Supporting Information

Scherzer *et al.* 10.1073/pnas.0802437105

SI Methods

Bioinformatics and Data Deposition. The PD Biomarker Study (PDBS) dataset is described in ref. 1 and was previously made available in the Gene Expression Omnibus (GEO) database (accession no. GSE6613). The MGH HG-U133A (2), MGH CodeLink (2), and Stanford (3) blood datasets were downloaded from GEO (accession nos. GSE1751, GSE1767, and GSE3649, respectively).

We used all samples from the Stanford data and only control samples without neurologic disease ("healthy controls") from PDBS (1), MGH HG-U133A (2), and MGH CodeLink (2) datasets. In the discovery set (PDBS), Spearman's rank correlation coefficients between *SNCA* probes and 22,283 probes on the array were calculated. In the three validation sets (MGH HG-U133A, MGH CodeLink, and Stanford), the Spearman's rank correlation coefficients between *SNCA* probes and *ALAS2*, *FECH*, and *BLVRB* were determined. Heatmaps were generated in R to visualize Spearman's rho and *P*-values. See Table S2.

Identification of Genes Correlated with Expression of *SNCA* in the Discovery Set. We hypothesized that transcripts whose levels are tightly correlated represent a transcriptionally controlled expression block. To test this hypothesis we deconvoluted the coexpression patterns of 22,283 transcripts across blood specimens from 22 control individuals (1). Pairwise correlations between *SNCA* expression and the expression of 14,500 genes targeted by 22,283 probe sets determined that the expression of three genes was tightly linked to *SNCA* expression.

SNCA expression signals was measured by four probes on the HG-U133A oligonucleotide microarray. These probes query distinct sequences of the *SNCA* mRNA. As expected, the *SNCA* expression signals measured by the four probes highly correlated. *SNCA* probe set 204466_s_at exhibited the best correlation with the three other *SNCA* target gene probe sets with Spearman's rank correlation coefficients ranging from 0.81 to 0.94 compared with the other *SNCA* probe sets (204467_s_at, 207827_x_at, and 211546_x_at). Correlations in-between the *SNCA* probes are independent of technical noise introduced by sample collection and processing. Thus, we used a correlation coefficient of ≥ 0.81 as a rigorous limit for detecting genes with correlated expression (correlation of the expression of a non-*SNCA* gene with *SNCA* expression is equal to or higher than the correlations observed between *SNCA* probe (204466_s_at) and the three other *SNCA* probes on the array (204467_s_at, 207827_x_at, and 211546_x_at). Thirty-five unique genes met this criterion (Table S1). Consistent with a specific role of *SNCA* in erythroid cell biology, six of the coexpressed genes were known erythroid cell-specific genes, i.e., hemoglobin delta (HBD), and erythroid associated factor (*ERAF*), Glycophorin A and B (*GYPA* and *GYPB*), as well as erythrocyte membrane protein band 4.2 (*EPB42*) and erythroid membrane protein band 3 (*SLC4A1*). Importantly, three of the coexpressed genes, 5-aminolevulinate synthase 2 (*ALAS2*), ferrochelatase (*FECH*), and biliverdin reductase B (*BLVRB*), encode critical steps in heme metabolism.

Biospecimens. Snap frozen human brain tissue and blood biospecimens of donors without neurologic disease were collected and analyzed. Superior frontal gyrus (sample size = 2; mean age ± standard deviation = 90 ± 4), substantia nigra (n = 7; 69 ± 17), venous whole blood (n = 6; 67 ± 9), PBMC (n = 2; 72 ± 14), packed red blood cells (n = 3; 34 ± 8), serum, and plasma (n = 14; 66 ± 11) were used for gene or protein expression analyses. Blood biospecimens were obtained from the Harvard PD Biomarker Study (1, 4). Postmortem brain biospecimens were provided by the Human Brain and Spinal Fluid Resource Center at the University of California, Los Angeles, and by Jennifer Chan of the Department of Pathology, Brigham and Women's Hospital. Erythroid $CD71^+$ cells from human bone marrow (n = 1; 23) isolated by immunomagnetic cell sorting were purchased from AllCells. Blood specimens were also collected from age-matched, wild-type and *Snca*-null mice (5). For protein analyses, sample aliquots were frozen at −80°C until use. Upon thawing at 4°C, a protease inhibitor mixture and Nonidet P-40 (0.5%) were added before serial dilution and loading onto 384-well-based ELISA plates (6). Protocols for these studies were approved by the Institutional Review Board of Brigham and Women's Hospital.

Quantitative Chromatin Immunoprecipitation (ChIP). G1E-ER-GATA-1 cells, which stably express GATA-1 as a fusion to the human estrogen receptor ligand-binding domain, were maintained as described in ref. 7. Real-time PCR quantitative ChIP analysis was conducted as described in refs. 8 and 9. Chromatin fragments averaged ≈400 bp. Cells were grown in medium containing 15% FBS with or without 1 μM β-estradiol (Sigma) for 24 h. Immunoprecipitated DNA was analyzed by real-time PCR (ABI Prism 7000; PE Applied Biosystems). Primers were designed by using PRIMER EXPRESS 1.0 (PE Applied Biosystems) to amplify 50- to 150-bp amplicons and were based on sequences in Ensembl (www.ensembl.org/Mus_musculus/geneview?gene=ENSMUSG00000025889). Samples from three or more independent immunoprecipitations were analyzed. Product was measured by SYBR green fluorescence in 15-μl reactions. The amount of product was determined relative to a standard curve of input chromatin. Dissociation curves showed that PCRs yielded single products. Primer sequences are available upon request. Rabbit anti-GATA-1 and GATA-2 antibodies were generated against the respective N-terminal fragments of bacterially expressed and purified murine GATA-1 and GATA-2 (7). Preimmune serum was used as a control.

