Cell, Volume *135*

Supplemental Data

Application of a Translational

Profiling Approach for the Comparative

Analysis of CNS Cell Types

Joseph P. Doyle, Joseph D. Dougherty, Myriam Heiman, Eric F. Schmidt, Tanya R. Stevens, Guojun Ma, Sujata Bupp, Prerana Shrestha, Rajiv D. Shah, Martin L. Doughty, Shiaoching Gong, Paul Greengard, and Nathaniel Heintz

Supplemental Experimental Procedures

Generation of Monoclonal Antibodies

All immunoprecipitations except for the *Drd1* and *Drd2* lines, which used the Goat Anti-EGFP described in the accompanying paper (Heiman et al., 2008), were done using two monoclonal anti-EGFP antibodies (clones 19C8 and 19F7) specifically generated for this purpose at the Monoclonal Antibody Core Facility at Memorial Sloan-Kettering cancer center. Mice were immunized with purified GST-EGFP fusion protein and several rounds of screening were performed to identify clones which functioned well in immunoprecipitation assays. Initially, monoclonal supernatants were tested by ELISA using 96 well plates coated with EGFP purified from transiently transfected 293T cells. Next, positive clones were screened in immunoprecipitation assays, again using the EGFP purified from transfected 293T cells. Finally, we identified four positive clones which strongly immunoprecipitated EGFP from cerebellar lysates from a transgenic mouse line expressing EGFP under the *Pcp2* BAC driver.

Quantification of laminar position of cortical pyramidal cells.

Anti-EGFP immunohistochemistry with DAB was performed on 20 micron sagittal sections from the *Etv1* TS88, *Glt25d2* DU9, and *Ntsr1* TS16 bacTRAP lines as described (Gong et al), and images were acquired as above. Corresponding digital images of the adult sagital sections were downloaded from gensat.org for lines expressing EGFP from the same BACs. Sections containing motor cortex (corresponding to Paxinos section 111) (Paxinos and Franklin, 2001) were imported into ImageJ (rsb.info.nih.gov/ij). The distance from the apical tip of the soma to the pial surface was measured using the 'straight line selection' tool. The apical tip was defined as the site at which the apical dendrite and the cell body converge. Only cells with a clearly visible apical dendrite and a uniformly stained soma were measured. At least 50 cells were measured from each image. All measurements were then converted from pixels to microns using the following scale: 1 pixel = 1.33 micons.

Microarray Normalization and Analysis

MIAME compliant raw data are available from GEO. Analyzed data are available in Table S5.

Replicate array samples were normalized within groups with quantile normalization (GCRMA), and between groups by global normalization to Affymetrix biotinylated spike-in controls. Data were filtered to remove those probesets with low

signal (<50) from analysis, as well as those probesets identified as monoclonal background (Table S2), and replicate samples were averaged. Each IP was then compared to the unbound samples from the same tissue to calculate a ratio of IP/UB as a measure of 'enrichment'. UB samples generally show little or no depletion of cell-specific RNA following IP, and UB samples from several different IPs from the same tissue were averaged for each comparison. UB samples from corpus striatum (*C*hat line) and neostriatum (*Drd1* and *Drd2* lines) were normalized together. IPs were globally normalized to UBs using the Affymetrix biotinylated spike in controls, to correct for any broad biases in scanning and hybridization. For each cell type, Table S5 contains the IP/UB values for all genes with fold change >2 and $p < .05$ by Welch's t-Test, with Benjamini and Hochberg FDR multiple testing correction, as calculated by Genespring GX version 7.3 (Agilent).

For three cell populations, we calculated a further 'corrected enrichment'. For the Bergmann glial line, *Sept4*, which also shows low level expression in mature oligodendrocytes, the corrected enrichment is the intersection of the IP/UB analysis described above with a comparison of the Bergman Glial line with the mature oligodendrocyte line, *Cmtm5*, using the same fold change and statistical criteria applied above. For unipolar brush cell line *Grp*, which also shows expression in some Bergmann glia, the corrected enrichment is the intersection of the IP/UB enrichment described above with a comparison of *Grp* with the Bergman glial line, using the same criteria. Finally, a substantial proportion of the RNA in the cerebellum is generated by granule cells, and thus the UB samples are highly enriched in granule cell RNA. Therefore, to identify granule cell genes, we calculated 'corrected enrichment' by comparing granule cell IP to the average of all other cerebellar cell type IPs, using the same criteria as above.

Hierarchical clustering was performed in Genespring using the 'condition tree' function with a smoothed correlation metric on the GCRMA normalized data for the 20% of probesets with the highest coefficient of variation.

Shannon entropy was calculated in excel from GCRMA normalized values with the following procedure: After excluding probesets with no signal > 100 in at least one sample, normalized expression measurements for each data set were categorized into 5 bins by log base 10 values (1-9, 10-99, 100-999, 1000-9999, 10000-99999), and Shannon entropy was calculated for each gene across 1) just the IP samples, 2) just the UB samples, and 3) across all samples using the following formula (Schneider, 2007).

$$H = -\sum_{i=1}^{M} P_i \log_2 P_i$$

The 10% of probesets with the highest and lowest entropy across all samples were analyzed using the BINGO plugin for the cytoscape software, using the full mouse Gene Ontologies, a p threshold of .01 on hypergeometic tests with Benjamin Hochberg FDR correction for multiple testing (Ashburner et al., 2000; Maere et al., 2005). Results from this analysis are substantially similar to those obtained using the EASE online implementation of Gene Ontologies and EASE statistic (Dennis et al., 2003).

Pearson's correlation with MBP was calculated in Genespring.

Comparative anlaysis of all cell types, and heatmaps (Figure 7), were generated with the R stastistical software. Data were normalized as above. Then, for each cell type the total probeset list was filtered to remove those probesets with signal less than 50, or IP/UB values less than the average plus two standard deviations of the IP/UB values of the relevant negative control genes, or 1, whichever was lesser. Then, with this filtered list of genes, a fold change was calculated for this cell type versus all other samples, iteratively. For each comparison, the fold changes were ranked from highest

to lowest, and these ranks were averaged across comparisons for a cell type. The top one hundred probesets from this average ranking were selected for further analysis.

To assess if the ribosomal immunoprecipitation is biased towards longer transcripts, we plotted signal intensity versus transcript length for all probesets as described in Heiman et al. (Figure S7). No positive correlation between signal intensity and length was detected for any sample.

