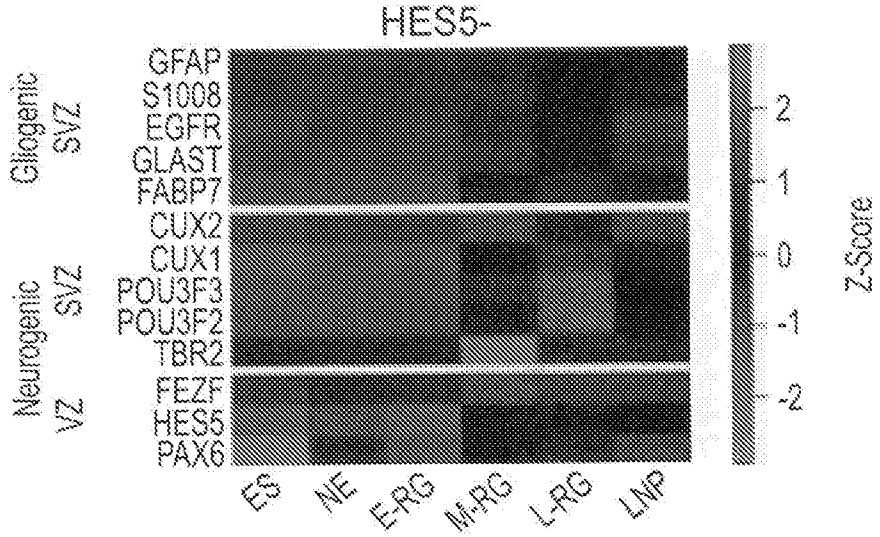


FIG. 4D

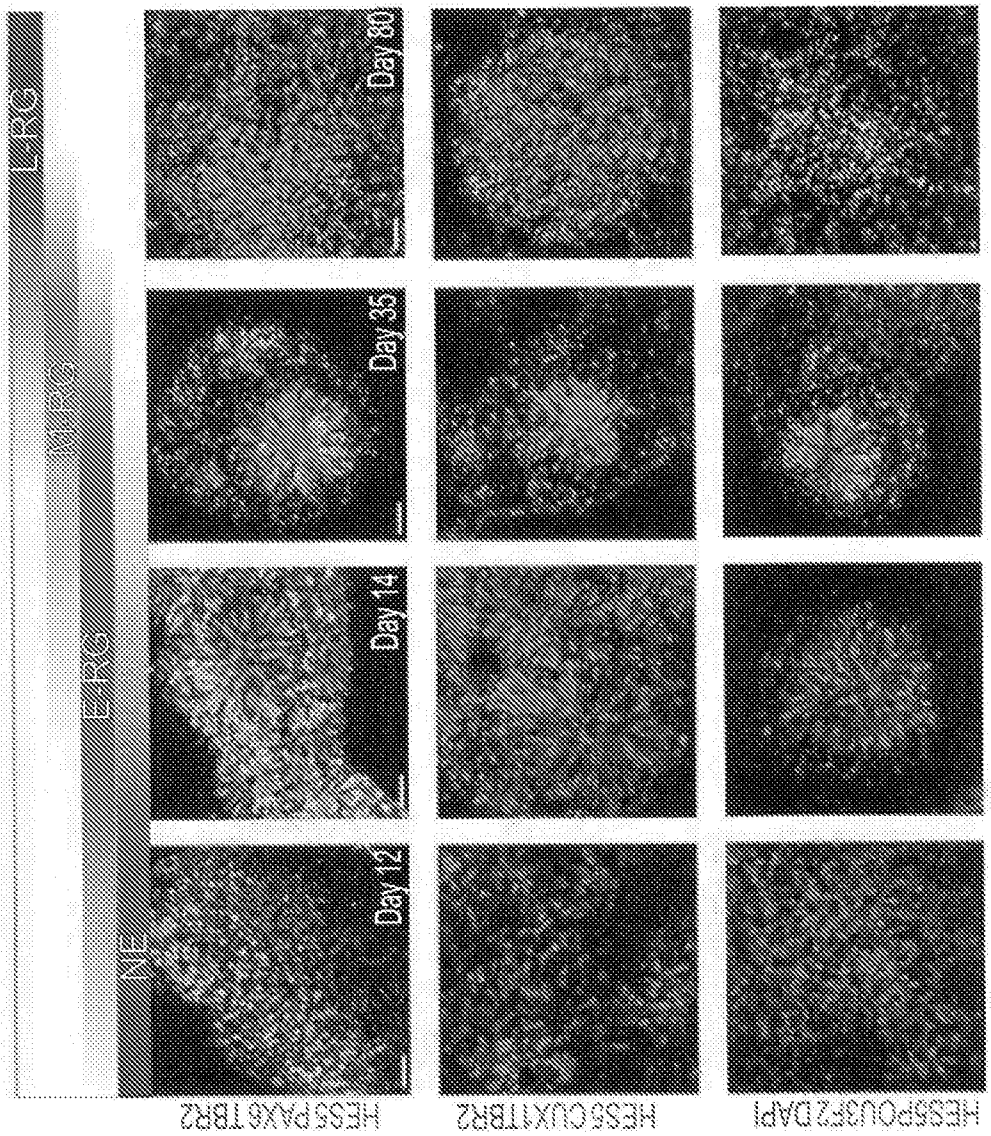

NE
EPC
LPC
Day 12
Day 14
Day 35
Day 80
HES5PAX6TBR2
HES5CUX1TBR2
HES5POU3F2DAPI

FIG. 4E

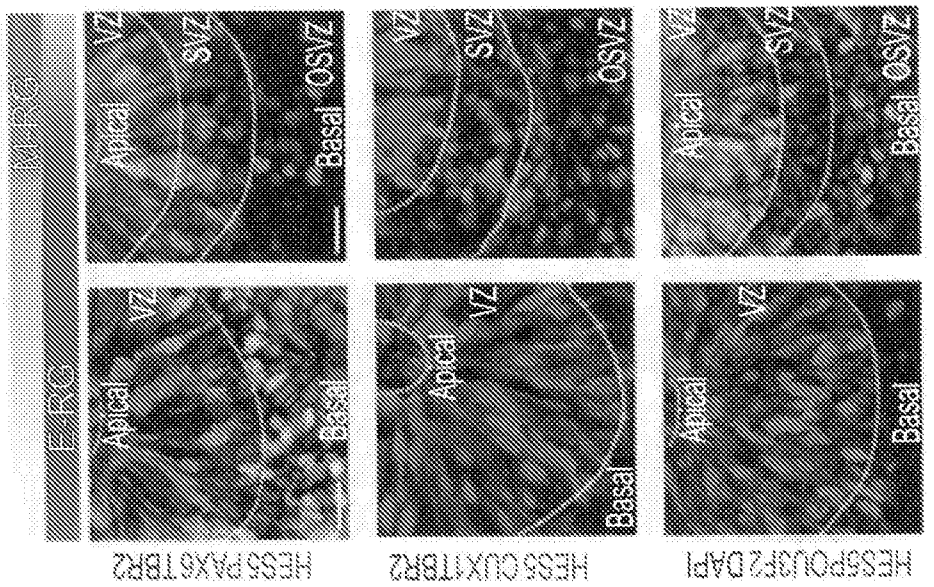

EPC
LPC
Apical
Basal
VZ
SVZ
OSVZ
HES5PAX6TBR2
HES5CUX1TBR2
HES5POU3F2DAPI

FIG. 5B

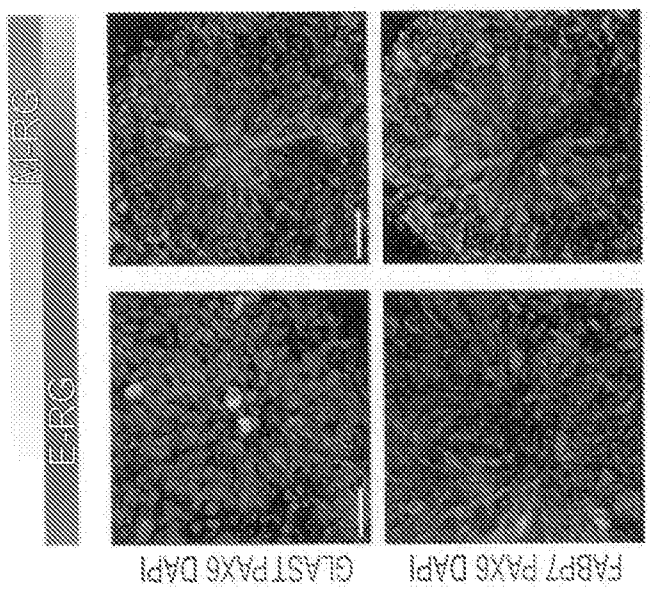

MNR2
E-RG
GLAST PAX6 DAPI
FABP7 PAX6 DAPI

FIG. 5A

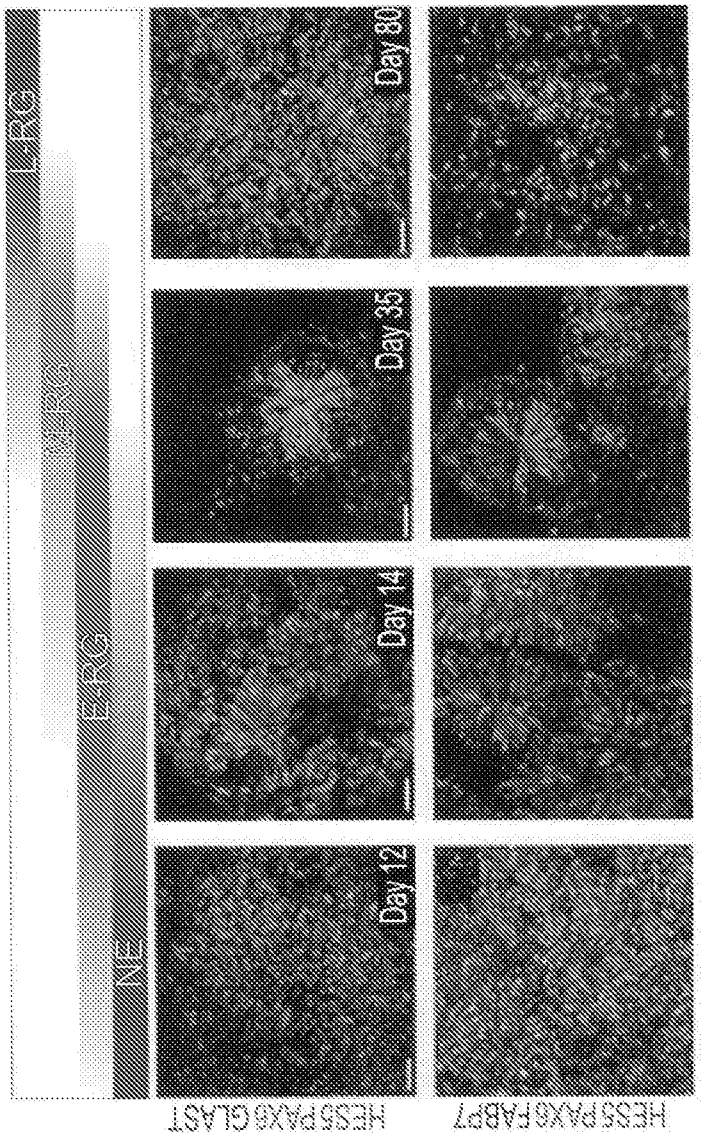

L-RG
MNR2
E-RG
NE
Day 80
Day 35
Day 14
Day 12
HES5 PAX6 GLAST
HES5 PAX6 FABP7

FIG. 5C

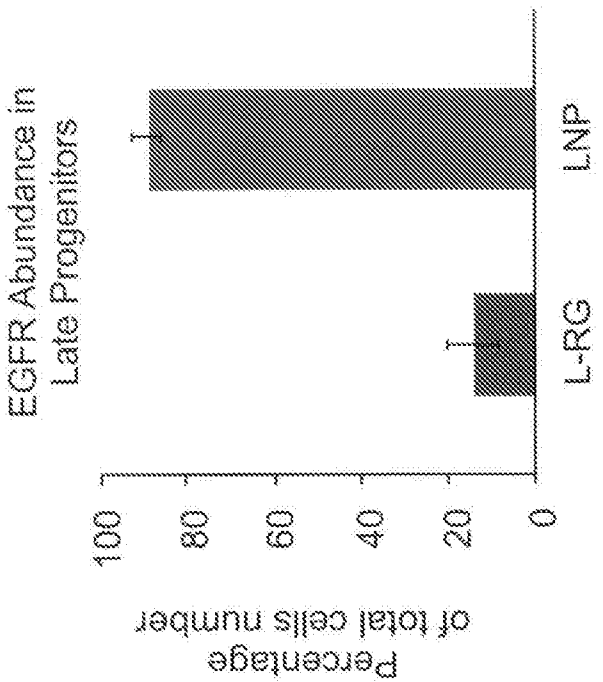

EGFR Abundance in
Late Progenitors
Percentage
of total cells number
100
80
60
40
20
0
L-RG
LNP

FIG. 5D

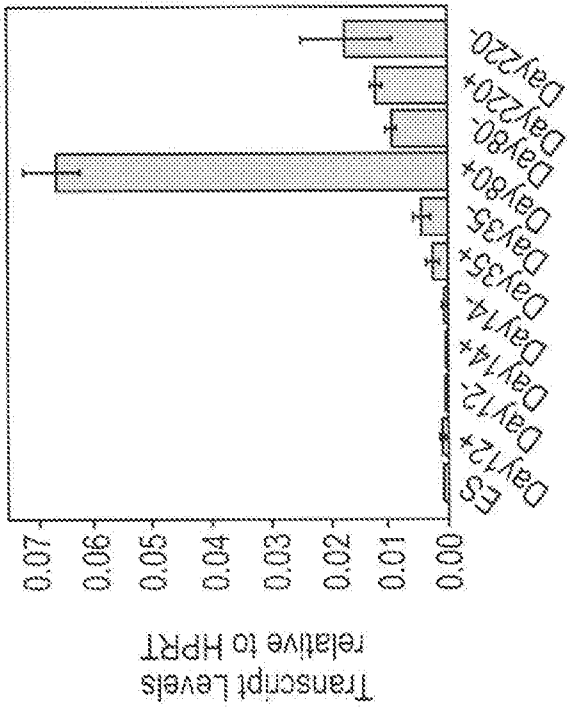

GFAP
Transcript Levels
relative to HPRT
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0.00
ES
Day12+
Day12-
Day14+
Day14-
Day35+
Day80+
Day80-
Day220+
Day220-

FIG. 6A

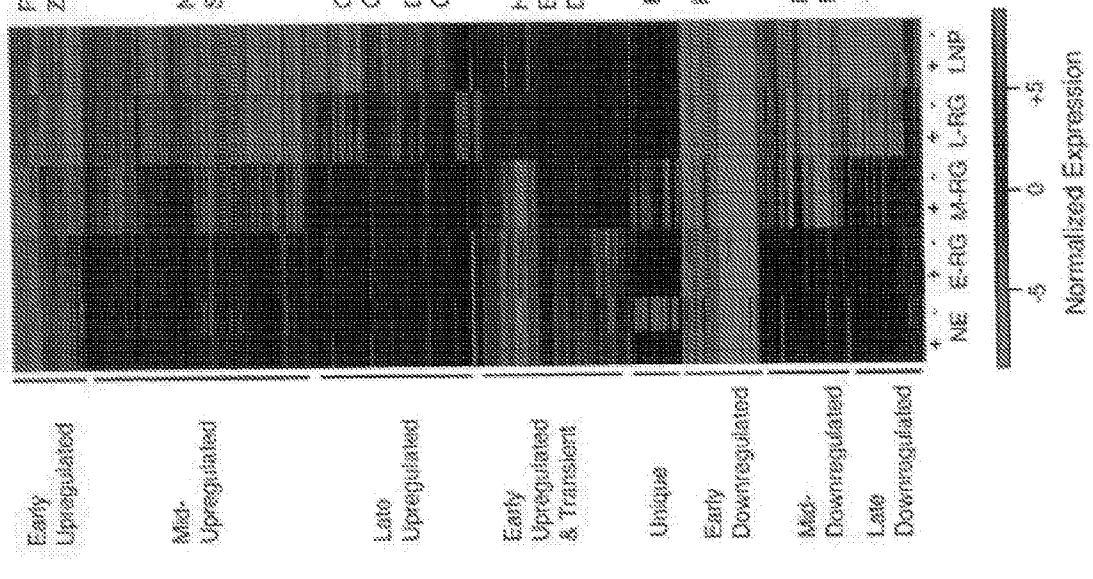

Early Upregulated
Mid-Upregulated
Late Upregulated
Early Upregulated & Transient
Unique
Early Downregulated
Mid-Downregulated
Late Downregulated
FOXG1, PAX6
ZIC1, SP8, ARX
NFIA, NFIB
SLITRK3
GABBR2, GRIA4
GRM3
DLX1, DLX2
OLIG1, OLIG2, LGALS1
HES5, NR2E1, DLL1
EMX2, MEIS2, LGR5
DACH1, PLAGL1, LGI1
EOMES, RSPO2
POU5F1
LIN28A, HMGA2
F11R
OTX2
NE E-RG M-RG L-RG LNP
-5 0 +5
Normalized Expression

