FIGURE LEGENDS

Figure 1. Successful gene targeting of GFP cassette into the Olig2 locus in the hESC line BG01. Homologous recombination was performed using a GFP vector that targets the Olig2 gene in BG01 (A). Pink line represents the homology arms and two light blue boxes represent the two exons of Olig2. EGFP cassette is shown in green. Neo (dark blue box) and Tk2 (yellow box) cassettes are designed for positive and negative selection respectively. Restriction enzyme *HindIII* (abbreviated as "H" in panel A) was used to digest the genomic DNA of hESC clones for Southern blot analysis. Six clones were identified to be correctly targeted to the Olig2 locus by Southern blot using a 5' flanking probe. R-Olig2, the representative clone is shown with a 6.2 kb band of targeted allele and a 3.9kb band for the wild type allele, as predicted (B). Floxed neo cassette was further removed by transient transfection of a Cre construct (C) and displayed absence of genomic integration of Cre fragment by PCR (D). To examine GFP expression, R-Olig2 cells were induced by Shh agonist and retinoic acid for 16 days before differentiated spheres were fixed, sectioned and stained with Olig2 antibody (E). All cells that were labeled with Olig2 antibody co-expressed GFP (F, G), which was directly visualized under fluorescence microscope. Panel H shows higher magnification of the boxed area in G. Bar, 50 µm.

Figure 2. GFP expression reflects endogenous Olig2 expression along the neural differentiation pathway. R-Olig2 cells were induced with Shh agonist and retinoic acid and the expression of transcription factors Olig2, Pax6 (A, D, G, J), HB9 (B, E, H, K)

and Nkx2.2 (C, F, I, L) was examined by immunocytochemistry along with GFP expression (direct visualization under fluorescence microscope) at d10, d16, d22 and d26 of differentiation. Pax6 was first detected at d10 (red, A) and starting at d16, a subset of Pax6+ cells co-expressed Olig2 and GFP (arrowheads, D, G, J). Although Olig2 and GFP expression was detected as early as d10, almost no expression was detected for motoneuron marker HB9 (B), until d19 of differentiation (arrow, E), when the majority of cells did not co-express HB9 and Olig2. Some HB9+ cells co-expressed GFP but not Olig2 (arrowheads, E, H, K). Nkx2.2 did not co-label with Olig2 or GFP at the early stages of differentiation (arrow, C, F), and the initiation of co-expression of Nkx2.2 and Olig2 and GFP could be observed starting from d22 (arrowheads, I, L). Insets show higher magnification images of boxed areas in corresponding panels. Detailed time course images are shown in Supplementary Figures 2-4. Bar, 50 μm.

Figure 3. Early GFP-expressing cells are likely motoneuron precursors which have the ability to make synapses. Twenty-one-day differentiated GFP+ spheres were harvested and co-cultured with C2C12 myoblasts for 5-10 days. GFP+ cells were able to cause acetylcholine receptor aggregation, as detected by staining using an antibody against α-BTX (arrows, panels A and D, B and E, and C and D show aggregations of three representative fields at the neuromuscular junction). Although cell bodies of the motoneurons were still weakly visible under green fluorescence microscopy, GFP expression in most neuron processes was downregulated as cells matured. Antibodies of GFP (A-C) and βIII tubulin (blue, D-F) were used to delineate the cell bodies and processes of these GFP+ cell-derived motoneurons. Images of *z* stacks were taken at 1

μm increments and A'-F' and A''-F'' are *xz* and *yz* views processed by Zeiss Axiovision software. GFP Ab: GFP antibody. Bar, 25 μm.

Figure 4. Late stage GFP+ cells are likely glial precursors and give rise to oligodendrocytes and astrocytes. Thirty-five-day differentiated GFP+ spheres were harvested and seeded on fibronectin and further induced by PDGFAA, T3 and Shh agonist for 5 days. Cells expressed GFP (B, E, direct fluorescence) and glial progenitor markers A2B5 (A) and PDGFRα (D). Merged images (C, F) show co-expression. When further induced, GFP+ cells (H, direct fluorescence) gave rise to GalC+ oligodendrocytes (G-I) and GFAP+ astrocytes (J-L). While a small subpopulation of GFAP+ cells still expressed GFP, which were visualized directly by fluorescence microscopy (Supplementary Figure 5), in most GFAP+ astrocytes, GFP expression was revealed by antibody staining (K, L). GFP Ab: GFP antibody. Insets in C and L show higher magnification images of the boxed area. Bar, 50 μm.

Figure 5. Early and late stage GFP+ cells show distinct gene expression profile. GFP+ cells derived from R-Olig2 at early (day 17) and late (day 38) stages were purified by FACS (A) for gene expression profile analysis. Cells maintained GFP expression after sorting (B, direct fluorescence) and showed distinct global gene expression profiles. When compared to undifferentiated R-Olig2 ES cells, GFP+ cells from early stage showed high expression of genes of the motoneuron lineage (C), while those from a later stage showed an upregulation of genes of the oligodendrocyte lineage (D). In the heatmaps, high expression levels relative to mean values (average signals of ES, d17 and

d38 of the corresponding genes) are colored red. Low expression levels are colored green. Black represents no significant change in the expression level between mean and sample. Fold increase of gene expression level of d17 and d38 GFP+ cells compared to undifferentiated R-Olig2 ES cells is listed in the tables in C and D.

Figure 6. Engrafted late stage GFP+ cells migrated and differentiated into oligodendrocyte progenitors and oligodendrocytes after transplantation into the rat spinal cords. Engrafted GFP+ cells survived, migrated and integrated into the host white matter regions of the cervical spinal cord robustly as identified by GFP (A) and an antibody against human nuclear antigen (hNA) (arrows, R-U). Specifically, at 2 weeks posttransplantation, engrafted cells co-expressed GFP and oligodendrocyte precursor markers PDGFRα (arrow, B-E) and NG2 (arrow, F-I). By 6 weeks posttransplantation, cells continued to express GFP, migrated further from the injection site (*, A) and started to myelinate the host tissue, as evidenced by co-expression of GFP and PLP (arrow, J-M"), and GFP and CNP (arrow, N-Q"). Some transplanted cells differentiated into astrocytes as detected by GFAP staining (V, inset shows higher magnification of the boxed area). For panels J-Q", *z* stack images were taken at 1 μm increments. M' and Q' and M''-Q'' are images of *xz* and *yz* views processed by Zeiss Axiovision software. GFP was directly visualized under fluorescence microscopy. Bar in A, 500 μm, B-E and J-V, 50 μm, F-I, 20 μm.

