SUPPLEMENTAL INFORMATION

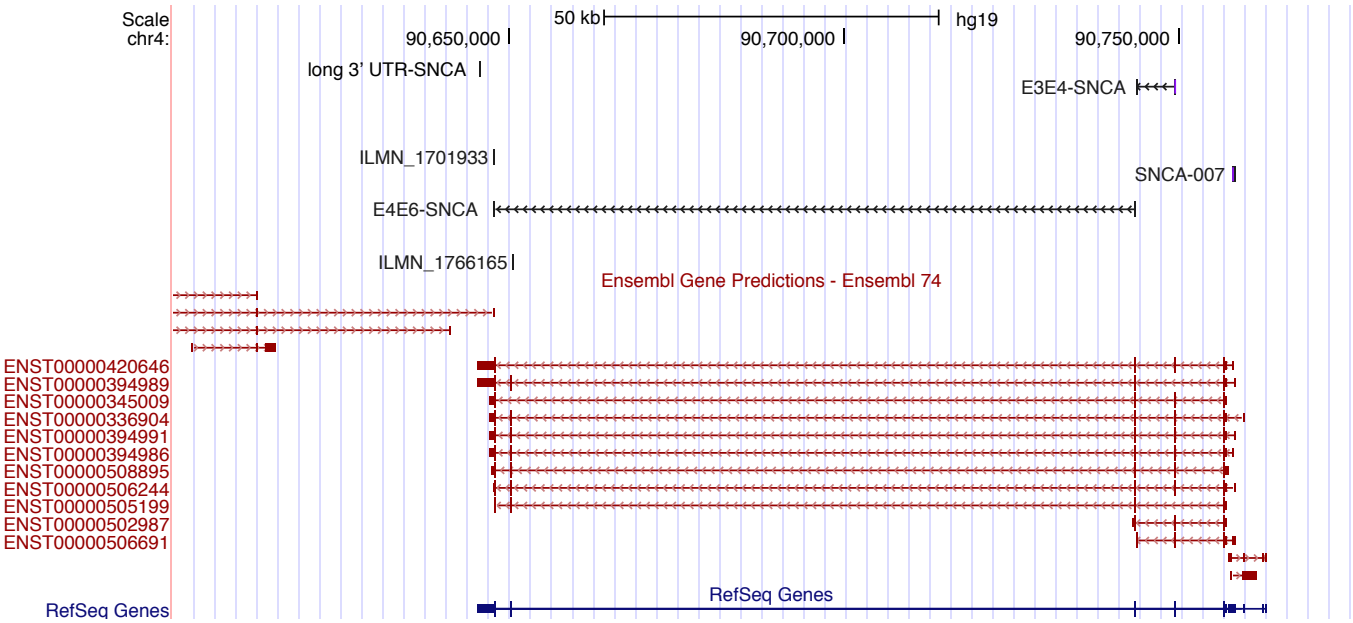
Association Between α-Synuclein Blood Transcripts and Early, Neuroimaging-Supported Parkinson's Disease

Joseph J. Locascio, PhD, Shirley Eberly, Zhixiang Liao, Ganqiang Liu, Ashley N. Hoesing, Karen Duong, Ana Trisini-Lipsanopoulos, Kaltra Dhima, Albert Y. Hung, Alice W. Flaherty, Michael A. Schwarzschild, Michael T. Hayes, Anne-Marie Wills, U. Shivraj Sohur, Nicte I. Mejia, Dennis J. Selkoe, David Oakes, Ira Shoulson, Xianjun Dong, Ken Marek, Bin Zheng, Adrian Ivinson, Bradley T. Hyman, John H. Growdon, Lewis R. Sudarsky, Michael G. Schlossmacher, Bernard Ravina, and C. R. Scherzer

List of Supplementary Materials:

A. Supplemental Data

1. Supplemental Figure S1. Target sequences of various *SNCA* probes used in this study, related to Figure 3.



The small vertical black lines indicate the genomic location of various probes used. Probes for digital transcript counting assays targeted (i) the boundaries of exon 3 and exon 4 (labeled E3E4-*SNCA*), (ii) specifically transcripts with an extended 3' untranslated region (labeled long-3'UTR-*SNCA*), (iii) transcripts that skip exon 5 (labeled E4E6-*SNCA* transcripts), or (iv) exon 1 of the rare short *SNCA*-007 transcript isoform. The genomic location of Illumina microarray probes used for technical confirmation of qPCR results in the PROBE study is also shown (ILMN_1701933 and ILMN_1766165). qPCR probes used in HBS and PROBE targeted the exon 3/exon 4 junction as shown for the *E3E4-SNCA* probe for digital transcript counting. The *SNCA* reference gene sequence is shown in blue at the bottom. Sequences of the eleven *SNCA* transcript isoforms annotated in the Ensembl genome Browser are shown in red with corresponding Ensembl IDs on the left hand side. This figure was created using the UCSC Genome Browser http://genome.ucsc.edu/.