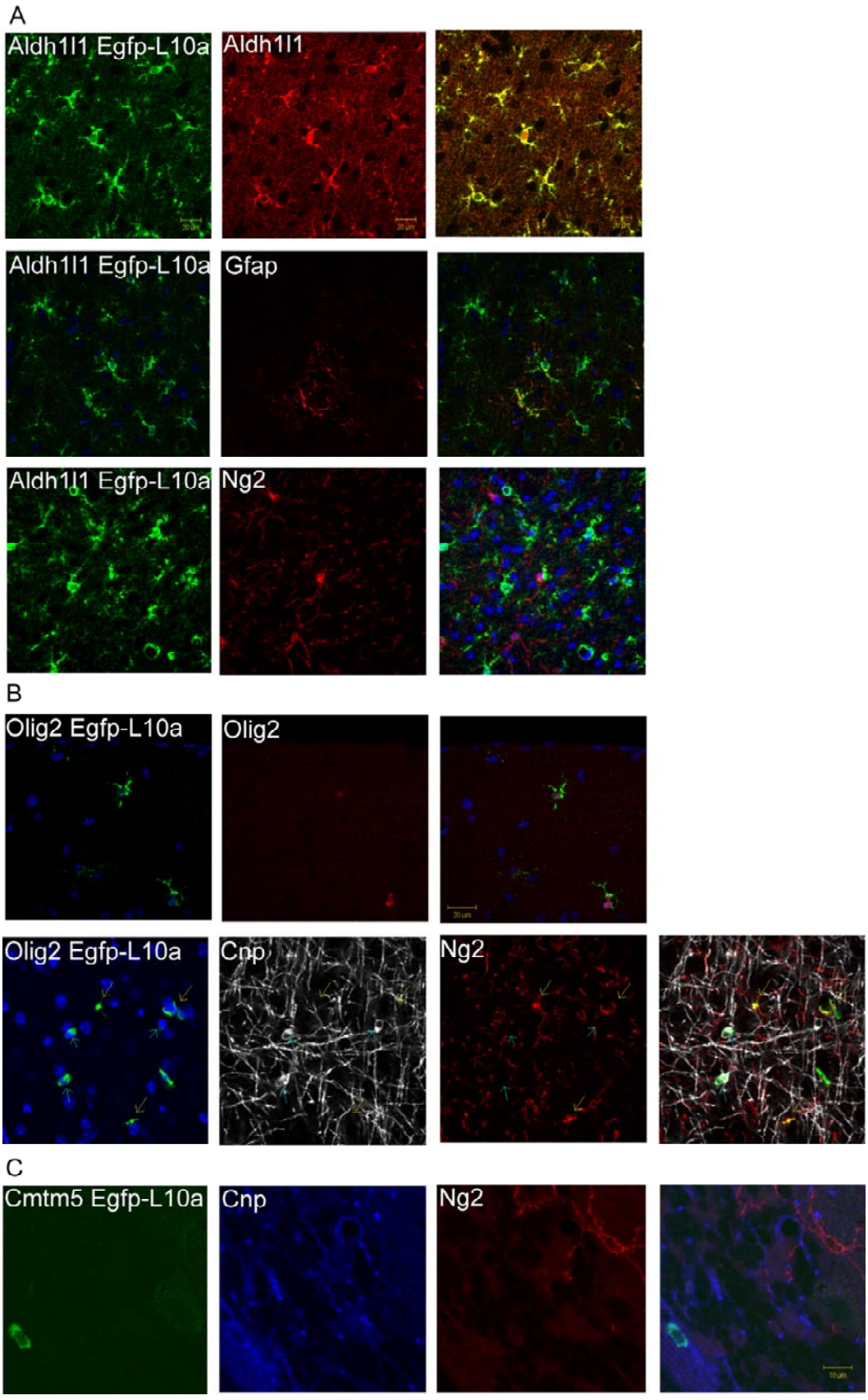

A
Aldh1l1 Egfp-L10a
Aldh1l1
Aldh1l1 Egfp-L10a
Gfap
Aldh1l1 Egfp-L10a
Ng2
B
Olig2 Egfp-L10a
Olig2
Olig2 Egfp-L10a
Cnp
Ng2
C
Cmtm5 Egfp-L10a
Cnp
Ng2

Figure S1. Immunofluorescent characterization of glial cell lines.
bacTRAP lines were created to examine the translational profile of the three major classes of glia across the CNS. A) Astrocytes were targeted using a BAC for the gene *Aldh1l1* that has previously been described as astrocyte specific (Anthony and Heintz, 2007; Cahoy et al., 2008), and all Aldh1l1 positive cells express the EGPP-L10a transgene (Row 1). EGFP-L10a was detected in both Gfap+ (reactive) and Gfap- astrocytes (Row 2), as well as Bergmann glia (not shown). It was not found in Ng2+ oligodendrocyte progenitors (Row 3), nor in Cnp+ myelinating oligodendrocytes (not shown). B) The *Olig2* line directed transgene expression in all cells with Olig2 positive nuclei (Row 1) and specifically into both the Ng2+ progenitors, also called synantocytes or polydendrocytes, (Butt et al., 2005), and Cnp+ mature oligodendrocyte lineage cells. (Row 2). C) In contrast, the *Cmtm5* line expressed the transgene specifically, albeit weakly, in mature (Cnp+) oligodendrocytes, but not Ng2+ progenitors. *Nuclei counterstained with DAPI (blue) in A and B.*

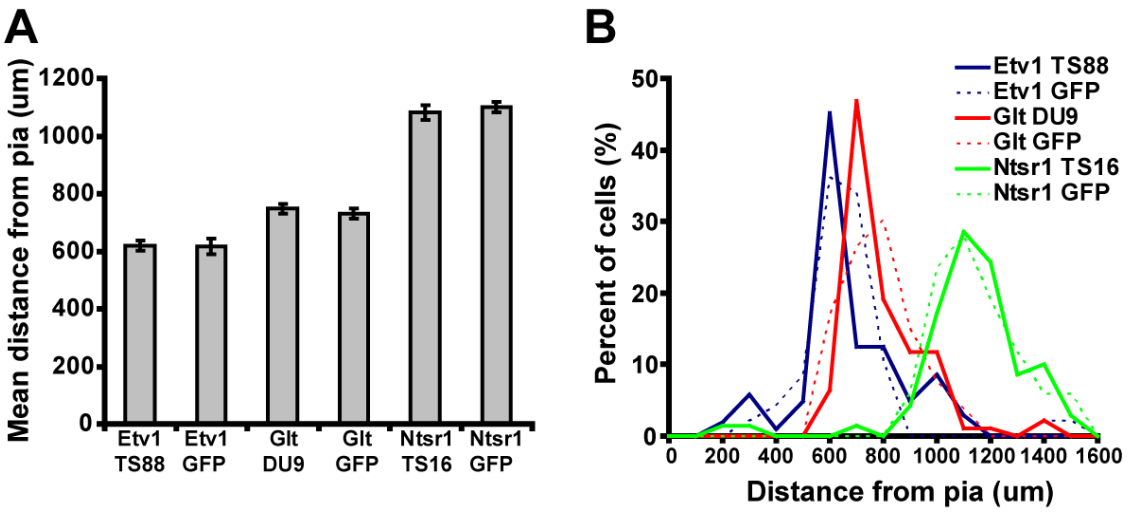
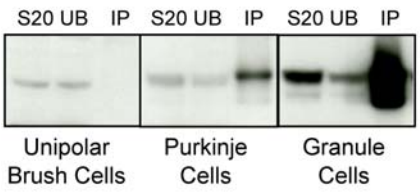


Figure S2. Transgenic EGFP and EGFP-L10a cortical pyramidal cell lines have same laminar distribution.
A) Graph (mean +/- SE) of the distance of EGFP+ cell somas from the pial surface for the bacTRAP EGFP-L10a and GENSAT EGFP lines for each pyramidal cell BAC driver. The depth of the cells was consistent between both lines for each driver. B) The percentage of cells in 100 micron bins is shown as a histogram of the distribution of cell depths for each line in A. The EGFP-L10a line and EGFP line for each driver had overlapping distributions of cell depths.