Supplementary Figure 1. Olig2-GFP knockin hESC line R-Olig2 maintains the expression of pluripotency markers and differentiates into three germ layers. R-Olig2

cells maintained a typical hESC morphology (A) and the expression of pluripotency markers such as Oct4(B), SSEA4 (C), and Tra1-60 (D). They did not express GFP at the undifferentiated stage (A). R-Olig2 retained the capacity of differentiating into βIII tubulin (E) or nestin expressing ectodermal cells (F), smooth muscle actin (SMA)+ mesodermal cells (G) and α-fetoprotein (AFP, H) expressing endodermal cells. Identical to its parental counterpart, R-Olig2 cells maintained a normal male karyotype (I). Bar, 100 μm, Bar in A, 50 μm.

Supplementary Figure 2. Examination of Pax6 expression by immunocytochemistry along the R-Olig2 differentiation time course. Cells cultured in suspension at different time points (day 10, 16, 22 and 26) were harvested, fixed, embedded, sectioned and stained for Pax6 (red) and Olig2 (blue). Merged images of Pax6, Olig2 and direct GFP fluorescence are shown in the column on the right. Bar, 50 μm.

Supplementary Figure 3. Examination of HB9 expression by immunocytochemistry along the R-Olig2 differentiation time course. Cells cultured in suspension at different time points (day 10, 19, 22 and 26) were harvested, fixed, embedded, sectioned and stained for HB9 (red) and Olig2 (blue). Merged images of HB9, Olig2 and direct GFP fluorescence are shown in the column on the right. Bar, 50 μm.

Supplementary Figure 4. Examination of Nkx2.2 expression by immunocytochemistry along the R-Olig2 differentiation time course. Cells cultured in suspension at different time points (day 16, 22 and 26) were harvested, fixed, embedded, sectioned and stained

for Nkx2.2 (red) and Olig2 (blue). Merged images of Nkx2.2, Olig2 and direct GFP fluorescence are shown in the column on the right. Bar, 50 µm.

Supplementary Figure 5. Residual GFP expression could be observed when cells differentiated into GFAP+ astrocytes. When late stage GFP+ cells were differentiated into GFAP+ astrocytes, a small subpopulation of GFAP+ cells still expressed GFP, which were visualized directly by fluorescence microscopy (arrow). Bar, 50 µm.

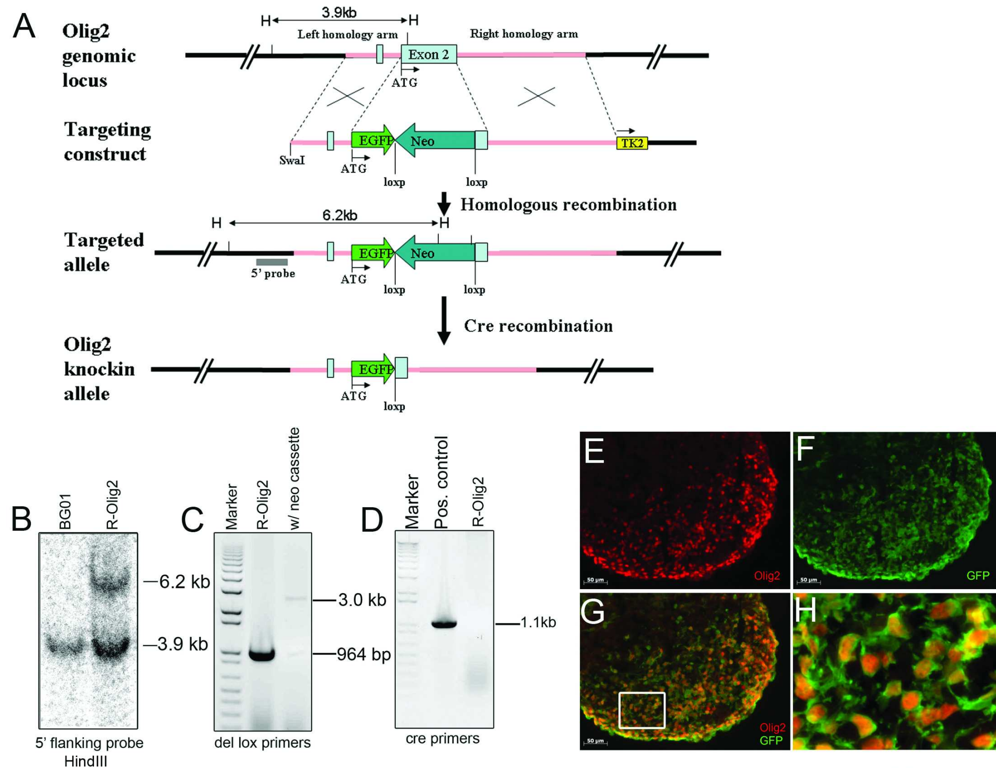


Figure 1 Liu

Figure 1. Successful gene targeting of GFP cassette into the Olig2 locus in the hESC line BG01. Homologous recombination was performed using a GFP vector that targets the Olig2 gene in BG01 (A). Pink line represents the homology arms and two light blue boxes represent the two exons of Olig2. EGFP cassette is shown in green. Neo (dark blue box) and Tk2 (yellow box) cassettes are designed for positive and negative selection respectively. Restriction enzyme HindIII (abbreviated as "H" in panel A) was used to digest the genomic DNA of hESC clones for Southern blot analysis. Six clones were identified to be correctly targeted to the Olig2 locus by Southern blot using a 5' flanking probe. R-Olig2, the representative clone is shown with a 6.2 kb band of targeted allele and a 3.9kb band for the wild type allele, as predicted (B). Floxed neo cassette was further removed by transient transfection of a Cre construct (C) and displayed absence of genomic integration of Cre fragment by PCR (D). To examine GFP expression, R-Olig2 cells were induced by Shh agonist and retinoic acid for 16 days before differentiated spheres were fixed, sectioned and stained with Olig2 antibody (E). All cells that were labeled with Olig2 antibody co-expressed GFP (F, G), which was directly visualized under fluorescence microscope. Panel H shows higher magnification of the boxed area in G. Bar, 50 μm.