RNA Isolation and Quality Control. RNA was extracted from snap-frozen human postmortem brain biospecimens, CD71+ immunopurified erythroid cells, and packed red blood cells by TRIzol (GIBCO/BRL) extraction similar to what we describe in ref. 10. RNA from mononuclear white blood cells was extracted by Ficoll technique. To separate packed red blood cells from plasma and buffy coat, fresh whole blood samples were immediately centrifuged after phlebotomy at 2,000 × g for 10 min. Venous whole blood was collected in PAXgene (Qiagen) tubes and immediately incubated at room temperature for 24 h, and RNA was extracted according to the PAXgene procedure including DNase treatment (1). RNA from untreated G1E-ER-GATA-1 cells and cells treated with 1 mM β-estradiol for up to 40 h was extracted with TRIzol. RNA quality was determined by spectrophotometry and by visual inspection of electropherograms using the RNA 6000 NanoChip kit on an Agilent 2100 Bioanalyzer (Agilent Technologies). RNA passing quality control criteria was used for further analysis.

Quantitative Real-Time PCR. For quantitative gene expression analysis in human biospecimens, TaqMan Assay-on-Demand primers and probes (Applied Biosystems) were designed using the manufacturer's "rules," including crossing of exon junctions

(primer/probe information is available upon request). Amplification products were analyzed for specificity by agarose gel electrophoresis. The comparative threshold cycle method was used for analysis (11). Glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) and *RPL13* ribosomal RNA controlled for RNA loading. Equal amplification efficiencies were confirmed for target and reference genes. Total RNA (3 μg) was reverse-transcribed into cDNA using TaqMan Reverse Transcription reagents and random hexamers. Real-time PCR using ABI Prism 7000 and 7300 and TaqMan kits was performed according to the manufacturers' protocols. Samples were loaded in duplicate or triplicate. Control reactions lacking template or reverse transcription showed no amplification. Quantitative gene expression in cultured G1E cells was performed as reported in ref. 7. Primer sequences are available upon request.

ELISA. Rabbit sera were raised and affinity-purified at Open Biosystems against recombinant, full-length human *SNCA* protein. Antibody (Ab) Syn-1 was purchased from BD Transduction Labs. Recombinant human *SNCA* protein was a gift from P. Lansbury. For ELISAs, 96- and 384-well MaxiSorp plates (Nunc) were coated with a capturing polyclonal Ab (hSA-2; mSA-1) diluted in coating buffer ($NaHCO_3$ with 0.2% NaN_3, pH 9.6). After washes with PBS/0.05% Tween-20 (PBS-T), plates were blocked for 2 h at 37°C in blocking buffer (1.125% fish skin gelatin in PBS-T). After four washes, samples were loaded and incubated at 4°C for 12 h. Syn-1 mAb (assaying Ab) was biotinylated using 200 μg of Sulfo-NHS-LC Biotin (Pierce), diluted in blocking buffer, and incubated for 2 h at 37°C. After four washes, ExtrAvidin phosphatase (Sigma) was applied for 1 h at 37°C. Color development was carried out with Fast-*p*-Nitrophenyl Phosphate (Sigma) and monitored at 405 nm for up to 60 min, as described in ref. 12. For serial dilutions of cell lysates, whole blood, plasma, and serum, blocking buffer was used as diluent, which was also used to create the blank and the corresponding standard curve of recombinant human *SNCA* protein. Saturation kinetics were examined for identification of time point(s) where standards and sample dilutions were in the log phase. Concentrations were interpolated from the standard curve established by serial dilutions of HPLC- and mass spectrometry-characterized, human wild-type α-synuclein (13).

Western blot analysis and immunohistochemistry were performed as described in ref. 14. Anti-DJ-1 antibody is characterized in ref. 15. Sections of routinely fixed human bone marrow (provided by G. Pinkus, Department of Pathology, Brigham and Women's Hospital) were rehydrated, treated for 10 min with Gram's iodine solution (2 g of potassium iodide, 1 g of iodine, 100 ml of dH_2O), washed twice in dH_2O, treated with 5% thiosulfate (5 g of Na-thiosulfate, 100 ml of dH_2O) for 5 min, washed three times in dH_2O, and processed by routine immunohistochemistry.

