npg

ORIGINAL ARTICLE

YY1 binds to *α-synuclein* 3′-flanking region SNP and stimulates antisense noncoding RNA expression

Ikuko Mizuta[1,2], Kazuaki Takafuji[3], Yuko Ando[1], Wataru Satake[1], Motoi Kanagawa[1], Kazuhiro Kobayashi[1], Shushi Nagamori[3], Takayuki Shinohara[1], Chiyomi Ito[1], Mitsutoshi Yamamoto[4], Nobutaka Hattori[5], Miho Murata[6], Yoshikatsu Kanai[3], Shigeo Murayama[7], Masanori Nakagawa[8] and Tatsushi Toda[1]

α-synuclein (*SNCA*) is an established susceptibility gene for Parkinson's disease (PD), one of the most common human neurodegenerative disorders. Increased SNCA is considered to lead to PD and dementia with Lewy bodies. Four single-nucleotide polymorphisms (SNPs) in *SNCA* 3′ region were prominently associated with PD among different ethnic groups. To examine how these SNPs influence disease susceptibility, we analyzed their potential effects on *SNCA* gene expression. We found that rs356219 showed allele-specific features. Gel shift assay using nuclear extracts from SH-SY5Y cells showed binding of one or more proteins to the protective allele, rs356219-A. We purified the rs356219-A–protein complex with DNA affinity beads and identified a bound protein using mass spectrometry. This protein, YY1 (Yin Yang 1), is an ubiquitous transcription factor with multiple functions. We next investigated *SNCA* expression change in SH-SY5Y cells by YY1 transfection. We also analyzed the expression of antisense noncoding RNA (ncRNA) *RP11-115D19.1* in *SNCA* 3′-flanking region, because rs356219 is located in intron of *RP11-115D19.1*. Little change was observed in *SNCA* expression levels; however, *RP11-115D19.1* expression was prominently stimulated by YY1. In autopsied cortices, positive correlation was observed among *RP11-115D19.1*, *SNCA* and *YY1* expression levels, suggesting their functional interactions *in vivo*. Knockdown of *RP11-115D19.1* increased *SNCA* expression significantly in SH-SY5Y cells, suggesting its repressive effect on *SNCA* expression. Our findings of the protective allele-specific YY1 and antisense ncRNA raised a novel possible mechanism to regulate *SNCA* expression.
Journal of Human Genetics (2013) **58,** 711–719; doi:10.1038/jhg.2013.90; published online 12 September 2013

Keywords: α-synuclein; association; ncRNA; Parkinson's disease susceptibility; SNP; YY1

INTRODUCTION

Parkinson's disease (PD) (OMIM 168600) is one of the most common human neurodegenerative disorders, affecting 1–2% of people aged ⩾65 years.[1] Clinical features of PD (parkinsonism) include resting tremor, bradykinesia, rigidity and postural instability. PD is characterized pathologically by the loss of dopaminergic neurons in the substantia nigra of the midbrain and by the presence of intracellular inclusions known as Lewy bodies.[2]

Linkage studies for Mendelian-inherited PD have identified autosomal dominant genes, including *α-synuclein* (*SNCA*) and *LRRK2*, as well as the autosomal recessive genes *parkin*, *PINK1*, *DJ-1*, *ATP13A2*, *PLA2G6* and *FBXO7*.[3,4] However, Mendelian-inherited PD is rare compared with the far more common sporadic PD, a complex disorder caused by multiple genetic and environmental factors.[5]

SNCA was the first-identified causal gene for Mendelian-inherited PD.[6] The missense mutation A53T was identified in the original autosomal-dominant family, followed by confirmation of SNCA protein as a major component of Lewy bodies, the pathological hallmark of PD in both Mendelian-inherited and sporadic cases.[6,7] To date, three missense mutations and multiplication of *SNCA* have been identified in familial PD. Missense mutations are thought to increase the aggregation of SNCA protein. In patients with triplication of the *SNCA* locus, a doubling of wild-type *SNCA* gene dosage by triplication has been shown to result in the doubling of mRNA and protein expression in blood and in the brain.[8,9] Duplication of *SNCA* has also been implicated in familial PD. Patients with *SNCA* duplications show much milder clinical features than do those with *SNCA* triplications, more closely resembling sporadic cases.[10,11]

SNCA is also a causal gene for Mendelian-inherited dementia with Lewy bodies (DLB) (OMIM 127750). DLB, usually sporadic, is characterized clinically by dementia and parkinsonism and

[1]Division of Neurology/Molecular Brain Science, Kobe University Graduate School of Medicine, Kobe, Japan; [2]Department of Neurology, Graduate School of Medical Science, Kyoto Prefectural University of Medicine, Kyoto, Japan; [3]Division of Bio-system Pharmacology, Department of Pharmacology, Graduate School of Medicine, Osaka University, Suita, Japan; [4]Department of Neurology, Kagawa Prefectural Central Hospital, Takamatsu, Japan; [5]Department of Neurology, Juntendo University School of Medicine, Tokyo, Japan; [6]Department of Neurology, National Center Hospital, National Center of Neurology and Psychiatry, Kodaira, Japan; [7]Department of Neuropathology, Tokyo Metropolitan Institute of Gerontology, Tokyo, Japan and [8]North Medical Center, Kyoto Prefectural University of Medicine, Kyoto, Japan
Correspondence: Professor T Toda, Division of Neurology/Molecular Brain Science, Kobe University Graduate School of Medicine, 7-5-1 Kusunoki-chou, Chuo-ku, Kobe 650-0017, Japan.
E-mail: toda@med.kobe-u.ac.jp

pathologically by widespread Lewy bodies. According to the overlapped features, PD and DLB are collectively called Lewy body diseases (LBD).