186x149mm (300 x 300 DPI)

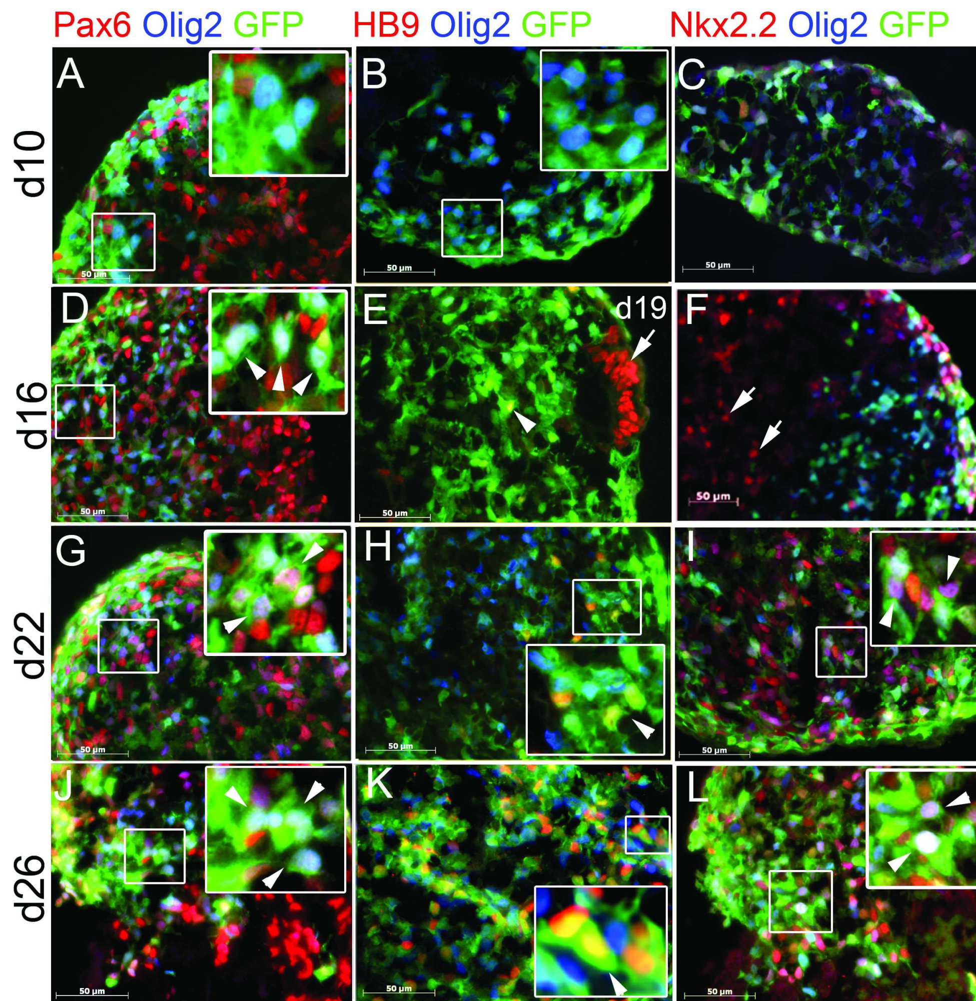


Figure 2 Liu

Figure 2. GFP expression reflects endogenous Olig2 expression along the neural differentiation pathway. R-Olig2 cells were induced with Shh agonist and retinoic acid and the expression of transcription factors Olig2, Pax6 (A, D, G, J), HB9 (B, E, H, K) and Nkx2.2 (C, F, I, L) was examined by immunocytochemistry along with GFP expression (direct visualization under fluorescence microscope) at d10, d16, d22 and d26 of differentiation. Pax6 was first detected at d10 (red, A) and starting at d16, a subset of Pax6+ cells co-expressed Olig2 and GFP (arrowheads, D, G, J). Although Olig2 and GFP expression was detected as early as d10, almost no expression was detected for motoneuron marker HB9 (B), until d19 of differentiation (arrow, E), when the majority of cells did not co-express HB9 and Olig2. Some HB9+ cells co-expressed GFP but not Olig2 (arrowheads, E, H, K). Nkx2.2 did not co-label with Olig2 or GFP at the early stages of differentiation (arrow, C, F), and the initiation of co-expression of Nkx2.2 and Olig2 and GFP could be observed starting from d22 (arrowheads, I, L). Insets show higher magnification images of boxed areas in corresponding panels. Detailed time course images are shown in Supplementary Figures 2-4. Bar, 50 μm.
202x212mm (300 x 300 DPI)