Northern Blot Analysis. Erythroblasts infected with the anemia-inducing strain of Friend virus were cultured in the presence of erythropoietin, as described in ref. 16. Cells were harvested every 8 h, and RNA was extracted by the guanidinium isothiocyanate method (Tel-Test). Seven micrograms of RNA was loaded per lane. The RNA was transferred to Hybond-N+ membrane (General Electric Healthcare) by Northern blot analysis. A digital image of the 28S ribosomal RNA was obtained, before transfer, as a loading control through the use of a transilluminator (Bio-Rad). The analysis software was Quantity One 1-D. The filter was probed with α-synuclein cDNA, labeled with the Random Primer DNA labeling kit (Roche) in Hybrisol I (Chemicon). The PCR primers used to generate the *SNCA* probe were 5′-ctttagccatggatgtgttca-3′ and 5′-gtgcaatgacattcttaggct-3′.

siRNA transfection. The human dopaminergic cell line SH-SY5Y was maintained in DMEM/F12, supplemented with 10% FCS, penicillin (100 units/ml), and streptomycin (100 μg/ml). Cells from passages 6–20 were used, and they were seeded at 40% confluency into 30-mm six-well dishes. The required amount of *GATA2* siRNA (Invitrogen) and 5 μl of Lipofectamine RNAiMAX (Invitrogen) were diluted into a final volume of 250 μl in Opti-MEM I (GIBCO), respectively, then combined, gently mixed, and incubated at room temperature for 25 min. 500 μl of this transfection solution was overlaid onto cells at a final concentration of 5, 10, 20, 40, 80, or 160 nM siRNA. Transfection of SH-SY5Y cells with Stealth RNAi Negative Control (Med GC Content with no significant homology to any known gene sequences from mouse, rat, and human) served as a negative control. After a 48-h incubation at 37°C in the presence of 5% CO_2, cells were lysed by TRIzol reagent, and total RNA was isolated by chloroform/isopropanol precipitation. To determine α-synuclein levels, cells were lysed in Nonidet P-40- and protease inhibitor-containing lysis buffer at 4°C (6) and centrifuged at 10,000 or 100,000 $\times$ *g* for 30 min. The protein concentration of each supernatant was equalized, and three independently prepared specimens of each transfection arm were diluted (1:6) in ELISA blocking buffer (13).

1. Scherzer CR, *et al.* (2007) Molecular markers of early Parkinson's disease based on gene expression in blood. *Proc Natl Acad Sci USA* 104:955–960.
2. Borovecki F, *et al.* (2005) Genome-wide expression profiling of human blood reveals biomarkers for Huntington's disease. *Proc Natl Acad Sci USA* 102:11023–11028.
3. Whitney AR, *et al.* (2003) Individuality and variation in gene expression patterns in human blood. *Proc Natl Acad Sci USA* 100:1896–1901.
4. Hennecke G, Scherzer CR (2008) RNA biomarkers of Parkinson's disease: Developing tools for novel therapies. *Biomarkers Med* 2:41–53.
5. Abeliovich A, *et al.* (2000) Mice lacking α-synuclein display functional deficits in the nigrostriatal dopamine system. *Neuron* 25:239–252.
6. Schlossmacher MG, Shimura H (2005) *Parkinson Disease: Assays for the Ubiquitin Ligase Activity of Neural Parkin* (Humana, Totowa, NJ).
7. Im H, *et al.* (2005) Chromatin domain activation via GATA-1 utilization of a small subset of dispersed GATA motifs within a broad chromosomal region. *Proc Natal Acad Sci USA* 102:17065–17070.
8. Kiekhaefer CM, Grass JA, Johnson KD, Boyer ME, Bresnick EH (2002) Hematopoietic-specific activators establish an overlapping pattern of histone acetylation and methylation within a mammalian chromatin domain. *Proc Natl Acad Sci USA* 99:14309–14314.
9. Johnson KD, Christensen HM, Zhao B, Bresnick EH (2001) Distinct mechanisms control RNA polymerase II recruitment to a tissue-specific locus control region and a downstream promoter. *Mol Cell* 8:465–471.
10. Scherzer CR, Jensen RV, Gullans SR, Feany MB (2003) Gene expression changes presage neurodegeneration in a Drosophila model of Parkinson's disease. *Hum Mol Genet* 12:2457–2466.
11. Applied Biosystems (2001) User Bulletin #2 ABI Prism 7700 Sequence Detection System, pp 1–35.
12. Tokuda T, *et al.* (2006) Decreased α-synuclein in cerebrospinal fluid of aged individuals and subjects with Parkinson's disease. *Biochem Biophys Res Commun* 349:162–166.
13. Mollenhauer B, *et al.* (2008) Quantification of α-synuclein in cerebrospinal fluid: A first cross-sectional study in patients with neurodegeneration and exploration of its potential source. Exp Neurol, doi:10.1016/j.expneurol.2008.06.004.
14. Schlossmacher MG, *et al.* (2002) Parkin localizes to the Lewy bodies of Parkinson disease and dementia with Lewy bodies. *Am J Pathol* 160:1655–1667.
15. Baulac S, LaVoie MJ, Strahle J, Schlossmacher MG, Xia W (2004) Dimerization of Parkinson's disease-causing DJ-1 and formation of high molecular weight complexes in human brain. *Mol Cell Neurosci* 27:236–246.
16. Koury MJ, Sawyer ST, Bondurant MC (1984) Splenic erythroblasts in anemia-inducing Friend disease: A source of cells for studies of erythropoietin-mediated differentiation. *J Cell Physiol* 121:526–532.