A popular hypothesis is that SNCA aggregation has a crucial role in neuronal loss and Lewy body formation and that increased SNCA leads to LBD. In this scenario, PD-associated *SNCA* polymorphisms might influence *SNCA* expression levels in sporadic PD. We and the others have previously reported that single-nucleotide polymorphisms (SNPs) in the *SNCA* 3′region were prominently associated with sporadic PD in both the Japanese and European populations.[12,13] Moreover, our recent genome-wide association study (GWAS) in the Japanese population,[14] as well as a GWAS analysis in individuals of European ancestry,[15] confirmed a strong association with the *SNCA* 3′region. In this report, we have examined how *SNCA* SNPs might influence LBD susceptibility.

MATERIALS AND METHODS

Genotyping
As described previously,[13] we recruited 882 unrelated sporadic PD patients (age: 64.9 ± 9.8 years; male/female ratio: 0.79) and 938 unrelated controls (age: 45.3 ± 16.3 years; male/female ratio: 1.10). The diagnosis of idiopathic PD was based on the presence of $\geqslant 2$ of the cardinal features of PD (tremor, rigidity, bradykinesia and postural instability), according to the criteria for sporadic PD.[16] Patients were evaluated by certified neurologists specializing in PD. The average age of onset was 57.4 ± 10.9 years. Forty-two patients showed early onset of PD (<40 years of age), and 51 patients had a positive family history of PD. Patients who carried *parkin* mutations were excluded. All patients and controls were of Japanese ancestry. Informed consent was obtained from each individual, and approval for the study was obtained from the University Ethical Committees. Genomic DNA was extracted from the whole blood using FlexGene (Qiagen GmbH, Hilden, Germany). The rs356219, rs356220 and rs356203 SNPs were genotyped using TaqMan (Applied Biosystems, Lifetechnologies Corp., Carlsbad, CA, USA). SNPAlyze software (DYNACOM, Chiba, Japan) was used for pairwise linkage disequilibrium analysis (Lewontin's coefficient, D′, and standardized coefficient, *r*).

Luciferase assay
DNA fragments of ~ 250 bp corresponding to the regions containing four *SNCA* SNPs (rs356219, rs356220, rs356165 and rs356203) were amplified by PCR using heterozygous genomic DNA as template and then cloned into the SalI site of the pGL3-promoter vector (Promega Corporation, Madison, WI, USA). The orientation and allele identity of the insert were determined by DNA sequence analysis. Primer sequences used for PCR are listed in Supplementary Table 1. The human neuroblastoma cell line SH-SY5Y was grown in Dulbecco's Modified Eagle's medium supplemented with 10% fetal bovine serum and antibiotics. We transfected cells (3×10^5 cells per well on 24-well plates) with 360 ng of each construct and 40 ng of pRL-TK vector (Promega), as an internal control for transfection efficiency, using Effectene (Qiagen). After 48 h, cells were solubilized, and luciferase activity was measured using the dual luciferase assay system (Promega).

Preparation of nuclear extract
Nuclear extracts from SH-SY5Y cells were prepared as previously described, with minor modifications.[17] In brief, cells were washed and re-suspended in buffer A (10 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES)-KOH pH7.8, 10 mM KCl, 0.1 mM EDTA pH8.0, 0.1% NP-40, 1 mM dithiothreitol (DTT) and protease inhibitor cocktail (NACALAI TESQUE, Inc., Kyoto, Japan)). Nuclei were pelleted and re-suspended in buffer C (50 mM HEPES-KOH pH7.8, 420 mM KCl, 0.1 mM EDTA pH8.0, 5 mM MgCl₂, 2% glycerol, 1 mM DTT and protease inhibitor cocktail). After vortexing every 5 min for 30 min on ice, the samples were centrifuged, and the supernatants were used as nuclear extract.

Gel shift assay
Gel shift assays were performed as previously described.[18] SH-SY5Y nuclear extract ($5-10$ μg protein) was incubated with 0.031 pmol of 33-bp oligonucleotides (Supplementary Table 2) labeled with digoxigenin-11-ddUTP in 20 μl of binding buffer (20 mM HEPES pH7.6, 1 mM EDTA, 10 mM (NH₄)₂SO₄, 1 mM DTT, 0.2% Tween 20 and 30 mM KCl) containing 1 μg of poly(dI-dC) for 20 min, using the DIG gel shift kit (Roche Diagnostics, Mannheim, Germany). For competition studies, nuclear extract was pre-incubated with 3.85 pmol of unlabeled oligonucleotide before addition of the labeled probe. Supershift assays were performed by incubating the protein–DNA complexes with 0.8 μg of rabbit immunoglobulin G or rabbit polyclonal anti-YY1 (anti-Yin Yang 1; Active Motif, Carlsbad, CA, USA) for 5 min. The protein–DNA complexes were separated on DNA retardation polyacrylamide gels (Invitrogen, Lifetechnologies Corp., Carlsbad, CA, USA) and transferred to Hybond-N+ Nylon membrane (GE Healthcare, Amersham Place, UK). All assays and gel electrophoresis procedures were performed at 4 °C. Signal detection was performed using the CSPD chemiluminescent detection system (Roche Diagnostics).

