

SUPPLEMENTARY METHODS

Enzymatic striatal cell dissociation. Four transgenic mice at Postnatal day 20 or 2 months were anesthetized with halothane, brains were removed and sectioned in the coronal plane at 500 μm on a Vibratome (Leica) in 26 mM NaHCO_3 , 1.25 mM NaH_2PO_4 , 130 mM NaCl , 3 mM KCl , 5 mM MgCl_2 , 10 mM Glucose, 1 mM CaCl_2 , pH 7.2-7.4, osmolarity 290-310, and bubbled with 5% CO_2 /95% O_2 . Total striatum was dissected from brain slices and dissociated for 45 minutes at 37°C with Papain enzyme (Papain Dissociation System, Worthington Biochem) in Earls Balanced Salt Solution (EBSS) with DNase according to the manufacture's protocol. Brain slices were triturated with three glass pipettes of decreasing tip diameter and dissociated cells were centrifuged at 900 rpm at room temperature for 5 minutes. To remove excess debris, cell pellets were resuspended in EBSS, DNase, and albumin ovomucoid inhibitor (AOI), and the cell suspension was subjected to centrifugation on an AOI discontinuous gradient (per Papain Dissociation System protocol) at 900 rpm at 4 °C. Cell pellets were re-suspended in a buffer media (L15- CO_2 without phenol, 1x Pen-Strep, 10 mM Hepes, 25 $\mu\text{g/ml}$ DNase, 1 mg/ml BSA), and filtered through a 70 μm mesh (BD Falcon, #352350). Cells were treated with propidium iodide (PI) (20 $\mu\text{g/ml}$) to label dead cells and sorted on a FACSVantage SE cell sorter (Becton Dickson) for FITC signals (detecting EGFP) and PE signals (detecting PI). Cells were kept on ice pre- and post-sorting.

For imaging of the FACS-sorted cells, pre-sort and post-sort cells were resuspended in Dulbecco's Modified Eagle Medium (DMEM)/F-12 with B27 and plated in eight well chamber slides coated with 1x Polyornithine and 1x fibronectin. Cells were allowed to settle for 6 hours at 37 °C, then fixed with 4% paraformaldehyde (PFA), followed by rinses in 0.1M PBS, and immunostained with anti-rabbit GFP antibody (Molecular Probes) prior to imaging using Zeiss Axioscope II microscope.

Amplification of cRNA. RNA was extracted from 5,000-10,000 sorted EGFP (+) cells using the PicoPure RNA Isolation kit (Arcturus) with a DNase I digestion according to the manufacture's protocol. Approximately 3 to 10 ng of total RNA was isolated from each FACS-sort. RNA quality and relative concentration were

confirmed on the Picochip using Agilent 2100 Bioanalyzer (Agilent Technologies). RNA amplification and labeling was performed using Agilent Low RNA Input Fluorescent Linear Amplification Kit. Two rounds of *in vitro* transcription (IVT) amplification were performed.

In the first round of IVT amplification, the RNA sample (3 ng in a 10.3 μ l volume) was first annealed to 1.2 μ l T7 promoter primer for 10 minutes at 65 $^{\circ}$ C, followed by 5 minutes on ice. Synthesis of cDNA was done at 40 $^{\circ}$ C for 2 hours with a reaction mixture (RNA annealed to T7 promoter primer --11.5 μ l, 5x First strand buffer--4 μ l, 0.1M DTT--2 μ l, 10 mM dNTPs--1 μ l, MMLV Reverse Transcriptase --1 μ l, and RNase out--0.5 μ l), followed by Reverse transcriptase inactivation for 15 minutes at 65 $^{\circ}$ C and 5 minutes on ice. To synthesize the first-round cRNA, *in vitro* transcription (IVT) was performed at 40 $^{\circ}$ C for 16 hours in a reaction mixture (cDNA solution--20 μ l, nuclease free water--12.1 μ l, 4x Transcription buffer--20 μ l, 0.1M DTT--6 μ l, NTPs--8 μ l, CTP--5.6 μ l, 50% PEG--6.4 μ l, RNase Out --0.5 μ l, Inorganic Phosphate--0.6 μ l, T7 RNA polymerase--0.8 μ l). First-round cRNA was purified using RNA Mini Elute Kit (Qiagen) and eluted with 10.5 μ l nuclease free water.