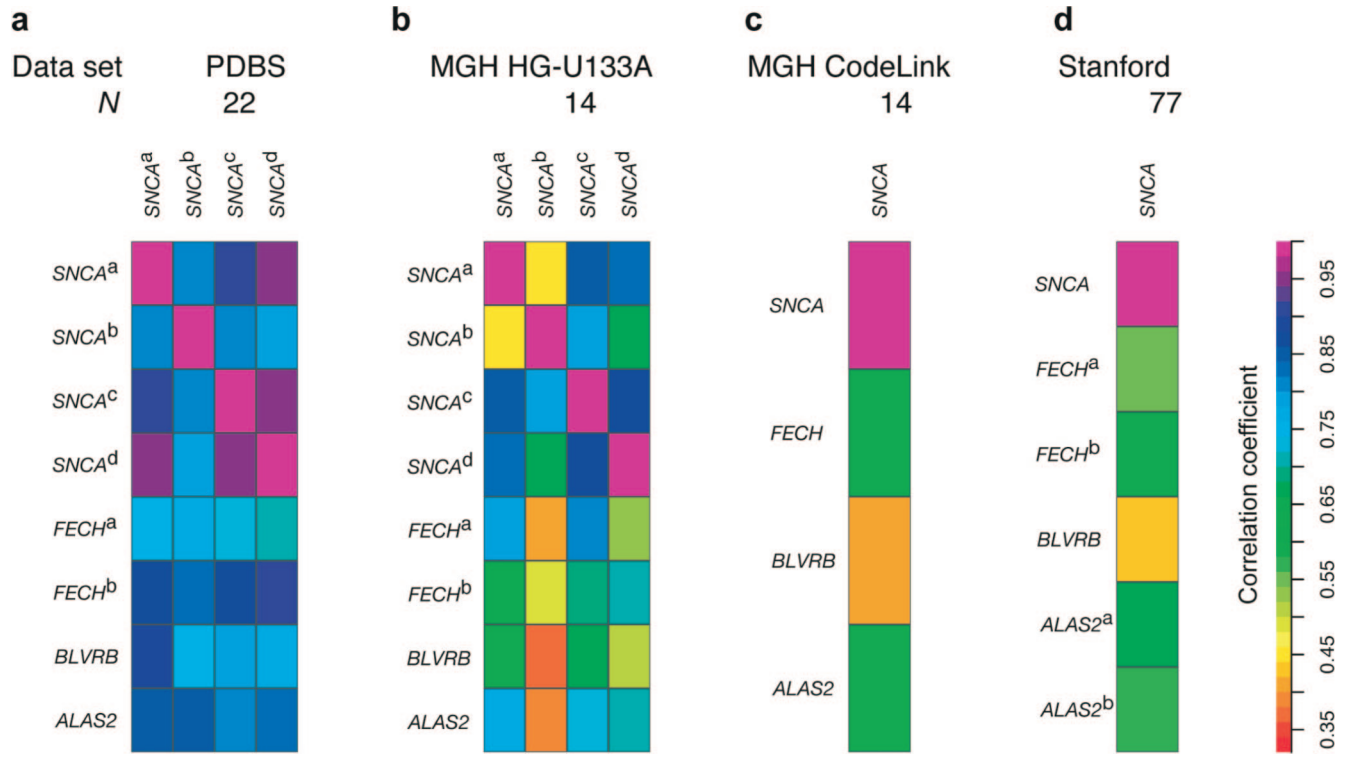

Fig. S1. Expression of *SNCA* and the three heme metabolism genes, *FECH*, *ALAS2*, and *BLVRB*, is tightly and significantly correlated across human blood samples in the discovery set and in three validation sets representing independent populations and assay platforms. Heatmaps visualize pairwise Spearman's rank correlation coefficients correlating expression of *SNCA* (*x* axis) and *SNCA*, *ALAS2*, *FECH*, and *BLVRB* expression (*y* axis). Strong correlations are color-coded purple or blue, and weak correlations are yellow and orange (see color bar and *SI Methods*). For genes that are detected by more than a single probe, correlations with each individual probe are shown. (*a*) Deconvolution of the coexpression patterns of 14,500 genes targeted by 22,283 probe sets across blood specimens from a discovery set of 22 normal individuals uncovered a tight and significant coexpression block of four genes, *SNCA*, *FECH*, *ALAS2*, and *BLVRB* (see text for details). Note that on the Affymetrix arrays *SNCA* is targeted by four probe sets and *FECH* is targeted by two probe sets. (*b*–*d*) If the coexpression of *SNCA* with *ALAS2*, *FECH*, and *BLVRB* is a robust and biologically relevant finding, it should be a universal signature in human blood. Rigorous validation of the coexpression block in three test datasets (MGH HG-U133A dataset, $n = 14$ healthy individuals; MGH CodeLink dataset, $n = 14$; and Stanford dataset, $n = 77$) confirmed that variation in *FECH*, *ALAS2*, *BLVRB*, and *SNCA* expression was tightly, reproducibly, and highly significantly linked. See the main text and *SI Methods* for detailed discussion.