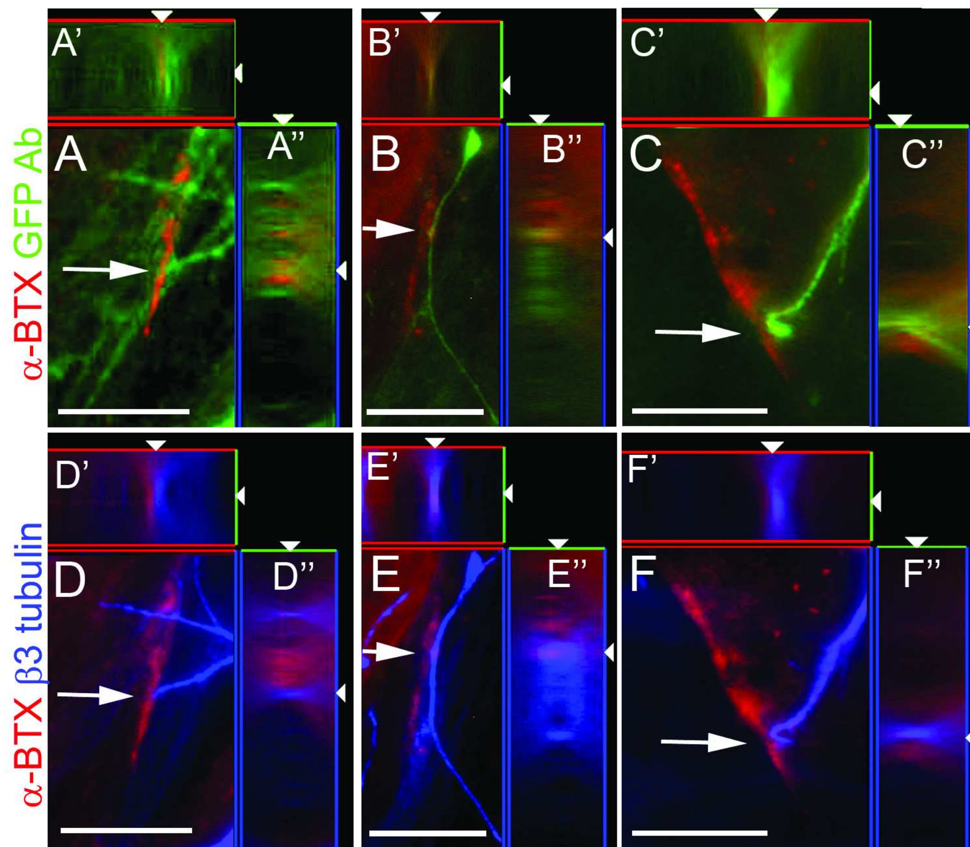


Figure 3. Early GFP-expressing cells are likely motoneuron precursors which have the ability to make synapses. Twenty-one-day differentiated GFP+ spheres were harvested and co-cultured with C2C12 myoblasts for 5-10 days. GFP+ cells were able to cause acetylcholine receptor aggregation, as detected by staining using an antibody against α-BTX (arrows, panels A and D, B and E, and C and D show aggregations of three representative fields at the neuromuscular junction). Although cell bodies of the motoneurons were still weakly visible under green fluorescence microscopy, GFP expression in most neuron processes was downregulated as cells matured. Antibodies of GFP (A-C) and βIII tubulin (blue, D-F) were used to delineate the cell bodies and processes of these GFP+ cell-derived motoneurons. Images of z stacks were taken at 1 μm increments and A′-F′ and A′′-F′′ are xz and yz views processed by Zeiss Axiovision software. GFP Ab: GFP antibody. Bar, 25 μm.
129x116mm (300 x 300 DPI)

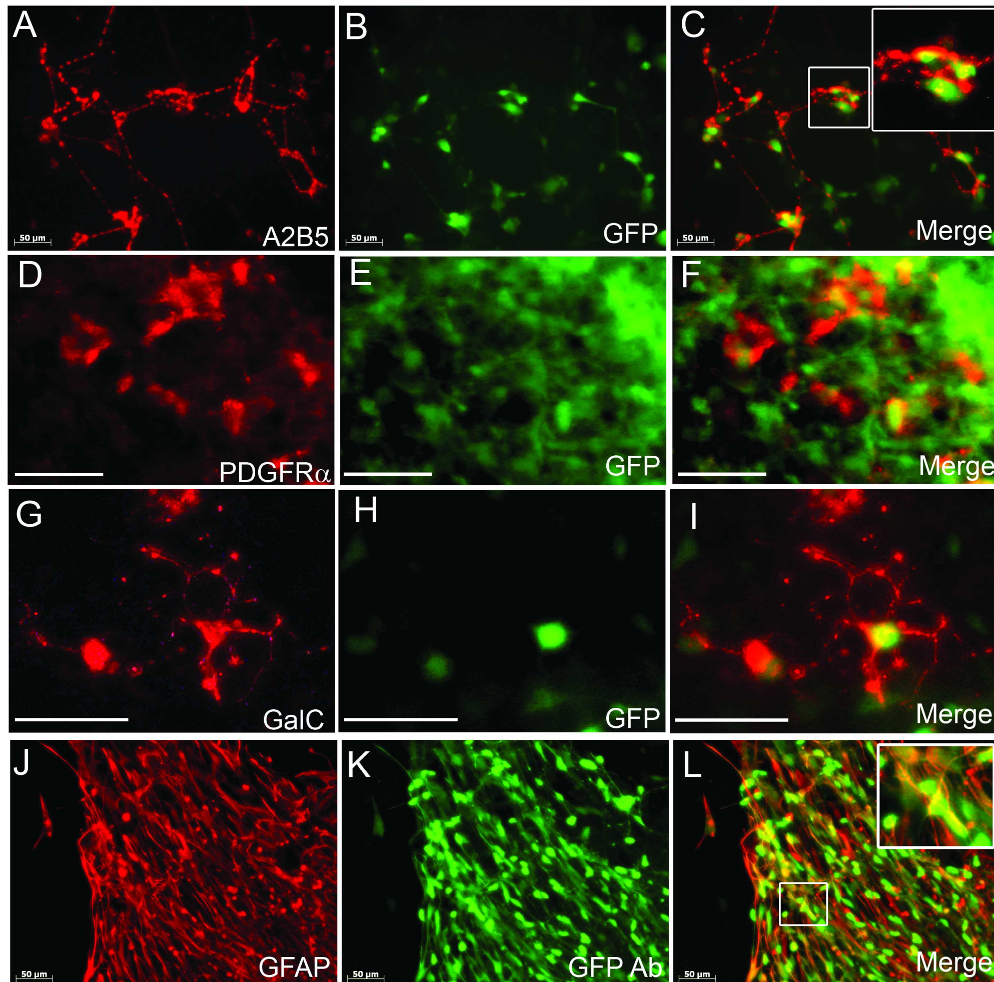
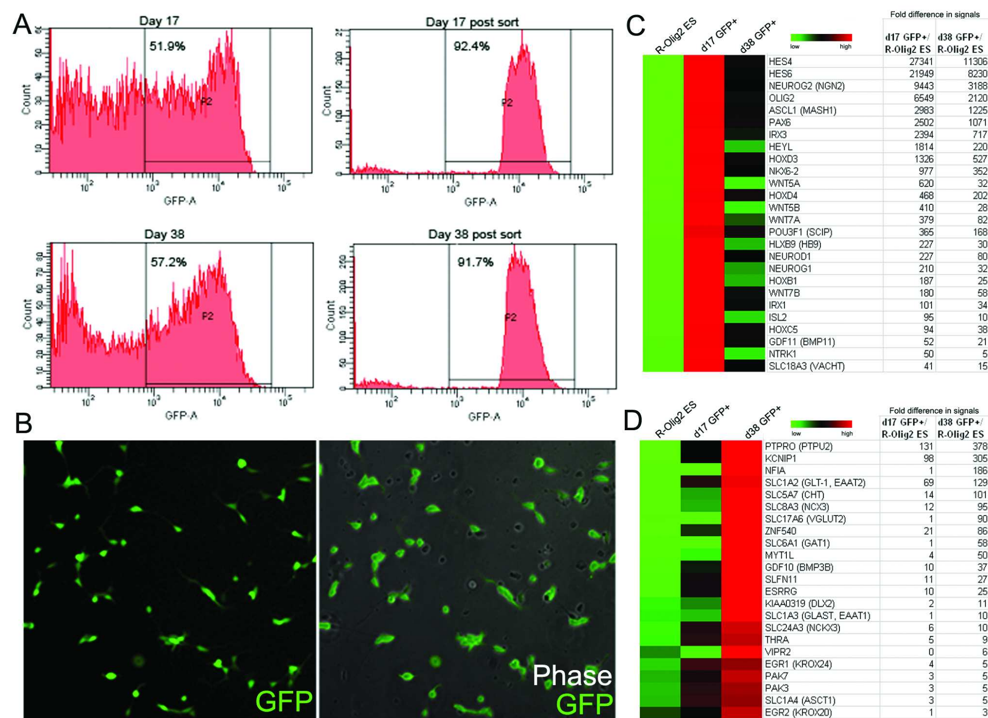