A)



B)

Line	IP #	ngs IP'd RNA	Amplified RNA
Unipolar Brush	IP1	61.6	6.8 ug/ul
Unipolar Brush	IP2	66.3	6.9 ug/ul
Unipolar Brush	IP3	62.1	6.2 ug/ul
Purkinje Cells	IP1	255.1	6.7 ug/ul
Purkinje Cells	IP2	358	7.2 ug/ul
Purkinje Cells	IP3	206	8.5 ug/ul
Granule Cells	IP1	4326.7	8.3 ug/ul
Granule Cells	IP2	4232.1	6.7 ug/ul
Granule Cells	IP3	6166	8.5 ug/ul

C) Quality Control: Bioanalyzer

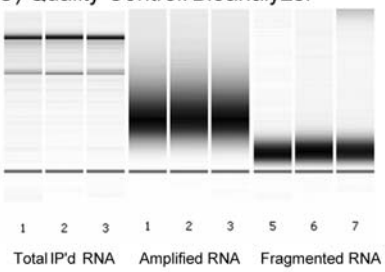


Figure S3. The TRAP methodology can consistently purify RNA efficiently from rare and common cell types.

A) Western blot for EGFP from S20 fraction, flow through (UB), and immunoprecipitate from rare (unipolar brush cells), common (Purkinje cells) and extremely common (granule cells) cell types of cerebellum. Note that even when protein is undetectable, there is sufficient good quality RNA for microarray amplification. B) Yield of RNA is consistent across replicates within each line. C) Total RNA is of good quality for all three cell types, with intact 18s and 28s bands, as well as mRNA which can be amplified and fragmented following standard protocols.

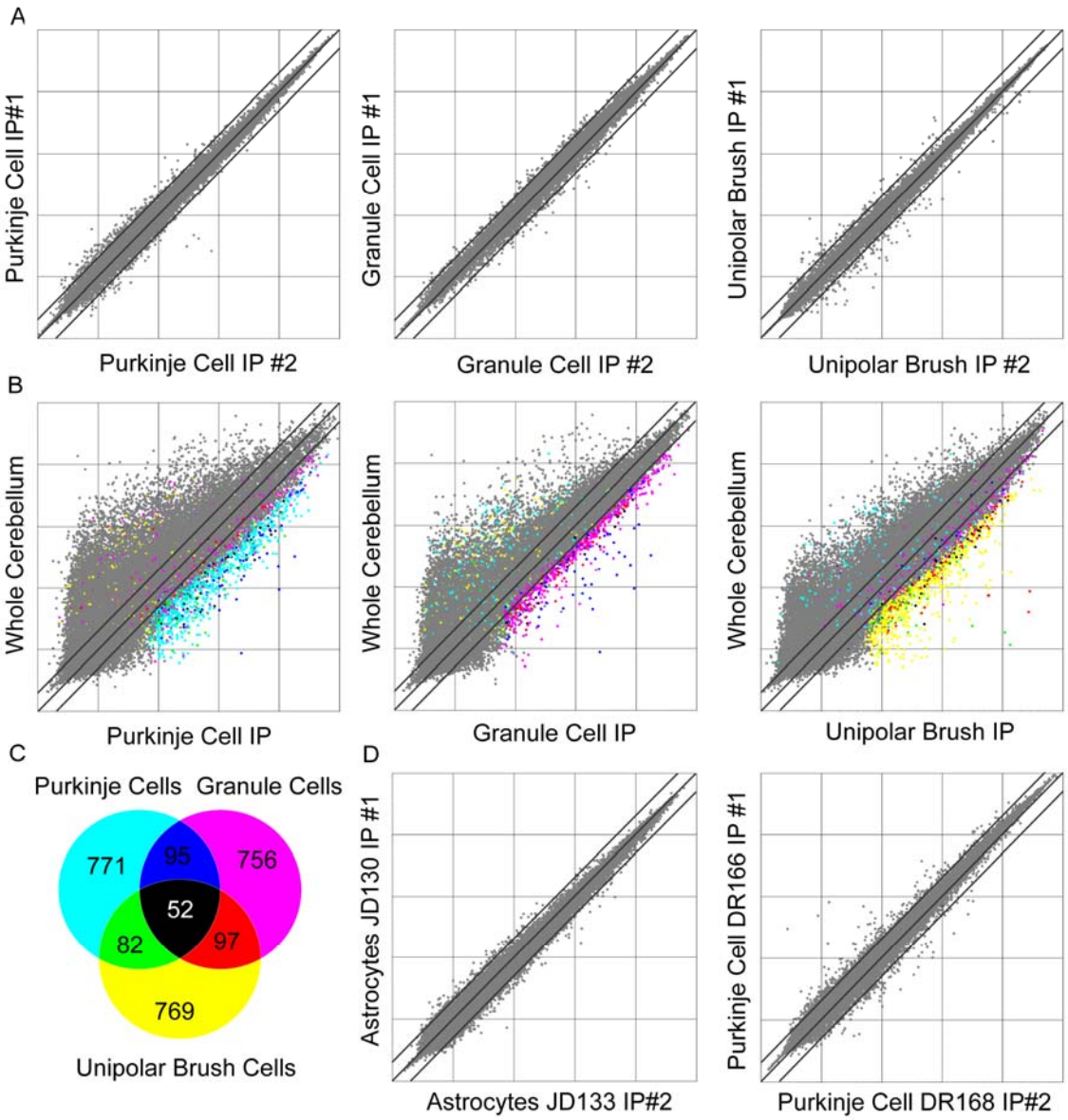


Figure S4. TRAP microarray data are highly reproducible and cell-type specific.
A) Scatterplots of microarray data reveal high reproducibility between replicate immunoprecipitations across three representative cell types of the cerebellum. Grey lines show positions of 0.5, 1, and 2 fold change cut offs. B) Scatterplots of same cell types compared to microarray data from whole cerebellum reveal thousands of probesets enriched in each cell type. Dot colors match C. C) Venn diagram of top 1000 enriched probesets (with signal value cutoff >100) for each cell type reveal that each cell type has a unique pattern of enriched genes. D) Separate BAC transgenic lines made with same driver show similar degree of reproducibility as IPs from the same line, thus location of BAC integration in the genome does not significantly perturb gene expression or immunoprecipitation. *Scatterplots on Log base 10 scale.*

Not Expressed in M.N.