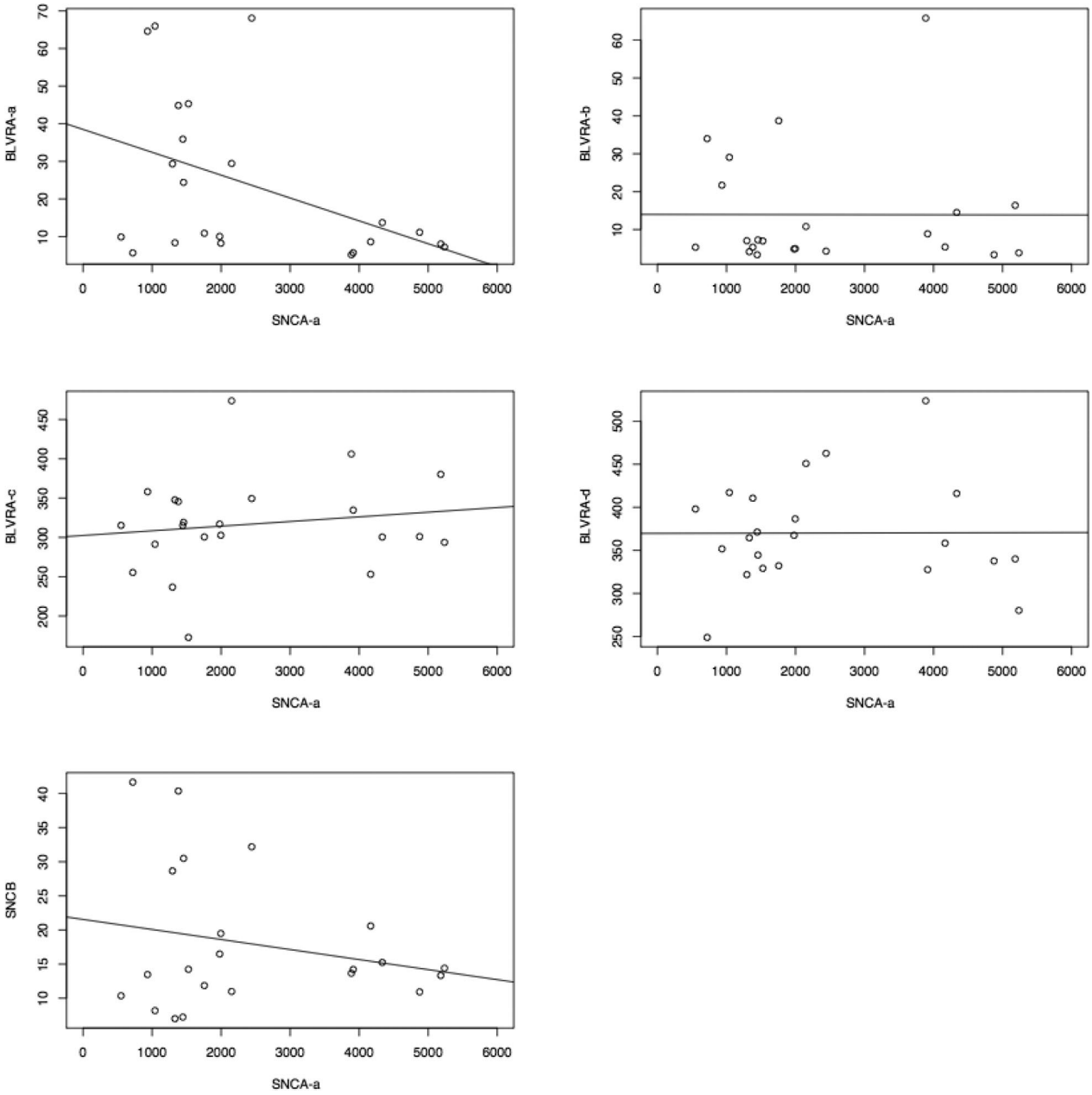

Fig. S2. Scatterplots correlating the expression of the pertinent negative control genes, *BLVR-A* (biliverdin reductase isoform A, the paralog of *BLVRB*) and *SNCB* (beta-synuclein, the paralog of *SNCA*) and *SNCA*. *BLVR-A* and *SNCB* expression is not correlated with *SNCA* expression.