Figure 4. Late stage GFP+ cells are likely glial precursors and give rise to oligodendrocytes and astrocytes. Thirty-five-day differentiated GFP+ spheres were harvested and seeded on fibronectin and further induced by PDGFAA, T3 and Shh agonist for 5 days. Cells expressed GFP (B, E, direct fluorescence) and glial progenitor markers A2B5 (A) and PDGFRα (D). Merged images (C, F) show co-expression. When further induced, GFP+ cells (H, direct fluorescence) gave rise to GalC+ oligodendrocytes (G-I) and GFAP+ astrocytes (J-L). While a small subpopulation of GFAP+ cells still expressed GFP, which were visualized directly by fluorescence microscopy (Supplementary Figure 5), in most GFAP+ astrocytes, GFP expression was revealed by antibody staining (K, L). GFP Ab: GFP antibody. Insets in C and L show higher magnification images of the boxed area. Bar, 50 μm.
233x238mm (300 x 300 DPI)



C

	Fold difference in signals	
	d17 GFP+/ R-Olig2 ES	d38 GFP+/ R-Olig2 ES
HES4	27341	11306
HES6	21949	8230
NEUROG2 (NGN2)	9443	3188
OLIG2	6549	2120
ASCL1 (MASH1)	2983	1225
PAX6	2502	1071
IRX3	2394	717
HEYL	1814	220
HOXD3	1326	527
NKX6-2	977	352
WNT5A	620	32
HOXD4	468	202
WNT5B	410	28
WNT7A	379	82
POU3F1 (SCIP)	365	168
HLXB9 (HB9)	227	30
NEUROD1	227	80
NEUROG1	210	32
HOXB1	187	25
WNT7B	180	58
IRX1	101	34
ISL2	95	10
HOXC5	94	38
GDF11 (BMP11)	52	21
NTRK1	50	5
SLC18A3 (VACHT)	41	15

D

	Fold difference in signals	
	d17 GFP+/ R-Olig2 ES	d38 GFP+/ R-Olig2 ES
PTPRO (PTPU2)	131	378
KCNIP1	98	305
NFIA	1	186
SLC1A2 (GLT-1, EAAT2)	69	129
SLC5A7 (CHT)	14	101
SLC8A3 (NCX3)	12	95
SLC17A6 (VGLUT2)	1	90
ZNF540	21	86
SLC6A1 (GAT1)	1	58
MYT1L	4	50
GDF10 (BMP3B)	10	37
SLFN11	11	27
ESRRG	10	25
KIAA0319 (DLX2)	2	11
SLC1A3 (GLAST, EAAT1)	1	10
SLC24A3 (NCKX3)	6	10
THRA	5	9
VIPR2	0	6
EGR1 (KROX24)	4	5
PAK7	3	5
PAK3	3	5
SLC1A4 (ASCT1)	3	5
EGR2 (KROX20)	1	3

Figure 5 Liu

Figure 5. Early and late stage GFP+ cells show distinct gene expression profile. GFP+ cells derived from R-Olig2 at early (day 17) and late (day 38) stages were purified by FACS (A) for gene expression profile analysis. Cells maintained GFP expression after sorting (B, direct fluorescence) and showed distinct global gene expression profiles. When compared to undifferentiated R-Olig2 ES cells, GFP+ cells from early stage showed high expression of genes of the motoneuron lineage (C), while those from a later stage showed an upregulation of genes of the oligodendrocyte lineage (D). In the heatmaps, high expression levels relative to mean values (average signals of ES, d17 and d38 of the corresponding genes) are colored red. Low expression levels are colored green. Black represents no significant change in the expression level between mean and sample. Fold increase of gene expression level of d17 and d38 GFP+ cells compared to undifferentiated R-Olig2 ES cells is listed in the tables in C and D.
174x132mm (300 x 300 DPI)