Expressed in M.N.

Enriched in M.N. (IP/UB >2)

Selected novel and expressed receptors

Gene	*ip/ub*	*rf*	*Title*
Acvr1	1.24		Activin A type 1
Acvr1b	0.74		Activin A type 1B
Acvr2	0.8		Activin A type 2
Agtr2	0.85	1	Angiotensin II type 2
Amhr2	28	**2**	Anti-Mullerian hormone, 2
Cd151	5.45		CD151 antigen
Cd24a	4.72		CD24a antigen
Gfra1	2.11	3	GDNF family alpha 1
Gfra4	1.03		GDNF family alpha 4
Ghr	0.82		Growth hormone
Gpr3	2.48		G-protein coupled 3
Gpr22	1.3		G protein-coupled 19
Gpr27	0.49		G protein-coupled 27
Gpr45	1.3		G-protein coupled 45
Gpr49	0.6		G-protein coupled 49
Gpr54	1.39		G-protein coupled 54
Gpr61	1.09		G-protein coupled 61
Gpr68	1.67		G protein-coupled 68
Gpr83	1.34		G-protein coupled 83
Gpr85	0.85		G protein-coupled 85
Gpr123	2.48		G protein-coupled 123
Gpr135	0.47		G protein-coupled 135
Gpr139	0.51		G protein-coupled 139
Ifngr	0.72		Interferon gamma
Ifngr2	2.8		Interferon gamma 2
Oprs1	2.66		Opioid sigma 1
Tnfrsf12a	2.62		TNFR superfamily, 12a
Tnfrsf1a	1.44		TNFR superfamily, 1a
Tnfrsf21	**1.68**		TNFR superfamily, 21
Vdr	8.54		Vitamin D

Acetylcholine neurotranmission genes

Gene	*ip/ub*	*rf*	*Title*
Chat	8.0	4	choline acetyl-transferase
Ache	2.1	5	acetylcholinesterase
Slc18a3	16.8	6	Vesicular ACH transporter
Slc5a7	26.4		Choline transporter

Peptide Transmitters

Calca	3.83	7	CGRP, alpha
Calcb	21.7	7	CGRP, beta

Cytokines and others

Ccl17	5.07		chemokine (cc) ligand 17
Cfi	30.8		complement component i
Cxcl12	6.63		chemokine (cxc) ligand 12
Cxcl5	2.9		chemokine (cxc) ligand 5
Egfl7	2.7		EGF-like domain 7
Fgf1	2.54	8	fibroblast growth factor 1
Il7	5.44		interleukin 7
Kitl	1.7		kit ligand

Homeobox transcription factors

Hlxb9	46.7	9	homeobox gene HB9
Hod	0.56		homeobox only domain
Hoxa5	0.43	10	homeo box A5
Hoxa7	0.53	10	homeo box A7
Hoxa9	0.54	10	homeo box A9
Hoxb5	0.31	10	homeo box B5
Hoxc5	3	10	homeo box C5
Hoxc8	0.97	10	homeo box C8
Hoxc9	0.67	10	homeo box C9
Hoxc10	0.46	10	homeo box C10
Hoxd9	0.88	10	homeo box D9
Irx1			Iroquois rltd homeobox 1
Irx3	8.12	11	Iroquois rltd homeobox 3
Irx5			Iroquois rltd homeobox 5
Isl1	20	9	insulin rltd protein 1
Isl2	5.67	9	insulin rltd protein 2

Figure S5. Selected Motor Neuron enriched and expressed receptors, ligands, and transcription factors.

Listed are a selection of additional known and novel receptors detected in MNs (upper panel) as well as selected ligands and transmitter-related genes enriched or relatively specific to MNs (middle panel). Evident are the genes related to acetylcholine neurotransmission, as well as known MN ligands *Fgf1*, and *Calca*, *Calcb*. Homeobox transcription factors expressed in MN are also shown (lower panel). Rostral Hox genes are not detected, while some caudal Hox genes known to drive differentiation of various motor pools remain expressed into adulthood. Known MN specific transcription factors, such as *Hlxb9* and *Isl2*, are detected as such with the TRAP methodology.

IP/UB: *Fold change versus whole spinal cord for expressed genes.* ***RF****: Expression data from published rodent literature.* ***1*** *(Mendelsohn et al., 1984),* ***2*** *(Wang et al., 2005),* ***3*** *(Oppenheim et al., 1995),* ***4*** *(Bjèorklund et al., 1984),* ***5*** *(Navaratnam and Lewis, 1970),* ***6*** *(Arvidsson et al., 1997),* ***7*** *(Gibson et al., 1984),* ***8*** *(Elde et al., 1991),* ***9*** *(Vult von Steyern et al., 1999)* ***7*** *(Nishi et al., 2001),* ***8*** *(Berthele et al., 1999),* ***9*** *(Towers et al., 2000),* ***10*** *(Akin and Nazarali, 2005),* ***11*** *(Jessell, 2000).*