In the second round of IVT amplification, the first-round cRNA was annealed to 1 μ l random hexamers for 10 minutes at 65 $^{\circ}$ C followed by 5 minutes on ice. The annealed first-round cRNA was used to synthesize first-strand cDNA in a reaction mixture (first-round cRNA annealed to random hexamers--11.5 μ l, 5x first strand buffer--4 μ l, 0.1M DTT--2 μ l, 10 mM dNTPs--1 μ l, MMLV RT--1 μ l, and RNase out--0.5 μ l) for 2 hours at 40 $^{\circ}$ C. This first-strand cDNA was annealed to 1.2 μ l T7 promoter primer for 10 minutes at 65 $^{\circ}$ C, followed by 5 minutes on ice. Second-strand cDNA was synthesized in a reaction mixture (first-strand cDNA annealed to T7 promoter primer--21.2 μ l, 5x first-strand buffer--8 μ l, 0.1M DTT--4 μ l, 10 mM dNTPs--2 μ l, MMLV reverse transcriptase--2 μ l, and RNase out--1 μ l) for 2 hours at 40 $^{\circ}$ C. The reaction was inactivated by 15 minutes at 65 $^{\circ}$ C followed by 5 minutes on ice. The cDNA was then split into two volumes of 20 μ l each and labeled with either CTP-Cy3 or CTP-Cy-5 (Perkin Elmer) in the second IVT amplification, by incubating for 5 hours at 40 $^{\circ}$ C in the following reaction mixture (2nd round amplified cDNA--20 μ l, CTP-Cy3 or CTP-Cy5--2.4 μ l, nuclease free water--15.3 μ l, 4x Transcription buffer--20 μ l, 0.1M DTT--6 μ l, NTPs--8 μ l, 50% PEG--6.4 μ l, RNase Out--0.5 μ l, Inorganic Phosphate--0.6 μ l, T7 RNA polymerase--0.8 μ l). Labeled second-round

cRNA was purified with RNeasy mini elute kit (Qiagen) and eluted in 30 μ l of nuclease-free water. The final labeled cRNA concentration was measured (Nanodrop technologies), and the RNA quality was checked on Bioanalyzer Nanochip (Agilent technologies).

Microarray data analysis. Data analysis of homotypic comparisons (biological replicates) was performed in the TIGR TM4 microarray statistical package^{S7}. The local background values were subtracted from the raw signal intensity values and transformed as log base 10. To assure the data quality the log2 intensities below the mean + 2.64 s.d. of negative control intensities in each array were excluded from further data analysis and cyclic loss normalization was carried out in all distinct pair-wise comparisons. In addition to homotypic comparisons presented in the Results, different amounts of input RNA were used to demonstrate the feasibility of using lower RNA amounts, as could be the case if one were to sort less abundant cell types. We performed microarray analyses using RNA input of 3 ng, 0.3 ng and 0.03 ng, corresponding to the RNA extracted from 3000, 300 and 30 cells approximately. Correlations between homotypic comparisons from samples hybridized onto different slides using as little as 0.03 ng input RNA were 0.95 to 0.97 (**Supplementary Fig.5** online). This clearly demonstrates that one can generate reproducible hybridizations over a wide range of input cell number or RNA, as others have demonstrated^{S8-S10} although one can see that there is a loss of low abundance messages at the lowest concentrations (*ibid*). Therefore, we used the upper amounts, corresponding to 3000 or more cells or 3 ng of input RNA for the experiments described here so as to ensure maximum transcript representation.

For identification of differential expression in heterotypic comparisons (different neuronal subtypes), genes were selected based on 2 criteria: the fold change (>2.0 fold) and significance of their expression changes ($p < 0.01$). Since, theoretically, no genes should be differentially expressed in homotypic comparisons, this allows one to measure the false positive level and derive a threshold that corresponds to a given level of confidence^{S11}. Based on 16 homotypic comparisons in P20 and adult ($n=12$), the 95% confidence interval was within the range of 2-fold differential expression (**Fig. 3**). The statistical significance of gene expression differences between neuronal subtypes was determined by pairwise comparisons (paired

T-test, $P < 0.01$) corrected for multiple comparisons using a false discovery control rate of 5%. P -values were computed using the maximum number of permutations of the data for each gene in TM4^{S7}. For the most stringent analysis all genes with expression changes above 2 fold in at least four out of five experiments were selected. We also used more relaxed criteria in which only three out of five experiments exceeded 2.0 fold expression change (this always included at least one D1/D2 experimental replicate) and identified an additional 10 genes (**Table 1**).

Semi-quantitative RT-PCR analyses. Reverse transcription (RT) reaction mixture included cRNA, 1 x Buffer RT, 0.5 mM dNTPs, 300 ng/ μ l random primers, 0.05 Units RNase (Promega), 0.2 Units/ μ l Omniscript RT and nuclease free water (all reagents except RNase were supplied in the Qiagen Omniscript kit) for a final volume 100 μ l. The RT reaction was incubated at 37°C overnight.