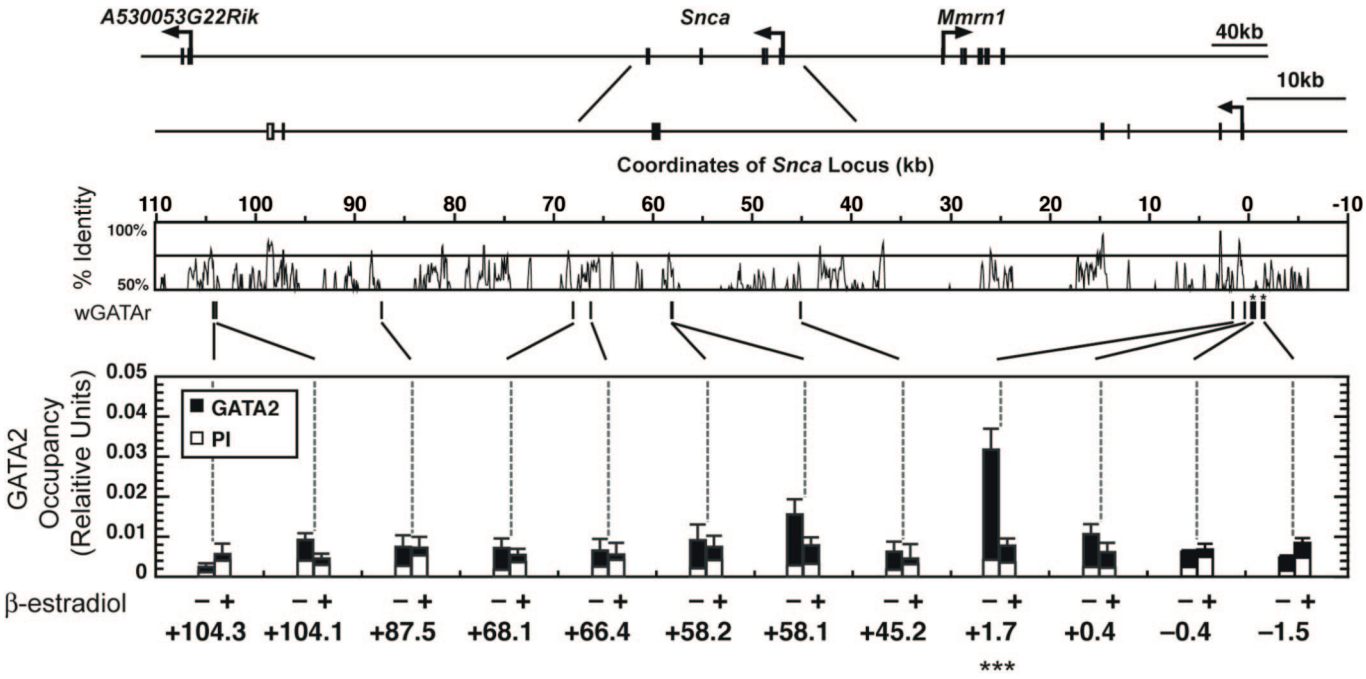

Fig. S3. Endogenous GATA-2 specifically occupies *Snca* intron-1 in the absence of GATA-1. The organization of the murine *Snca* locus with respect to neighboring genes on chromosome 6 and Vista plot are shown as in Fig. 4. Endogenous GATA-2 specifically occupies a WGATAR motif in intron-1 of the *SNCA* locus in the absence of GATA-1 (***). Endogenous GATA-2 occupancy is shown, measured by quantitative ChIP analysis in G1E-ER-GATA-1 cells, in which ER-GATA-1 was uninduced. The bar graphs depict relative GATA-2 occupancy in uninduced G1E-ER-GATA-1 cells at each of the 12 sites described in Fig. 4 (mean ± standard error, at least three independent experiments) measured by ChIP. Preimmune serum (PI) was used as a control.

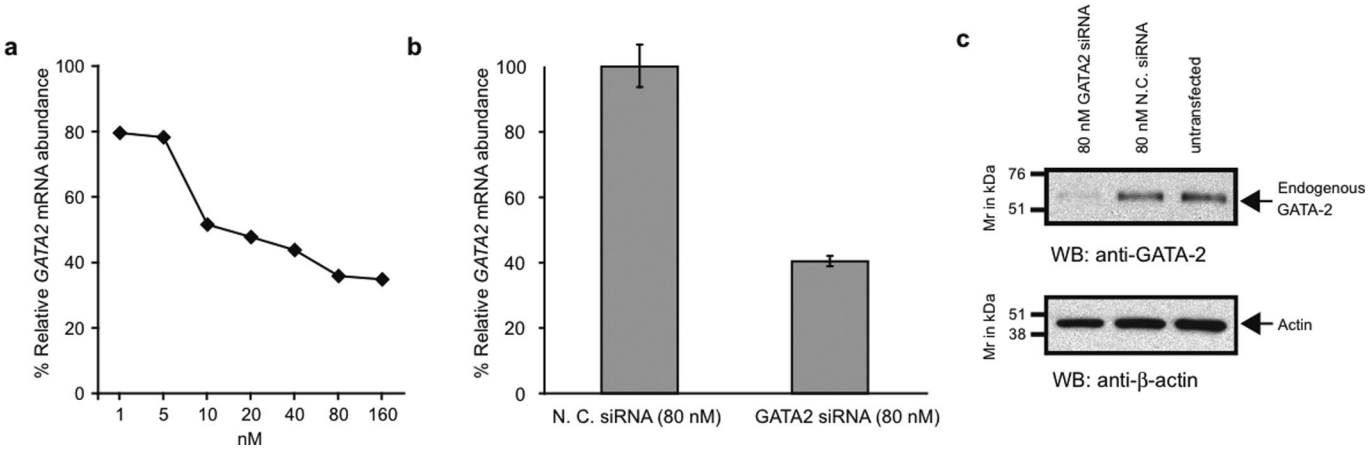

Fig. S4. (*a*) Quantitative PCR shows a dose-dependent reduction of *GATA2* mRNA abundance after transfection with 1–160 nM *GATA2* siRNA. (*b*) Transfection with 80 nM *GATA2* siRNA levels is sufficient to reliably knock down *GATA2* mRNA abundance to 40% of the abundance in cells transfected with negative control siRNA. (*c*) GATA-2 protein levels are substantially reduced after transfection with 80 nM *GATA2* siRNA by Western blot analysis compared with negative control transfected and untransfected cells.