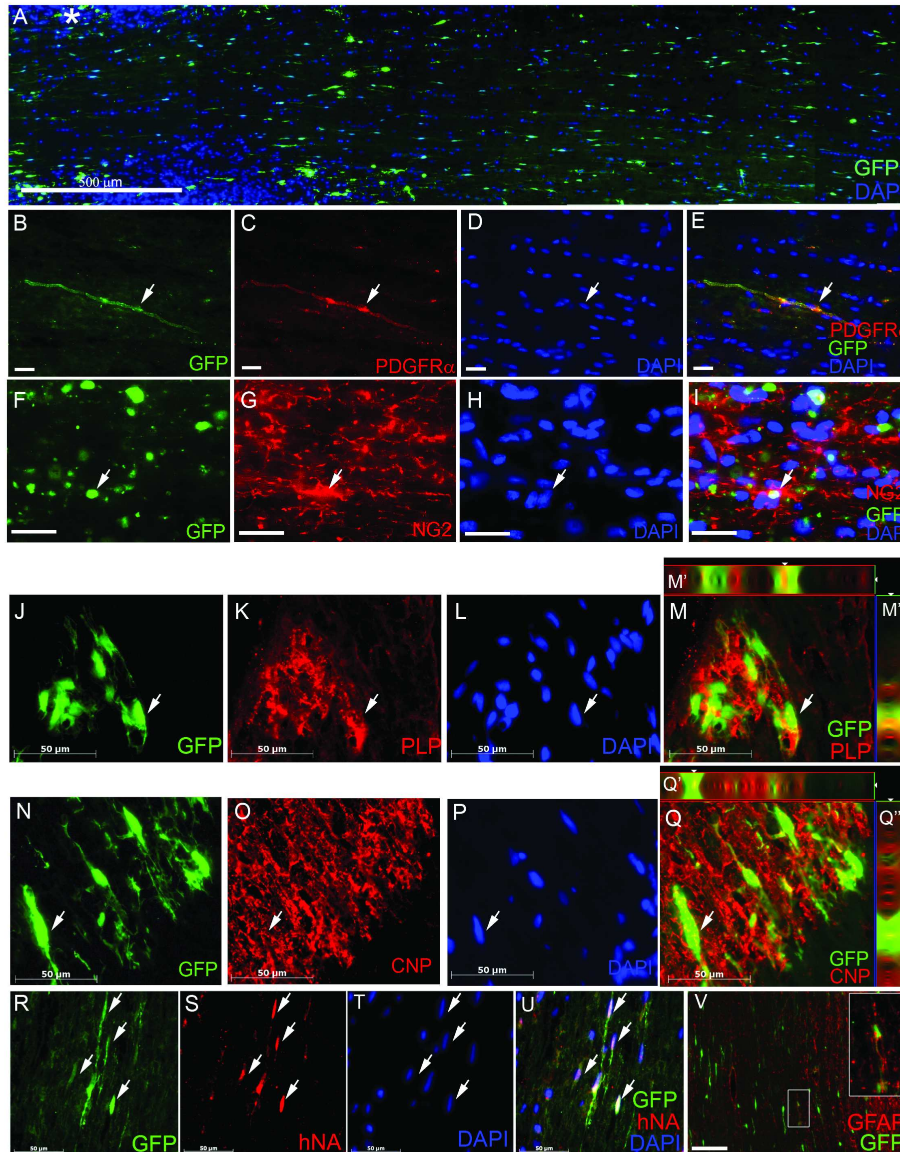
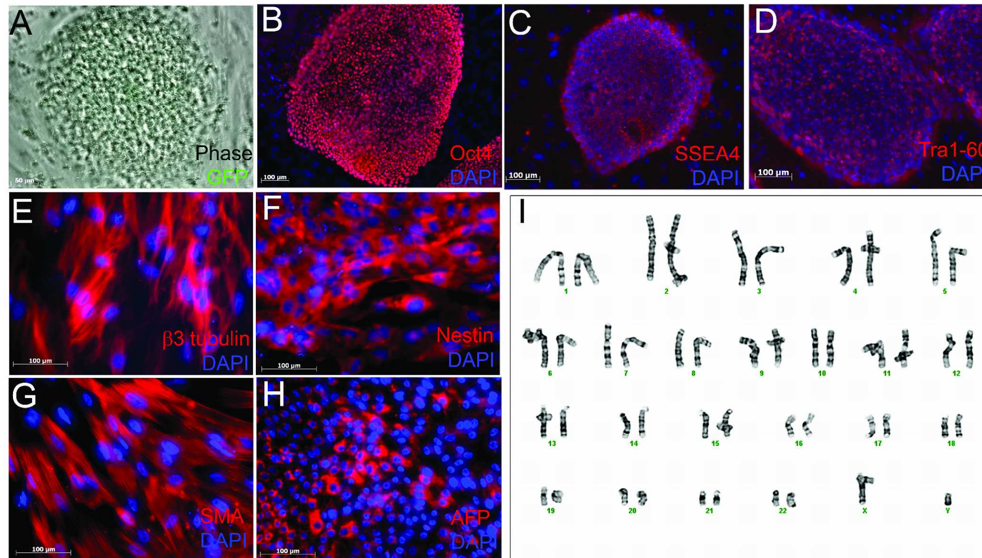