The cDNA quantity was normalized for each sample using a β -*actin* dilution series. The PCR master mix contained 1 μ M oligonucleotide primers with 0.75 units Taq (Qiagen), 1x buffer (Qiagen), and 0.25 mM dNTPs (Roche). In a subset of PCR reactions, we also used 1 μ M oligonucleotide primers plus 1x Hot Start Taq Master Mix (Qiagen). One cycle PCR was performed using 37 and 42 cycles for genes with moderate or low levels of expression, and 27 and 32 cycles for genes with robust expression (such as *Penk1* and *Tac I*). For the genes that were not readily detected after one cycle of PCR, we also performed nested PCR with an initial 20 cycles using the outer primers, then 1 μ l of sample from the first PCR reaction was used in a second PCR reaction with the inner primers for 20 and 30 cycles. All PCR reactions had an annealing temperature of 58°C. All the PCR primers used in this study are listed in **Supplementary Table 3** online.

Quantitative RT-PCR. Two P20 samples of D1 or D2 cDNA were used for Taqman singleplex PCR. All reagents were supplied by Applied Biosystems. PCR master mix contained 1x Taqman Universal PCR Master Mix (No Amp Erase UNG), 1x Gene Expression Assay mix, and 1 μ l cDNA for a total volume of 20 μ l. The following Gene Expression Assays were used: *Zfp521/Evi3*_Mm00521009_m1, *Ebf1*- Mn00395519_m1,

Tac1- Mm00436880_m1, *Penk1*-Mm01212875_m1, *Drd2*- Mm00438541_m1, and (*β -actin*) Actb-Mm00607939_s1. Samples were run in three replicates for each Gene Expression Assay. PCR reactions were carried out on a 7700 Thermal Cycler (Applied Biosystems) with 40 cycles of 95 °C for 15 seconds and 60 °C for 1 minute. C_T values for each gene were normalized to C_T values for *β -actin* to obtain a relative expression level for each replicate and the three replicates were averaged together. The qPCR ratio^{S6} of D1/D2 and D2/D1 were compared to assess enrichment in each cell population (see **Supplementary Table 2** online).

Double *in situ* hybridization. Probes for *in situ* hybridization were PCR-amplified from striatal cDNA and subcloned into pBluescript (Stratagene). Antisense radioactive cRNA probes were generated and labeled with ³⁵S UTP (Perkin Elmer) by *in vitro* transcription (IVT) using Stratagene RNAmass High Yield Transcription kit (Stratagene) according to manufacture's protocol. Briefly, IVT reaction mixture included 1x Transcription Buffer, 0.06 M DTT, 0.15 mM rATP, 0.15 mM rCTP, 0.15 mM rGTP, 0.0006 mM rUTP, 93.75 μ Ci ³⁵S UTP (Perkin Elmer), 0.6 U/ μ l RNase block, 6U/ μ l T7 polymerase, and 600ng of linearized plasmid DNA. The reaction mixture was incubated at 37°C for 2 hours. Antisense non-radioactive digoxigenin (DIG) probe (for *Penk1*) was generated using the same reaction mix except with 1x DIG RNA labeling mix (Roche) substituting rNTPs and ³⁵S UTP in the previous reaction. Both radioactive and non-radioactive cRNA probes were digested with DNase I (Qiagen) and cleaned with RNeasy mini elute kit (Qiagen).

To perform the ISH, brains from P20 mice were dissected and frozen on dry ice, sectioned in the coronal plane at 20 μ m thickness on a Leica CM 1850 cryostat, and mounted directly onto slides followed by fixation in 4% PFA for 20 minutes, and rinsed once in 0.1 M Phosphate buffer (PB). Sections were rinsed in 7.5 mg/ml glycine in 0.1M PB, followed with a 0.1M PB rinse, and an Acetic anhydride (2.5 μ l/ml) /TEA (13 μ l/ml) pH 8.0 rinse. Sections were then dehydrated through a series of graded alcohols for 1 minute each (50% EtOH, 70% EtOH, 95% EtOH, 100% EtOH, chloroform- 5 minutes, 100% EtOH, and 95% EtOH), and subsequently blow dried prior to hybridization. Sections were hybridized overnight at 60°C in hybridization

solution containing 50% formamide, 35% Denhardt's solution, 0.14 x SSC, 10% dextran sulfate, 0.3 mg/ml salmon sperm DNA, 0.15 mg/ml Yeast tRNA, 40 mM DTT, S^{35} radioactive probes, and 600 ng/ml DIG-Penk1 probe.