FIG. 6B

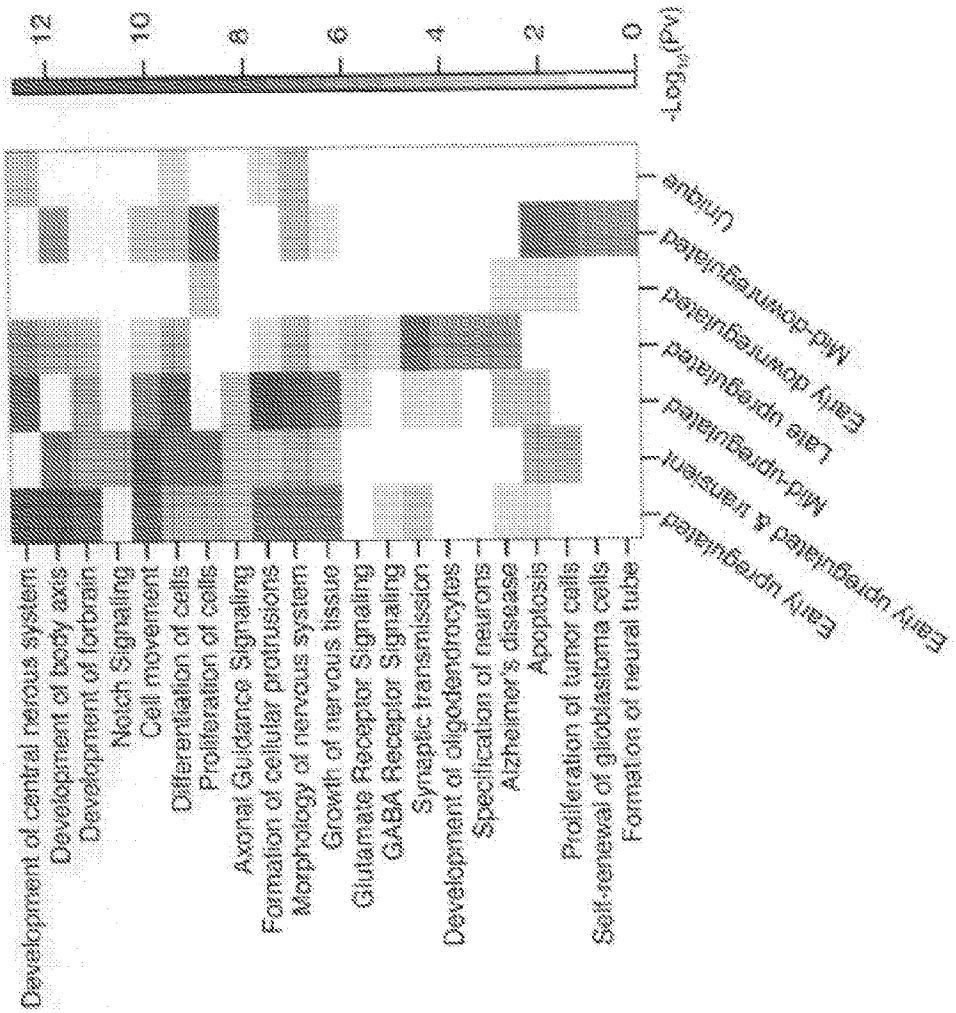
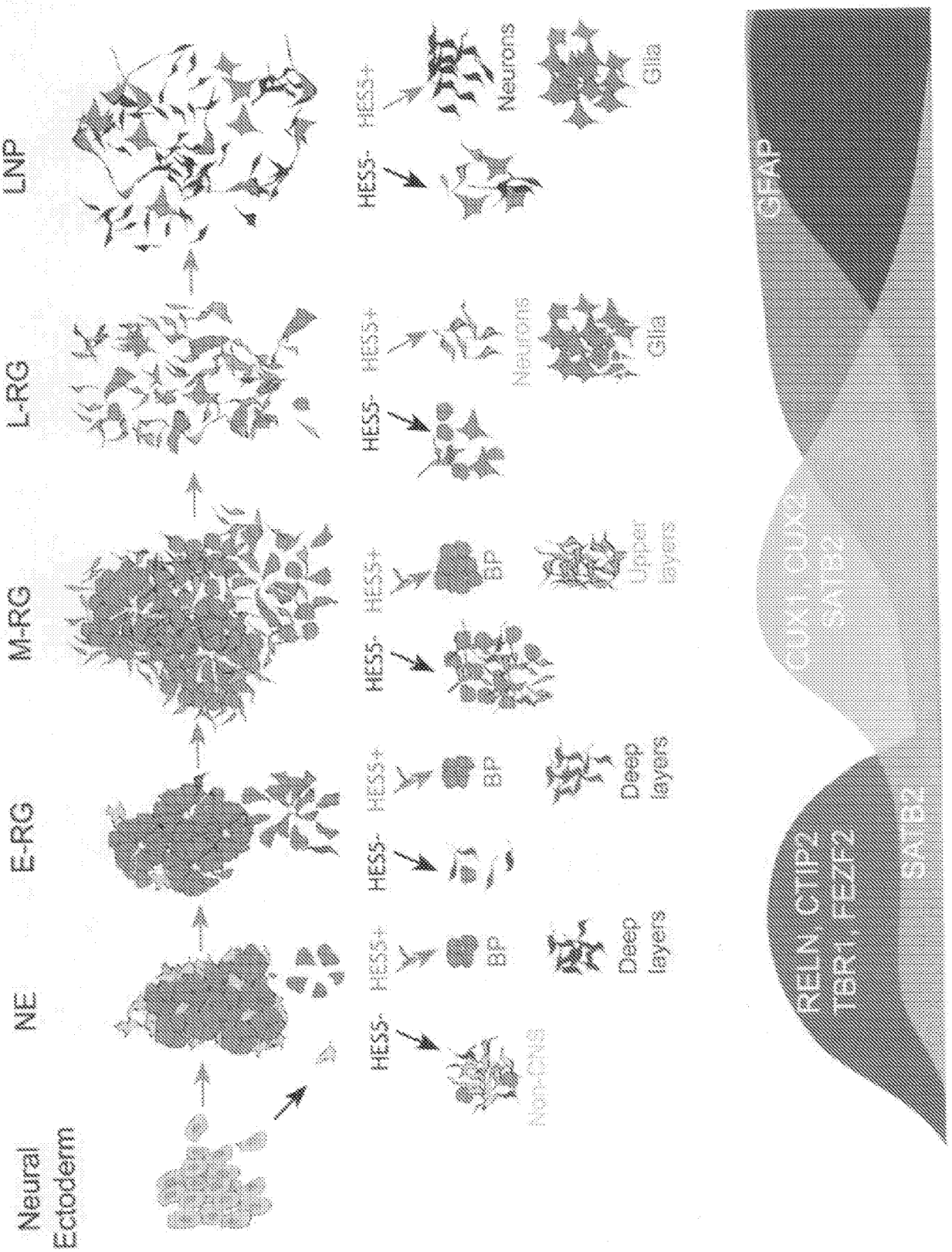

Development of central nervous system
Development of body axis
Development of forbrain
Notch Signaling
Cell movement
Differentiation of cells
Proliferation of cells
Axonal Guidance Signaling
Formation of cellular protrusions
Morphology of nervous system
Growth of nervous tissue
Glutamate Receptor Signaling
GABA Receptor Signaling
Synaptic transmission
Development of oligodendrocytes
Specification of neurons
Alzheimer's disease
Apoptosis
Proliferation of tumor cells
Self-renewal of glioblastoma cells
Formation of neural tube
Early upregulated
Early upregulated & transient
Mid-upregulated
Late upregulated
Early downregulated
Mid-downregulated
Unique
0
2
4
6
8
10
12
-Log10(Pv)

FIG. 7


Neural Ectoderm
NE
E-RG
M-RG
L-RG
LNP
HES5-
HES5+
BP
Deep layers
Upper layers
Neurons
Glia
RELN, CTIP2, TBR1, FZD3
SATB2
GFAP

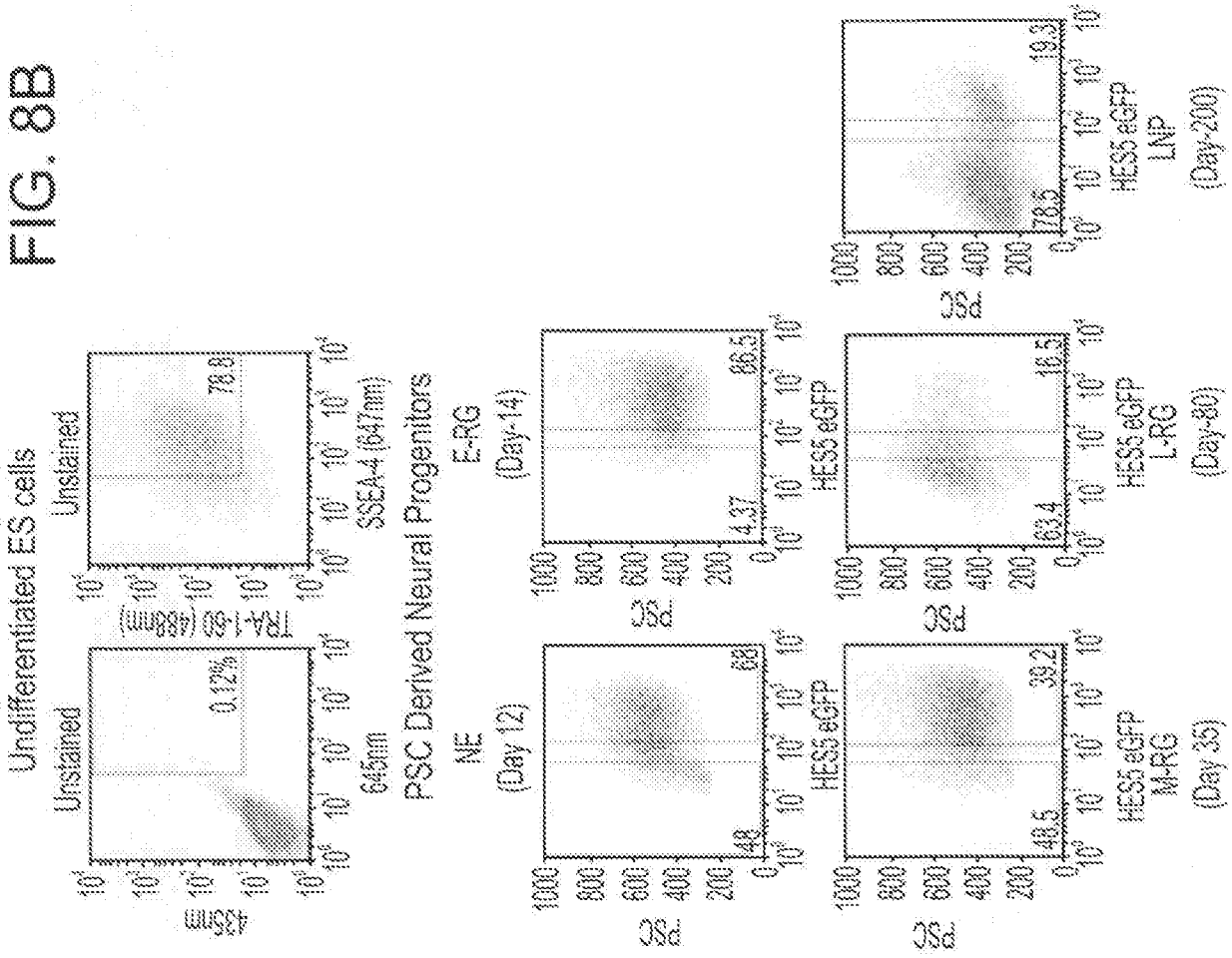


FIG. 8B
Undifferentiated ES cells
Unstained
435nm
645nm
0.12%
Unstained
TRA-1-60 (488nm)
78.8
SSEA-4 (647nm)
PSC Derived Neural Progenitors
NE
(Day 12)
E-RG
(Day-14)
L-RG
(Day-80)
M-RG
(Day 35)
LNP
(Day-200)
HES5 eGFP
PSC
48
68
4.37
66.5
63.4
16.5
49.5
39.2
78.5
19.3

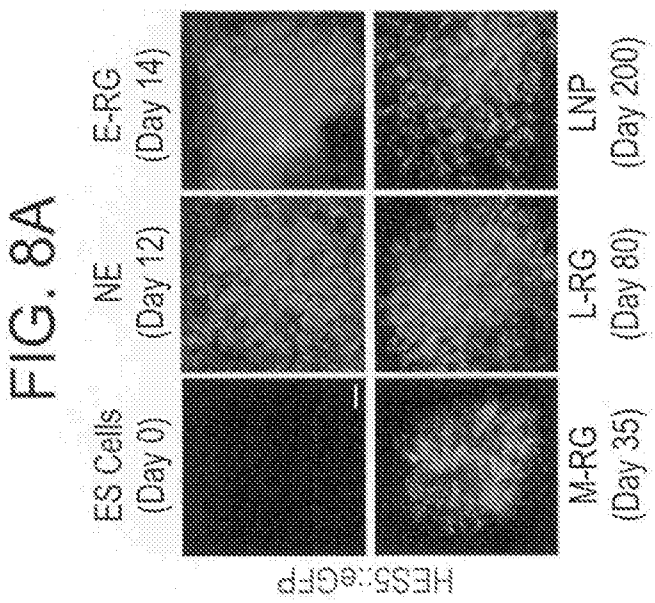


FIG. 8A
ES Cells
(Day 0)
NE
(Day 12)
E-RG
(Day 14)
M-RG
(Day 35)
L-RG
(Day 80)
LNP
(Day 200)
HES5.1eGFP

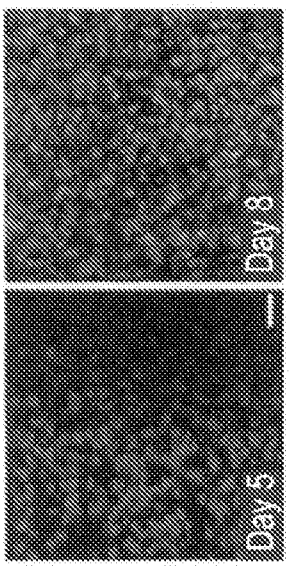


FIG. 8C
Day 5
Day 8
HES5PAX6SOX1

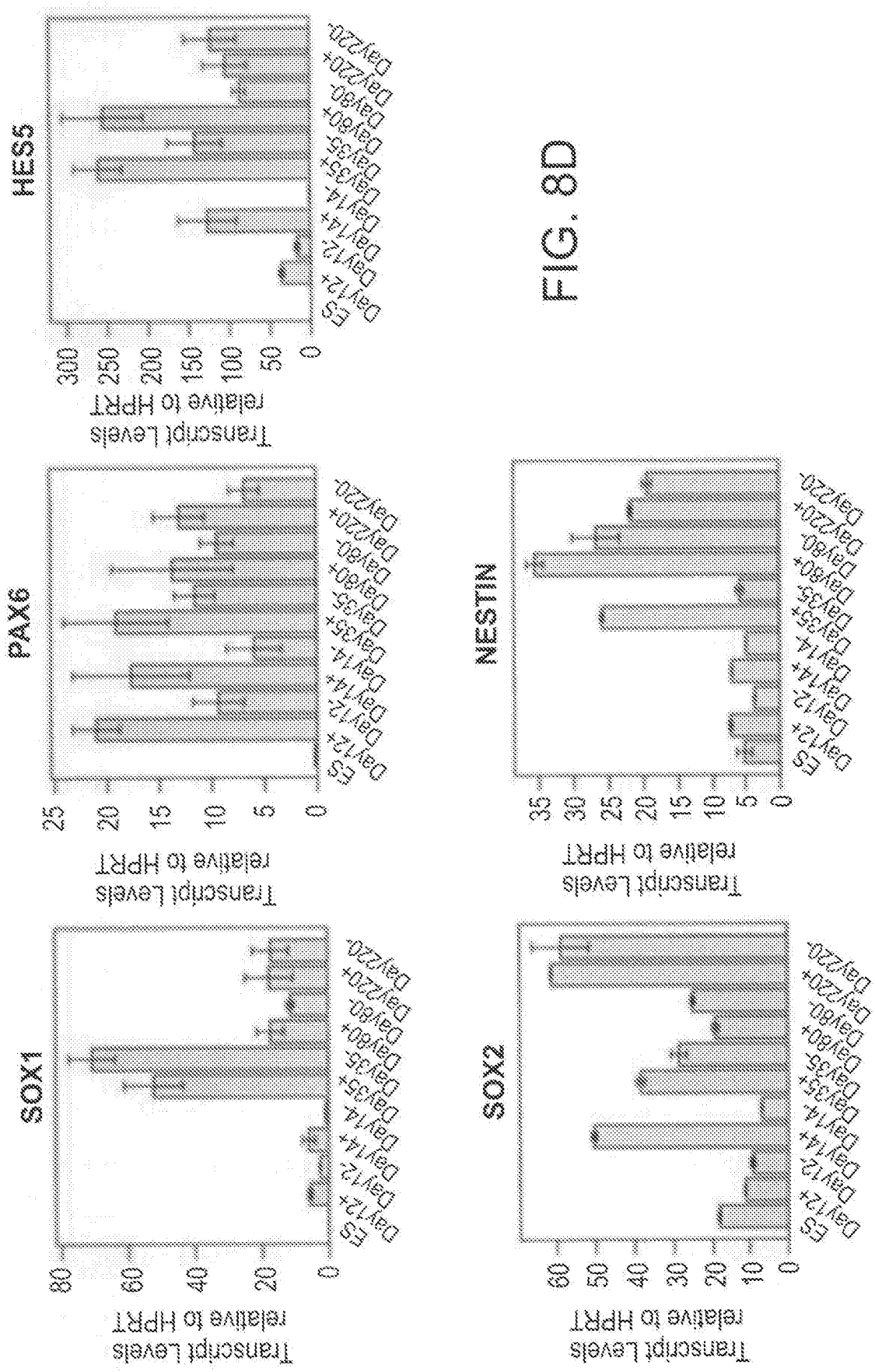

SOX1
Transcript Levels relative to HPRT
0
20
40
60
80
ES
Day12+
Day12-
Day14+
Day35+
Day80+
Day220+
PAX6
Transcript Levels relative to HPRT
0
5
10
15
20
25
NESTIN
Transcript Levels relative to HPRT
0
5
10
15
20
25
30
35
SOX2
Transcript Levels relative to HPRT
0
10
20
30
40
50
60
HES5
Transcript Levels relative to HPRT
0
50
100
150
200
250
300

FIG. 8D

FIG. 8E

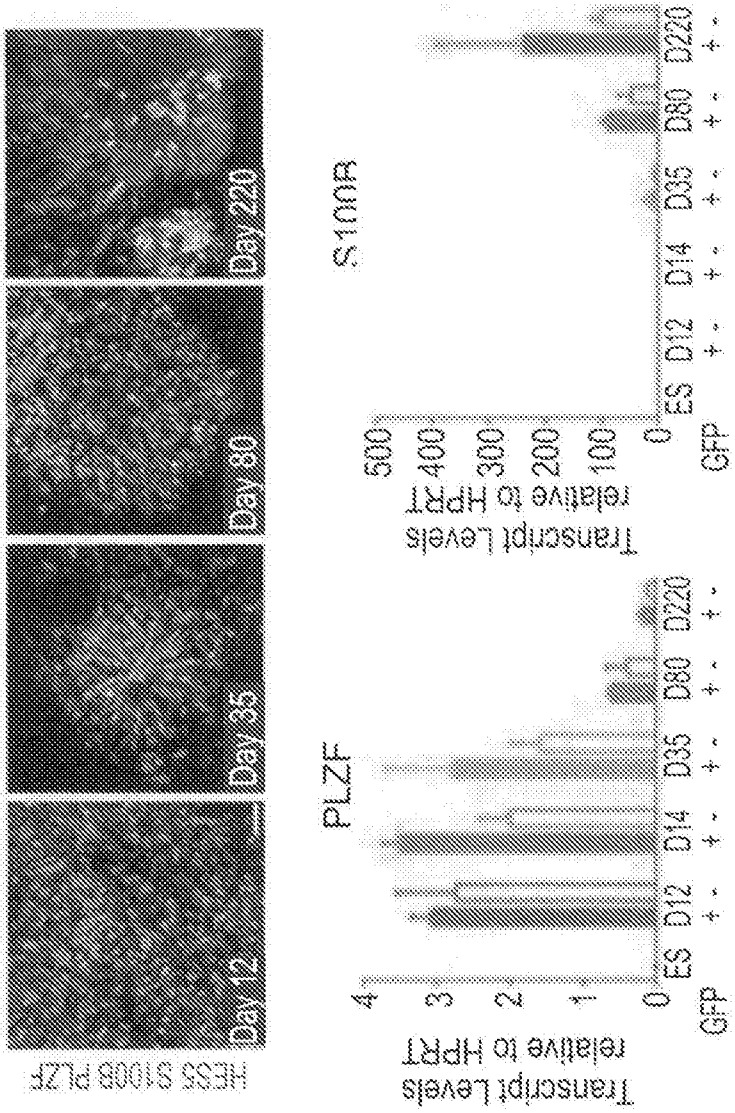

HES5 S100B PLZF
Day 12
Day 35
Day 80
Day 220
PLZF
S100B
EGFR
Transcript Levels relative to HPRT
GFP
ES D12 D14 D35 D80 D220