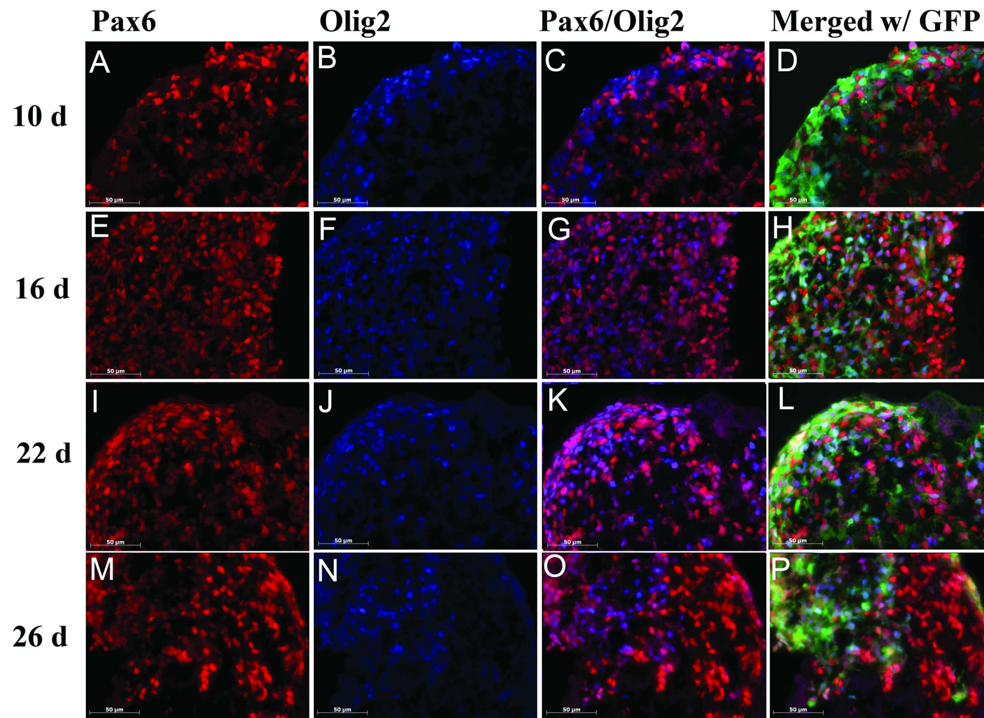
Figure 6 Liu

Figure 6. Engrafted late stage GFP+ cells migrated and differentiated into oligodendrocyte progenitors and oligodendrocytes after transplantation into the rat spinal cords. Engrafted GFP+ cells survived, migrated and integrated into the host white matter regions of the cervical spinal cord robustly as identified by GFP (A) and an antibody against human nuclear antigen (hNA) (arrows, R-U). Specifically, at 2 weeks posttransplantation, engrafted cells co-expressed GFP and oligodendrocyte precursor markers PDGFRα (arrow, B-E) and NG2 (arrow, F-I). By 6 weeks posttransplantation, cells continued to express GFP, migrated further from the injection site (*, A) and started to myelinate the host tissue, as evidenced by co-expression of GFP and PLP (arrow, J-M”), and GFP and CNP (arrow, N-Q”). Some transplanted cells differentiated into astrocytes as detected by GFAP staining (V, inset shows higher magnification of the boxed area). For panels J-Q”, z stack images were taken at 1 µm increments. M’ and Q’ and M’’-Q’’ are images of xz and yz views processed by Zeiss Axiovision software. GFP was directly visualized under fluorescence microscopy.

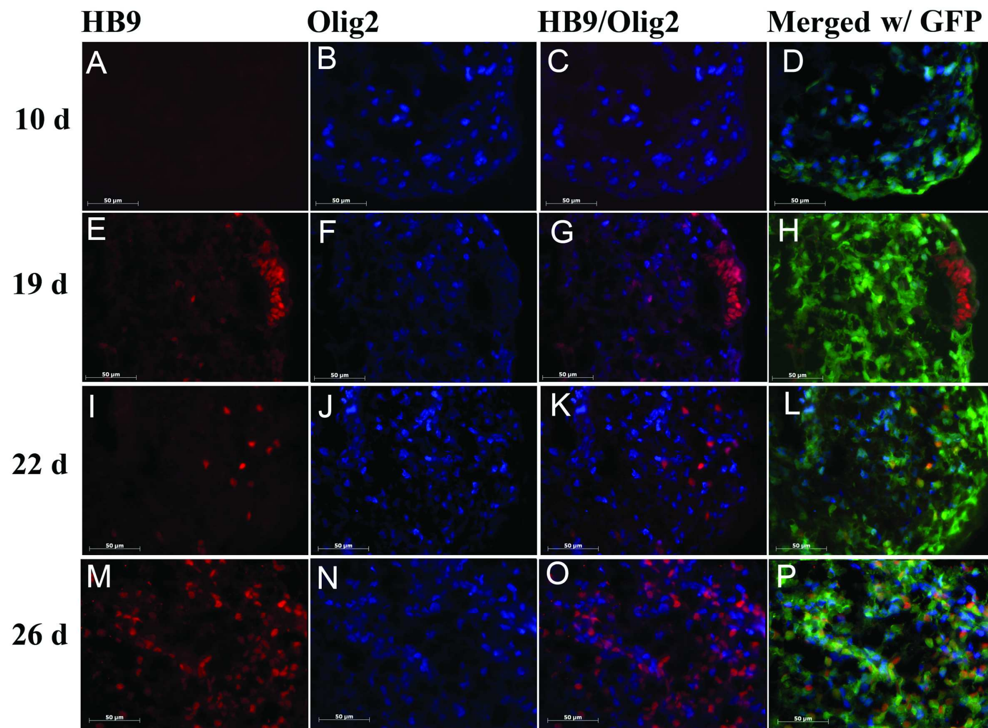
Bar in A, 500 μm, B-E and J-V, 50 μm, F-I, 20 μm.
190x249mm (300 x 300 DPI)