2. Supplemental Table S1. The PROBE study - patient characteristics, related to Figure 1

	Patients with Parkinson's disease (*N* = 93)	**Controls (*N* = 49)**	***P* value**
Age at phlebotomy (years, SD)	62.8 (9.5)	61.4 (9.6)	0.42
Gender			
Male (N, %)	62 (67%)	26 (53%)	0.11
Female (N, %)	31 (33%)	23 (47%)	
Clinical findings (N, %)			
Resting tremor	80 (86%)		
Bradykinesia	93 (100%)		
Rigidity	92 (99%)		
Postural imbalance	32 (34%)		
Asymmetry	93 (100%)		
Medications (N, %)			
Levodopa†	74 (79%)		
Dopamine replacement medications††	88 (95%)		
De novo†††	5 (5%)		
Disease severity			
Disease Duration (years) (mean, SD)	5.5 (0.9)		
Hoehn and Yahr (Mean, SD)	2.0 (0.5)		
UPDRS Part I - mentation, behavior, mood	1.7 (1.8)		
UPDRS Part II - ADLs	8.5 (5.5)		
UPDRS Part III - motor examination	20.4 (11.1)		
UPDRS Total Score	30.6 (16.2)		
Hematological Variables (mean, SD)			
White Blood Cells	6.1 (1.5)	6.3 (1.3)	0.30
Red Blood Cells	4.7 (0.4)	4.7 (0.4)	0.76
Platelets	241.1 (59.5)	264.0 (62.0)	0.04
Hemoglobin	14.6 (1.2)	14.6 (1.3)	0.93

† number of subjects on levodopa (alone or in combination with other PD medications); †† number of subjects on any kind of dopamine replacement medication (e.g. levodopa, dopamine agonists, etc.); †††, number of subjects not treated with levodopa or other dopamine replacement medications).

3. Supplemental Table S2. The PPMI study - patient characteristics, related to Figure 3.

	Patients with Parkinson's disease (*N* = 202)	**Controls (*N* = 138)**	***P* value**
Age at phlebotomy (mean, SD)	59.6 (7.4)	59.8 (11.7)	0.9
Gender			
Male (N, %)	120 (59.4%)	84 (60.9%)	0.8
Female (N,%)	82 (40.6%)	54 (39.1%)	
Clinical findings (N, %)			
Resting tremor	134 (66.3%)		
Bradykinesia	201 (99.5%)		
Rigidity	193 (95.5%)		
Postural imbalance	16 (7.9%)		
Medications (N, %)			
Levodopa or any dopamine replacement medications	4 (2%)		
De novo	198 (98%)		
Disease severity (mean, SD)			
Disease duration (years)	0.6 (0.6)		
Hoehn & Yahr	1.6 (0.5)		
MDS-UPDRS Part I Score	5.71 (3.94)		
MDS-UPDRS Part II Score	5.84 (4.16)		
MDS-UPDRS Part III Score	21.27 (8.52)		
MDS-UPDRS Total Score	32.82 (12.99)		
Modified Schwab & England ADL	92.64 (5.99)		
MOCA	27.3 (2.2)		
Geriatric Depression Scale	2.32 (2.51)		
SCOPA-AUT	9.28 (5.98)		
UPSIT Total Score	22.46 (7.96)		
Dopamine Transporter Binding Ratios with I123 Ioflupane (mean, SD)			
Right Caudate	2.0 (0.6)	3.6 (0.6)	<0.0001
Left Caudate	2.0 (0.6)	3.0 (0.7)	<0.0001
Right Putamen	0.8 (0.4)	2.1 (0.6)	<0.0001
Left Putamen	0.8 (0.3)	2.1 (0.6)	<0.0001

4. Supplemental Table S3. Baseline characteristics of the 25% patients with PD with the highest *SNCA* expression values and the 25% of PD patients with the lowest *SNCA* expression values at enrollment.

HBS	**PD with high *SNCA* expression**	**PD with low *SNCA* expression**	***P*-value**
Total number (N)	55	55	
Number of men (N, %)	30 (54.55%)	38 (69.09%)	0.17
Age at enrollment (years, SD)	68.65 (8.84)	65.89 (9.27)	0.11
Disease duration at enrollment (years, SD)	4.35 (4.85)	4.73 (4.34)	0.66
UPDRS total score (mean, SD)	26.71 (13.51)	30.55 (14.78)	0.15
Hoehn and Yahr (mean, SD)	2.02 (0.6)	2.09 (0.56)	0.54
MMSE (mean, SD)	28.37 (1.95)	28.53 (2.28)	0.26
Levodopa equivalent drug dose (mean, SD)	482.4 (352.32)	574.06 (515.96)	0.42
Years of education (mean, SD)	15.13 (1.67)	14.96 (2.08)	0.65
PROBE	**PD with high *SNCA* expression**	**PD with low *SNCA* expression**	***P*-value**
Total Number (N)	16	16	
Number of Men (N, %)	9 (56.25%)	12 (75%)	0.46
Age at enrollment (years, SD)	61.62 (8.87)	66.59 (9.29)	0.13
Disease duration at enrollment (years, SD)	5.64 (0.62)	5.71 (1.12)	0.82
UPDRS total score (mean, SD)	22.06 (9.64)	29.07 (17.47)	0.25
Hoehn and Yahr (mean, SD)	1.94 (0.31)	2.09 (0.38)	0.39
MMSE (mean, SD)	29.19 (0.98)	28.44 (2.78)	0.95
Years of education (mean, SD)	15.38 (3.12)	15.75 (1.81)	0.68