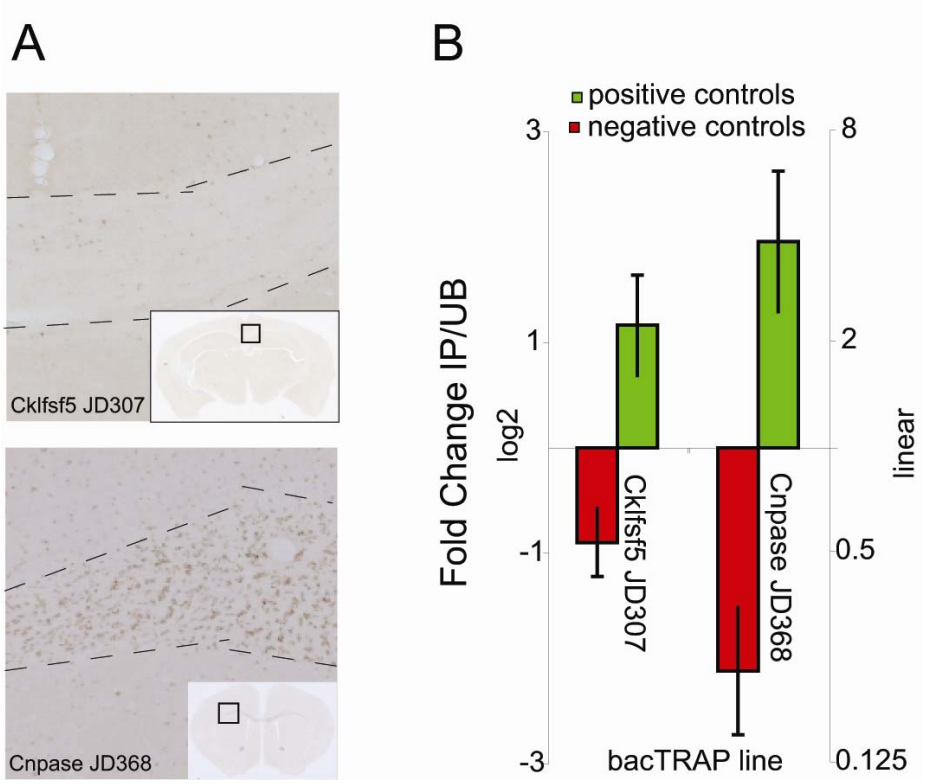
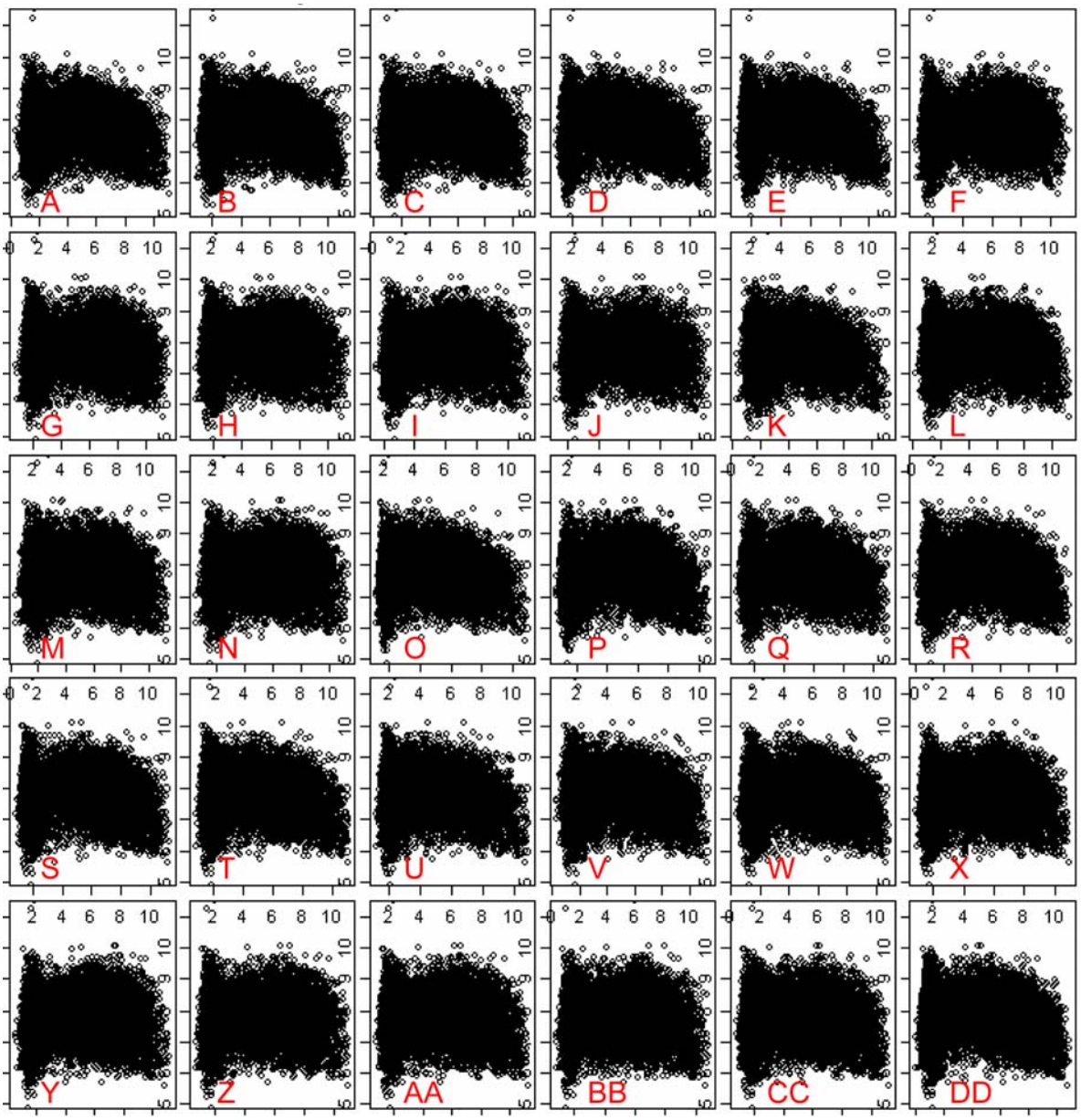


Figure S6. Higher transgene expression yields better signal to noise.
A) DAB immunohistochemistry for EGFP-L10a reveals differential intensity of transgene expression in two bacTRAP lines targeting the same cell population. B) Quality of TRAP microarray data, as assessed by average IP/UB fold change for positive and negative controls, is better for the line with higher transgene expression.



Panel	Corr.	Cell Type	Panel	Corr.	Cell Type
A	-0.13	Brain Stem Motor Neurons	**P**	-0.07	Corticothalamic Neurons
B	-0.11	Whole Tissue, Brain Stem	**Q**	-0.07	Cort+ Interneurons
C	-0.12	Cerebellar Astroglia	**R**	-0.11	Pnoc+ Neurons
D	-0.12	Bergmann Glia	**S**	-0.10	Mature Oligodendrocytes from Cortex
E	-0.11	Granule Cell Layer Interneurons	**T**	-0.11	Mixed Oligodendroglia from Cortex
F	-0.11	Granule Cell	**U**	-0.11	Cck+ Neurons
G	-0.12	Mature Oligodendrocytes from Cerebellum	**V**	-0.08	Whole Tissue, Cortex
H	-0.12	Mixed Oligodendroglia from Cerebellum	**W**	-0.13	Basal Forebrain Cholingergic Neurons
I	-0.11	Purkinje Cells	**X**	-0.10	Unbound RNA of the Basal Forebrain
J	-0.13	Stellate and Basket Cells	**Y**	-0.09	Motor Neurons of the Spinal Cord
K	-0.13	Unipolar Brush Cells	**Z**	-0.09	Whole Tissue, Spinal Cord
L	-0.09	Whole Tissue, Cerebellum	**AA**	-0.13	Corpus Striatum Cholinergic Neurons
M	-0.13	Cortical Astrocytes	**BB**	-0.08	Drd1 Positive Medium Spiny Neurons
N	-0.07	Corticospinal, Corticopontine Neurons	**CC**	-0.07	Drd2 Positive Medium Spiny Neurons
O	-0.09	Corticostriatal Neurons	**DD**	-0.10	Whole Tissue, Basal Ganglia

Figure S7. Microarray signal intensity does not increase with transcript length.
A-DD) To assess whether longer mRNA's were more efficiently immunoprecipitated than shorter transcripts, transcript length (Y-axis, log base 2), was plotted versus GCRMA normalized signal intensity (X-axis, log base 2) for all 30 samples. None of the immunoprecipitated nor the unbound (whole tissue RNA) samples show positive correlations between length and signal.