Ultraviolet (UV) cross-linking
The size of the rs356219-A binding protein was determined by UV cross-linking. We incubated 0.05 pmol of ^{32}P-labeled 33 bp rs356219-A oligonucleotide with SH-SY5Y nuclear extract and 6.16 pmol of unlabeled 33 bp competitor (rs356219-G or -A) for 15 min at 4 °C in the solution used for the gel shift assays. DNA–protein complexes were transferred to a flat-bottom 96-well plate, placed on a transilluminator and irradiated (312 nm, 15 min at 4 °C). Samples were run on 10% NuPAGE Novex Bis-Tris gels (Invitrogen), dried and autoradiographed.

Purification of the rs356219-A binding protein using affinity magnetic beads
Oligonucleotides containing rs356219-A with 5′-TCGA-3′ linkers were ligated in a head-to-tail manner and inserted into the SalI site of the pGL3-promoter vector (Promega). A fragment with 14 tandem repeats of rs356219-A was amplified from the plasmid by PCR using a biotin-labeled forward primer (5′-GGTAAAATCGATAAGGATCC-3′) and a non-labeled reverse primer (5′-TT GAAGGCTCTCAAGGGCAT-3′). The biotinylated PCR product (30 pmol) was bound to 1 mg of Dynabeads M-280 streptavidin (Dynal, Life Technologies, Carlsbad, CA, USA). Fifty microliters of SH-SY5Y nuclear extract, prepared with Buffer C containing 200 mM KCl, was precleared in 1 ml of binding buffer in the presence of 1 mg of unbound beads and 144 pmol of 33 bp non-labeled competitor rs356219-A or -G. Following magnetic separation, the supernatant was incubated with 1 mg of probe-bound Dynabeads for 30 min at 4 °C. The beads were washed five times and then eluted sequentially, twice with 40 μl elution buffer (20 mM HEPES-KOH pH7.6, 1 mM EDTA, 10 mM (NH₄)₂SO₄, 1 mM DTT, 0.2% Tween 20, protease inhibitor cocktail and 10% glycerol) containing 200 mM KCl and twice with 40 μl elution buffer containing 300 mM KCl. Proteins in each elute were analyzed by electrophoresis on 10% NuPAGE Novex Bis-Tris gel (Invitrogen). Binding activity of each elute was evaluated by gel shift assay.

In-gel digestion
Protein bands were excised from a silver-stained gel. Proteins were digested with sequencing grade-modified trypsin (Roche Diagnostics) as described elsewhere.[19] Tryptic peptides were extracted from gel pieces in 50% acetonitrile with 5% formic acid and dried using a vacuum centrifuge. Dried peptides were resuspended in 5% acetonitrile containing 0.1% trifluoroacetic acid and desalted using StageTips with a C18 Empore disk membrane (3M, Minneapolis, MN, USA) according to the published procedure.[20]

Liquid chromatography-tandem mass spectrometry (LC-MS/MS) and data analysis
LC-MS/MS analysis was performed using a Paradigm MS4 nanoHPLC system (Michrom BioResources, Inc., Auburn, CA, USA) coupled to a LTQ linear ion trap mass spectrometer (Thermo Electron Corp., Waltham, MA, USA) with a

nanoelectrospray ionization source (AMR Inc., Tokyo, Japan). Tryptic peptides were injected by a HTC-PAL autosampler (CTC Analytics, Zwingen, Switzerland) and enriched on a C18 trap column (300 µm I.D. × 5 mm length, CERI, Tokyo, Japan) at a flow rate of 6 µl min^{-1}. The samples were subsequently separated by a C18 reverse phase column (100 µm I.D. × 150 mm length, Nikkyo Technos, Tokyo, Japan) at a flow rate of 1 µl min^{-1}, with a linear gradient from 2% to 65% mobile phase B, that is; from 98% to 35% of mobile phase A. The mobile phase B consisted of 95% acetonitrile with 0.1% formic acid, whereas the mobile phase A consisted of 2% acetonitrile with 0.1% formic acid. LC-MS/MS analysis was carried out using a data-dependent triple-play mode. Automated gain control values were set at 1.5×10^4, 1.5×10^3 and 5.0×10^3 for Full-MS, Zoom-MS and MS/MS, respectively. A spray voltage of 2.4 kV was applied. The MS scan range was *m/z* 300–2000. Peptides and proteins were identified by Mascot software ver. 2.2 (Matrix Science, London, UK), screened against the most recent version of the human IPI database from EMBL-EBI (http://www.ebi.ac.uk/IPI/IPIhuman.html). Maximum tolerance was set to 1.2 Da for MS data, 0.5 Da for MS/MS data and strict trypsin specificity allowing for up to one missed cleavage. Carbamidomethylation of cysteine and oxidation of methionine were allowed as variable modifications.

Transfection of YY1

SH-SY5Y cells were transfected with pCMV6-XL5 expressing human YY1 (OriGene, Rockville, MD, USA) or blank vector by using an electroporator CUY21 *Pro-Vitro* (NEPA GENE, Chiba, Japan). Twenty-four hours after transfection, the cells were harvested. For western blotting, whole-cell lysates were separated by electrophoresis on 10% NuPAGE Novex Bis-Tris gel and blotted onto PVDF membrane (Immobilon-P, Merck Millipore, Darmstadt, Germany). The membranes were incubated with anti-YY1 (sc-1703, Santa Cruz Biotechnology, Santa Cruz, CA, USA) or anti-Actin (sc-10731, Santa Cruz Biotechnology), followed by reaction with horseradish peroxidase-conjugated secondary antibodies. Signal detection was performed using the ECL detection system (Thermo Fisher Scientific, Rockford, IL, USA).