B. Supplemental Experimental Procedures

A. Harvard Biomarker Study

Study population. A case-control study nested in the ongoing, longitudinal Harvard Biomarker Study (HBS) was performed. 222 cases with PD and 183 controls without neurologic disease enrolled in HBS prior to 9/16/2010 and meeting stringent RNA quality-control criteria (below) were included for evaluation. 21 of the 222 cases with PD were *de novo* patients who had not received treatment with dopamine replacement medications prior to and at the time of phlebotomy.

HBS is a Harvard-wide, longitudinal, case-control study designed to accelerate the discovery and validation of molecular markers that track or predict progression of early-stage Parkinson's disease (PD) and memory impairment. More than 2,000 individuals with early-stage PD (defined as a Modified Hoehn and Yahr stage ≤ 3 upon enrollment), individuals with memory impairment (without PD), and controls without neurologic disease have been enrolled from Massachusetts General Hospital and Brigham & Women's Hospital. Inclusion criteria for cases with PD were age ≥ 21; diagnosis of PD according to UK brain bank criteria or movement disorders specialist diagnosis (Hughes et al., 2002); MMSE score > 21 or legal guardian or next of kin present to provide informed consent; able to provide informed consent. The UKPDSBB criteria (modified as described in Ref. Ding et al., 2011) required that patients have bradykinesia, at least one of the other cardinal features (rigidity, rest tremor, postural instability), lack exclusion criteria, and have at least 3 of 8 supportive features (such as asymmetric onset, levodopa responsiveness, and chronic progression of clinical features). Exclusion criteria for cases with PD in HBS were diagnosis of a blood or bleeding disorder, known hematocrit < 30, known active ulcer or active colitis. To increase diagnostic certainty, subjects with a diagnosis change on longitudinal follow-up were excluded from analysis.

Inclusion criteria for healthy controls were non-blood relatives (largely spouses) of cases with PD or patients with AD enrolled in HBS, no current diagnosis or history of a neurological disease, ability to provide informed consent, and age ≥ 21 (≥ 30 for spouses of AD patients). Exclusion criteria for non-blood relatives of patients with PD: same as for cases. Using this definition the controls were comparable to the PD

cases in that were drawn from the same source population and could be identified as a case, if they had the disease. The study was approved by the Institutional Review Boards of Brigham and Women's Hospital and Massachusetts General Hospital.

Clinical assessments of severity. Clinical assessments included the Unified Parkinson Disease Rating Scale (UPDRS), a 40-item assessment of impairment and disability in 4 subscales (Fahn, 1987). Part I addresses cognitive and behavioral aspects of PD. Part II addresses activities of daily living. Part III deals with the motor examination and part IV addresses complications of therapy including wearing off. This scale is well established in terms of reliability and validity (Siderowf et al., 2002) and has been used in multiple PD studies (Ramaker et al., 2002). Clinical disease severity was assessed in the operationally defined "on state" based on expert judgment at the time of the visit as in Refs. (Alves et al., 2005; Jankovic and Kapadia, 2001; Louis et al., 1999). Patients thus were not necessarily at the peak of optimal response (true on state) or in their true off state (at least 24 hours and as much as 4 weeks after the last dose). The Hoehn and Yahr assessment is a five-point scale that describes the subject's disability and impairment in PD, with stage one being the least disabled and stage five indicating the most severe disability (Hoehn and Yahr, 1967). The Mini-Mental State Examination (MMSE), is a brief standardized assessment of cognition in adults (Cockrell and Folstein, 1988). It has many potential uses, including screening for cognitive impairment, estimating the severity of cognitive impairment, and following the progression of cognitive changes over time. It has been used in several PD studies. Similar assessments of clinical severity were performed in PROBE and in PPMI.

Sample collection. HBS was designed to minimize bias from sample processing by collecting, handling and analyzing specimens of cases and controls in a standardized and blinded manner according to rules of evidence (Hennecke and Scherzer, 2008; Ransohoff, 2004, 2005). Phlebotomy and biospecimens processing were performed in a standardized manner. Cases and controls were processed in parallel to avoid bias due to "run order" of samples.