Abbreviation	Full Gene Name	BAC
Sept4	Septin 4	RP23-21N23
Aldh1l1	Aldehyde dehydrogenase 1 family, member L1	RP23-7M9
Cmtm5	CKLF-like MARVEL transmembrane domain containing 5	RP24-317F19
Cck	Cholecystokinin	RP23-234I17
Chat	Choline Acetyltransferase	RP23-431D9
Cort	Cortistatin	RP23-281A14
Drd2	Dopamine receptor 2	RP23-161H15
Drd1	Dopamine receptor D1A	RP23-47M2
Grp	Gastrin-Releasing Peptide	RP23-179M10
Grm2	Glutamate Receptor, Metabotropic 2	RP23-335E12
Glt25d2	Glycosyltransferase 25 Domain-containing 2	RP23-160M1
Lypd6	LY6/PLAUR Domain containing 6	RP23-14O24
NeuroD1	Neurogenic Differentiation 1	RP24-151C22
Ntsr1	Neurotensin Receptor 1	RP23-314D14
Olig2	Oligodendrocyte transcription factor 2	RP23-356P18
Pnoc	Prepronociceptin	RP23-264L8
Pcp2	Purkinje Cell Protein 2	RP24-186D18
Etv1	Ets1 Variant Gene 1	RP23-250K4

Table S1. List of BACs utilized for transgenesis
Included are the list of abbreviations used throughout the paper for each gene, and the ID for the BAC clone that was modified.

Probeset	Symbol	Name
1428909_at	1200015M12Rik	novel
1428720_s_at	2010309G21Rik	novel
1453238_s_at	3930401B19Rik	novel
1443928_at	AI462064	novel
1446633_at	Apg7l	autophagy 7-like (S. cerevisiae)
1446418_at	Apg7l	autophagy 7-like (S. cerevisiae)
1428612_at	Apg7l	autophagy 7-like (S. cerevisiae)
1428611_at	Apg7l	autophagy 7-like (S. cerevisiae)
1428610_at	Apg7l	autophagy 7-like (S. cerevisiae)
1435640_x_at	C85067	novel
1452538_at	Igh-4	immunoglobulin heavy chain 4 (serum IgG1)
1451963_at	Igh-4	immunoglobulin heavy chain 4 (serum IgG1)
1447333_at	Igh-4	immunoglobulin heavy chain 4 (serum IgG1)
1442544_at	Igh-4	immunoglobulin heavy chain 4 (serum IgG1)
1427870_x_at	Igh-4	immunoglobulin heavy chain 4 (serum IgG1)
1427869_at	Igh-4	immunoglobulin heavy chain 4 (serum IgG1)
1427758_x_at	Igh-4	immunoglobulin heavy chain 4 (serum IgG1)
1427756_x_at	Igh-4	immunoglobulin heavy chain 4 (serum IgG1)
1426196_at	Igh-4	immunoglobulin heavy chain 4 (serum IgG1)
1425324_x_at	Igh-4	immunoglobulin heavy chain 4 (serum IgG1)
1425247_a_at	Igh-4	immunoglobulin heavy chain 4 (serum IgG1)
1424305_at	Igj	immunoglobulin joining chain
1452417_x_at	Igk-V28	immunoglobulin kappa chain variable 8 (V8)
1427660_x_at	Igk-V28	immunoglobulin kappa chain variable 8 (V8)
1427455_x_at	Igk-V28	immunoglobulin kappa chain variable 8 (V8)
1459776_x_at	Lmna	lamin A
1459775_at	Lmna	lamin A
1457670_s_at	Lmna	lamin A
1456182_x_at	Lmna	lamin A
1455892_x_at	Lmna	lamin A
1452212_at	Lmna	lamin A
1443737_at	Lmna	lamin A
1426868_x_at	Lmna	lamin A
1425472_a_at	Lmna	lamin A
1421654_a_at	Lmna	lamin A
1450009_at	Ltf	lactotransferrin
1456456_x_at	Mela	melanoma antigen
1433438_x_at	Mela	melanoma antigen
1452556_at	Mnda	myeloid cell nuclear differentiation antigen
1427798_x_at	Mnda	myeloid cell nuclear differentiation antigen
1427797_s_at	Mnda	myeloid cell nuclear differentiation antigen
1416957_at	Pou2af1	POU domain, class 2, associating factor 1
1451206_s_at	Pscdbp	pleckstrin homology, Sec7 and coiled-coil domains, binding protein
1435697_a_at	Pscdbp	pleckstrin homology, Sec7 and coiled-coil domains, binding protein
1444064_at	Samhd1	SAM domain and HD domain, 1
1434438_at	Samhd1	SAM domain and HD domain, 1
1420273_x_at	Samhd1	SAM domain and HD domain, 1
1420272_at	Samhd1	SAM domain and HD domain, 1
1418131_at	Samhd1	SAM domain and HD domain, 1
1436727_x_at	Sptlc1	serine palmitoyltransferase, long chain base subunit 1
1436726_s_at	Sptlc1	serine palmitoyltransferase, long chain base subunit 1
1422691_at	Sptlc1	serine palmitoyltransferase, long chain base subunit 1
1422690_at	Sptlc1	serine palmitoyltransferase, long chain base subunit 1
1420570_x_at	Tcl1b3	T-cell leukemia/lymphoma 1B, 3
1435137_s_at		novel
1427932_s_at		novel

Table S2. List of Probesets excluded from analysis

To identify mRNAs which interact with monoclonal antibodies or protein G dynabeads in the absence of EGFP, the TRAP protocol was performed on a wildtype mouse brain and compared to unbound whole brain RNA. The excluded probesets listed here are those for genes found to be highly enriched (>6 fold) in the wildtype IP/UB as well as enriched in multiple IPs from diverse regions and cell types.

PNOC

Gene ID	Genbank	Open Biosystems #	Gene Name	Enzymes
B230118H07Rik	BC025075	MMM1013-7513748	RIKEN cDNA B230118H07 gene	Sal I cut, T7 polymerase
Chd4	BC058578	MMM1013-9201124	chromodomain helicase DNA binding protein 4	Sal I cut, T7 polymerase
Cidea	AA061879	EMM1002-1115733	cell death-inducing DNA fragmentation factor, alpha subunit-like effector A	Eco RI cut, T3 polymerase
Crabp1	BC065787	MMM1013-9202137	cellular retinoic acid binding protein I	Eco RI cut, T3 polymerase
Ddef1	BC094581	MMM1013-98479313	development and differentiation enhancing	Eco RI cut, T3 polymerase
Igf1	BC012409	MMM1013-65370	insulin-like growth factor 1	Sal I cut, T7 polymerase