Overexpression/knockdown of *RP11-115D19.1*

For overexpression experiment, *RP11-115D19.1-003* cDNA (469 bp) was amplified from SH-SY5Y cDNA and *RP11-115D19.1-005* cDNA (567 bp) was synthesized (Genscript, Piscataway, NJ, USA). Each fragment was cloned into pCMV6-XL5 and transfected to SH-SY5Y cells as described above. For repression experiment, siRNA targeting to *RP11-115D19.1* (5′-CATGCTTC-CAGAGAATGCATATTCT-3′) was designed from the common region between *RP11-115D19.1-003* and *-005*. The siRNA and negative control (Lo GC duplex) were purchased (Stealth RNAi, Invitrogen) and transfected as described above. Twenty-four hours after transfection, the cells were harvested.

Real-time reverse transcription-PCR (RT-PCR)

As described previously, autopsied frontal cortices were obtained from the Brain Bank for Aging Research (Tokyo Metropolitan Geriatric Hospital/Tokyo Metropolitan Institute of Gerontology) and Department of Neurology, Juntendo University School of Medicine, Tokyo, Japan. The samples contained 21 cases (age, 82.6 ± 7.1 (s.d.) years; 11 males and 10 females) with Lewy body pathology defined by the third Consensus Guideline for Dementia with Lewy Bodies, comprising PD with and without dementia and DLB and 18 control subjects (age, 81.2 ± 5.2 years; 12 males and 6 females) without parkinsonism or dementia and without neurodegenerative pathological changes. Total RNA was extracted from SH-SY5Y cells or tissue using RNeasy (Qiagen), and cDNA was prepared using High Capacity RNA-to-cDNA kit (Applied Biosystems). Real-time RT-PCR was carried out on StepOnePlus real-time PCR system (Applied Biosystems) using Fast SYBR Green Master Mix (Applied Biosystems). First-strand cDNA was amplified using primers specific for *SNCA* (forward; 5′-CAGAAGCAGCAGGAAAGACA-3′, reverse; 5′-CCACTGCTCCT CCAACATTT-3′, product size; 132 bp), *RP11-115D19.1-003* (forward; 5′-TAA AACCTGCAAATTCACATCTTC-3′, reverse; 5′-AAGTAGGTAAGTAGGGCAG TGCAT-3′ product size; 133 bp), *RP11-115D19.1-005* (forward; 5′-CCATGCTT CCAGAGAATGCA-3′, reverse; 5′-GTGCTTCCCTTTCACTGAAG-3′, product size; 144 bp), *GAPDH* (forward; 5′-CATCTTCCAGGAGCGAGATC-3′, reverse; 5′-TGCAAATGAGCCCCAGCCTT-3′, product size; 114 bp), *NF* (*neurofilament L*, forward; 5′-AAGAACACCGACGCCGTGCG-3′, reverse; 5′-TGCCATTTCACTCTTTGTGG-3′, product size; 222 bp), and *YY1* (forward; 5′-TGGCAAAGCTTTTGTTGAGA-3′, reverse; 5′-ATGTGTGCGCAAATTGA AGT-3′, product size; 130 bp). For quantification, we used a relative standard curve method and amplified cDNA corresponding to 100 ng (*RP11-115D19.1*) or 1.7 ng (*SNCA*, *GAPDH*, and *NF*) RNA per well. Standard curves were generated from amplification of diluted series of cDNA from cortices (*SNCA*, *GAPDH*, and *NF*), YY1-transfected SH-SY5Y cells (*RP11-115D19.1*) or plasmids (pCMV-YY1, pCMV-RP11-115D19.1-003, pCMV-RP11-115D19.1-005). *SNCA*, *RP11-115D19.1* and *YY1* expression levels were normalized to those of *GAPDH* (SH-SY5Y cells) or *NF* (cortices). The values were determined in triplicate or duplicate. Genotyping of rs356219 of SH-SY5Y and cortices was performed using restriction fragment length polymorphism and sequencing.

RESULTS

Identification of four *SNCA* SNPs associated with PD in both the Japanese and European populations

We previously identified six SNPs (rs3857053, rs356165, rs7684318, rs3775424, rs3796661 and rs2737029), located in the 3′-flanking, 3′-UTR and intron 4 of *SNCA*, that were prominently associated with PD.[13] Independently, analysis of a European population also reported

Table 1 PD-associated *SNCA* SNPs reported in the Japanese and German studies

SNP ID	Position (NCBI build 36)	Region	Japanese study[a] P-value	MAF (Case/control)	German study[b] P-value	MAF (Case/control)	MAF of Europeans in the dbSNP database HapMap-CEU	EGP_CEPH_PANEL	AFD_EUR_PANEL
rs356219	90637601	3′-Flanking	7.3×10^{-9}	0.34/0.43	0.00467	0.46/0.38	0.425		0.354
rs356220	90641340	3′-Flanking	8.7×10^{-10}	0.33/0.43	0.00485	0.46/0.38	0.427		
rs3857053	90645674	3′-Flanking	1.1×10^{-9}	0.33/0.43				0.071	
rs356165	90646886	3′-UTR	2.0×10^{-9}	0.33/0.43	0.00555	0.46/0.38	0.433	0.412	
rs7684318	90655003	Intron 4	5.0×10^{-10}	0.33/0.43					0.024
rs3775424	90665256	Intron 4	5.4×10^{-10}	0.33/0.43			0.033		
rs356203	90666041	Intron 4	1.4×10^{-7}	0.34/0.43	0.00548	0.46/0.38		0.45	
rs3796661	90687507	Intron 4	2.7×10^{-9}	0.33/0.42			0.033		0.021
rs2737029	90711770	Intron 4	1.7×10^{-11}	0.32/0.42	0.09994	0.46/0.42	0.45		0.396

Abbreviations: MAF, minor allele frequency; NCBI, National Center for Biotechnology Information; PD, Parkinson's disease; *SNCA*, α-*synuclein*; SNP, single-nucleotide polymorphism; UTR, untranslated region.
[a]From our previous study[13] and present study (rs356219, rs356220 and rs356203).
[b]From the previous German study[12] (replication sample set II).