FIG. 8F

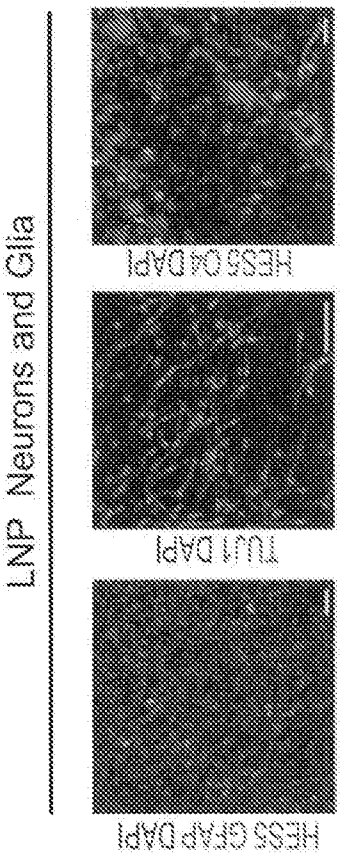

LNP Neurons and Glia
HES5 GFAP DAPI
TUJ1 DAPI
HES5 OLIG DAPI

FIG. 9A

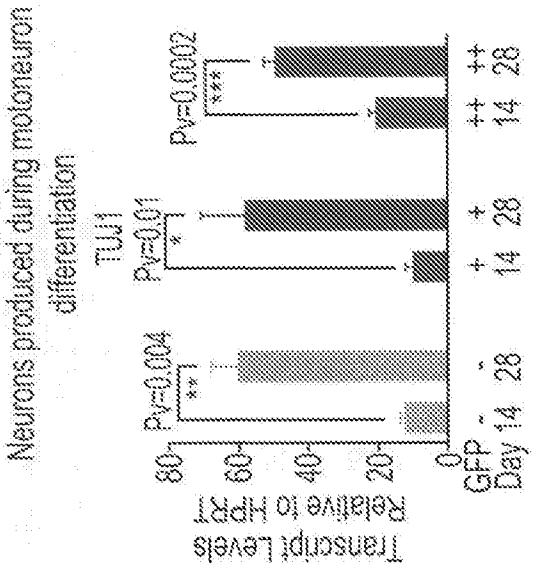


FIG. 9B

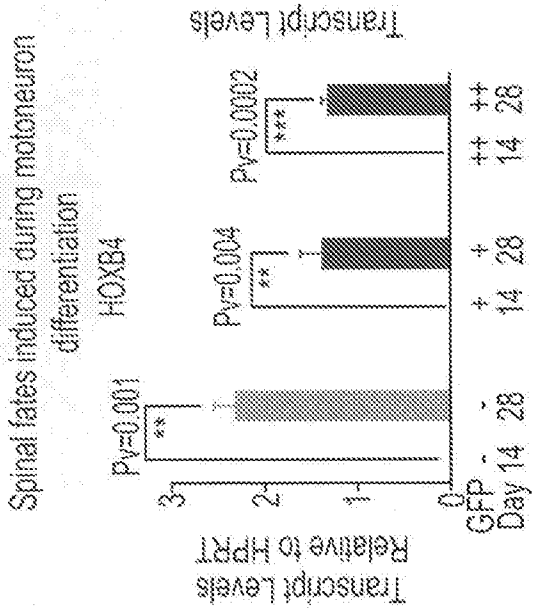


FIG. 9C

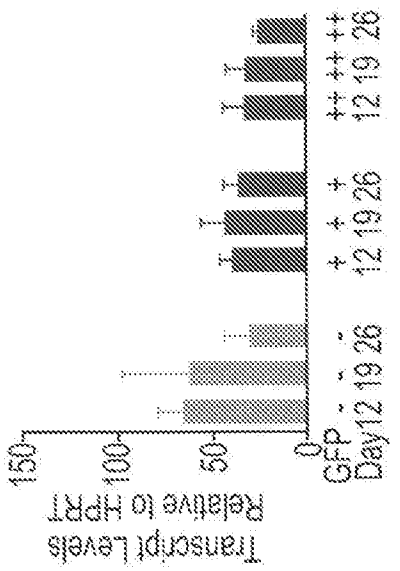


FIG. 9D

FIG. 9E

FIG. 9F

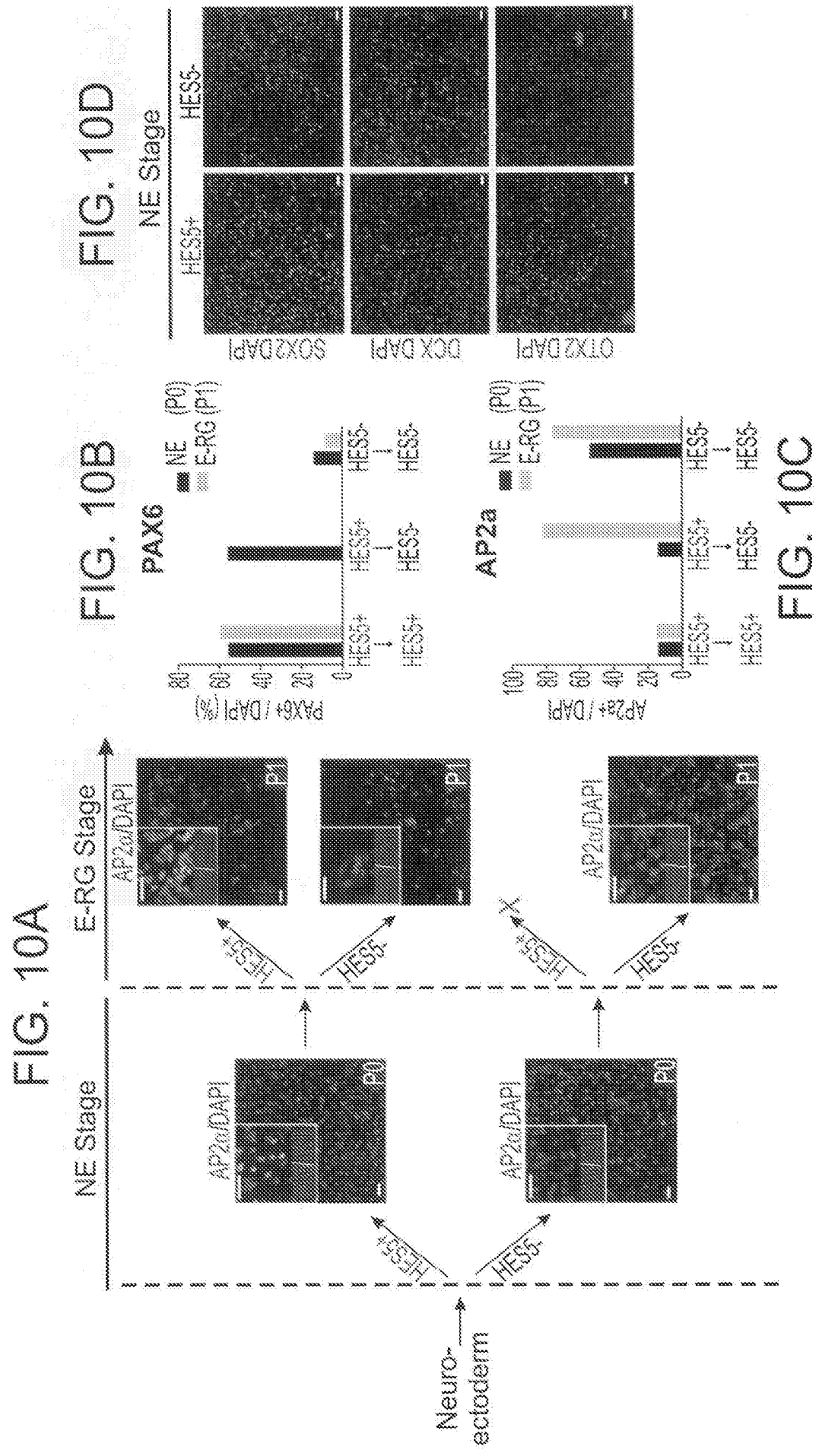


FIG. 10A
NE Stage
E-RG Stage
Neuro-ectoderm
AP2α/DAPI
P0
P1
HES5+
HES5-
FIG. 10B
PAX6
NE (P0)
E-RG (P1)
PAX6+ / DAPI (%)
HES5+ → HES5+
HES5+ → HES5-
HES5- → HES5-
FIG. 10C
AP2a
AP2a+ / DAPI
FIG. 10D
NE Stage
SOX2 DAPI
DCX DAPI
OTX2 DAPI

FIG. 10E

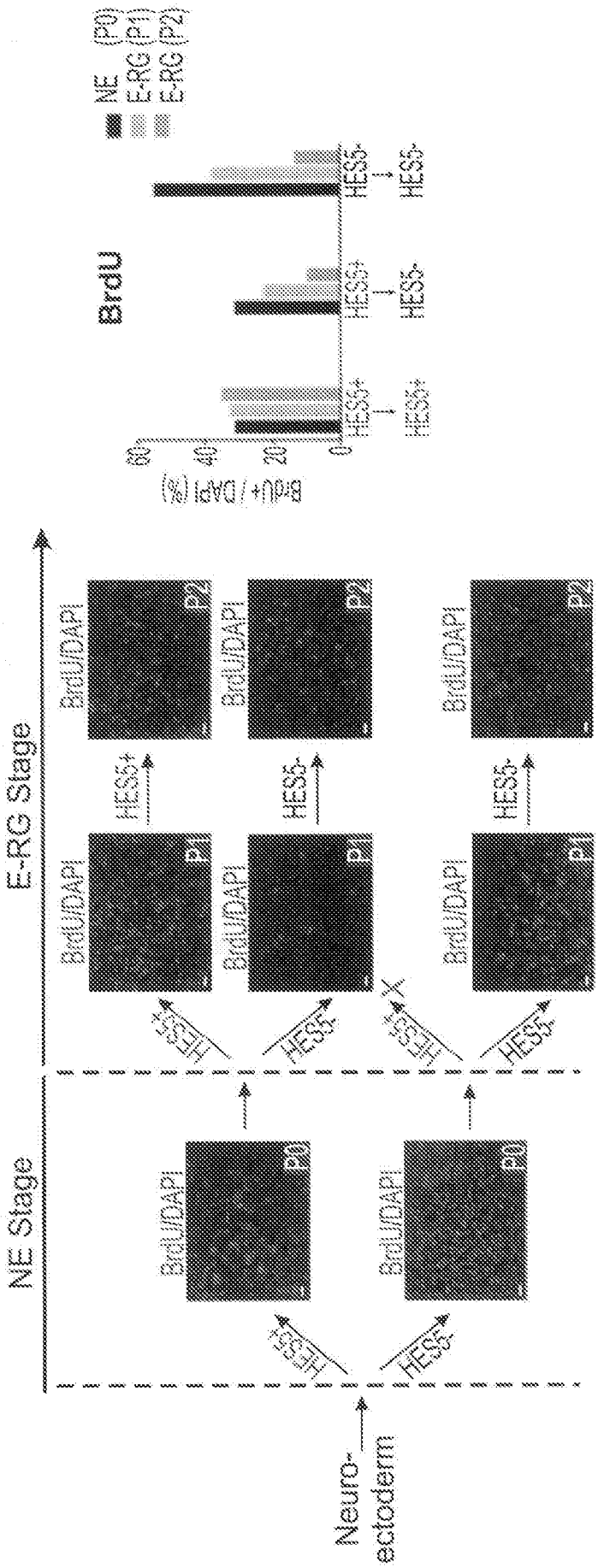

NE (P0)
E-RG (P1)
E-RG (P2)
BrdU
HES5+
HES5-
BrdU+/DAPI (%)
60
40
20
0
NE Stage
E-RG Stage
Neuro-ectoderm
BrdU/DAPI
P0
P1
P2

FIG. 11A

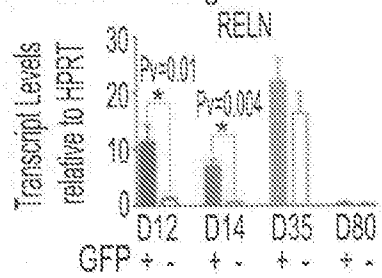
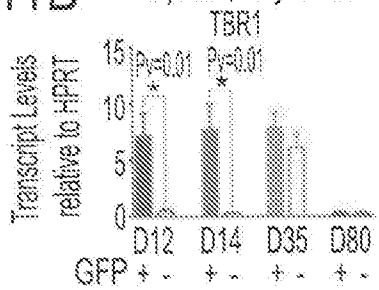
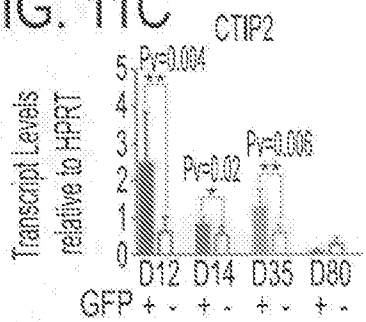


FIG. 11B



Subcerebral, Layer V

FIG. 11C



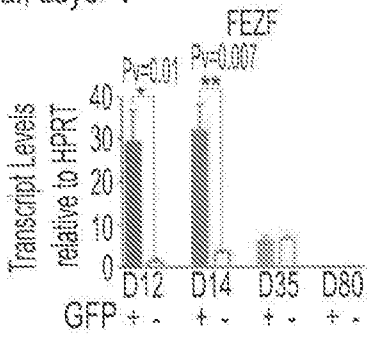


FIG. 11D

FIG. 11E

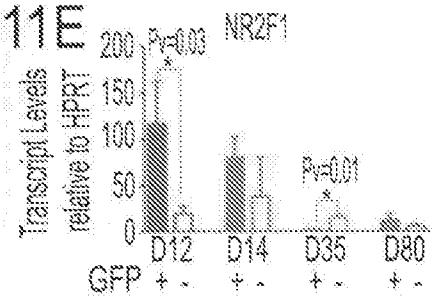


FIG. 11F

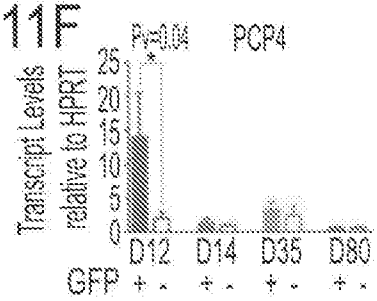
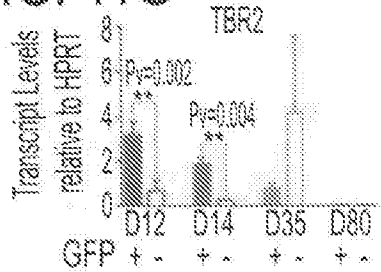


FIG. 11G

SVZ,Callosal, Layer II-IV



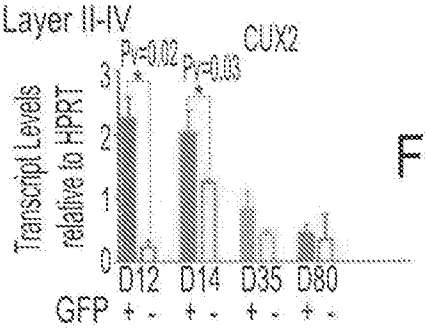


FIG. 11H

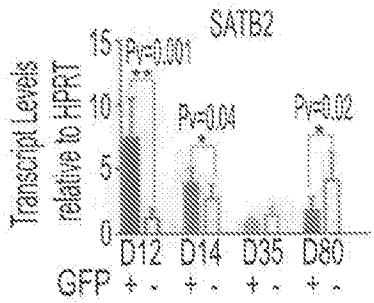


FIG. 11I

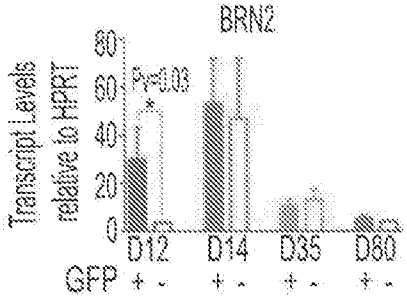


FIG. 11J

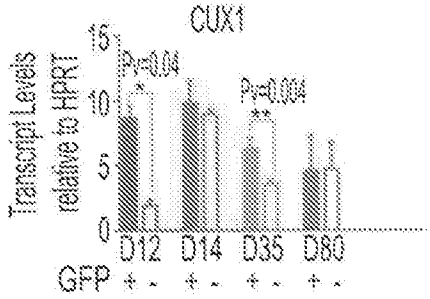


FIG. 11K

FIG. 12A

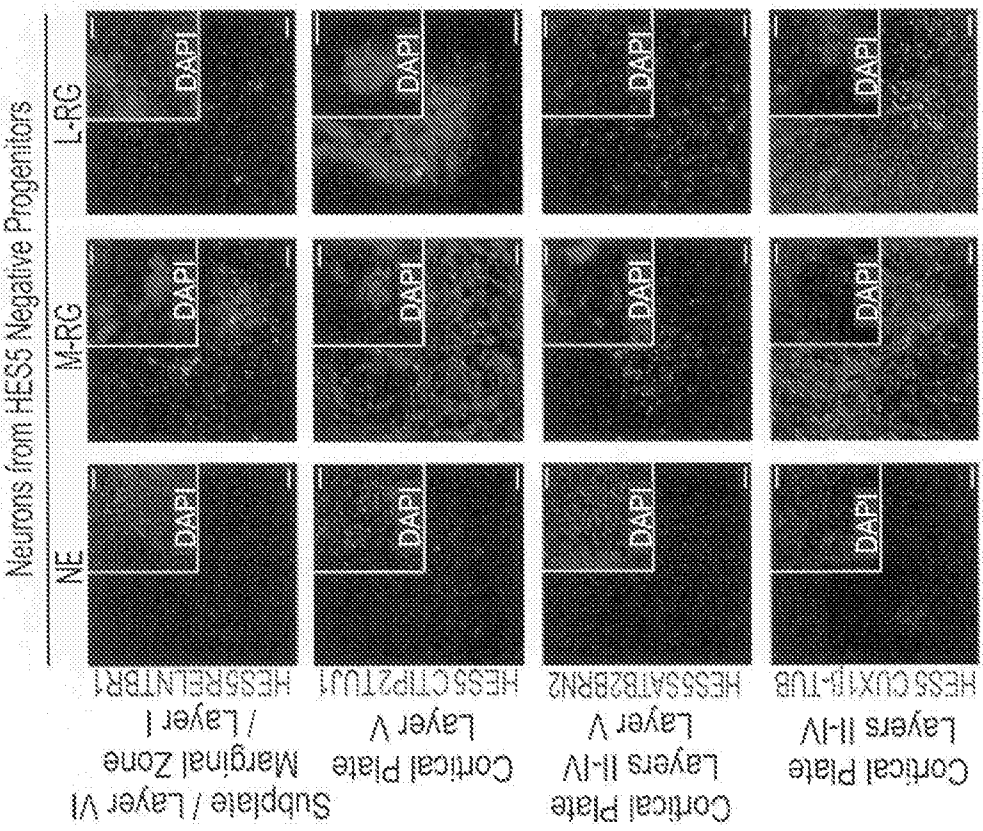


FIG. 12B

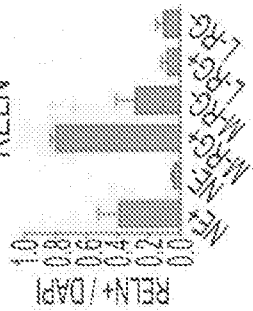


FIG. 12C

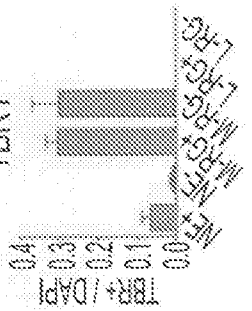


FIG. 12D

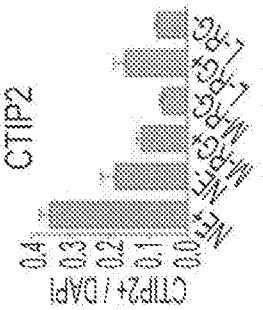


FIG. 12E

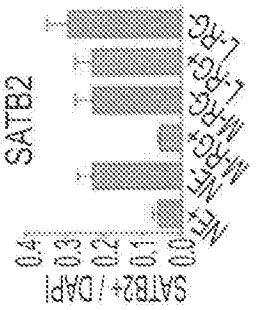


FIG. 12F

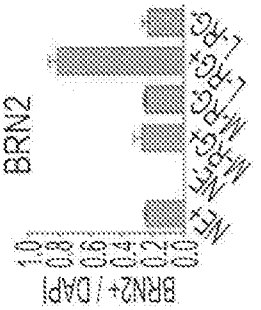
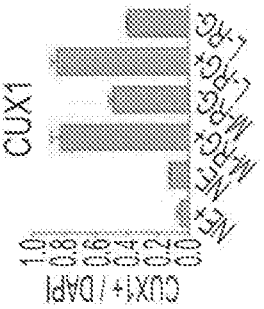