GRM2

Gene ID	Genbank	Open Biosystems #	Gene Name	Enzymes
Lypd1	CA328245	EMM1002-6960067	RIKEN cDNA 2700050C12 gene	Eco RI cut, T3 polymerase
Slc6a5, GlyT2	BM941867	EMM1032-584726	solute carrier family 6 (neurotransmitter transporter, glycine), member 5	Eco RI cut, T3 polymerase
Amn	AA023455	EMM1002-1077429	amnionless	Eco RI cut, T3 polymerase
Penk1	AA098193	EMM1002-1183311	preproenkephalin 1	Eco RI cut, T7 polymerase
Gcgr	BC031885	MMM1013-7511930	glucagon receptor	Sal I cut, T7 polymerase
Ceacam10	BC003346	MMM1013-62963	CEA-related cell adhesion molecule 10	Sal I cut, T7 polymerase

Table S3. List of plasmids for *in situ* hybridization studies
ISH was conducted on 11 genes enriched in either *Grm2* positive granule cell layer interneurons, or *Pnoc* positive neurons of cortex, using plasmids ordered from Open Biosystems. Plasmids were sequenced to confirm the accuracy and orientation of the ESTs. Plasmids were linearized with either Sal I or Eco RI and Dig labeled RNA was transcribed with either T3 or T7 polymerase to create anti-sense cRNA probe.

Gene	Forward Primer	Reverse Primer	Primerbank ID or Ref
B-Actin	agagggaaatcgtgcgtgac	caatagtgatgacctggccgt	(Overbergh et al., 1999)
Gfap	gtaaagactgtggagatgcgggatggtgagg	gtgctggtgtgggtgggaactgag	
Chat	ccattgtgaagcggtttggg	gccaggcggttgtttagataca	26338049a1
Slc18a3	gtgaagataggcgtgctatttgc	gactgtggaggcgaacatgac	11096330a2
Cnp1	tgcttgatgataccaaccacg	gctgggcacagtctagtcg	6753476a3
Pcp2	tgcagggcgatcggatggaggag	tgaggggtgagcaggggttgagg	
Rilp	ctgatgcggcaacctcagat	ttgagcaagaacacgttggct	30185833a1
Popdc3	tgactgaacacccactctgc	actgccacccataaaacctact	31745187a1
Kcnn1	ttgaaaagcgtaaacggctca	cagagcaaaagagcagagtga	14161696a1
Foxq1	aaattggaggtgttcgtccca	tccccgtctgagcctaagg	31560693a1
Cox7a1	gctctggtccggtcttttagc	gtactgggaggtcattgtcgg	6753504a1
Tpm2	aagtcgctgatagcctcagag	ggtctggtgtatctccacgttc	50190a1
Crygs	cagacttccgctcgtacctaa	tcgccctggggtaagatgt	6753532a1
Cfi	cttggctctccacttgagttc	ggagcgatgcgtgtatttctg	6671744a1

Table S4. List of qRT-PCR primers
qRT-PCR primer sequences were mostly from PrimerBank (Wang and Seed, 2003) or literature (Overbergh et al., 1999). All are listed in 5' to 3' direction.

Probeset	Correlation	Symbol	Name
1419646_a_at	1	Mbp	myelin basic protein
1417275_at	0.968	Mal	myelin and lymphocyte protein, T-cell differentiation protein
1451718_at	0.964	Plp	proteolipid protein (myelin)
1440902_at	0.958	Galnt5	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 5
1433785_at	0.958	Mobp	myelin-associated oligodendrocytic basic protein
1436578_at	0.955	A330104H05Rik	novel
1428792_at	0.946	Bcas1	breast carcinoma amplified sequence 1
1439506_at	0.934	C11ORF9 homolog	novel
1426960_a_at	0.931	Fa2h	fatty acid 2-hydroxylase
1433543_at	0.926	Anln	anillin, actin binding protein (scraps homolog, Drosophila)
1448768_at	0.921	Mog	myelin oligodendrocyte glycoprotein
1418086_at	0.917	Ppp1r14a	protein phosphatase 1, regulatory (inhibitor) subunit 14A
1434094_at	0.914	6330530A05Rik	novel
1420760_s_at	0.911	Ndrg1	N-myc downstream regulated 1
1447807_s_at	0.905	Plekhh1	pleckstrin homology domain containing, family H (with MyTH4 domain) member 1
1425546_a_at	0.905	Trf	transferrin
1436974_at	0.904	A230069A22Rik	novel
1450241_a_at	0.904	Evi2a	ecotropic viral integration site 2a
1418472_at	0.903	Aspa	aspartoacylase (aminoacylase) 2
1452834_at	0.9	2600010E01Rik	novel
1429909_at	0.894	4833411O04Rik	novel
1453009_at	0.893	1110060I01Rik	novel
1440813_s_at	0.89	Plxnb3	plexin B3
1437171_x_at	0.888	Gsn	gelsolin
1423871_at	0.887	BC014795	novel
1434399_at	0.887	Galnt6	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 6
1435854_at	0.885	Tmem10	transmembrane protein 10
1416371_at	0.878	Apod	apolipoprotein D
1418406_at	0.877	Pde8a	cAMP-specific cyclic nucleotide phosphodiesterase PDE8
1416003_at	0.875	Cldn11	claudin 11
1434606_at	0.872	Erbb3	v-erb-b2 erythroblastic leukemia viral oncogene homolog 3 (avian)
1451932_a_at	0.872	Tsrc1	thrombospondin repeat containing 1
1418980_a_at	0.871	Cnp1	cyclic nucleotide phosphodiesterase 1
1416318_at	0.87	Serpinb1a	serine (or cysteine) proteinase inhibitor, clade B, member 1a
1424468_s_at	0.868	D330037A14Rik	novel

Table S6. Genes correlating with myelin basic protein

Correlation of genes with a Mbp probeset (1419646_a_at) identifies known and putative novel myelination genes. Four of the first eleven genes correlated with Mbp represent genes for the known myelin components *Mal, Plp, Mobp*, and *Mog*. This coexpression suggests that novel genes highly correlated with *Mbp* (*C11orf9, A330104H05Rik, Bcas1*) may also be involved in myelination. *Pearson's correlation. Only the first probeset for each gene is shown for those genes with multiple probesets on the array. In yellow are genes identified in an independent proteomic screen of myelin components (Vanrobaeys et al., 2005).*

Symbol	Antibody Name	Product Number	Source
EGFP	Goat Anti-EGFP	NA	(Heiman et al.)
EGFP	Chicken Anti-EGFP	AB19370	Abcam, Cambridge Ma.
NEUN	Neuronal Nuclei	MAB377	Chemicon, Temecula, Ca.
OLIG2	Oligodendrocyte transciption factor 2	AB9610	Chemicon, Temecula, Ca.
CSPG4	Ng2 Proteoglycan	AB53420	Chemicon, Temecula, Ca.
CHAT	Choline Acetyl-Transferase	AB143	Chemicon, Temecula, Ca.
GRM2/GRM3	metabotropic glutamate receptor 2&3	AB1553	Chemicon, Temecula, Ca.
GRM1	metabotropic glutamate receptor 1	AB1551	Chemicon, Temecula, Ca.
CNP1	Cnpase	MAB326R	Chemicon, Temecula, Ca.
CALB1	Calbindin	300	Swant, Bellinzona, CH
CALB2	Calretinin	6B3	Swant, Bellinzona, CH
PVALB	Parvalbumin	Pv28	Swant, Bellinzona, CH
GFAP	Glial Fibrillary acidic protein	Z0334	Dako, Denmark
S100	S100	Z0311	Dako, Denmark
GLUL	Glutamine Synthetase	G2781	Sigma-Aldrich, St Louis, Mo.
SLC18a3	Vesicular Acetylcholine Transporter	sc-7717	Santa Cruz Biotech, Santa Cruz, Ca.
ALDH1L1	10-Formyltetrahydrofolate Dehydrogenase	NA	a gift from Dr. Robert Cook
-	GABA	A2052	Sigma-Aldrich, St Louis, Mo.