RNA isolation and quality control. Venous blood was collected in PAXgene tubes (Qiagen, Valencia, CA) at the baseline visit. Blood and PAXgene reagents were mixed by gently inverting 8-10 times and incubated upright at room temperature for 24 hours. PAXgene tubes were placed at 4°C and RNA was extracted within five days according to the PreAnalytix PAXgene Blood RNA kit manual extraction protocol including DNase treatment. RNA quality was determined with the RNA 6000 Nano Chip on an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA) and the RNA Integrity Number package(Imbeaud et al., 2005). Only RNA samples meeting stringent Q/C parameters were included for analysis. These were (a) RNA Integrity Number (RIN)(Auer et al., 2003) ≥ 7.3 by Agilent Bioanalyzer (excellent mRNA integrity); (b) ratio of absorbances at 260 nm/280 nm > 1.8 (no protein contamination); and (c) 28S/18S ribosomal RNA ratio ≥ 1.0 (intact ribosomal RNA)**.**

Quantitative real-time polymerase chain reaction. TaqMan Assay-on-Demand primers and probes were designed according to the manufacturer's guidelines including crossing of exon junctions. 500 ng of RNA from each blood sample was reverse-transcribed using the High-Capacity cDNA ReverseTranscription Kit (Applied Biosystems). qPCR was performed on 384-well plates on an ABI 79000HT System with TaqMan Universal PCR Master Mix (Applied Biosystems). Expression values were analyzed using the comparative threshold cycle method (Applied Biosystems, 2001). The geometric mean of four rationally selected endogenous reference genes was used to control for RNA loading. The reference gene selection method is described below. The operator performing the assay was blinded to the diagnosis. No template (negative) controls containing water substituted for template were included (NTC). To control for plate-to-plate variation and drift equal amounts of RNA derived from Human Universal Reference RNA (UR) were spotted at the beginning, end, and within the entire experiment. Control reactions lacking template or reverse transcription showed no amplification. Target and reference genes showed similar amplification efficiencies in a dilution series. Probe information is listed below. To estimate the precision (or re-test reliability) in our

experiment, 5% of RNA samples were randomly re-sampled during the entire experiment.

Endogenous reference gene selection. Reference genes, or housekeeping genes, should be stably expressed in the tissue and condition of interest. To find such genes, Vandesompele et al. proposed an algorithm to identify the most stable genes in a given expression data set (Vandesompele et al., 2002). The stability of each gene, measured by a M-value, is calculated by averaging all pair wise variations of the gene with other genes. A gene with a higher M-value is considered as less stable than a gene with a lower M-value. To find suitable, stably expressed reference genes for analysis of human blood gene expression from older individuals, we examined a qPCR dataset exploring gene expression values of 69 genes across 246 human blood samples collected and processed using our standard operating procedure (see above) using the Vandesompele et al. method (Vandesompele et al., 2002). The four genes with the lowest M-values in this data set were *MON1B* (0.1998), *SRCAP* (0.1998), *GLT25D1* (0.2213), as well as the traditional "housekeeping gene" *GUSB* (0.2438).

For qPCR analysis we used the geometric mean of these four most stably expressed genes to normalize for input RNA using the Comparative CT method (Applied_Biosystems, 2001). We then used log-transformed deltaCT values (DCT) for all statistical and significance analyses. For presentation purposes only, in Table 2 and Figure 1, linearizing transformations for qPCR variables (DCT values) were performed as follows, e.g. $2^{-\exp(\text{mean log-transformed DCT})}$. In Table 3, for qPCR variables (DCT values) a linearizing transformation was applied to the ranges of each quartile of values shown, e.g. $2^{-\text{lowest DCT}}$ and $2^{-\text{highest DCT}}$.

No.	Gene	Target/Reference	ABI Assay ID	Probe Sequence
1	SNCA	Target	Hs00240906_m1	GTGGCAACAGTGGCTGAGAAGACCA
2	MON1B	Reference	Hs00208876_m1	ACTGGCCTGGGTGACCTCCAAATTC
3	GUSB	Reference	Hs99999908_m1	TGACTGAACAGTCACCGACGAGAGT
4	SRCAP	Reference	Hs01064942_m1	TGCCATCAAGGCGGAAGAGCGGCAT
5	GLT25D1	Reference	Hs00430696_m1	GACTGGGACCTCATCTATGTGGGCC

B. Blood α-Synuclein, Gene Expression, and Smell Testing as Diagnostic and Prognostic Biomarkers in Parkinson's Disease Study (PROBE)

Study population. PROBE is a nation-wide, multi-center, case-control biomarker study nested in the parent observational PostCEPT study (Ravina et al., 2009). The PD patients enrolled in PROBE were initially recruited in a negative phase III clinical trial of a mixed lineage kinase inhibitor (PRECEPT) (2007). A total of 537 of 709 eligible PRECEPT trial cases (76%) were then enrolled in the observational PostCEPT study for longitudinal clinical follow-up and as a resource for biological marker assessments (Ravina et al., 2009). These analyses include 93 PD cases from PostCEPT who were subsequently re-enrolled in the present PROBE study, as well as 49 new controls who were recruited into PROBE as described below.