Table S1. *SNCA* expression highly correlates with the expression of 35 unique genes in the discovery set

		Spearman rank correlation coefficient				*P*-value			
	Gene symbol	*SNCA*[a]	*SNCA*[b]	*SNCA*[c]	*SNCA*[d]	*SNCA*[a]	*SNCA*[b]	*SNCA*[c]	*SNCA*[d]
1	*ADIPOR1*	0.86	0.71	0.76	0.78	7.6e-07	1.5e-04	3.0e-05	1.3e-05
2	*ALAS2*	**0.84**	**0.85**	**0.80**	**0.82**	**4.4e-07**	**5.1e-07**	**4.5e-06**	**9.3e-07**
3	*BLVRB*	**0.89**	**0.74**	**0.79**	**0.78**	**1.6e-06**	**6.2e-05**	**9.2e-06**	**1.5e-05**
4	*BNIP3L*	0.82	0.73	0.71	0.78	1.1e-06	9.5e-05	1.5e-04	1.6e-05
5	*BPGM*	0.89	0.86	0.86	0.91	1.6e-06	6.2e-07	5.9e-07	2.0e-06
6	*CA1*	0.89	0.82	0.84	0.87	1.7e-06	1.2e-06	4.4e-07	1.1e-06
7	*EPB42*	0.91	0.83	0.90	0.85	2.0e-06	7.4e-07	1.8e-06	4.8e-07
8	*EPN1*	0.86	0.69	0.78	0.85	7.0e-07	2.4e-04	1.2e-05	4.4e-07
9	*ERAF*	0.91	0.79	0.84	0.92	1.9e-06	9.2e-06	4.9e-07	2.0e-06
10	*FBXO7*	0.87	0.78	0.77	0.75	1.1e-06	1.3e-05	2.4e-05	4.5e-05
11	*FECH*[b]*	**0.88**	**0.84**	**0.86**	**0.91**	**1.3e-06**	**4.7e-07**	**7.6e-07**	**1.9e-06**
12	*GLRX5*	0.90	0.84	0.83	0.92	1.8e-06	4.9e-07	8.6e-07	2.0e-06
13	*GMPR*	0.88	0.88	0.83	0.80	1.2e-06	1.2e-06	7.4e-07	6.1e-06
14	*GSPT1*	0.91	0.83	0.81	0.84	1.9e-06	6.0e-07	2.7e-06	4.6e-07
	GSPT1	0.87	0.81	0.78	0.75	9.9e-07	2.7e-06	1.6e-05	4.3e-05
15	*GYPA*	0.85	0.65	0.73	0.76	4.9e-07	6.2e-04	7.7e-05	2.7e-05
16	*GYPB*	0.81	0.70	0.67	0.76	1.9e-06	1.8e-04	3.9e-04	2.7e-05
	GYPB	0.91	0.86	0.87	0.90	1.9e-06	6.5e-07	1.1e-06	1.8e-06
17	*HBD*	0.94	0.81	0.86	0.89	2.1e-06	2.4e-06	6.7e-07	1.5e-06
18	*HBLD2*	0.82	0.76	0.83	0.88	1.3e-06	2.7e-05	5.4e-07	1.4e-06
19	*IL8*	0.88	0.85	0.80	0.84	1.2e-06	5.5e-07	4.5e-06	5.1e-07
20	*LGALS3*	0.84	0.85	0.83	0.81	4.9e-07	4.8e-07	8.0e-07	2.7e-06
21	*MKRN1*	0.82	0.83	0.76	0.74	1.2e-06	6.9e-07	3.0e-05	6.2e-05
	MKRN1	0.84	0.79	0.71	0.73	4.4e-07	8.2e-06	1.7e-04	9.8e-05
22	*MPP1*	0.85	0.78	0.71	0.72	5.1e-07	1.1e-05	1.6e-04	1.3e-04
23	*NUDT4*	0.85	0.86	0.75	0.80	4.9e-07	6.7e-07	5.1e-05	3.6e-06
24	*PIP5K2A*	0.88	0.74	0.83	0.87	1.4e-06	6.9e-05	5.4e-07	1.1e-06
25	*PPM1A*	0.88	0.77	0.83	0.85	1.2e-06	2.1e-05	5.4e-07	5.5e-07
26	*PRDX2*	0.87	0.73	0.79	0.89	1.1e-06	8.5e-05	6.5e-06	1.6e-06
27	*PSMF1*	0.82	0.73	0.78	0.74	1.5e-06	9.8e-05	1.4e-05	6.9e-05
28	*RIOK3*	0.84	0.79	0.84	0.81	4.5e-07	9.2e-06	4.4e-07	3.0e-06
	RIOK3	0.82	0.66	0.70	0.79	1.1e-06	5.7e-04	1.9e-04	7.8e-06
29	*RNF10*	0.83	0.87	0.82	0.80	6.0e-07	9.9e-07	1.6e-06	4.8e-06
30	*RNF14*	0.83	0.77	0.76	0.82	8.0e-07	2.4e-05	3.0e-05	1.1e-06
31	*RPIA*	0.86	0.80	0.83	0.79	7.3e-07	5.7e-06	8.0e-07	8.7e-06
32	*SELENBP1*	0.91	0.84	0.86	0.83	1.9e-06	4.5e-07	7.6e-07	8.6e-07
	SNCA[a]	**1.00**	**0.81**	**0.92**	**0.94**	**1.2e-06**	**2.6e-06**	**2.0e-06**	**2.0e-06**
	SNCA[b]	**0.81**	**1.00**	**0.80**	**0.79**	**2.6e-06**	**1.2e-06**	**4.5e-06**	**7.7e-06**
	SNCA[c]	**0.92**	**0.80**	**1.00**	**0.94**	**2.0e-06**	**4.5e-06**	**1.2e-06**	**2.0e-06**
	SNCA[d]	**0.94**	**0.79**	**0.94**	**1.00**	**2.0e-06**	**7.8e-06**	**2.0e-06**	**1.2e-06**
33	*SNX3*	0.83	0.57	0.65	0.72	5.7e-07	3.0e-03	6.2e-04	1.0e-04
34	*TMOD1*	0.85	0.79	0.79	0.75	4.8e-07	7.3e-06	9.7e-06	4.2e-05
35	*TRAK2*	0.87	0.76	0.72	0.80	1.1e-06	3.2e-05	1.1e-04	4.2e-06