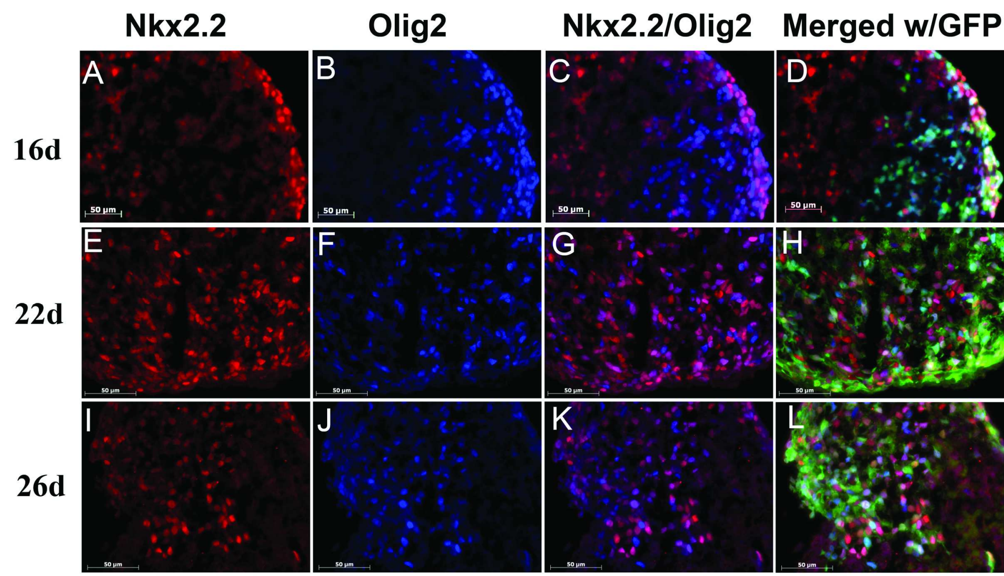
Supplementary Figure 1. Olig2-GFP knockin hESC line R-Olig2 maintains the expression of pluripotency markers and differentiates into three germ layers. R-Olig2 cells maintained a typical hESC morphology (A) and the expression of pluripotency markers such as Oct4(B), SSEA4 (C), and Tra1-60 (D). They did not express GFP at the undifferentiated stage (A). R-Olig2 retained the capacity of differentiating into βIII tubulin (E) or nestin expressing ectodermal cells (F), smooth muscle actin (SMA) + mesodermal cells (G) and α-fetoprotein (AFP, H) expressing endodermal cells. Identical to its parental counterpart, R-Olig2 cells maintained a normal male karyotype (I). Bar, 100 μm, Bar in A, 50 μm.
177x106mm (600 x 600 DPI)



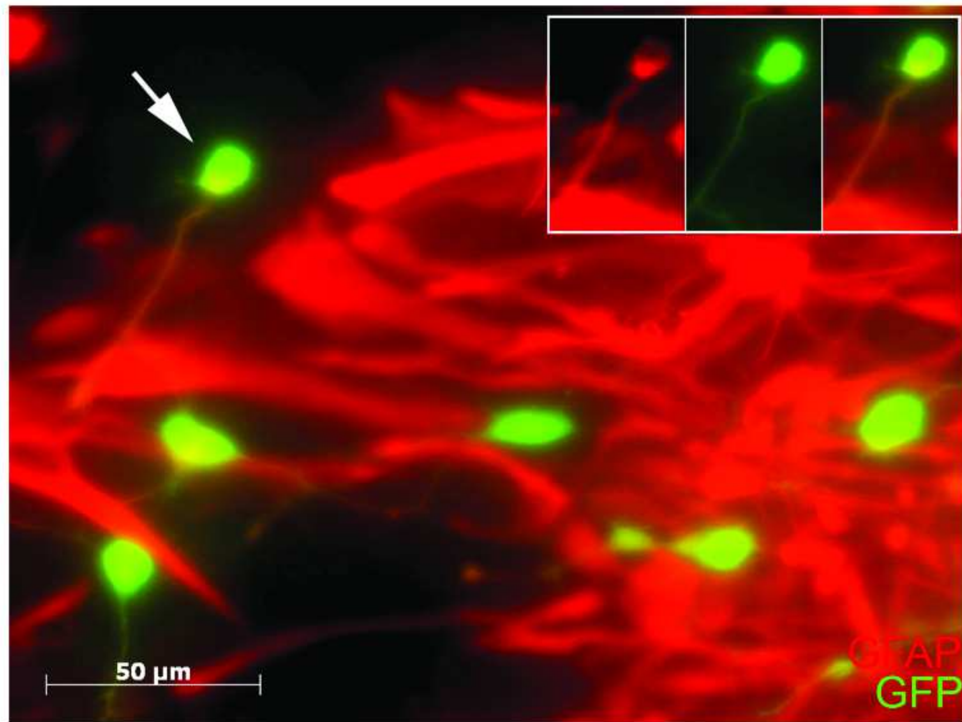
Supplementary Figure 2. Examination of Pax6 expression by immunocytochemistry along the R-Olig2 differentiation time course. Cells cultured in suspension at different time points (day 10, 16, 22 and 26) were harvested, fixed, embedded, sectioned and stained for Pax6 (red) and Olig2 (blue). Merged images of Pax6, Olig2 and direct GFP fluorescence are shown in the column on the right. Bar, 50 μm.
203x153mm (300 x 300 DPI)



Supplementary Figure 3. Examination of HB9 expression by immunocytochemistry along the R-Olig2 differentiation time course. Cells cultured in suspension at different time points (day 10, 19, 22 and 26) were harvested, fixed, embedded, sectioned and stained for HB9 (red) and Olig2 (blue). Merged images of HB9, Olig2 and direct GFP fluorescence are shown in the column on the right. Bar, 50 μm.
203x157mm (300 x 300 DPI)



Supplementary Figure 4. Examination of Nkx2.2 expression by immunocytochemistry along the R-Olig2 differentiation time course. Cells cultured in suspension at different time points (day 16, 22 and 26) were harvested, fixed, embedded, sectioned and stained for Nkx2.2 (red) and Olig2 (blue). Merged images of Nkx2.2, Olig2 and direct GFP fluorescence are shown in the column on the right. Bar, 50 μm.
203x126mm (300 x 300 DPI)



Supplementary Figure 5 Liu

Supplementary Figure 5. Residual GFP expression could be observed when cells differentiated into GFAP+ astrocytes. When late stage GFP+ cells were differentiated into GFAP+ astrocytes, a small subpopulation of GFAP+ cells still expressed GFP, which were visualized directly by fluorescence microscopy (arrow). Bar, 50 μm.
66x54mm (300 x 300 DPI)