PD inclusion criteria were a diagnosis of PD according to UK brain bank criteria for PD; willing and able to provide informed consent. Healthy controls were defined as absent any known neurological disease. Operationally controls included any non-blood relatives of PD cases, including spouses and in-laws. Using this definition the controls are comparable to the PD cases in that they are drawn from the same population and could be identified as a case if they had disease. To prevent a lower average age in the control group we restricted the lower age limit to 45 for healthy controls, which is 2 standard deviations below the average age in the PD cohort. For healthy controls we have also included a MMSE to exclude those with impaired cognition such as early AD. The cutoff of 27 represents the lower quartile for a group in this age range with 9 or more years of education (Crum et al., 1993). Definitive diagnosis requires histopathology. Exclusion criteria for all groups included the use of anticoagulants (e.g., warfarin or heparin) or a bleeding disorder (acquired or inherited); blood disorder (e.g. leukemia) or a history of anemia with a documented hematocrit of <30; known pregnancy; nasal trauma, sinusitis, or other nasal pathology that would interfere with smell testing. Pregnant women were excluded because of potential effects of pregnancy on gene expression results. Patients with coagulopathies were excluded for safety reasons. Subjects with a history of severe anemia that may confound the α-synuclein assays were also excluded.

Clinical assessments of severity. Clinical measures were similar to what we described for HBS above. In addition, the Modified Schwab and England Activities of Daily Living Scale was performed, which reflects the speed, ease, and independence with which an individual performs daily activities, or personal chores, such as eating, toileting, and dressing (Fahn, 1987). This scale uses a rating scale from 0% to 100%; with 100% representing complete independence in performing daily activities and 0% representing a vegetative, bedridden state. [^{123}I] β-CIT with SPECT, targeting the dopamine transporter was performed on cases (not on controls). It has been used in numerous PD studies including the PRECEPT study (Ravina et al., 2009). [^{123}I] β-CIT, which is a tropane derivative, binds to the DAT and with lower affinity to the serotonin transporter. DAT expression within dopaminergic neurons is reduced in PD brains compared to controls (Counihan and Penney, 1998; Uhl et al., 1994). The clinical assessments were completed as part of the "parent" PostCEPT study (Ravina et al., 2009).

RNA isolation and quality control. Briefly, venous whole blood was collected in three PAXgene (Qiagen, Valencia, CA) tubes at the baseline visit (2.5 ml per tube, total volume 7.5 ml). Blood and PAXgene reagents were mixed by gently inverting 8-10 times and incubated at room temperature for 24 hours. PAXgene tubes were then stored frozen at -20° C and batch-shipped on dry ice to Brigham & Women's Hospital. RNA was isolated as described for HBS and quality-controlled by Agilent Bioanalyzer(Scherzer et al., 2008). Only PROBE RNAs with preserved ribosomal peaks on electropherogram and RNA integrity numbers ≥ 7 were forwarded for further analyses.

Quantitative real-time polymerase chain reaction was performed as described for the HBS study.

Microarray procedures were performed as described (Zheng et al., 2010). 350 ng of total RNA were reverse transcribed for each subject and hybridized to Illumina HumanHT-12v3 Expression BeadChips targeting more than 25,000 annotated genes with 48,803 probes derived from the National Center for

Biotechnology Information Reference Sequence (NCBI) RefSeq (Build 36.2, Rel 22) and the UniGene (Build 99) databases, and scanned on a BeadArray Reader. Data were processed, normalized by "average normalization", and quality-controlled using GenomeStudio. The PROBE microarray data set has been submitted to the GEO database (accession number GSE57475).

C. Parkinson's Progression Markers Initiative (PPMI)

A case-control study nested in the ongoing, longitudinal PPMI study was performed. 358 RNA samples meeting stringent RNA quality-control criteria (below) were obtained from PPMI. 17 of these were biological replicates and one sample failed NanoString analysis. Thus, a total of 340 unique samples of 202 unique cases and 138 unique controls were used for final analysis. The PPMI study is an ongoing international multicenter study involving institutions in Europe, Australia, and the United States as previously reported (Parkinson Progression Marker Initiative, 2011; Kang et al., 2013) and http://www.ppmi-info.org/study-design/. PPMI was approved by the institutional review board of all participating sites. The present, nested biomarker study was approved by the Institutional Review Board of Brigham and Women's Hospital. Written informed consent was obtained from all participants before inclusion in the study. The clinical aspects of the PPMI study, including subject selection, recruitment, site selection, study assessments have been described (Parkinson Progression Marker Initiative, 2011; Kang et al., 2013). Briefly, subject selection criteria required presence of an asymmetric resting tremor or asymmetric bradykinesia or two of the following signs (bradykinesia, resting tremor and rigidity), diagnosis within two years of enrollment, and no treatment with PD medications. Only cases with dopamine transporter (DAT) imaging deficits consistent with PD were included in the current study. Healthy subjects had no significant neurologic dysfunction, no first-degree family member with PD and Montreal Cognitive Assessment (MoCA) > 26. The use of an imaging biomarker for eligibility in a PD biomarker study enhances diagnostic certainty and allowed enrollment of participants very early in the disease course when only one single asymmetric sign is clinically detectable. **Clinical assessments of severity** were performed as reported in Ref. Kang et al., 2013.

Data used in the preparation of this article were obtained from the Parkinson's Progression Markers Initiative (PPMI) database (www.ppmi-info.org/data). For up-to-date information on the study, visit www.ppmi-info.org.