714

strong association for four SNPs (rs356219, rs356220 rs356165, rs356203) in the 3′-flanking region of SNCA.[12] Of these SNPs, only rs356165 was common to the two studies. We genotyped the remaining SNPs of the European population (rs356219, rs356220 and rs356203) and confirmed their prominent association with the Japanese population (Table 1, Figure 1a). The remaining SNPs reported in the Japanese population study were rare in the European population (Table 1, Figure 1a). From these findings, we concluded that rs356219, rs356220, rs356165 and rs356203 were associated prominently with PD in both the populations. Our functional analysis focused on these four SNPs.

Allele-specific effect of rs356219 in luciferase and gel shift assays

To determine whether the four SNPs could affect SNCA expression levels, we constructed plasmids containing genomic fragments with each SNP, placed downstream from a luciferase transcriptional unit (Figure 1b). For rs356219, the clone containing the protective A allele showed 1.6-fold greater luciferase activity than did the disease-associated G allele. We observed no allele-specific differences for the remaining three SNPs, however. Next, we used gel shift assays to examine the possibility that proteins might bind to the SNPs in an allele-specific manner (Figure 1c). The protective rs356219-A allele showed an intense shift band relative to the disease allele rs356219-G. Alleles of rs356220 and rs356165 showed similar shift band intensities. For rs356203, the intensity of the C disease allele was stronger than that of the protective T allele. We focused our subsequent analyses on rs356219, which was the only SNP that showed allele-specific features in both the assays.

Identification of the rs356219-A binding protein

We performed UV-crosslinking and estimated the size of the protein bound to the rs356219-A allele at ~55 kDa (Figure 2a). Next, we purified the protein by binding to rs356219-A affinity beads in the presence of non-biotinylated rs356219-G fragments as a competitor. Control experiments used non-biotinylated rs356219-A fragments as competitor. After washing, the binding protein was eluted sequentially, twice with elution buffer containing 200 mM KCl and twice with buffer containing 300 mM KCl. Electrophoresis and gel shift assay located the ~55 kDa band and binding activity within the second (200 mM KCl) and third (300 mM KCl) elutes. Both were absorbed by the rs356219-A competitor in the control experiments (Figure 2b). Subsequent LC-MS/MS analysis showed that peptides derived from a transcription factor YY1 were identified from the ~55 kDa bands (Figure 2c). We concluded that the rs356219-A binding protein was YY1. This observation was confirmed by super-shift assays using anti-YY1 antibody (Figure 2d, Supplementary Figure 1).

Effects of YY1 transfection on SNCA and noncoding RNA (ncRNA) RP11-115D19.1 expression in SH-SY5Y cells

Luciferase assay and gel shift assay suggested that YY1 bound to rs356219 protective allele and stimulated transcription (Figures 1b and c and 2). Recent Ensemble Genome Browser showed antisense ncRNA named RP11-115D19.1 in SNCA 3′-flanking region. The SNP rs356219 locates intron of its two spliced isoforms (RP11-115D19.1-003 and -005; Figure 3a), suggesting that rs356219 might influence transcription of SNCA and/or RP11-115D19.1. If rs356219 influences SNCA promoter, increased luciferase activity in the protective allele (Figure 1b) appears incompatible with the hypothesis that increased SNCA leads to PD. To address this issue, we investigated YY1-induced expression of SNCA and RP11-115D19.1 in SH-SY5Y neuroblastoma
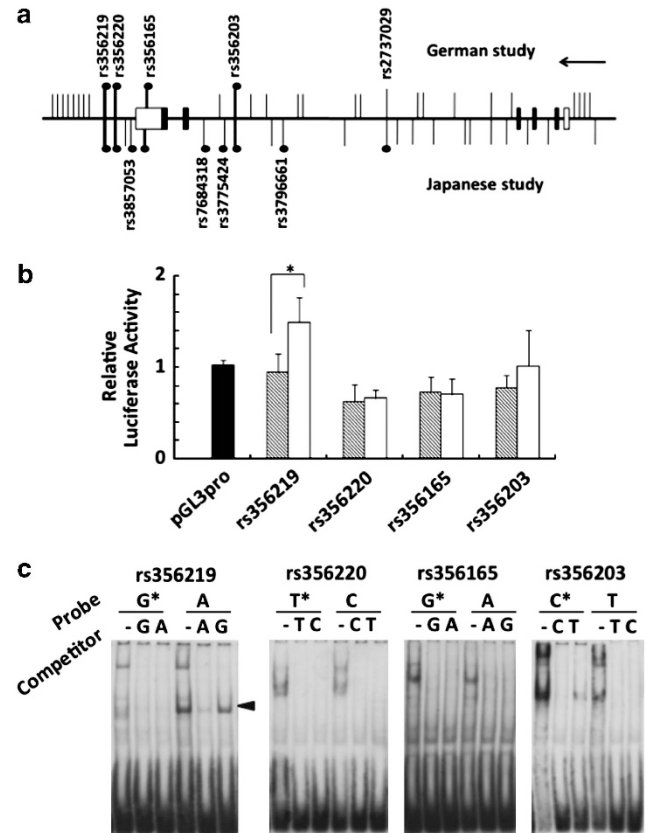