Table S7. Antibodies used for immunohistochemistry
A list of antibodies, product numbers, and sources used in this paper.

Supplemental References

Akin, Z. N., and Nazarali, A. J. (2005). Hox genes and their candidate downstream targets in the developing central nervous system. Cell Mol Neurobiol *25*, 697-741.
Anthony, T. E., and Heintz, N. (2007). The folate metabolic enzyme ALDH1L1 is restricted to the midline of the early CNS, suggesting a role in human neural tube defects. J Comp Neurol *500*, 368-383.
Arvidsson, U., Riedl, M., Elde, R., and Meister, B. (1997). Vesicular acetylcholine transporter (VAChT) protein: a novel and unique marker for cholinergic neurons in the central and peripheral nervous systems. J Comp Neurol *378*, 454-467.
Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., *et al.* (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet *25*, 25-29.
Berthele, A., Boxall, S. J., Urban, A., Anneser, J. M., Zieglgansberger, W., Urban, L., and Tolle, T. R. (1999). Distribution and developmental changes in metabotropic glutamate receptor messenger RNA expression in the rat lumbar spinal cord. Brain Res Dev Brain Res *112*, 39-53.
Bjèorklund, A., Hèokfelt, T., and Kuhar, M. J. (1984). Classical transmitters in the CNS (Amsterdam; New York: Elsevier).
Butt, A. M., Hamilton, N., Hubbard, P., Pugh, M., and Ibrahim, M. (2005). Synantocytes: the fifth element. J Anat *207*, 695-706.
Cahoy, J. D., Emery, B., Kaushal, A., Foo, L. C., Zamanian, J. L., Christopherson, K. S., Xing, Y., Lubischer, J. L., Krieg, P. A., Krupenko, S. A., *et al.* (2008). A transcriptome database for astrocytes, neurons, and oligodendrocytes: a new resource for understanding brain development and function. J Neurosci *28*, 264-278.
Dennis, G., Jr., Sherman, B. T., Hosack, D. A., Yang, J., Gao, W., Lane, H. C., and Lempicki, R. A. (2003). DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol *4*, P3.
Elde, R., Cao, Y. H., Cintra, A., Brelje, T. C., Pelto-Huikko, M., Junttila, T., Fuxe, K., Pettersson, R. F., and Hokfelt, T. (1991). Prominent expression of acidic fibroblast growth factor in motor and sensory neurons. Neuron *7*, 349-364.
Gibson, S. J., Polak, J. M., Bloom, S. R., Sabate, I. M., Mulderry, P. M., Ghatei, M. A., McGregor, G. P., Morrison, J. F., Kelly, J. S., Evans, R. M., and et al. (1984). Calcitonin gene-related peptide immunoreactivity in the spinal cord of man and of eight other species. J Neurosci *4*, 3101-3111.
Heiman, M., Schaefer, A., Gong, S., Peterson, J., Day, M., Ramsey, K., Suárez-Fariñas, M., Schwarz, C., Stephan, D. A., Surmeier, J., *et al.* (2008). Development of a translational profiling approach for the molecular characterization of individual CNS cell types. Cell.
Jessell, T. M. (2000). Neuronal specification in the spinal cord: inductive signals and transcriptional codes. Nat Rev Genet *1*, 20-29.
Maere, S., Heymans, K., and Kuiper, M. (2005). BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. Bioinformatics *21*, 3448-3449.
Mendelsohn, F. A., Quirion, R., Saavedra, J. M., Aguilera, G., and Catt, K. J. (1984). Autoradiographic localization of angiotensin II receptors in rat brain. Proc Natl Acad Sci U S A *81*, 1575-1579.

Navaratnam, V., and Lewis, P. R. (1970). Cholinesterase-containing neurones in the spinal cord of the rat. Brain Res *18*, 411-425.
Nishi, M., Hinds, H., Lu, H. P., Kawata, M., and Hayashi, Y. (2001). Motoneuron-specific expression of NR3B, a novel NMDA-type glutamate receptor subunit that works in a dominant-negative manner. J Neurosci *21*, RC185.
Oppenheim, R. W., Houenou, L. J., Johnson, J. E., Lin, L. F., Li, L., Lo, A. C., Newsome, A. L., Prevette, D. M., and Wang, S. (1995). Developing motor neurons rescued from programmed and axotomy-induced cell death by GDNF. Nature *373*, 344-346.
Overbergh, L., Valckx, D., Waer, M., and Mathieu, C. (1999). Quantification of murine cytokine mRNAs using real time quantitative reverse transcriptase PCR. Cytokine *11*, 305-312.
Paxinos, G., and Franklin, K. B. J. (2001). The mouse brain in stereotaxic coordinates, 2nd edn (San Diego, CA: Academic).
Schneider, T. D. (2007). Information Theory Primer (Frederick, MD: National Cancer Institute).
Towers, S., Princivalle, A., Billinton, A., Edmunds, M., Bettler, B., Urban, L., Castro-Lopes, J., and Bowery, N. G. (2000). GABAB receptor protein and mRNA distribution in rat spinal cord and dorsal root ganglia. Eur J Neurosci *12*, 3201-3210.
Vanrobaeys, F., Van Coster, R., Dhondt, G., Devreese, B., and Van Beeumen, J. (2005). Profiling of myelin proteins by 2D-gel electrophoresis and multidimensional liquid chromatography coupled to MALDI TOF-TOF mass spectrometry. J Proteome Res *4*, 2283-2293.
Vult von Steyern, F., Martinov, V., Rabben, I., Nja, A., de Lapeyriere, O., and Lomo, T. (1999). The homeodomain transcription factors Islet 1 and HB9 are expressed in adult alpha and gamma motoneurons identified by selective retrograde tracing. Eur J Neurosci *11*, 2093-2102.
Wang, P. Y., Koishi, K., McGeachie, A. B., Kimber, M., Maclaughlin, D. T., Donahoe, P. K., and McLennan, I. S. (2005). Mullerian inhibiting substance acts as a motor neuron survival factor in vitro. Proc Natl Acad Sci U S A *102*, 16421-16425.