Collection and RNA processing. Briefly, from each subject venous blood was collected in two PAXgene tubes (Qiagen, Valencia, CA) at the baseline visit. Blood and PAXgene reagents were mixed by gently inverting 8-10 times and incubated upright at room temperature for 24 hours. PAXgene tubes were then stored frozen at -80° C and batch-shipped on dry ice to the Coriell Institute for Medical Research from US sites and to BioRep, Italy, from European sites. RNA was extracted at these core facilities following the PAXgene procedure. RNA quality was determined by using an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA) and the RNA Integrity Number package (Imbeaud et al., 2005). Only RNAs meeting two stringent Q/C parameters were included in the analysis. These were (1) RIN $\geq$ 7.3 by Agilent Bioanalyzer (high mRNA integrity), (2) ratio of absorbances at 260 nm/280 nm $\geq$1.8 (no protein contamination).

NanoString assay performance. Probes were designed according to the manufacturer's design principles(Geiss et al., 2008), including screening for inter- and intra-reporter and capture probe interactions, and selection for probes with optimal melting temperatures(Geiss et al., 2008). Direct counts of the target RNAs were measured in 125 ng of RNA by digital expression analysis based on NanoString technology (without reverse transcription into cDNA). Probes for the target and control RNAs were multiplexed and assayed according to the manufacturer's protocol on the nCounter Digital Analyzer. The laboratory running the assay was blinded to the diagnosis. No-template (negative) controls containing water substituted for template were run and no signal was detected. To avoid run-order bias, samples of cases or controls were randomly assigned to plates. To control for plate-to-plate variation and drift equal amounts of Human Universal Reference RNA were included at the beginning, within, and at the end of the entire experiment.

Transcript counts (excluding *SNCA*-007) measured in Reference RNA were highly correlated with R^2 > 0.999 both within one plate and in between different plates, thus excluding drift as a potential source of bias in the experiment. Furthermore, 5% of participants' samples were randomly re-sampled to verify the re-test reliability (technical precision). The average R^2 was 0.98 for these correlations.

Data processing. For data processing, firstly values for NTC, replicate samples, and plate-to-plate controls were removed. Secondly, positive control scaling was applied as recommended by the manufacturer. Thirdly, reference gene normalization was applied to control for RNA loading. The geometric mean of the six reference genes was used to calculate the normalization factor. Our previous low-density array qPCR analysis of 246 human blood samples highlighted four RNAs with low variation in abundance (*GUSB*, *MON1B*, *SRCAP*, *GLT25D1*) (see Harvard Biomarker Study methods for detail) and these were included to control for input RNA together with two additional, standard reference genes, the ribosomal protein gene *RPL13* and *UBC* (Scherzer et al., 2008; Zheng et al., 2010).

Probe	Accession	Target Region	Target Sequence	Flags
RPL13	NM_000977.2	760-860	GGGCCTGGGATGGGGCTTCACTGCTGTGACTTCCTCCTGCCAGGGGATTTGGGGCTTTCTTGAAAG ACAGTCCAAGCCCTGGATAATGCTTTACTTTCTG	HK
UBC	NM_021009.3	1875-1975	TGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACC ATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGA	HK
GUSB	NM_000181.1	1350-1450	CGGTCGTGATGTGGTCTGTGGCCAACGAGCCTGCGTCCCACCTAGAATCTGCTGGCTACTACTTGA AGATGGTGATCGCTCACACCAAATCCTTGGACCC	HK
MON1B	NM_014940.2	2880-2980	ACTGGCCATTACCCTAGTTCTGCCCTTGTTTGTGGAGTTACAGCCTCAAGGTTGTAGCATGTGTGCT GGCAATCAGGGCCGCAGTGTGTTCTGCGCCTGC	HK
SRCAP	NM_006662.2	3025-3125	ACACTAGCCACAGGCCATTTCATGAGCGTCATCAACATTTTGATGCAGCTGAGAAAAGTTTGCAAT CATCCAAATCTGTTCGACCCTCGACCGGTTACCT	HK
GLT25D1	NM_024656.2	3313-3413	CTGTGCTTTACCTCCTTGCCCTTGTGTCTCAGGTGTGGTCCCTGCCTGCTTGATGAAGTTGCTCTGT TCAAGCCTTTGGTGGGATCATGTGTTTGGGGGC	HK
SNCA-007	ENST00000506691.1	66-166	CCTTTCCACCCTCGTGAGCGGAGAACTGGGAGTGGCCATTCGACGACAGAGAGAGGCGGGGAGG AGTCGGAGTTGTGGAGAAGCAGAGGGACTCAGTGTG	Target
SNCA-3UTR-1	NM_000345.3	1467-1567	ACTGGTTCCTTAAGTGGCTGTGATTAATTATTGAAAGTGGGGTGTTGAAGACCCCAACTACTATTG TAGAGTGGTCTATTTCTCCCTTCAATCCTGTCAA	Target
SNCA-E3E4	NM_000345.3	370-470	TGTTCTCTATGTAGGCTCCAAAACCAAGGAGGGAGTGGTGCATGGTGTGGCAACAGTGGCTGAGA AGACCAAAGAGCAAGTGACAAATGTTGGAGGAGCA	Target
SNCA-E4E6	NM_007308.1	280-380	GCAGCCACTGGCTTTGTCAAAAAGGACCAGTTGGGCAAGGAAGGGTATCAAGACTACGAACCTGA AGCCTAAGAAATATCTTTGCTCCCAGTTTCTTGAG	Target
SNCA-3UTR-2	NM_000345.3	2879-2979	GCCTTTTCAGGAAGCTTGTCTCATATTCACTCCCGAGACATTCACCTGCCAAGTGGCCTGAGGATC AATCCAGTCCTAGGTTTATTTTGCAGACTTACAT	Target