Figure 1 Transcriptional regulatory activity affected by SNCA SNPs. (a) Schematic representation of the SNCA gene with relative positions of SNPs genotyped in the Japanese and European population studies. The SNCA locus (horizontal line), coding regions (black boxes), 5′- and 3′-UTR (white boxes) and transcription orientation (arrow) are shown. The relative positions of SNPs (vertical lines) genotyped in the Japanese study (lower) and the German study (upper) are represented. Of the SNPs strongly associated with each study (solid circle), four SNPs are commonly associated with PD in both the studies (thick vertical lines). (b) Luciferase assay. Luciferase activities relative to the control pGL3-promoter construct (black bar) were compared between disease allele (shaded) and protective allele (white) at each of the four SNPs ($n = 7$). Luciferase activity of the protective rs356219-A allele was significantly increased relative to that of the G disease allele (*$P = 0.001$ by Student's t-test). (c) Gel shift assay. Binding of SH-SY5Y nuclear protein to alleles at the four SNPs are presented. Shift bands were compared between the disease allele (*) and the protective allele at each SNP. The intense shift band observed for the protective allele rs356219-A was competed by non-labeled competitor allele A but not by the disease allele G (arrowhead).

cells, whose genotype of rs356219 was heterozygous. YY1 transfection showed 23-fold increase in RP11-115D19.1-003, 24-fold increase in RP11-115D19.1-005 and twofold increase in SNCA expression levels (Figure 3b). Other two independent transfection experiments replicated the prominent stimulation of RP11-115D19.1 by YY1; however, they showed little changes of SNCA expression (data not shown). We analyzed allele-specific expression level of SNCA by using 3′-UTR SNP rs356165 as a marker and found little difference between the alleles, both in YY1 and blank transfection (Figure 3c). We could not analyze allele-specific expression of RP11-115D19.1, because no SNP was found in RP11-115D19.1 exons in SH-SY5Y. Although the role of RP11-115D19.1 on SNCA expression remains uncovered, increased luciferase activity in the

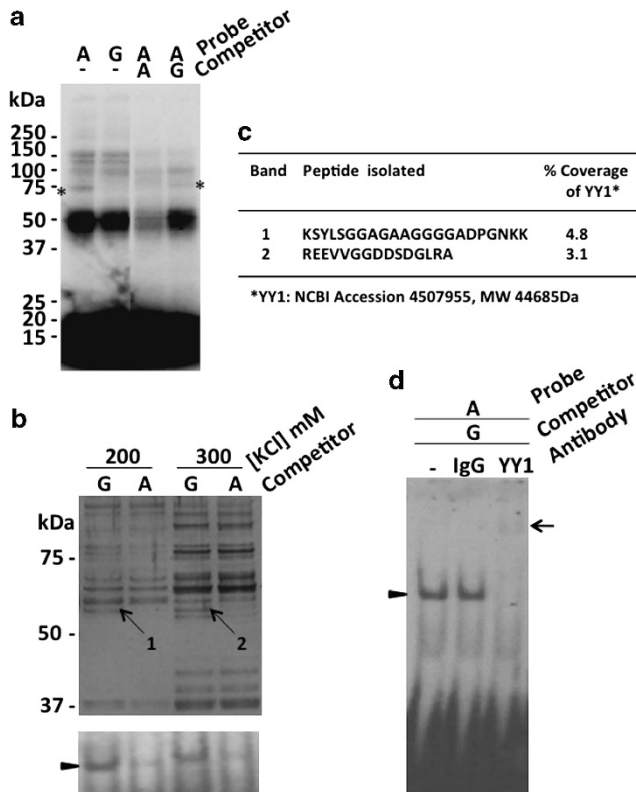**Figure 2** Purification and identification of the rs356219-A binding protein. (**a**) Estimation of the molecular size of the rs356219-A binding protein by UV cross-linking analysis. SH-SY5Y nuclear extract was incubated with the probe, rs356219-A (protective allele) or rs356219-G (disease allele) in the presence or absence of non-labeled competitor. Following UV irradiation, the cross-linked protein–probe complexes were analyzed by SDS-PAGE. The size of allele A and protein complex was estimated as ∼70 kDa (*). As the mean size of the free probe was estimated as ∼15 kDa, the protein of interest was estimated as ∼55 kDa. (**b**) Purification of the rs356219-A binding protein using DNA affinity beads. Nuclear extract from SH-SY5Y cells was incubated with affinity beads bound to tandem repeats of 33 bp region containing the A allele, in the presence of free competitor G allele (for purification) or control allele A. After the binding reaction, beads were washed and eluted twice with buffer containing 200 mM KCl and then twice with buffer containing 300 mM KCl. Aliquots of each elution were analyzed by SDS-PAGE (upper panel) and gel shift assays using rs356219-A as a probe (lower panel). The second and third elutions are shown. The rs356219-A binding band of ∼55 kDa was observed (bands 1 and 2, arrows) and correlated with the intensity of the gel shift band (arrowhead). (**c**) Summary of mass spectrometry analysis of the two ∼55 kDa bands. Peptides derived from YY1 were identified from both bands. (**d**) Supershift assay using anti-YY1 antibody. After binding of the rs356219-A probe and SH-SY5Y nuclear extract in the gel shift assay, anti-YY1 antibody was added and analyzed by electrophoresis. The shift band (arrowhead) was supershifted by anti-YY1 antibody (arrow), confirming that the shift band contains YY1.

protective allele is not contradictory to YY1-stimulated expression of *RP11-115D19.1*.