On the second day of hybridization, brain sections were treated with 20 ug/ml RNaseA in 0.1M Tris PH 8.0, 0.5 M NaCl, 1mM EDTA at 45 °C for 35 minutes, and underwent a graded series of 20-minute SSC rinses containing 2.5 mg/ml sodium thiosulfate (2 times 4x SSC at 60 °C, 4 times 2x SSC at room temperature, 2 times 0.2x SSC at 60 °C, one 0.1 x SSC at 60 °C, and one 0.1x SSC at room temperature). Sections were then rinsed in 0.1M Tris/ 0.15M NaCl pH 8.0, and were incubated in anti-DIG-HPA antibody (1:5000 dilution, Roche) overnight at room temperature. DIG signal was detected with NBT/BCIP (Roche). Sections were then rinsed four times for 30 minutes in 1 x SSC to prevent nonspecific reaction of NBT/BCIP with emulsion, blow dried, and exposed to film for 4-5 days then subsequently dipped in Ilford K.5D emulsion (Polysciences, Inc). After 4-5 weeks, sections were exposed and analyzed with a Zeiss Axioscope II microscope.

Immunofluorescence. Mouse brains from P0 were directly removed and immersed in 4% paraformaldehyde (PFA); and all other ages were perfused with 0.1M PBS and 4% PFA, post-fixed in 4% PFA overnight, cryoprotected in 30% sucrose, and sectioned in the sagittal plane at 35µm to 50µm thickness. For immunofluorescent staining, sections were incubated overnight at room temperature in 1:1000-1:2000 of rabbit anti-met-Enkephalin (Penk1)^{S12} in 0.1M PBS with 3% normal goat serum (NGS), 3% non-fat dry milk, 0.3% Tween. A secondary reaction was done using biotinylated goat anti-rabbit antibody in 0.1M PBS with 3%NGS, 0.3% Tween (Chemicon) followed by streptavidin Alexa Fluor 594 (Molecular Probes) in 0.1M PBS. For detection of Tac1/Substance P, we used rat anti- Substance P antibody (1:400 dilution Chemicon) in 0.1M PBS with 3% NGS, and 0.3% Triton-X., followed by anti-rat Alexa Fluor 594 (Molecular Probes) in 0.1M PBS. For detection of EGFP, we used a rabbit anti-GFP primary antibody (1:400 dilution, Molecular Probes), and a goat-anti-rabbit antibody conjugated to Alexa Fluor 488. Images were acquired using a LSM 510 Meta confocal (Zeiss) or using a Zeiss Axioscope II microscope.

Supplementary References

- S1. Garel, S., Marin, F., Grosschedl, R. & Charnay, P. Ebf1 controls early cell differentiation in the embryonic striatum. *Development* **126**, 5285-5294 (1999)
- S2. Girault, J.A., Horiuchi, A., Gustafson, E.L., Rosen, N.L. & Greengard, P. Differential Expression of ARPP-16 and ARPP-19, two highly related cAMP regulated phosphoproteins, one of which is specifically associated with dopamine-innervated brain regions. *J. Neuroscience* **10**, 1124-1133 (1990).
- S3. Grafstein-Dunn, E., Young, K.H., Cockett, M.I. & Khawaja, X.Z. Regional Distribution of regulators of G-protein signaling (RGS) 1, 2, 13, 14, 16, and GAIP messenger ribonucleic acids by in situ hybridization in rat. *Mol. Brain Res.* **88**, 113-123 (2001).
- S4. Ingi, T, *et al.* Dynamic regulation of RGS2 suggests a novel mechanism in G-protein signaling and neuronal plasticity. *J. Neuroscience* **18**, 7178-7188 (1998).
- S5. Betty, M., Harnish, S.W., Rhodes, K.J. & Cockett, M.I.. Distribution of heterotrimeric G-protein β and γ subunits in the rat brain. *Neuroscience* **85**, 475-486 (1998)
- S6. Zhang, Y, *et al.* Identification of genes expressed in *C. elegans* touch receptor neurons. *Nature* **418**, 331-5 (2002)
- S7. Saeed, A.I. *et al.* TM4: a free, open-source system for microarray data management and analysis. *Biotechniques* **34**, 374-378 (2003).
- S8. Tietjen, I. *et al.* Single-cell transcriptional analysis of neuronal progenitors. *Neuron* **38**, 161-175 (2003).
- S9. Kamme, F. *et al.* Single-cell microarray analysis in hippocampus CA1: demonstration and validation of cellular heterogeneity. *J Neurosci.* **23**, 3607-3615 (2003).
- S10. Arlotta, P. *et al.* Neuronal subtype-specific genes that control corticospinal motor neuron development in vivo. *Neuron* **45**, 207-221 (2005).
- S11. Svenningsson, P. *et al.* DARPP-32: an integrator of neurotransmission. *Annu Rev Pharmacol Toxicol.* **44**, 269-296 (2004).
- S12. Tecott, L.H. *et al.* Developmental expression of proenkephalin mRNA and peptides in rat striatum. *Brain Res Dev Brain Res.* **49**, 75-86 (1989).