FIG. 12G



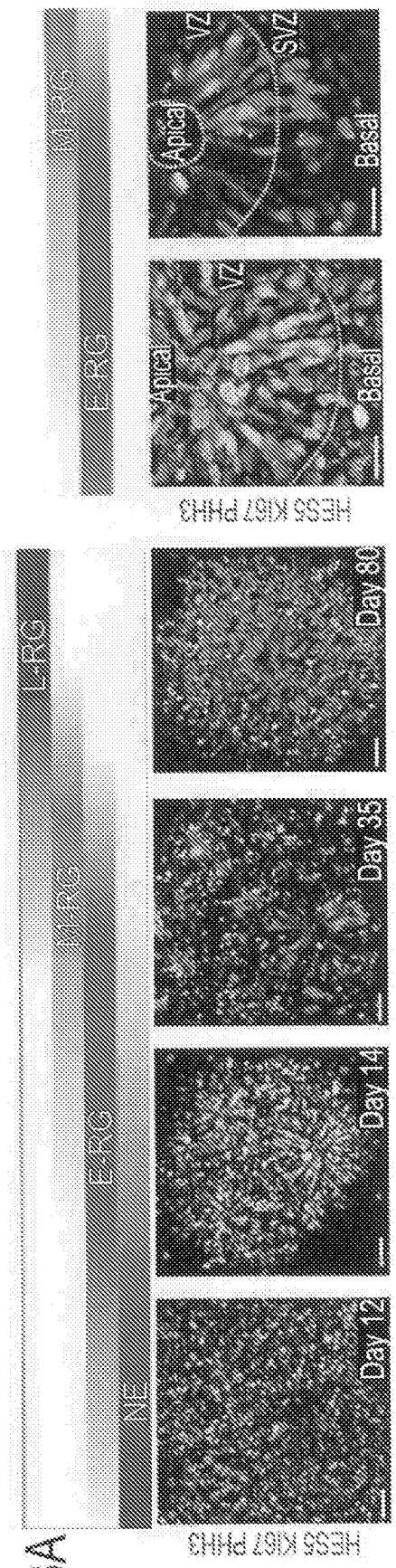


FIG. 13A

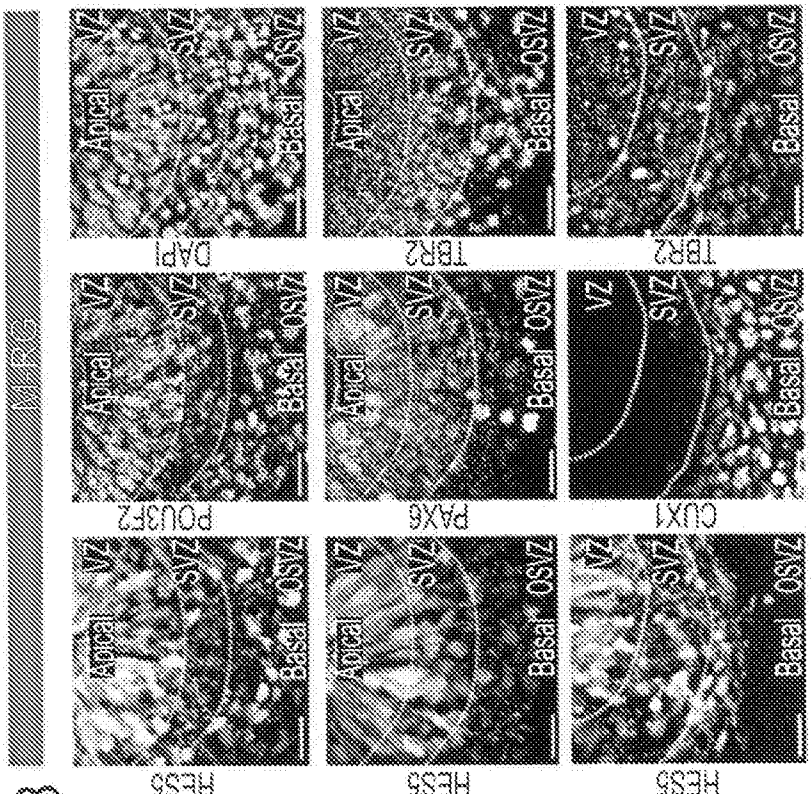


FIG. 13B

FIG. 13C FEZF

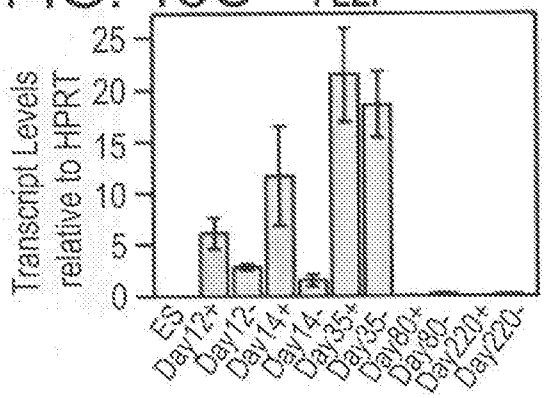


FIG. 13D TBR2

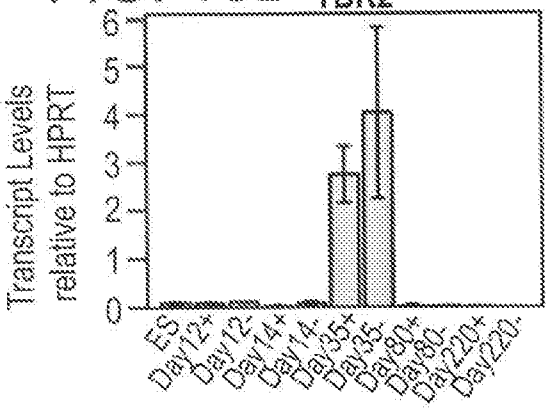


FIG. 13E POU3F2

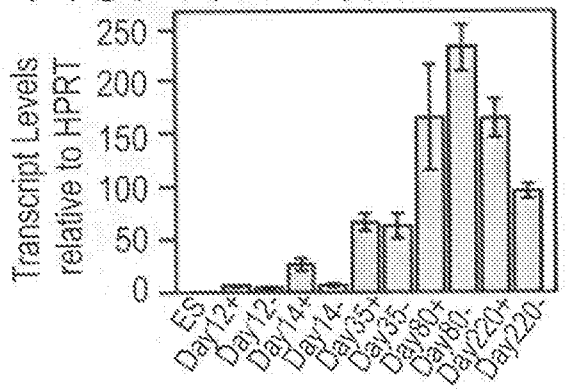


FIG. 13F POU3F3

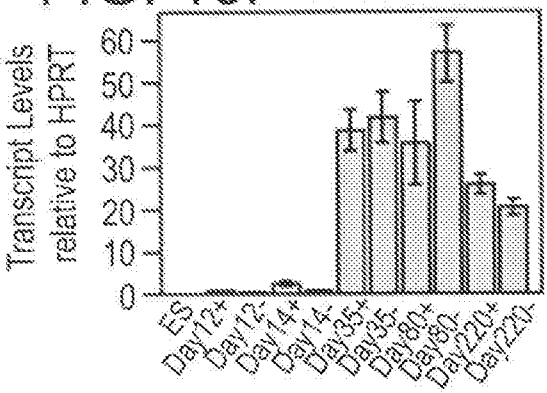


FIG. 13G CUX1

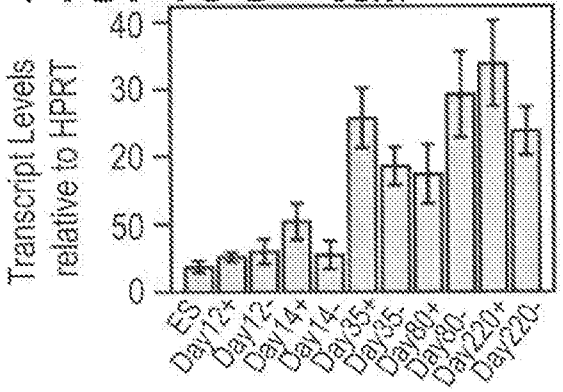


FIG. 13H CUX2

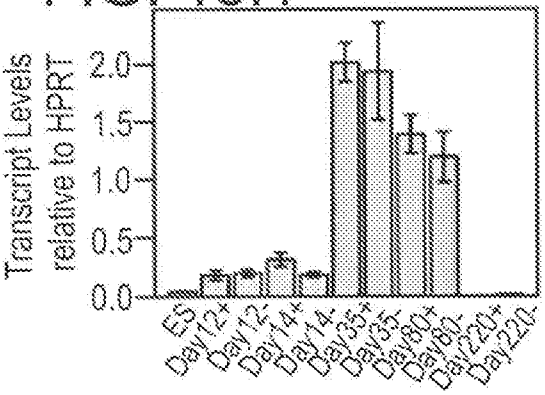


FIG. 13I FABP7

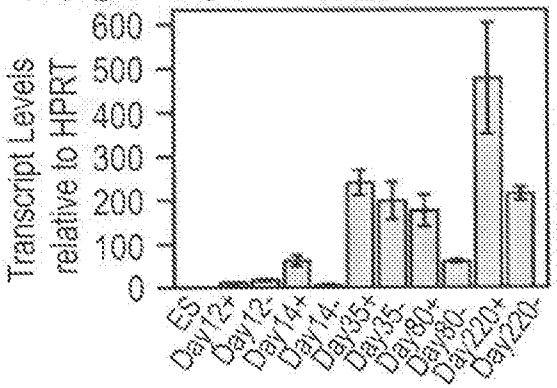
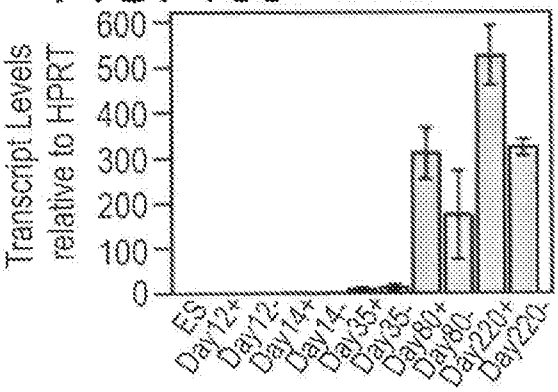


FIG. 13J GLAST



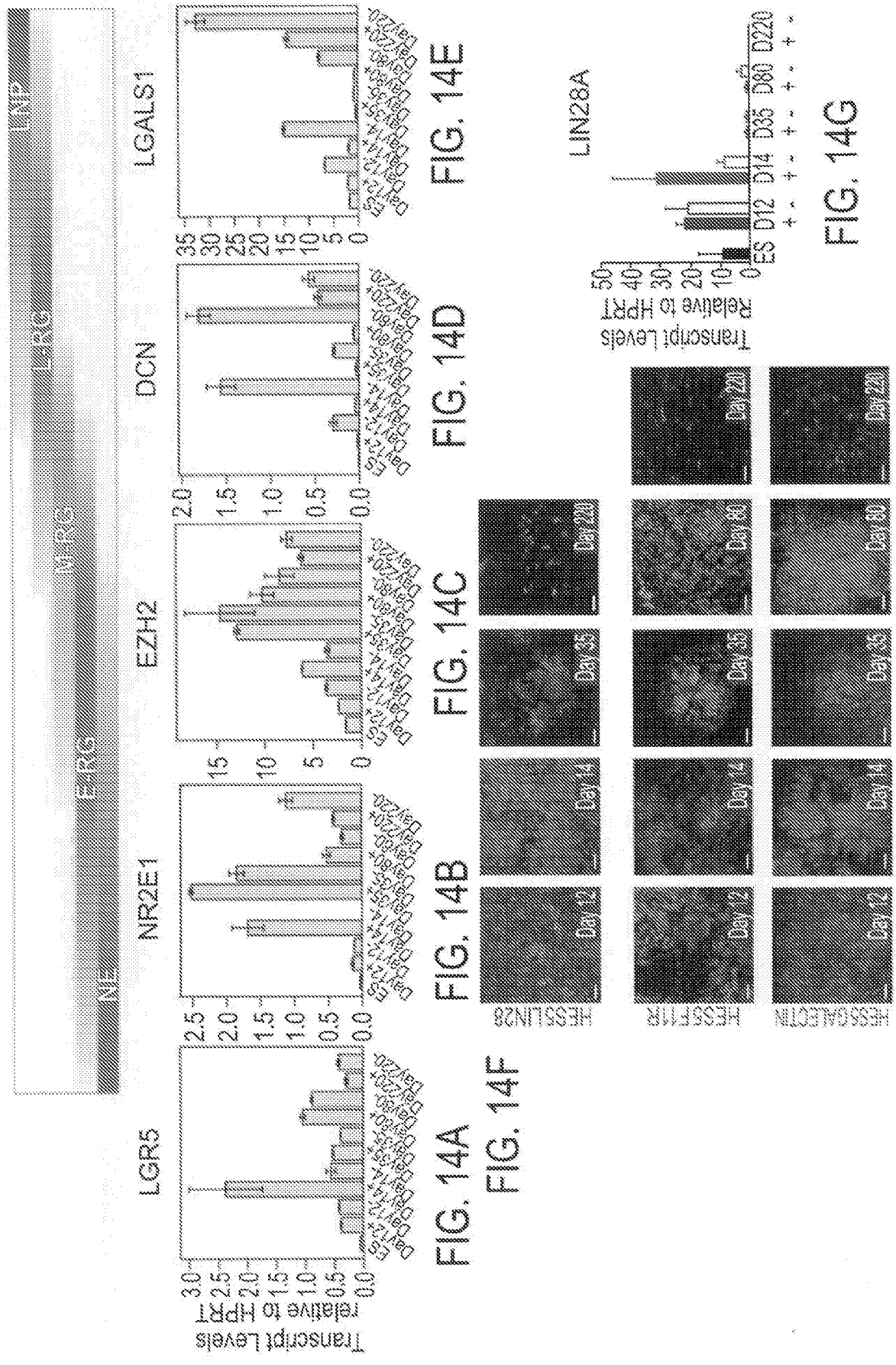


FIG. 14A FIG. 14B FIG. 14C FIG. 14D FIG. 14E FIG. 14F FIG. 14G

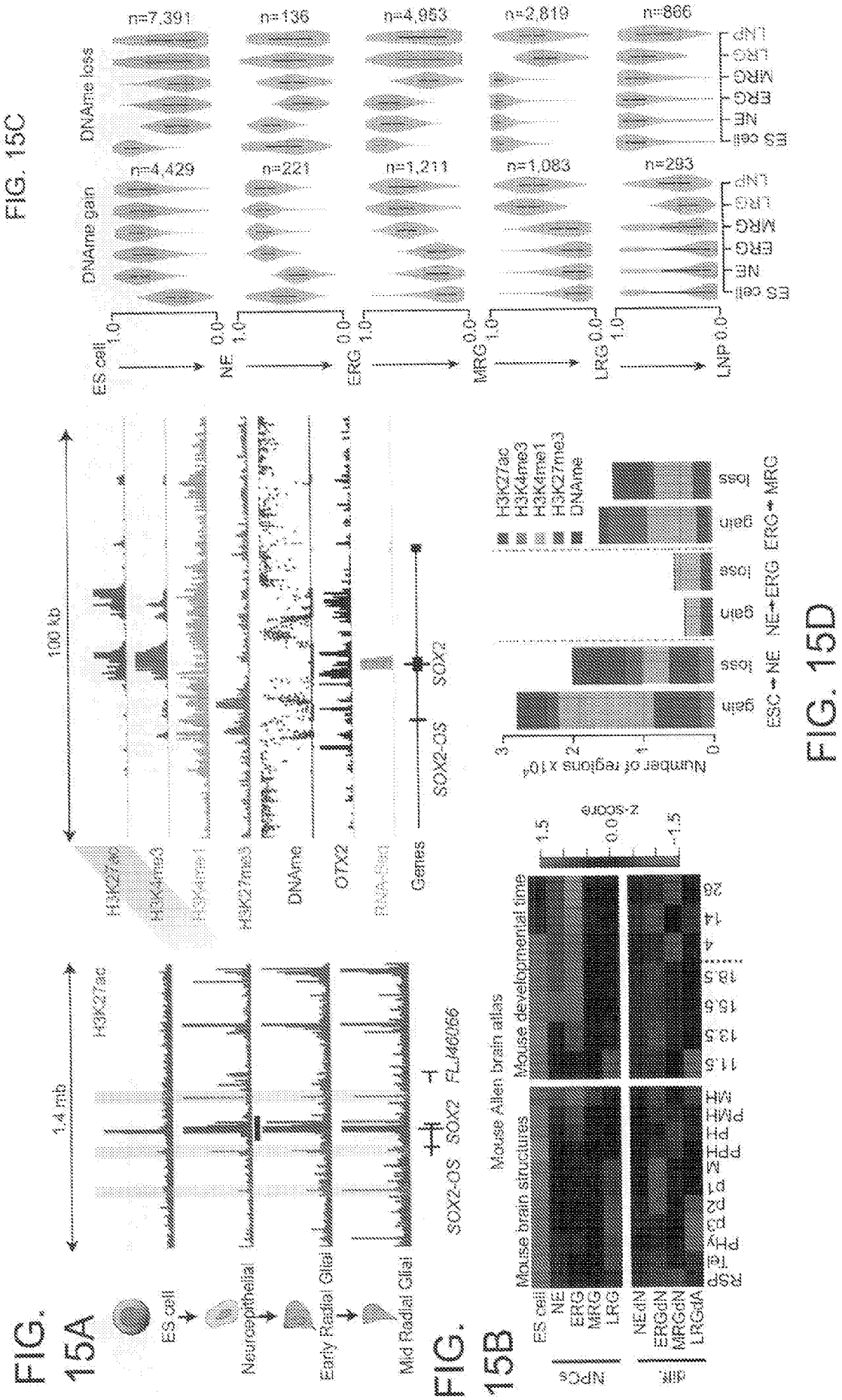


FIG. 15A
ES cell
Neuroepithelia
Early Radial Glial
Mid Radial Glial
1.4 mb
H3K27ac
SOX2-OS
SOX2
FLJ46066
100 kb
H3K27ac
H3K4me3
H3K4me1
H3K27me3
DNAme
OTX2
RNA-Seq
Genes
SOX2-OS
SOX2
FIG. 15B
Mouse Allen brain atlas
Mouse brain structures
Mouse developmental time
z-score
NPCs
diff.
ES cell
NE
ERG
MRG
LRG
NEdiN
ERGdiN
MRGdiN
LRGdiN
RSP
Tel
PHy
P3
P2
P1
M
PPH
PH
PMH
MH
11.5
13.5
15.5
18.5
4
14
26
FIG. 15C
DNAme gain
DNAme loss
ES cell
NE
ERG
MRG
LRG
LNP
n=4,429
n=7,391
n=221
n=136
n=1,211
n=4,953
n=1,083
n=2,819
n=293
n=866
FIG. 15D
H3K27ac
H3K4me3
H3K4me1
H3K27me3
DNAme
Number of regions x 10^4
gain
loss
ESC → NE
NE ↔ ERG
ERG → MRG