SNCA, ncRNA RP11-115D19.1 and YY1 expression in relation to rs356219

We found the possibility that rs356219 may influence expression of *RP11-115D19.1* from *in vitro* and cellular experiments (Figures 1–3). To extend the findings, we investigated *SNCA* and *RP11-115D19.1* expression in relation to rs356219 genotypes in autopsied brain
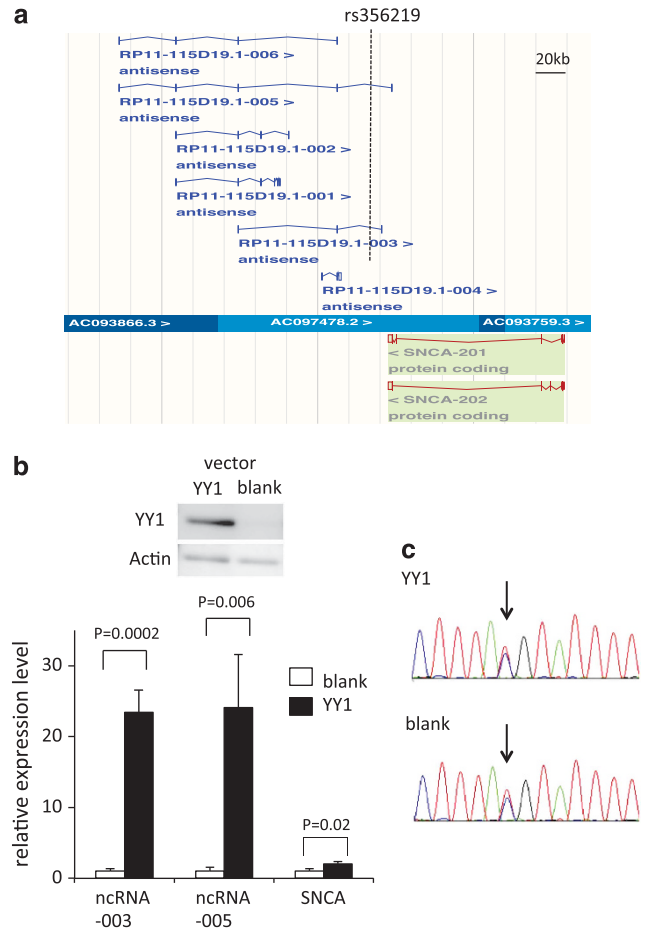

Figure 3 *SNCA* and its 3′-flanking ncRNA expression in YY1-transfected SH-SY5Y cells. (**a**) Genomic structure of *SNCA* and its 3′-flanking region (Ensemble Genome Browser). SNP rs356219 locates in intron of *RP11-115D19.1-003* and *-005* (dotted line). (**b**) Quantification of *SNCA* mRNA and two spliced isoforms of *RP11-115D19.1* ncRNA (*-003* and *-005*) in SH-SY5Y cells transfected by YY1 expression vector or blank vector. Overexpression of YY1 was confirmed by western blot analysis (upper). Three YY1 plates and three blank plates were quantified. Mean values of three YY1 plates (black bar) relative to the mean of three blank plates (white bar) are shown with s.d. (lower). YY1 transfection showed 23-fold increase in *RP11-115D19.1-003*, 24-fold increase in *RP11-115D19.1-005* and twofold increase in *SNCA* expression levels. (**c**) Semiquantitative analysis of allele-specific expression of *SNCA* 3′-UTR. Sequencing chromatograms of SH-SY5Y RT-PCR products of *SNCA* 3′-UTR containing rs356165 (arrow) are shown. Little difference between the heights of allele-C and allele-T was observed, both in the YY1 and blank transfection.