D. Statistical Analysis Methods

Baseline characteristics of PD and healthy control participants were compared. Independent group t-tests were used when normality assumptions appeared to be met (a Satterthwaite approximate t-test was used when group variances were significantly heterogeneous); otherwise nonparametric tests such as the Wilcoxon Rank Sum test were employed. Unadjusted associations between categorical variables were assessed with a chi-square test of association when expected frequencies were large enough to meet assumptions; otherwise the Fisher Exact Test was employed.

We ran graphical analyses of our data, and because of positive skewing in some of our *SNCA* variables (e.g. for NanoString, *SNCA* counts) which might violate assumptions of some significance tests, we log transformed them, which produced distributions which were for our purposes essentially normal. To test for unadjusted mean differences between controls and cases on *SNCA* measures, independent group t-tests were employed.

To further assess diagnostic group differences on the *SNCA* variables, controlling for WBC, RBC, and gender to eliminate bias and/or increase precision, we ran general linear models (GLM) in which the *SNCA* variables were each the dependent variable in separate analyses. We checked residuals from the model for adherence to assumptions of homoscedascity and normality.

Quartiles for each *SNCA* variable were calculated, and logistic regression models to predict PD vs. controls were used to estimate the odds ratios for each quartile compared to the highest quartile (reference quartile). Both unadjusted and stepwise adjusted analyses were performed.

Meta-analysis

Because analysis was performed only across three data sets we decided to simply perform a fixed effect analysis using general linear models, where “study” was included as factor, crossed with other terms. The inferring of results from this fixed effect analysis must thus be done cautiously as it is unclear how representative the three studies are for the population of *all* studies, which could have been done to address

the association between *SNCA* blood transcripts and PD. *SNCA* variables were first log-transformed, and then z-scores were calculated. We used z-scores to evaluate the association of *SNCA* blood transcript with PD across the three studies representing 500 cases and 363 controls (total of 863). *SNCA* variables were standardized for each study and adjusted for study, WBC, RBC, gender and age. For HBS and PROBE qPCR platform results were included, while for PPMI digital expression counts were included. For PPMI we first conservatively included expression measures for the probe targeting most annotated *SNCA* transcripts (E3E4-*SNCA*). In a second iteration, we included the long 3'UTR-*SNCA* results for PPMI.

Logistic regression models were then used to estimate the odds ratios for PD for individuals with *SNCA* transcripts in the lowest quartile of values compared to individuals in the highest quartile across the three studies.

Longitudinal mixed fixed and random effects analysis of cognitive decline (assessed using the MMSE) adjusting for age, gender, disease duration upon enrollment, and years of education as fixed terms, and intercepts as well as slopes across time for subjects as random terms, was performed as we described in Ref. (Locascio and Atri, 2011). The 25% patients with PD with the highest *SNCA* transcript levels at enrollment ("*SNCA* high expressors") and the 25% of patients with the lowest *SNCA* transcript levels at enrollment ("*SNCA* low expressors") were compared.

C. Supplemental References

The Parkinson Study Group PRECEPT Investigators (2007). Mixed lineage kinase inhibitor CEP-1347 fails to delay disability in early Parkinson disease. Neurology 69, 1480-1490.

Parkinson Progression Marker Initiative (2011). The Parkinson Progression Marker Initiative (PPMI). Prog Neurobiol 95, 629-635.

Alves, G., Wentzel-Larsen, T., Aarsland, D., and Larsen, J.P. (2005). Progression of motor impairment and disability in Parkinson disease: a population-based study. Neurology 65, 1436-1441.

Applied Biosystems (2001). User Bulletin #2 ABI Prism 7700 Sequence Detection System, 1-35.

Auer, H., Lyianarachchi, S., Newsom, D., Klisovic, M.I., Marcucci, G., and Kornacker, K. (2003). Chipping away at the chip bias: RNA degradation in microarray analysis. Nat Genet 35, 292-293.

Cockrell, J.R., and Folstein, M.F. (1988). Mini-Mental State Examination (MMSE). Psychopharmacol Bull 24, 689-692.

Counihan, T.J., and Penney, J.B., Jr. (1998). Regional dopamine transporter gene expression in the substantia nigra from control and Parkinson's disease brains. J Neurol Neurosurg Psychiatry 65, 164-169.