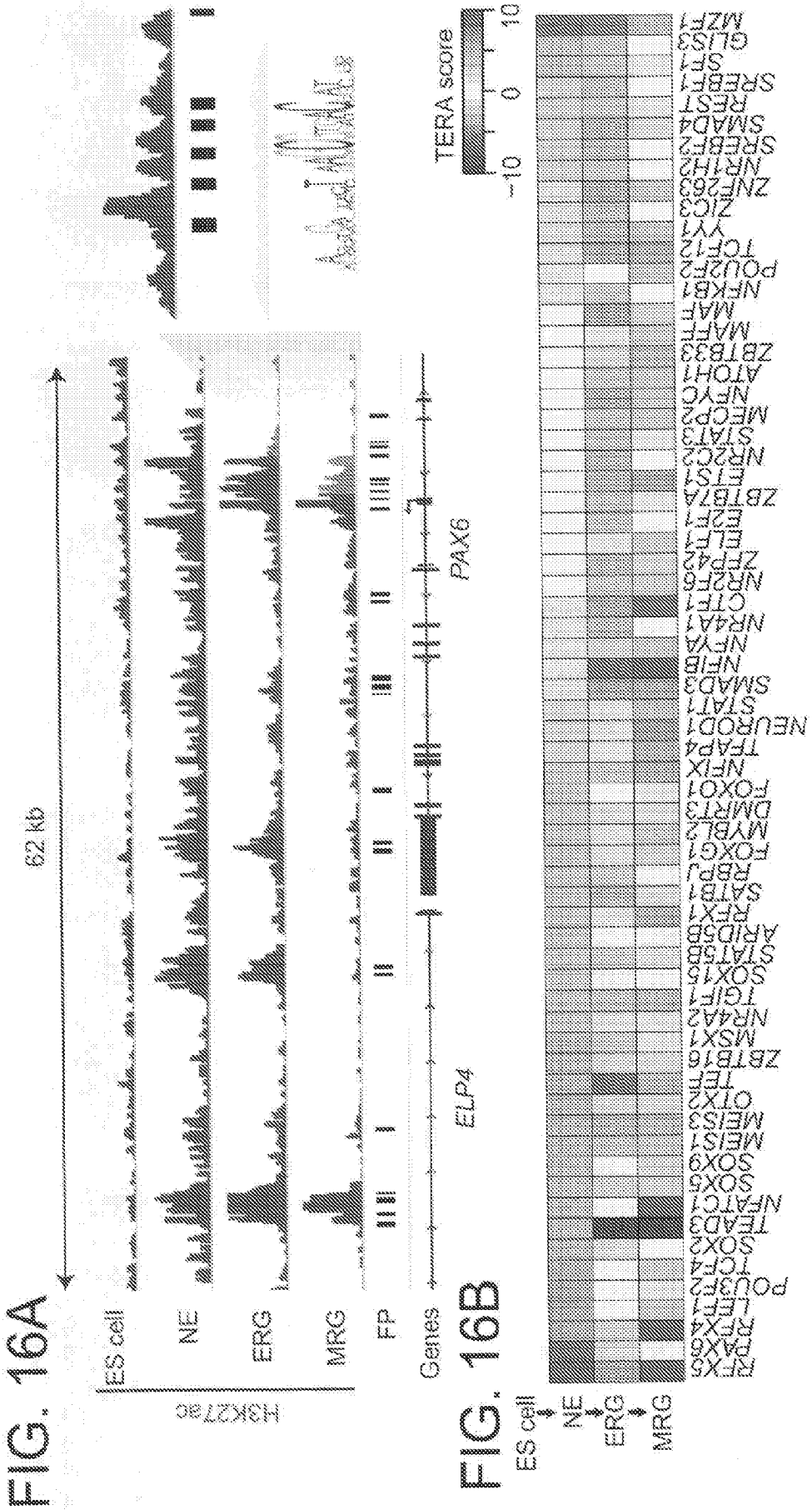


FIG. 16A

FIG. 16B

FIG. 17A

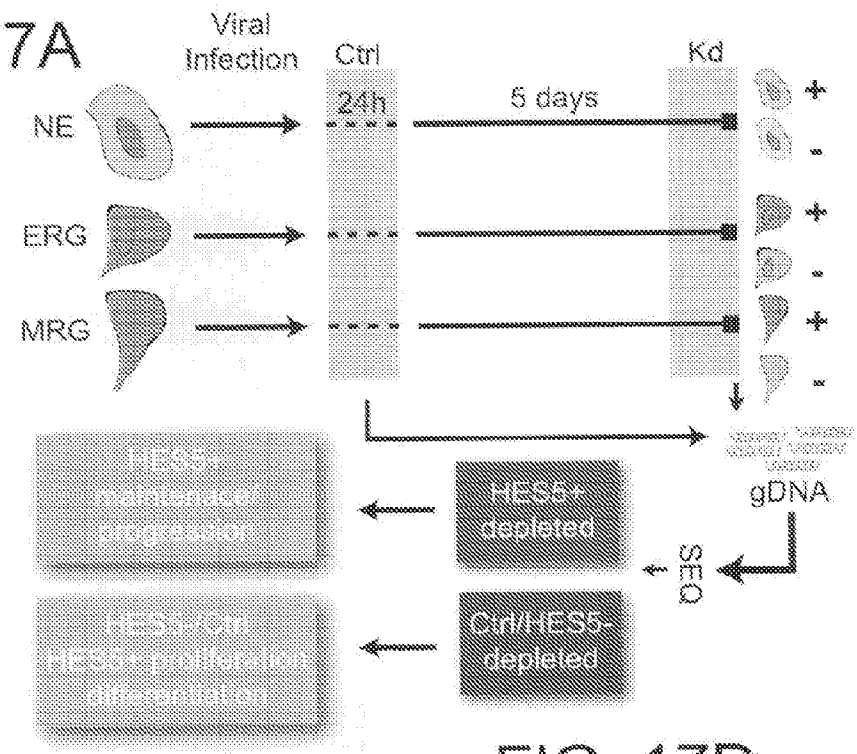


FIG. 17B

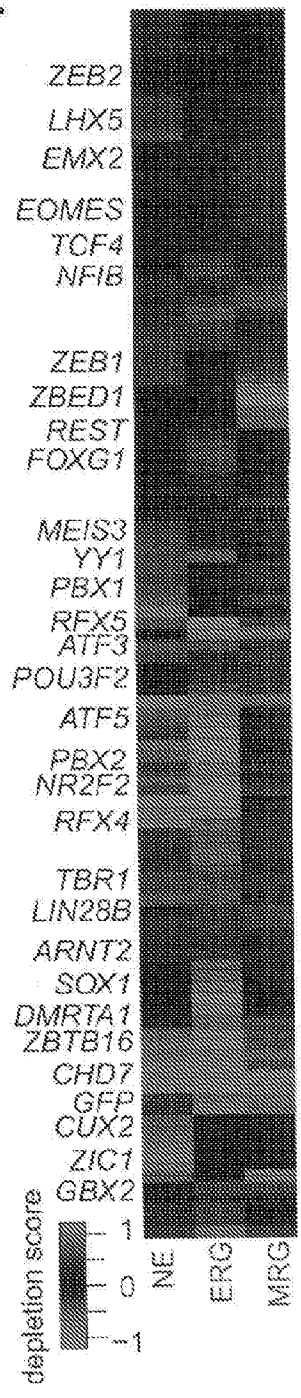


FIG. 17C

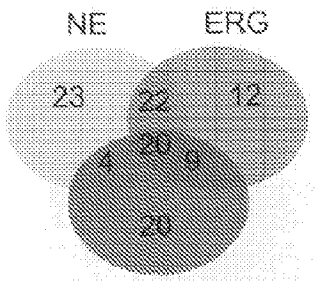
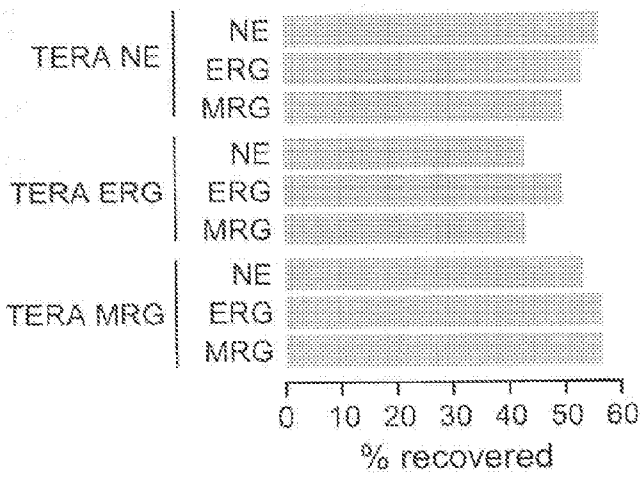


FIG. 17D



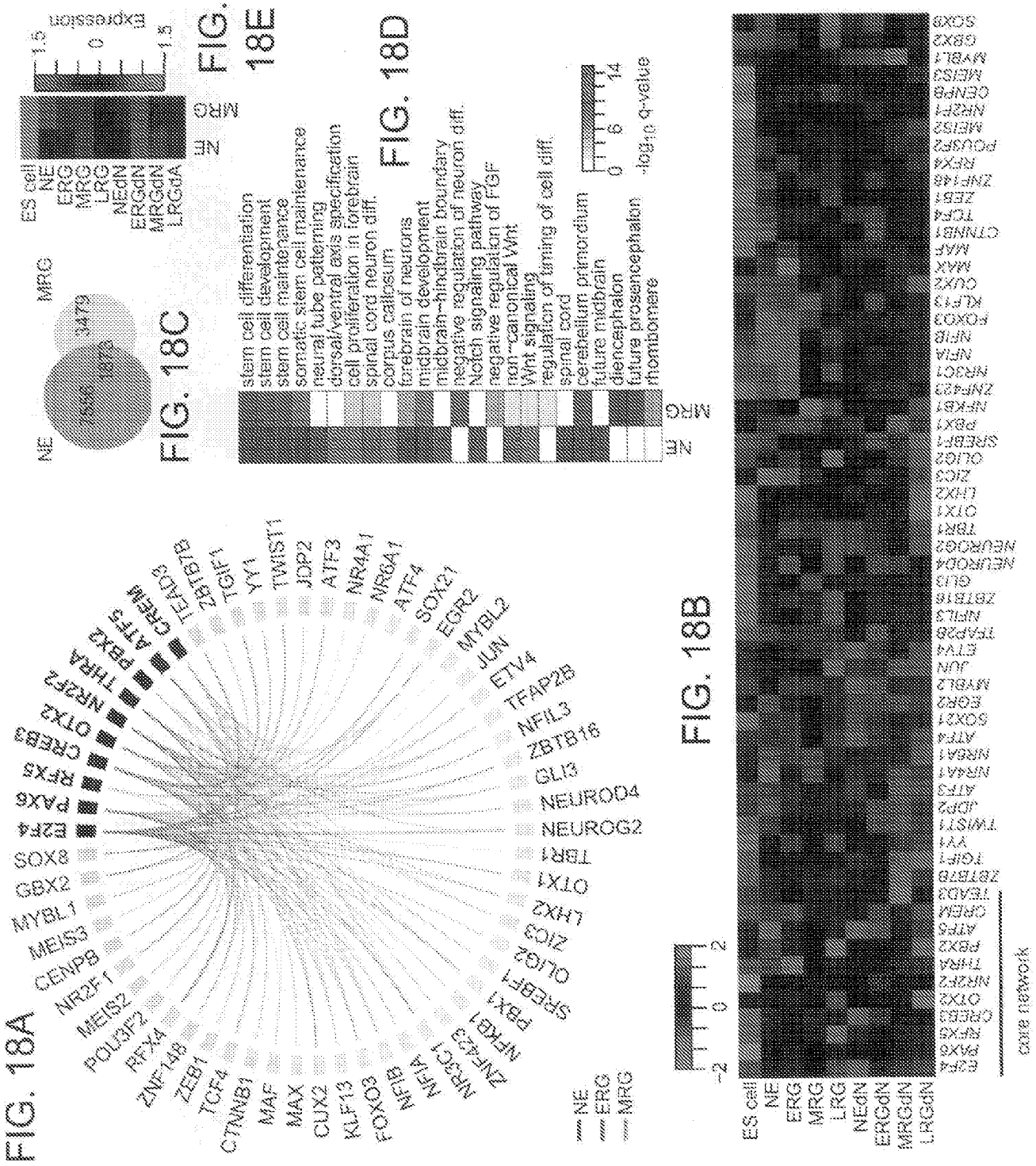


FIG. 18A
FIG. 18B
FIG. 18C
FIG. 18D
FIG. 18E
ES cell
NE
ERG
MRG
LRG
NEdN
ERGdN
MRGdN
LRGdN
core network
NE
MRG
1479
1523
1673
stem cell differentiation
stem cell development
stem cell maintenance
somatic stem cell maintenance
neural tube patterning
dorsal/ventral axis specification
cell proliferation in forebrain
spinal cord neuron diff.
corpus callosum
forebrain cell of neurons
midbrain development
midbrain–hindbrain boundary
negative regulation of neuron diff.
Notch signaling pathway
negative regulation of FGF
non-canonical Wnt
Wnt signaling
regulation of timing of cell diff.
spinal cord
cerebellum primordium
future midbrain
diencephalon
future prosencephalon
rhombomere
-log10 q-value
Expression

FIG. 19A

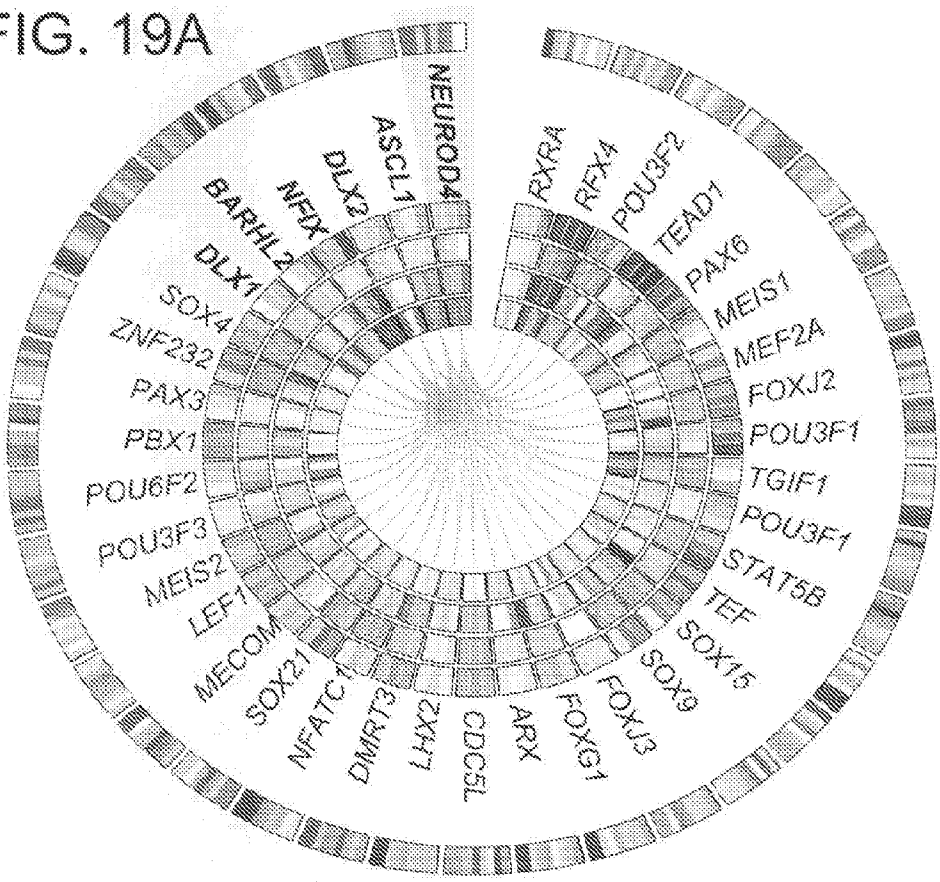


FIG. 19B

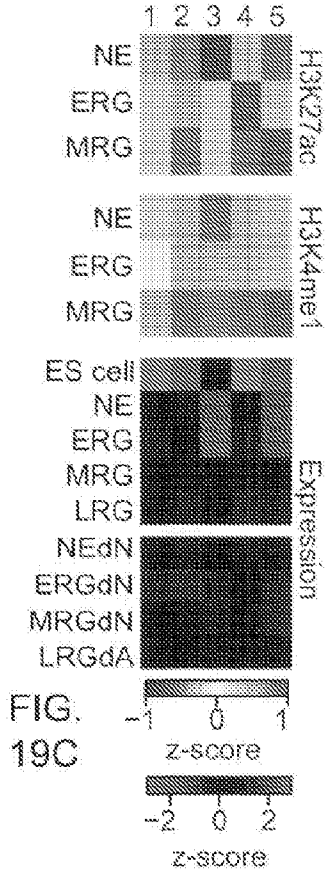


FIG. 19C

FIG. 19D

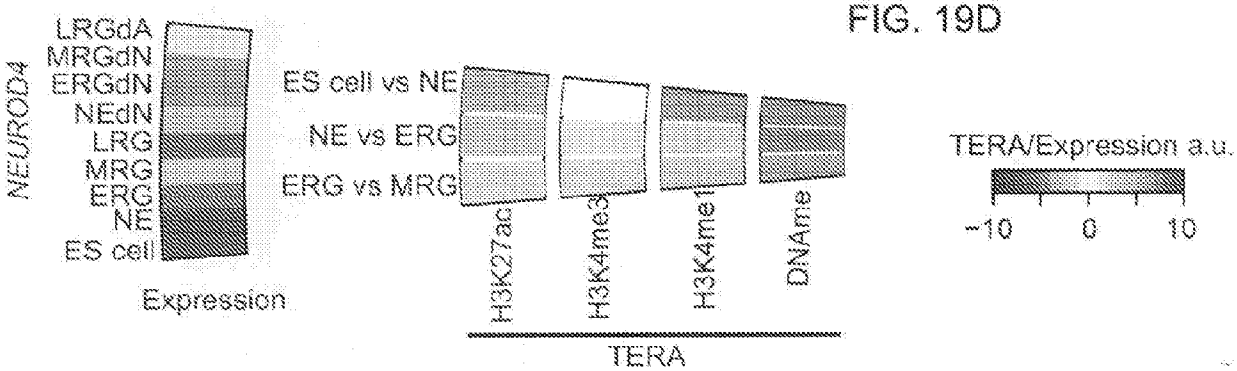
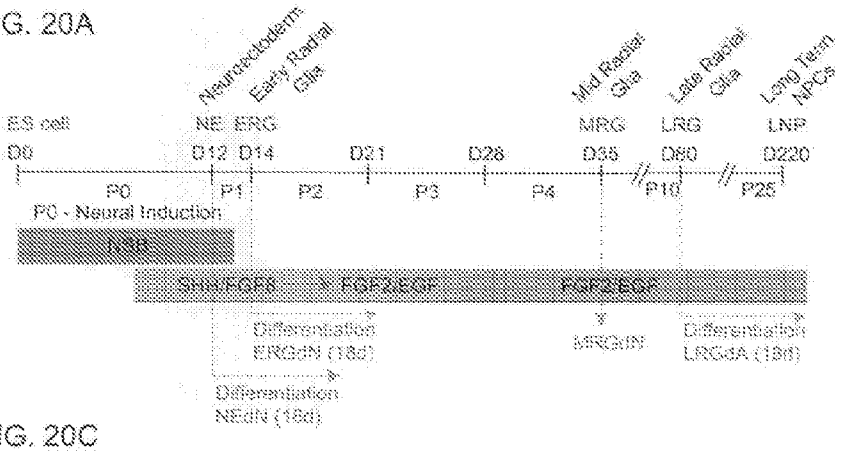