SNCA[a], 204466_s_at; *SNCA*[b], 204467_s_at; *SNCA*[c], 207827_x_at; *SNCA*[d], 211546_x_at; *FECH*[b], 203116_s_at.

*Only probes meeting the $R \geq 0.81$ threshold are listed. The *FECH*[a] probe did not meet this threshold and is not listed here.

Table S2. Expression of *SNCA* and three heme metabolism genes is tightly correlated

Blood dataset	Platform	No. of samples	No. of probes
PDBS	Affymetrix HG-U133A	22	22,283
MGH HG-U133A	Affymetrix HG-U133A	14	22,283
MGH CodeLink	CodeLink Uniset 20k	14	19,745
Stanford	Custom	77	37,402

Table S3. Validation of the correlation in *SNCA* and *FECH*, *ALAS2*, and *BLVRB* expression

			Probe level				Gene level
Study	Dataset	Platform	Correlated probes	Corr. coeff.	*P*-value	Probe-level replication	Replication
			Validation of the correlation in *SNCA* and *FECH* expression				
1	MGH HG-U133A	Affymetrix	*SNCA*[a] and *FECH*[a]	0.78	0.00068	Yes	Yes
			SNCA[b] and *FECH*[a]	0.42	0.067	No	
			SNCA[c] and *FECH*[a]	0.8	0.00043	Yes	
			SNCA[d] and *FECH*[a]	0.52	0.027	Yes	
			SNCA[a] and *FECH*[b]	0.65	0.0056	Yes	
			SNCA[b] and *FECH*[b]	0.48	0.039	Yes	
			SNCA[c] and *FECH*[b]	0.69	0.0028	Yes	
			SNCA[d] and *FECH*[b]	0.71	0.0022	Yes	
2	MGH CodeLink	GE HealthCare	*SNCA* and *FECH*[a]	0.63	0.0093	Yes	Yes
3	Stanford	Custom cDNA array	*SNCA* and *FECH*[a]	0.54	0.00000015	Yes	Yes
			SNCA and *FECH*[b]	0.64	0.00000000018	Yes	
			Validation of the correlation in *SNCA* and *ALAS2* expression				
1	MGH HG-U133A	Affymetrix	*SNCA*[a] and *ALAS2*	0.76	0.0011	Yes	Yes
			SNCA[b] and *ALAS2*	0.38	0.089	No	
			SNCA[c] and *ALAS2*	0.73	0.0019	Yes	
			SNCA[d] and *ALAS2*	0.71	0.003	Yes	
2	MGH CodeLink	GE HealthCare	*SNCA* and *ALAS2*	0.59	0.015	Yes	Yes
3	Stanford	Custom cDNA array	*SNCA* and *ALAS2*	0.67	0.000000000016	Yes	Yes
			Validation of the correlation in *SNCA* and *BLVRB* expression				
1	MGH HG-U133A	Affymetrix	*SNCA*[a] and *BLVRB*	0.64	0.0079	Yes	Yes
			SNCA[b] and *BLVRB*	0.36	0.1	No	
			SNCA[c] and *BLVRB*	0.67	0.005	Yes	
			SNCA[d] and *BLVRB*	0.5	0.03	Yes	
2	MGH CodeLink	GE HealthCare	*SNCA* and *BLVRB*	0.41	0.07	No	No
3	Stanford	Custom cDNA array	*SNCA* and *BLVRB*	0.42	0.000066	Yes	Yes

SNCA[a], 204466_s_at; *SNCA*[b], 204467_s_at; *SNCA*[c], 207827_x_at; *SNCA*[d], 211546_x_at; *FECH*[a], 203115_at; *FECH*[b], 203116_s_at. Corr. coeff., Spearman's rank correlation coefficient.