Crum, R.M., Anthony, J.C., Bassett, S.S., and Folstein, M.F. (1993). Population-based norms for the Mini-Mental State Examination by age and educational level. Jama 269, 2386-2391.

Ding, H., Sarokhan, A.K., Roderick, S.S., Bakshi, R., Maher, N.E., Ashourian, P., Kan, C.G., Chang, S., Santarlasci, A., Swords, K.E.*, et al.* (2011). Association of SNCA with Parkinson: replication in the Harvard NeuroDiscovery Center Biomarker Study. Mov Disord 26, 2283-2286.

Fahn, S., Marsden CD, Calne DB. (1987). Recent Developments in Parkinson's disease (Florham Park, NJ: Macmillan Health Care Information).

Geiss, G.K., Bumgarner, R.E., Birditt, B., Dahl, T., Dowidar, N., Dunaway, D.L., Fell, H.P., Ferree, S., George, R.D., Grogan, T.*, et al.* (2008). Direct multiplexed measurement of gene expression with color-coded probe pairs. Nat Biotechnol 26, 317-325.

Hennecke, G., and Scherzer, C.R. (2008). RNA biomarkers of Parkinson's disease: developing tools for novel therapies. Biomarkers in Medicine 2, 41-53.

Hoehn, M.M., and Yahr, M.D. (1967). Parkinsonism: onset, progression and mortality. Neurology 17, 427-442.

Hughes, A.J., Daniel, S.E., Ben-Shlomo, Y., and Lees, A.J. (2002). The accuracy of diagnosis of parkinsonian syndromes in a specialist movement disorder service. Brain 125, 861-870.

Imbeaud, S., Graudens, E., Boulanger, V., Barlet, X., Zaborski, P., Eveno, E., Mueller, O., Schroeder, A., and Auffray, C. (2005). Towards standardization of RNA quality assessment using user-independent classifiers of microcapillary electrophoresis traces. Nucleic Acids Res 33, e56.

Jankovic, J., and Kapadia, A.S. (2001). Functional decline in Parkinson disease. Arch Neurol 58, 1611-1615.

Kang, J.H., Irwin, D.J., Chen-Plotkin, A.S., Siderowf, A., Caspell, C., Coffey, C.S., Waligorska, T., Taylor, P., Pan, S., Frasier, M.*, et al.* (2013). Association of Cerebrospinal Fluid beta-Amyloid 1-42, T-tau, P-tau181, and alpha-Synuclein Levels With Clinical Features of Drug-Naive Patients With Early Parkinson Disease. JAMA Neurol.

Locascio, J.J., and Atri, A. (2011). An overview of longitudinal data analysis methods for neurological research. Dement Geriatr Cogn Dis Extra 1, 330-357.

Louis, E.D., Tang, M.X., Cote, L., Alfaro, B., Mejia, H., and Marder, K. (1999). Progression of parkinsonian signs in Parkinson disease. Arch Neurol 56, 334-337.

Ramaker, C., Marinus, J., Stiggelbout, A.M., and Van Hilten, B.J. (2002). Systematic evaluation of rating scales for impairment and disability in Parkinson's disease. Mov Disord 17, 867-876.

Ransohoff, D.F. (2004). Rules of evidence for cancer molecular-marker discovery and validation. Nat Rev Cancer 4, 309-314.

Ransohoff, D.F. (2005). Bias as a threat to the validity of cancer molecular-marker research. Nat Rev Cancer 5, 142-149.

Ravina, B., Tanner, C., Dieuliis, D., Eberly, S., Flagg, E., Galpern, W.R., Fahn, S., Goetz, C.G., Grate, S., Kurlan, R.*, et al.* (2009). A longitudinal program for biomarker development in Parkinson's disease: a feasibility study. Mov Disord 24, 2081-2090.

Scherzer, C.R., Grass, J.A., Liao, Z., Pepivani, I., Zheng, B., Eklund, A.C., Ney, P.A., Ng, J., McGoldrick, M., Mollenhauer, B.*, et al.* (2008). GATA transcription factors directly regulate the Parkinson's disease-linked gene alpha-synuclein. Proc Natl Acad Sci U S A 105, 10907-10912.

Siderowf, A., McDermott, M., Kieburtz, K., Blindauer, K., Plumb, S., and Shoulson, I. (2002). Test-retest reliability of the unified Parkinson's disease rating scale in patients with early Parkinson's disease: results from a multicenter clinical trial. Mov Disord 17, 758-763.

Uhl, G.R., Walther, D., Mash, D., Faucheux, B., and Javoy-Agid, F. (1994). Dopamine transporter messenger RNA in Parkinson's disease and control substantia nigra neurons. Ann Neurol 35, 494-498.

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., and Speleman, F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol 3, RESEARCH0034.

Zheng, B., Liao, Z., Locascio, J.J., Lesniak, K.A., Roderick, S.S., Watt, M.L., Eklund, A.C., Zhang-James, Y., Kim, P.D., Hauser, M.A.*, et al.* (2010). PGC-1alpha, a potential therapeutic target for early intervention in Parkinson's disease. Sci Transl Med 2, 52ra73.