FIG. 20A



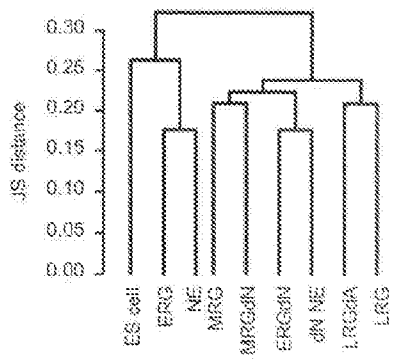
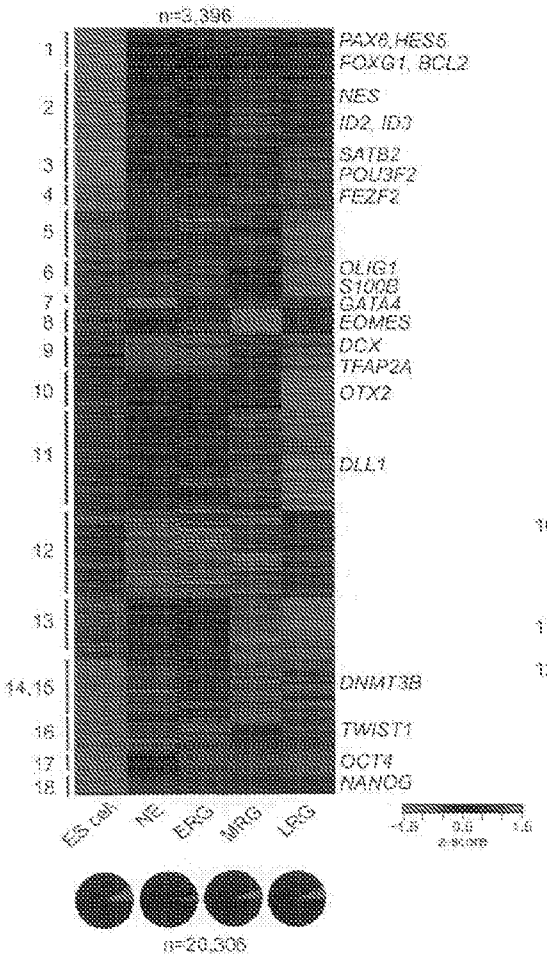


FIG. 20B

FIG. 20C



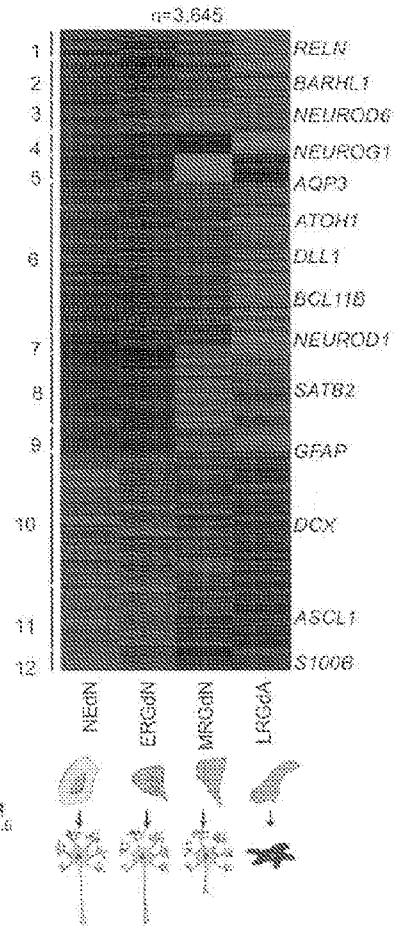


FIG. 20D

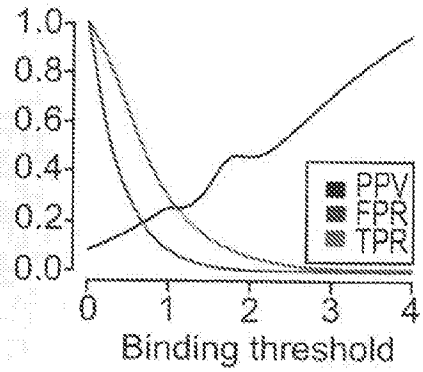


FIG. 21A

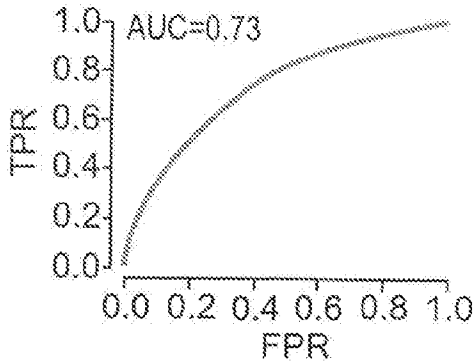


FIG. 21B

FIG. 21C

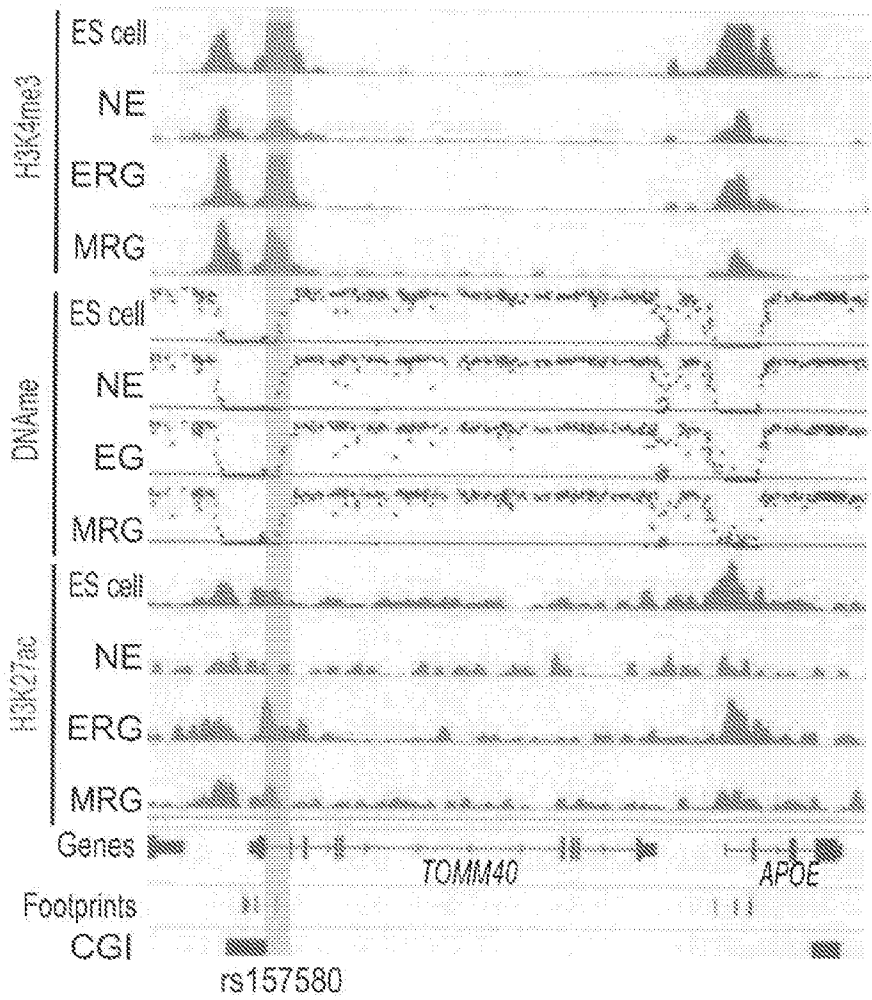


FIG. 21D

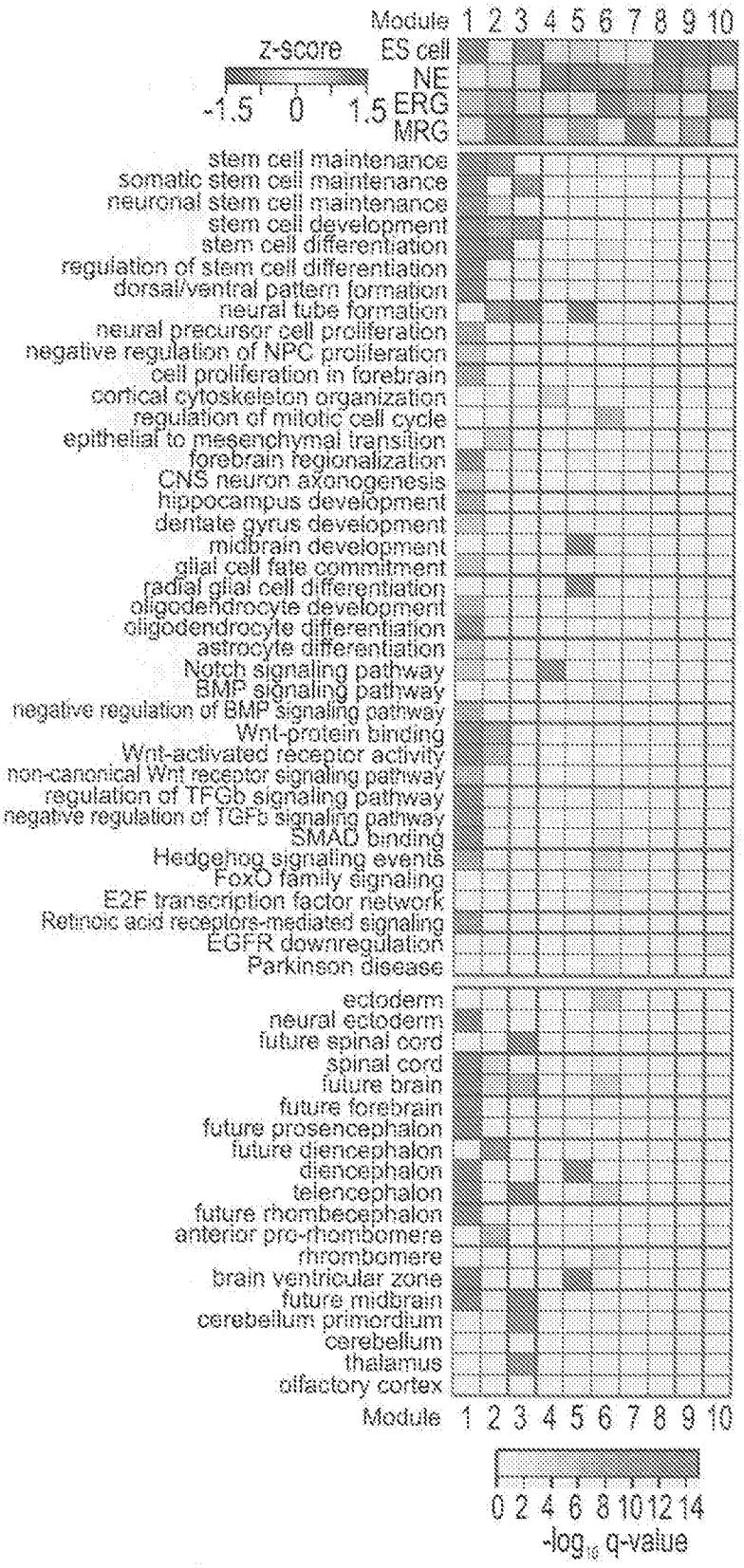
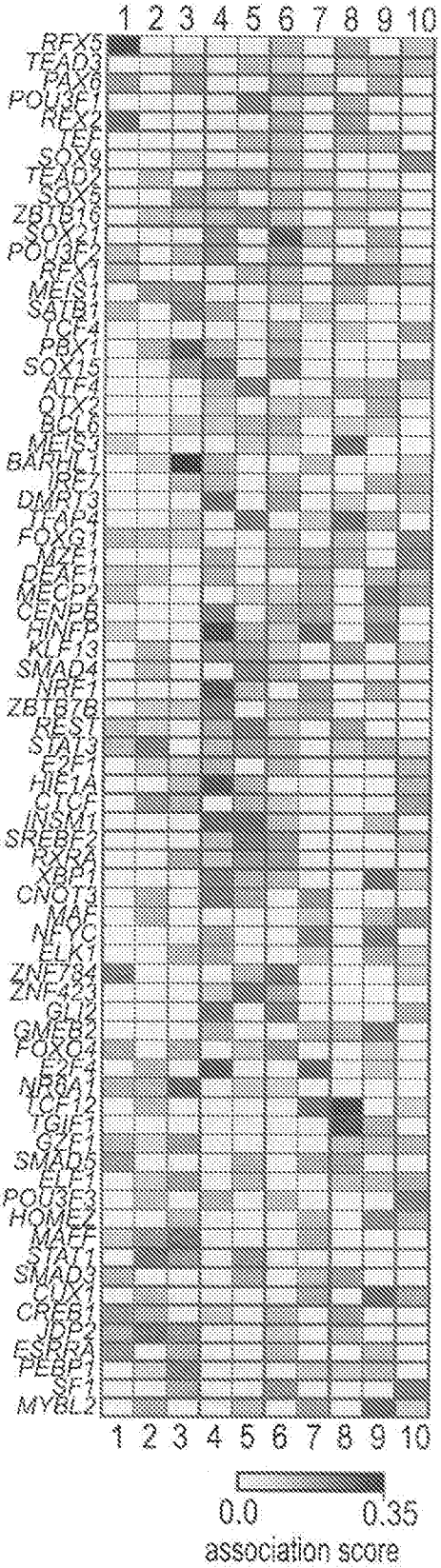


FIG. 21E



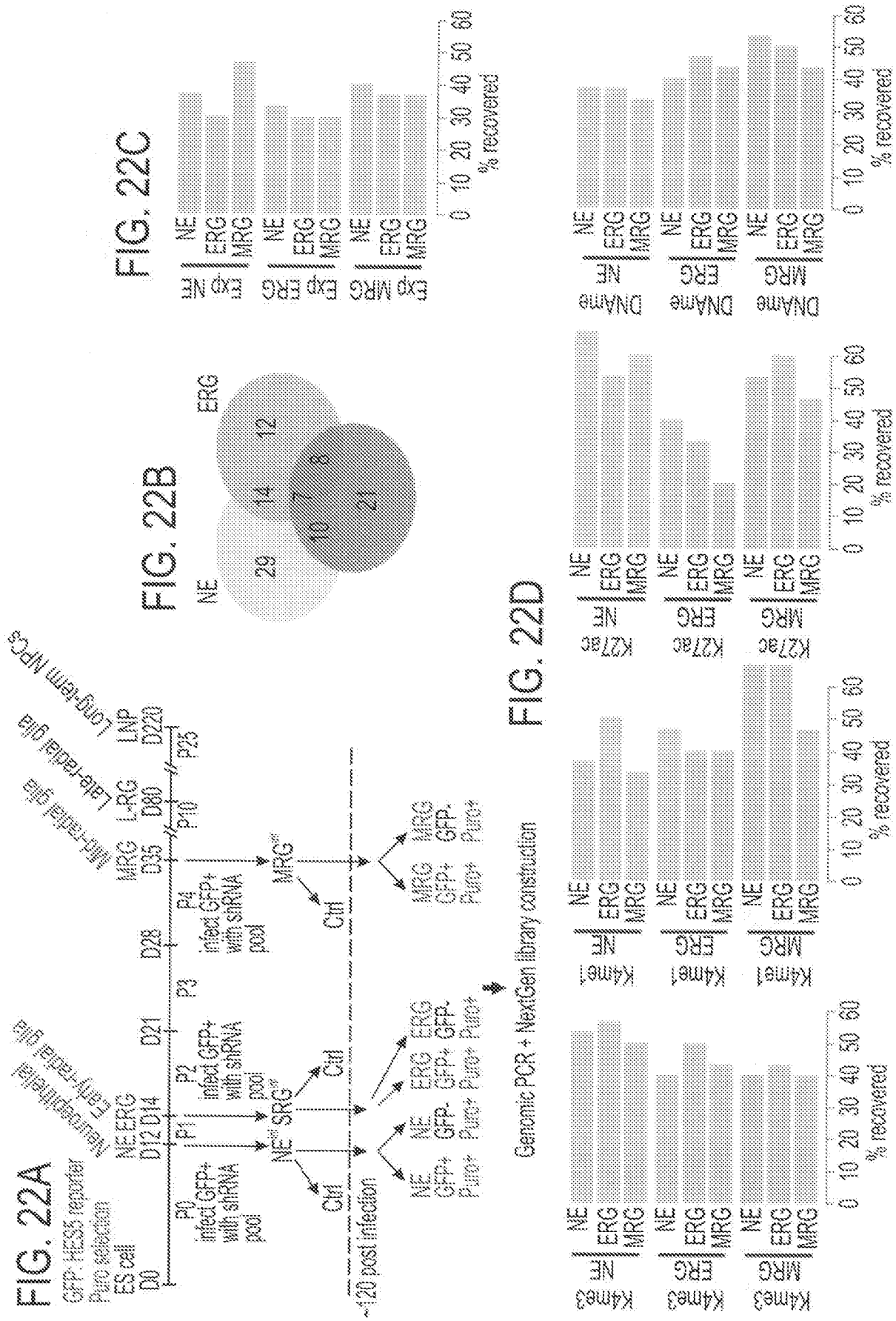


FIG. 22A
GFP::HES5 reporter
Puro selection
ES cell
D0
P0
infect GFP+ with shRNA pool
D12 D14
NE ERG
P1
NE SRG
D21
P2
infect GFP+ with shRNA pool
Ctrl
~120 post infection
NE GFP+ Puro+
NE GFP- Puro+
ERG GFP+ Puro+
ERG GFP- Puro+
Genomic PCR + NextGen library construction
D28
P3
P4
infect GFP+ with shRNA pool
MRG
MRG GFP+ Puro+
MRG GFP- Puro+
D35
P10
L-RG
D80
P25
LNP
D220
Neuroepithelial
Early-radial glia
Mid-radial glia
Late-radial glia
Long-term NPCs
FIG. 22B
NE
ERG
29
14
12
7
8
10
21
FIG. 22C
Exp NE
Exp ERG
Exp MRG
NE
ERG
MRG
0 10 20 30 40 50 60
% recovered
FIG. 22D
K4me3
K4me1
K27ac
DNAme
NE
ERG
MRG
0 10 20 30 40 50 60
% recovered

FIG. 22E

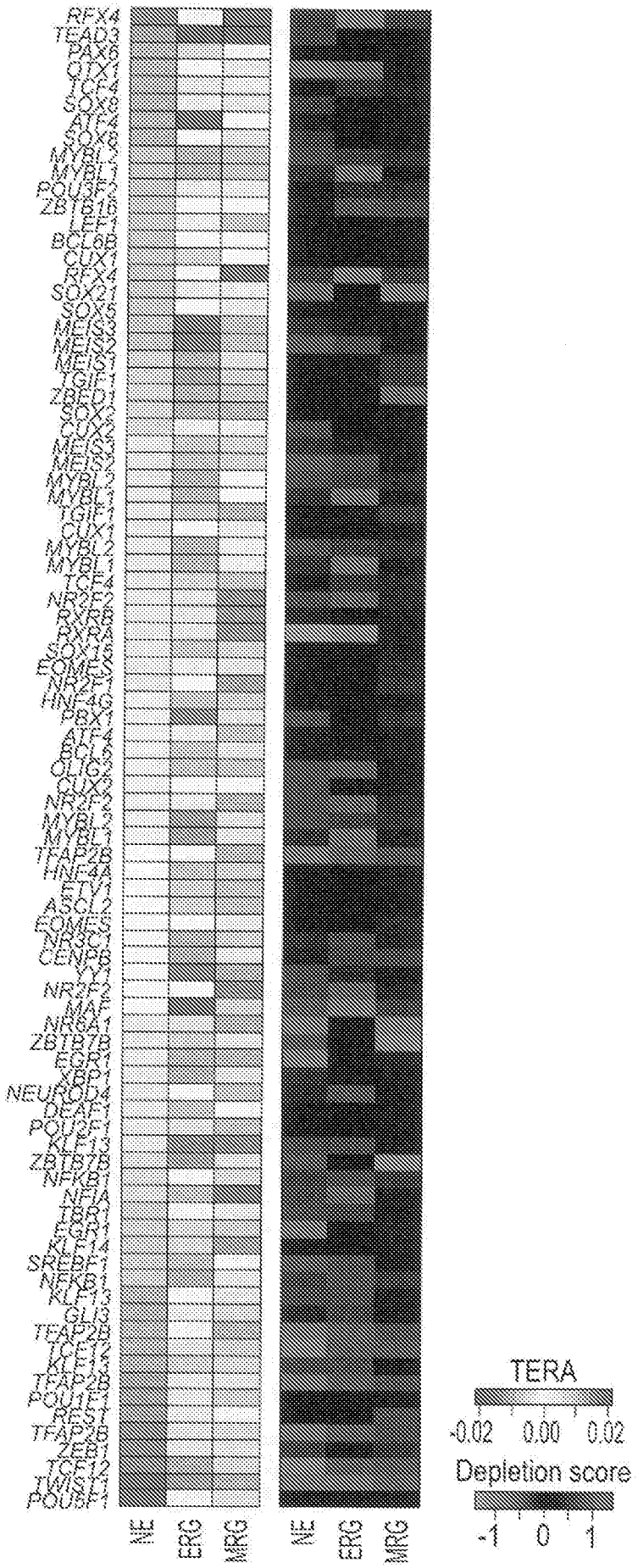

RFX4
TEAD3
PAX6
OTX1
TCF4
SOX8
ATF4
SOX9
MYBL2
MYBL1
POU3F2
ZBTB18
LEF1
BCL6B
CUX1
RFX4
SOX21
SOX5
MEIS3
MEIS2
MEIS1
TGIF1
ZBED1
SOX2
CUX2
MEIS3
MEIS2
MYBL2
MYBL1
TGIF1
CUX1
MYBL2
MYBL1
TCF4
NR2F2
RXRB
RXRA
SOX13
EOMES
NR2F1
HNF4G
PBX1
ATF4
BCL6
OLIG2
CUX2
NR2F2
MYBL2
MYBL1
TFAP2B
HNF4A
ETV1
ASCL2
EOMES
NR3C1
CENPB
YY1
NR2F2
MAF
NR6A1
ZBTB7B
EGR1
XBP1
NEUROD4
DEAF1
POU2F1
KLF13
ZBTB7B
NFKB1
NFIA
TBR1
EGR1
KLF14
SREBF1
NFKB1
KLF13
GLI3
TEAP2B
TCF12
KLF13
TFAP2B
POU1F1
RFX1
TFAP2B
ZEB1
TCF12
TWIST1
POU5F1
NE
ERG
MRG
NE
ERG
MRG
TERA
-0.02
0.00
0.02
Depletion score
-1
0
1

FIG. 22F

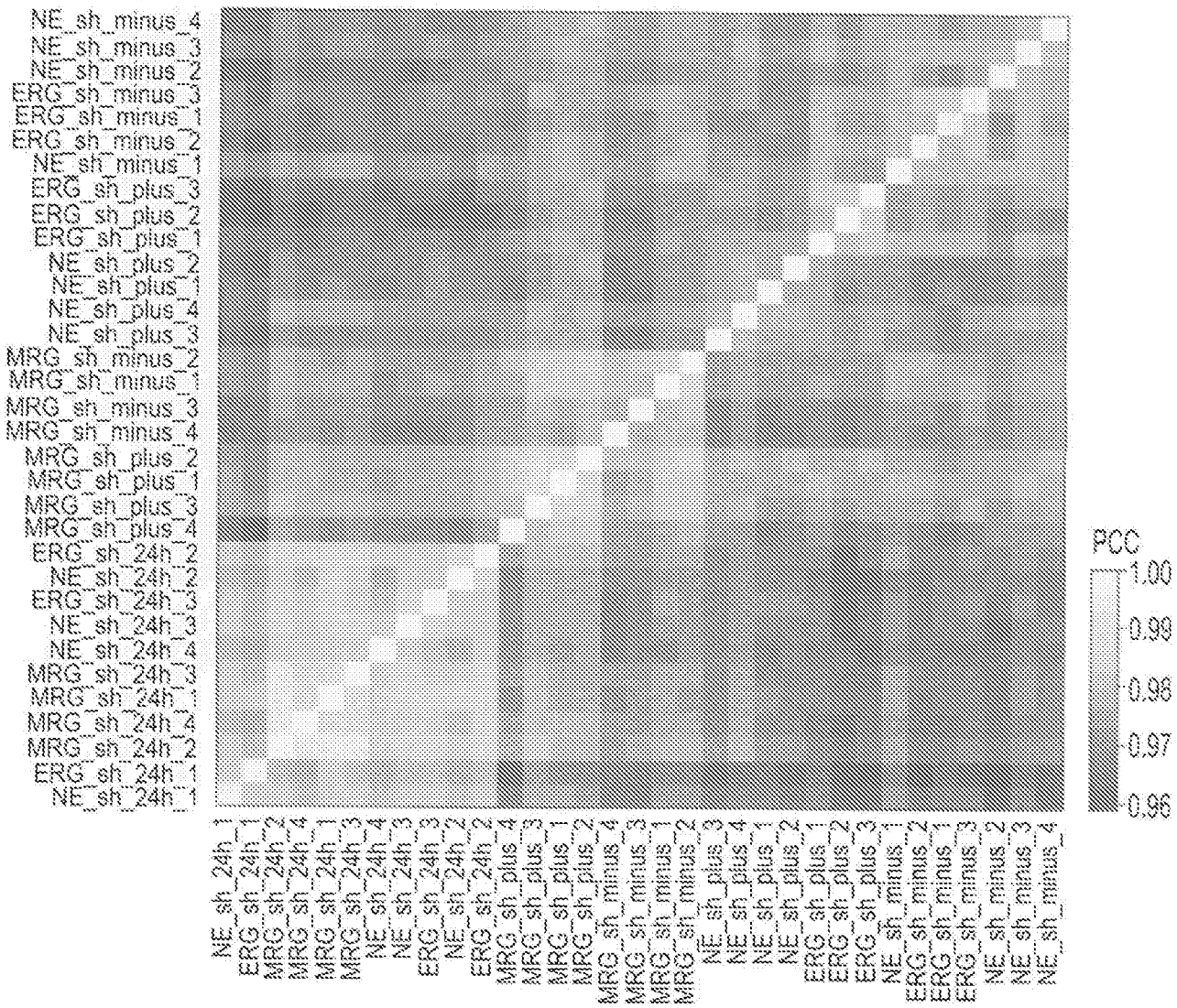
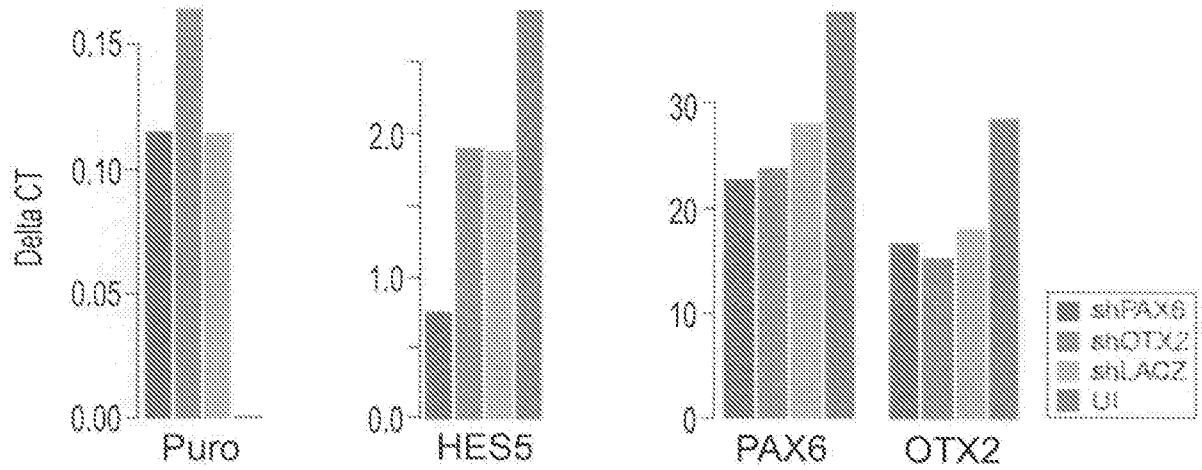


FIG. 22G



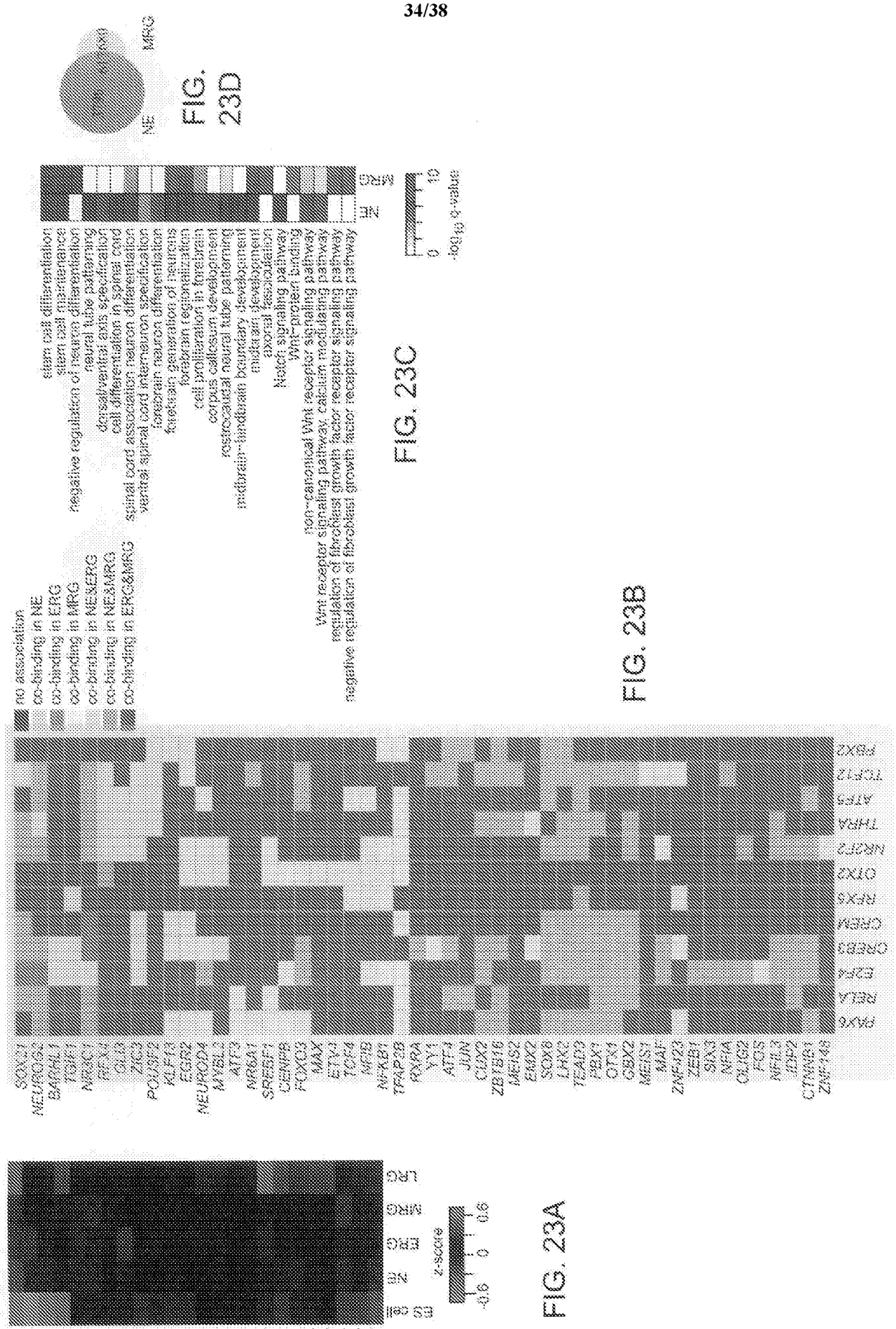

PLAGL1
PAX6
NR2F2
ARNT2
NCOR2
FOXM1
CHD7
THRA
CREB1
PLXNB2
TCF12
RELA
PBX2
RPA2
RFX5
E2F4
OTX2
TSC22D3
ATF5
KLF10
RCOR2
TOX3
CREM
ZFHX4
ES
NE
ERG
MRG
LRG
z-score
FIG. 23A
SOX21
NEUROG2
BARHL1
TCF3
MYEC1
RFX4
GLI3
ZIC3
POU3F3
KLF13
EGR2
NEUROD4
MYBL1
ATF3
NR6A1
SREBF1
CENPB
FOXO3
MAX
ETV1
TCF4
MTB
NFKB1
TFAP2B
RXRA
YY1
ATF4
JUN
CUX2
ZBTB16
MEIS2
EMX2
SOX6
LHX2
TEAD3
PBX1
OTX1
GBX2
MEIS1
MAF
ZNF423
ZEB1
SIX3
NFIA
OLIG2
FOS
NFIB
IRF2
CTNNB1
ZNF148
PAX6
RELA
E2F4
CREB3
CREM
RFX5
OTX2
NR2F2
THRA
ATF5
TCF12
PBX2
FIG. 23B
no association
co-binding in NE
co-binding in ERG
co-binding in MRG
co-binding in NE&ERG
co-binding in NE&MRG
co-binding in ERG&MRG
stem cell differentiation
stem cell maintenance
negative regulation of neuron differentiation
neural tube patterning
dorsal/ventral axis specification
cell differentiation in spinal cord
spinal cord association neuron differentiation
ventral spinal cord interneuron specification
forebrain neuron differentiation
forebrain generation of neurons
forebrain regionalization
cell proliferation in forebrain
corpus callosum development
rostrocaudal neural tube patterning
midbrain-hindbrain boundary development
midbrain development
axonal fasciculation
Notch signaling pathway
Wnt-protein binding
non-canonical Wnt receptor signaling pathway
Wnt receptor signaling pathway, calcium modulating pathway
regulation of fibroblast growth factor receptor signaling pathway
negative regulation of fibroblast growth factor receptor signaling pathway
NE
MRG
0
10
-log10 q-value
FIG. 23C
NE
ERG
MRG
FIG. 23D

FIG. 24A

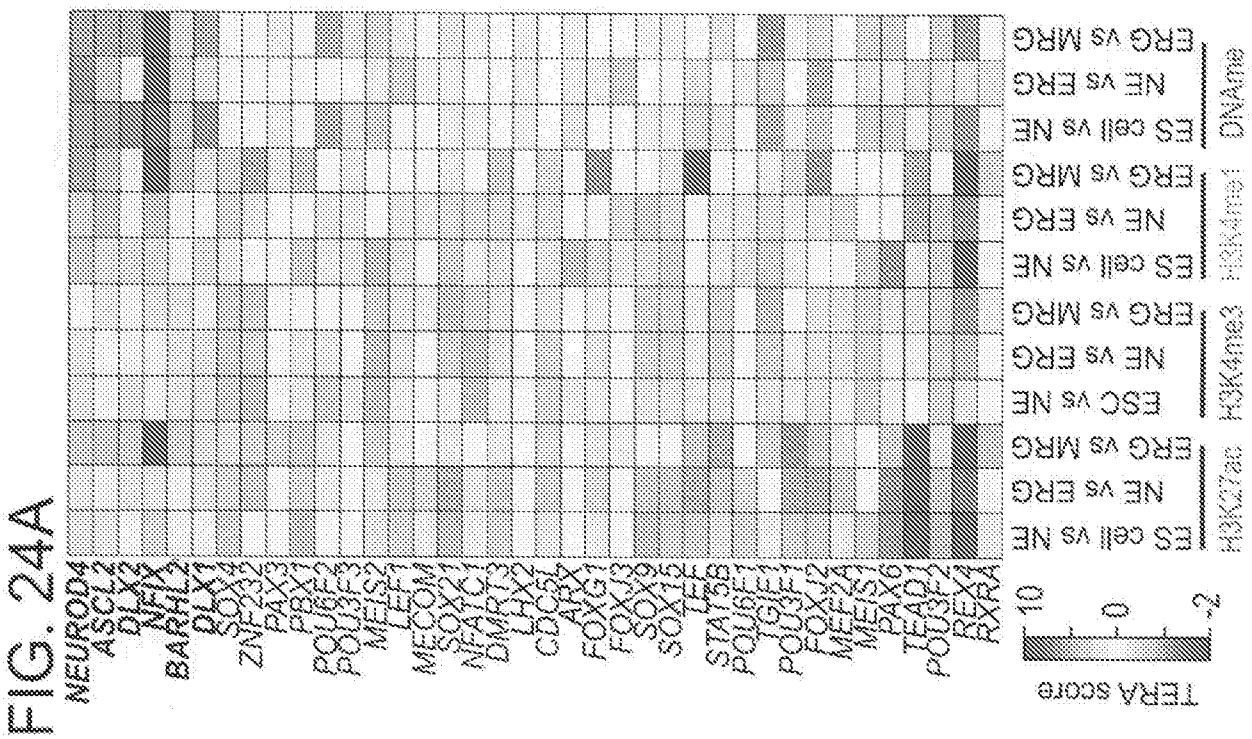


FIG. 24B

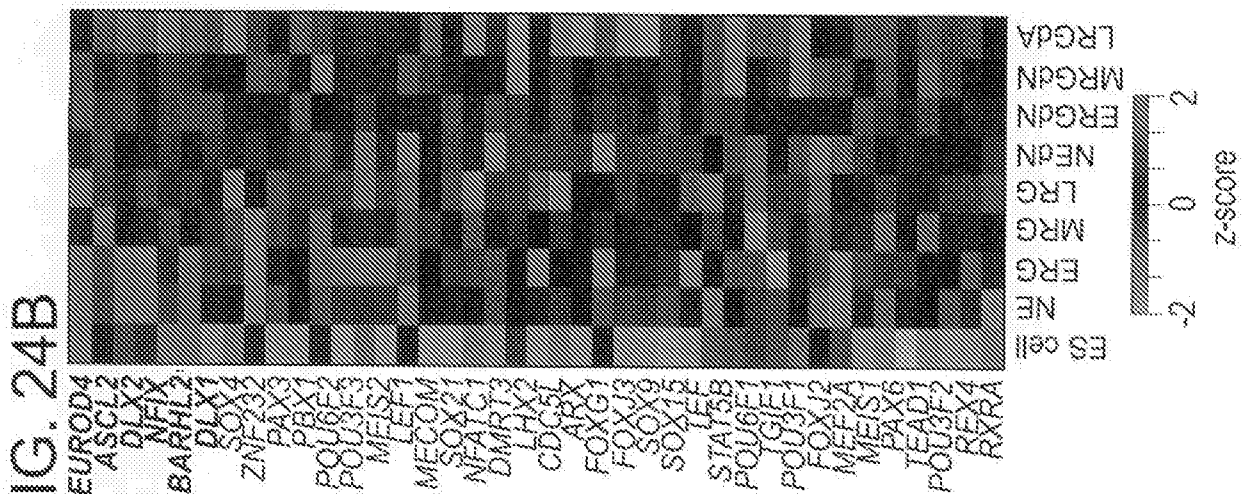
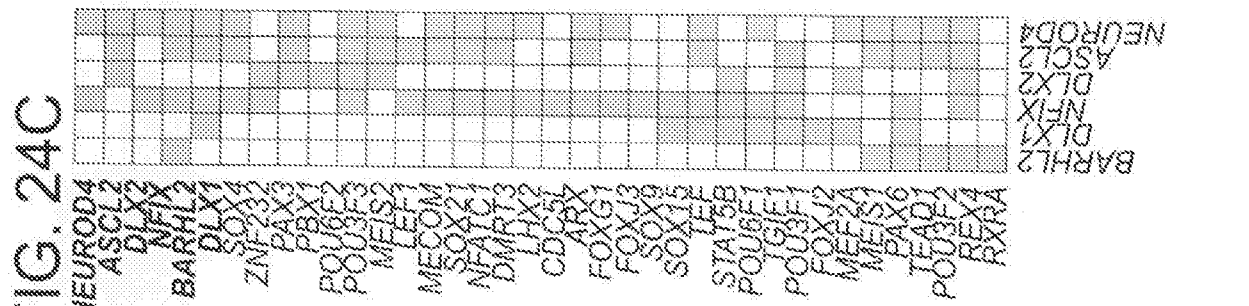


FIG. 24C



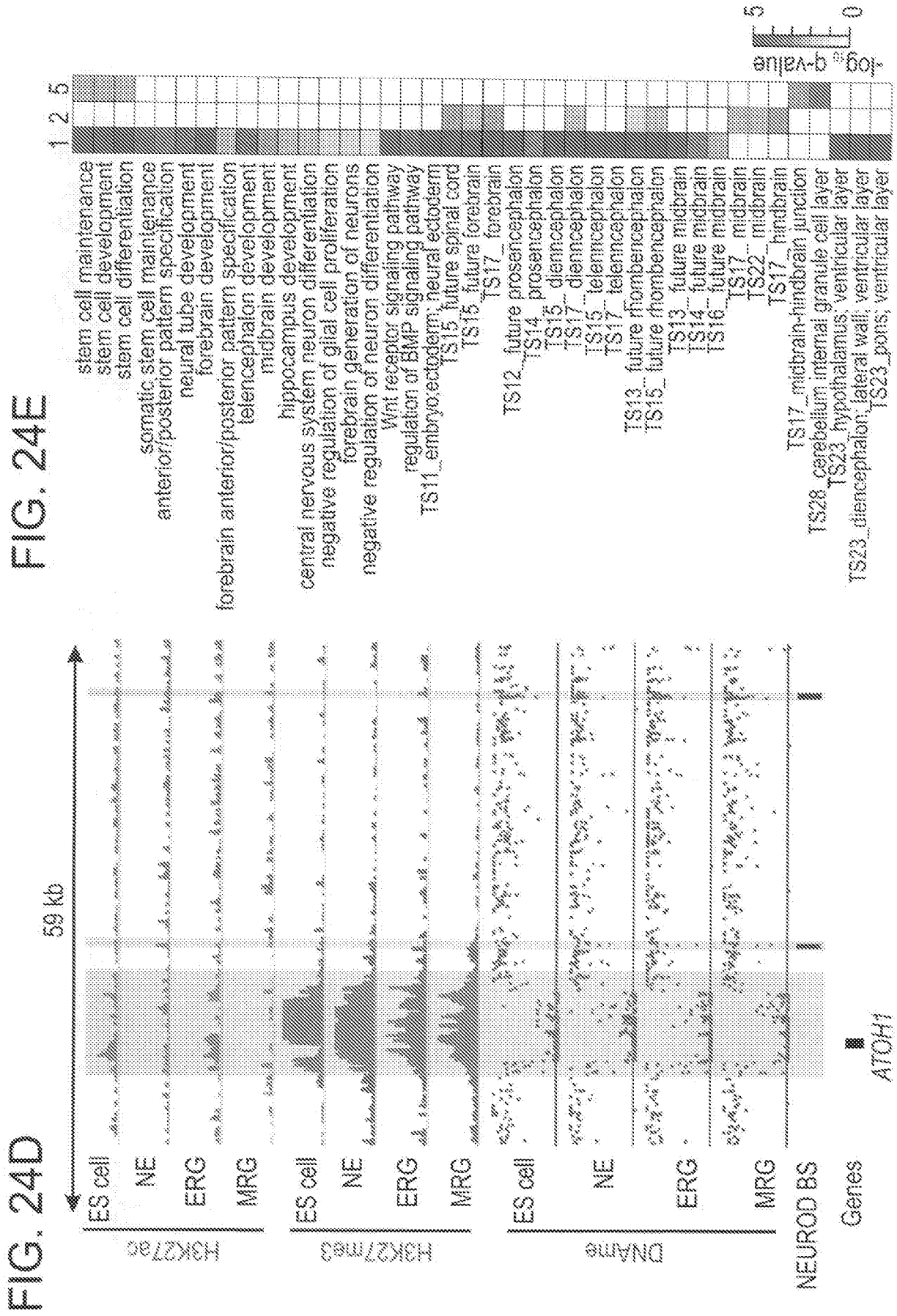


FIG. 24D
59 kb
H3K27ac
H3K27me3
DNAme
ES cell
NE
ERG
MRG
Genes
ATOH1
NEUROD BS
FIG. 24E
1 2 5
stem cell maintenance
stem cell development
stem cell differentiation
somatic stem cell maintenance
anterior/posterior pattern specification
neural tube development
forebrain development
forebrain anterior/posterior pattern specification
telencephalon development
midbrain development
hippocampus development
central nervous system neuron differentiation
negative regulation of glial cell proliferation
forebrain generation of neurons
negative regulation of neuron differentiation
Wnt receptor signaling pathway
regulation of BMP signaling pathway
TS11_embryo;ectoderm; neural ectoderm
TS15_future spinal cord
TS15_future forebrain
TS17_forebrain
TS12_future prosencephalon
TS14_prosencephalon
TS15_dienncephalon
TS17_dienncephalon
TS15_telenncephalon
TS17_telenncephalon
TS13_future rhombencephalon
TS15_future rhombencephalon
TS13_future midbrain
TS14_future midbrain
TS16_future midbrain
TS17_midbrain
TS22_midbrain
TS17_hindbrain
TS17_midbrain-hindbrain junction
TS28_cerebellum internal granule cell layer
TS23_hypothalamus; ventricular layer
TS23_diencephalon; lateral wall; ventricular layer
TS23_pons; ventricular layer
5
0
-log10 q-value

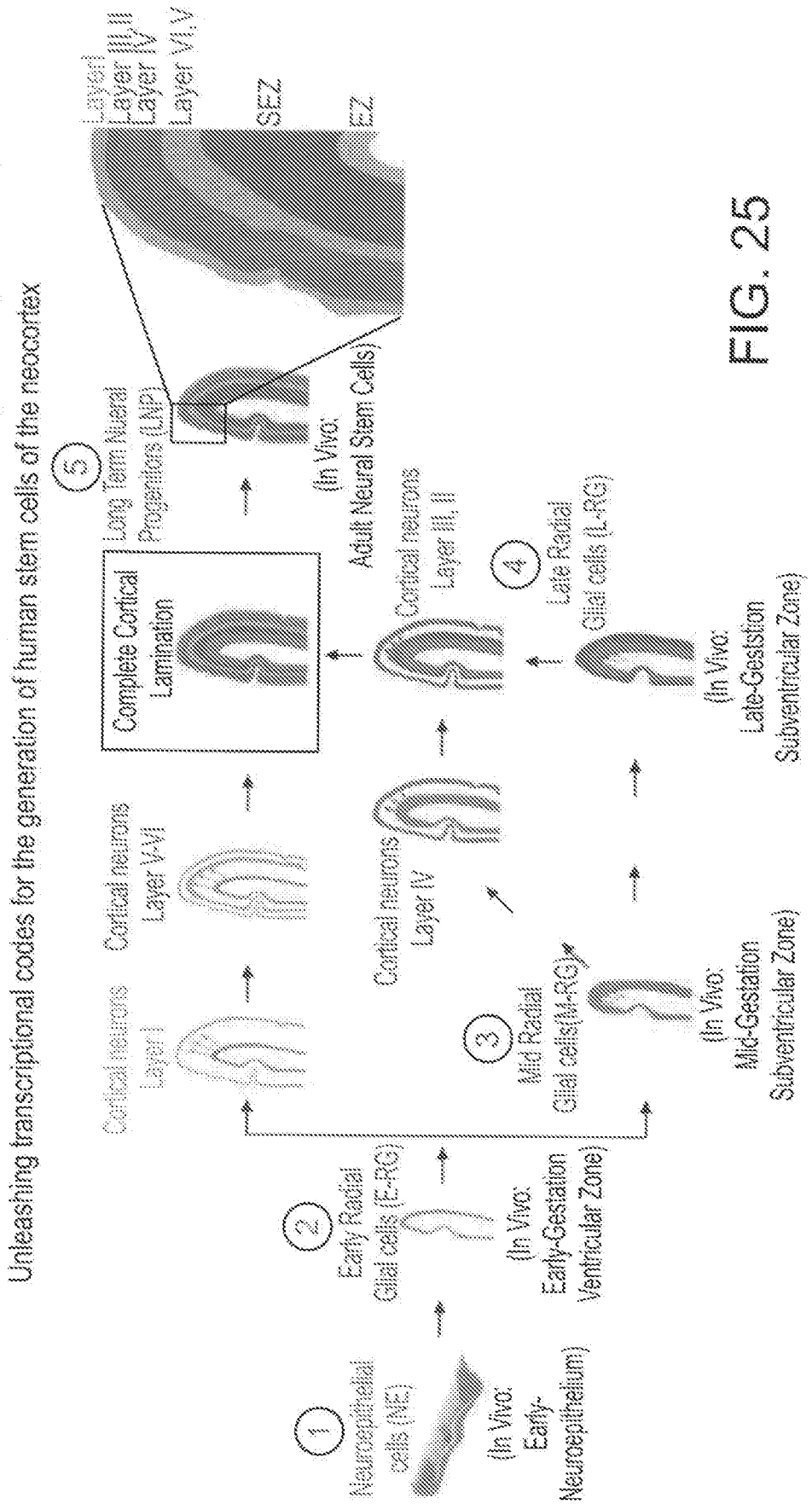

Unleashing transcriptional codes for the generation of human stem cells of the neocortex
1
Neuroepithelial cells (NE)
(In Vivo: Early-Neuroepithelium)
2
Early Radial Glial cells (E-RG)
(In Vivo: Early-Gestation Ventricular Zone)
3
Mid Radial Glial cells(M-RG)
(In Vivo: Mid-Gestation Subventricular Zone)
4
Late Radial Glial cells (L-RG)
(In Vivo: Late-Gestation Subventricular Zone)
5
Long Term Neural Progenitors (LNP)
(In Vivo: Adult Neural Stem Cells)
Cortical neurons Layer I
Cortical neurons Layer V-VI
Cortical neurons Layer IV
Cortical neurons Layer III, II
Complete Cortical Lamination
Layer I
Layer III, II
Layer IV
Layer VI, V
SEZ
EZ

FIG. 25

Formula 1

$$C \sim \text{Poisson}\left(\max\left[1, e_{ij}\right]\lambda_x\right)$$

FIG. 26A

Formula 2

$$e_{ij} = \frac{f\left(C_{ij}+1\right)}{2L_{ij}} + \frac{f\left(C_{ij}+1\right)}{2R_{ij}}$$

FIG.26B

Formula 3

$$P_{ij} = \frac{\left(\sum_K s_{ij}^{C_K} \geq s_{ij}^{C}\right)}{K}$$

FIG. 26C

Formula 4

$$P_i = \frac{\left(\sum_K s_i \geq s_{ij}^{C}\right)}{K}$$

FIG. 26D

PATENT COOPERATION TREATY

PCT

INTERNATIONAL SEARCH REPORT

(PCT Article 18 and Rules 43 and 44)

Applicant's or agent's file reference	**FOR FURTHER ACTION**	see Form PCT/ISA/220 as well as, where applicable, item 5 below.
64732		
International application No.	International filing date *(day/month/year)*	(Earliest) Priority Date *(day/month/year)*
PCT/IL2015/051253	23 December 2015 (23-12-2015)	23 December 2014 (23-12-2014)
Applicant		
RAMOT AT TEL-AVIV UNIVERSITY LTD.		

This international search report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This international search report consists of a total of 7 sheets.

[X] It is also accompanied by a copy of each prior art document cited in this report.

1. **Basis of the report**
   a. With regard to the **language,** the international search was carried out on the basis of:
      [X] the international application in the language in which it was filed
      [ ] a translation of the international application into _____, which is the language of a translation furnished for the purposes of international search (Rules 12.3(a) and 23.1(b))
   b. [ ] This international search report has been established taking into account the **rectification of an obvious mistake** authorized by or notified to this Authority under Rule 91 (Rule 43.6*bis*(a)).
   c. [X] With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, see Box No. I.

2. [ ] **Certain claims were found unsearchable** (See Box No. II)

3. [ ] **Unity of invention is lacking** (see Box No III)

4. With regard to the **title,**
   [X] the text is approved as submitted by the applicant
   [ ] the text has been established by this Authority to read as follows:

5. With regard to the **abstract,**
   [X] the text is approved as submitted by the applicant
   [ ] the text has been established, according to Rule 38.2, by this Authority as it appears in Box No. IV. The applicant may, within one month from the date of mailing of this international search report, submit comments to this Authority

6. With regard to the **drawings,**
   a. the figure of the **drawings** to be published with the abstract is Figure No. 1A
      [ ] as suggested by the applicant
      [X] as selected by this Authority, because the applicant failed to suggest a figure
      [ ] as selected by this Authority, because this figure better characterizes the invention
   b. [ ] none of the figures is to be published with the abstract

INTERNATIONAL SEARCH REPORT

International application No.
PCT/IL2015/051253

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:

   a. [ ] forming part of the international application as filed:

      [ ] in the form of an Annex C/ST.25 text file.

      [ ] on paper or in the form of an image file.

   b. [ ] furnished together with the international application under PCT Rule 13*ter*.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.

   c. [X] furnished subsequent to the international filing date for the purposes of international search only:

      [X] in the form of an Annex C/ST.25 text file (Rule 13*ter*.1(a)).

      [ ] on paper or in the form of an image file (Rule 13*ter*.1(b) and Administrative Instructions, Section 713).

2. [X] In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No
PCT/IL2015/051253

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12N5/0793 C12N5/079
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal, WPI Data, BIOSIS, EMBASE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	PENG JIANG ET AL: "hESC-derived Olig2+ progenitors generate a subtype of astroglia with protective effects against ischaemic brain injury", NATURE COMMUNICATIONS, vol. 4, 23 July 2013 (2013-07-23), XP055275855, DOI: 10.1038/ncomms3196 abstract ----- -/--	1,21,23

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
3 June 2016	09/06/2016

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Loubradou-Bourges, N

3

Form PCT/ISA/210 (second sheet) (April 2005)

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WOO SUN-MI ET AL: "Notch signaling is required for maintaining stem-cell features of neuroprogenitor cells derived from human embryonic stem cells", BMC NEUROSCIENCE, BIOMED CENTRAL, LONDON, GB, vol. 10, no. 1, 17 August 2009 (2009-08-17), page 97, XP021056519, ISSN: 1471-2202, DOI: 10.1186/1471-2202-10-97 abstract; figures 1,2	2-19,24, 46-48
X	STUART M CHAMBERS ET AL: "Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling", NATURE BIOTECHNOLOGY, vol. 27, no. 3, 1 March 2009 (2009-03-01), pages 275-280, XP055007827, ISSN: 1087-0156, DOI: 10.1038/nbt.1529 abstract figure 1	39,40
X	BEATA SURMACZ ET AL: "Directing Differentiation of Human Embryonic Stem Cells Toward Anterior Neural Ectoderm Using Small Molecules", STEM CELLS, vol. 30, no. 9, 1 September 2012 (2012-09-01), pages 1875-1884, XP055059808, ISSN: 1066-5099, DOI: 10.1002/stem.1166 abstract; figure 1	39,40
X	ELKABETZ YECHIEL ET AL: "Human ES cell-derived neural rosettes reveal a functionally distinct early neural stem cell stage", GENES AND DEVELOPMENT, COLD SPRING HARBOR LABORATORY PRESS, PLAINVIEW, NY, US, vol. 22, no. 2, 1 January 2008 (2008-01-01), pages 152-165, XP002563720, ISSN: 0890-9369, DOI: 10.1101/GAD.1616208 cited in the application abstract figure 6	1-38, 41-48

-/--

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	P. NOISA ET AL: "Notch signaling regulates the differentiation of neural crest from human pluripotent stem cells", JOURNAL OF CELL SCIENCE, vol. 127, no. 9, 25 February 2014 (2014-02-25), pages 2083-2094, XP055275213, GB ISSN: 0021-9533, DOI: 10.1242/jcs.145755 abstract figure 1	1-38, 41-48
X	MIRZA PELJTO ET AL: "Programming embryonic stem cells to neuronal subtypes", CURRENT OPINION IN NEUROBIOLOGY, vol. 21, no. 1, 20 October 2010 (2010-10-20), pages 43-51, XP028280690, ISSN: 0959-4388, DOI: 10.1016/J.CONB.2010.09.012 [retrieved on 2010-09-28] abstract; figure 2	1-38, 41-48
X	GASPARD N ET AL: "Mechanisms of neural specification from embryonic stem cells", CURRENT OPINION IN NEUROBIOLOGY, LONDON, GB, vol. 20, no. 1, 1 February 2010 (2010-02-01), pages 37-43, XP026915209, ISSN: 0959-4388, DOI: 10.1016/J.CONB.2009.12.001 [retrieved on 2010-01-14] abstract; figures 1-2	1-38, 41-48
A	CAROLIN MUSSMANN ET AL: "HES5 Is a Key Mediator of Wnt-3a-Induced Neuronal Differentiation", STEM CELLS AND DEVELOPMENT, vol. 23, no. 12, 15 June 2014 (2014-06-15) , pages 1328-1339, XP055275342, NL ISSN: 1547-3287, DOI: 10.1089/scd.2013.0557 abstract figures 1-6	1-48

-/--

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	PAOLO MALATESTA ET AL: "Radial glia and neural stem cells", CELL AND TISSUE RESEARCH, SPRINGER, BERLIN, DE, vol. 331, no. 1, 11 September 2007 (2007-09-11), pages 165-178, XP019564032, ISSN: 1432-0878 abstract; figures 1-5	1-48
A	WO 2006/095175 A1 (GH THE UNIV EDINGBUR [GB]; LOWELL SARAH ELIZABETH [GB]; SMITH AUSTIN G) 14 September 2006 (2006-09-14) claims 1-14 page 7, line 11 - line 20	1-48
X,P	MICHAEL J. ZILLER ET AL: "Dissecting neural differentiation regulatory networks through epigenetic footprinting", NATURE, vol. 518, no. 7539, 24 December 2014 (2014-12-24), pages 355-359, XP055275479, United Kingdom ISSN: 0028-0836, DOI: 10.1038/nature13990 the whole document	1-48
X,P	REUVEN EDRI ET AL: "Analysing human neural stem cell ontogeny by consecutive isolation of Notch active neural progenitors", NATURE COMMUNICATIONS, vol. 6, 23 March 2015 (2015-03-23), page 6500, XP055275483, DOI: 10.1038/ncomms7500 the whole document	1-48

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/IL2015/051253

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2006095175 A1	14-09-2006	EP 1877546 A1	16-01-2008
		JP 2008532509 A	21-08-2008
		US 2008248568 A1	09-10-2008
		WO 2006095175 A1	14-09-2006