tissues. We previously reported that *SNCA* expression levels tended to be positively correlated with the number of the disease allele of rs7684318 in autopsied cortices.[13] Because rs356219 and rs7684318 are in a tight linkage disequilibrium group (Supplementary Table 3), our previous finding is to be replicated in relation to rs356219. Actually, rs356219 genotypes were completely correlated with those of rs7684318 in our autopsied samples. Moreover, we synthesized new cDNAs and quantified *SNCA* expression levels by real-time RT-PCR to compare among rs356219 genotypes. Although the difference among genotypes did not reach significance, we confirmed similar tendency with our previous report; mean of *SNCA* mRNA was lowest in AA (homozygote of protective alleles) either in cases, controls and combined (Supplementary Figure 2a). We also analyzed *RP11-*
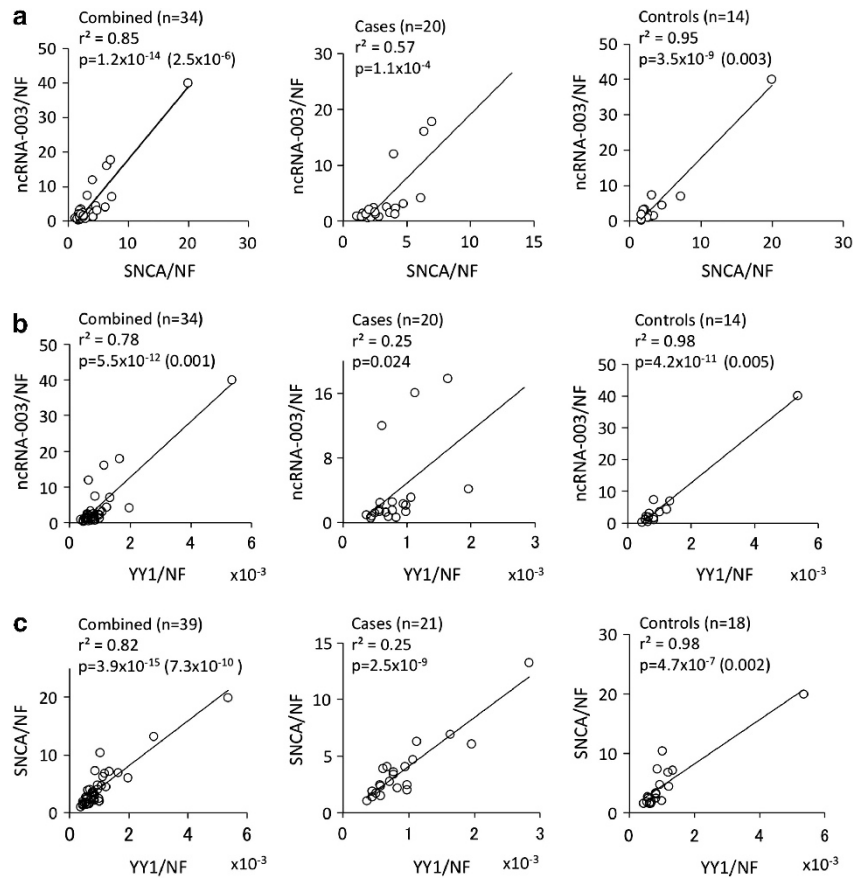
Figure 4 Correlation of SNCA, its 3′-flanking ncRNA and YY1 expression levels in autopsied frontal cortices. Scatterplots of *SNCA* versus *RP11-115D19.1-003* (a), *YY1* versus *RP11-115D19.1-003* (b) and *YY1* versus *SNCA* (c) expression levels (standardized by NF in arbitrary units). Plots are shown in all the samples (combined), cases and controls separately. Linear regression lines are indicated with r^2 and *P* values (Pearson's correlation analysis). Even after excluding the outliner in the controls, the strong correlations remain (*P*-values are in parenthesis).

115D19.1 expression in relation to rs356219. *RP11-115D19.1-005* expression levels in 38 out of the 39 cortices were too low to be quantified (data not shown); however, those of *RP11-115D19.1-003* could be quantified in 34 cortices. Compared with *SNCA*, mean values of *RP11-115D19.1-003* expression levels showed little difference among rs356219 genotypes (Supplementary Figure 2b). However, it was notable that *RP11-115D19.1-003* expression levels were strongly and positively correlated with those of *SNCA* (Figure 4a). Even after excluding the outliner in the controls, the strong correlation remained ($P = 2.5 \times 10^{-6}$ in combined and $P = 0.003$ in controls). We also examined *YY1* mRNA expression levels in cortical tissues. *YY1* expression levels were normalized by those of *NF*, because previous immunohistochemical study in an adult rodent brain reported that YY1 were strongly expressed in neurons and not detected in astrocytes.[21] *YY1* expression levels showed no significant difference among rs356219 genotype (Supplementary Figure 2c); however, positive correlation between *YY1* and *RP11-115D19.1-003* (Figure 4b) was compatible to the YY1-induced ncRNA expression shown in cellular YY1-overexpression experiment (Figure 3).

Effects of overexpression/knockdown of ncRNA *RP11-115D19.1* on *SNCA* expression in SH-SY5Y cells
Positive correlation between *SNCA* and *RP11-115D19.1* in the brain suggests regulatory roles of *RP11-115D19.1* on the *SNCA* expression. To address this issue, we investigated *SNCA* expression levels in SH-

SY5Y cells transfected by *RP11-115D19.1* expressing vectors or siRNA designed from a common region of *RP11-115D19.1-003* and -005. Overexpression of *RP11-115D19.1* did not affect *SNCA* expression levels significantly (Figure 5a). On the contrary, siRNA-mediated knockdown (∼90%) of *RP11-115D19.1* increased *SNCA* expression levels significantly (1.2-fold, $P = 0.0014$) (Figure 5b), which was replicated in three independent knockdown experiments (data not shown). These findings suggest that *RP11-115D19.1* may have repressive effect on the *SNCA* expression.

DISCUSSION
Recent GWAS have provided new information regarding PD susceptibility genes.[14,15] Our GWAS analysis in the Japanese population confirmed strong associations at *SNCA* and *LRRK2* and identified novel susceptibility loci, including *PARK16* (1q32) and *BST1* (4p15).[14] GWAS in individuals of European ancestry confirmed strong associations at *SNCA*, *MAPT* and *LRRK2*, and also confirmed *PARK16* as a novel locus.[15] Following meta-analysis of GWAS in the European ancestry populations confirmed *BST1* susceptibility.[22] Susceptibility associated with *MAPT* was not consistent between the populations, in part because of ethnic differences in allele frequencies. For example, a strong association for the H1/H2 haplotype of *MAPT* (OMIM 157140) has been reported in Caucasians,[23] but this observation has not been replicated in Asians. The frequency of the H2 haplotype is
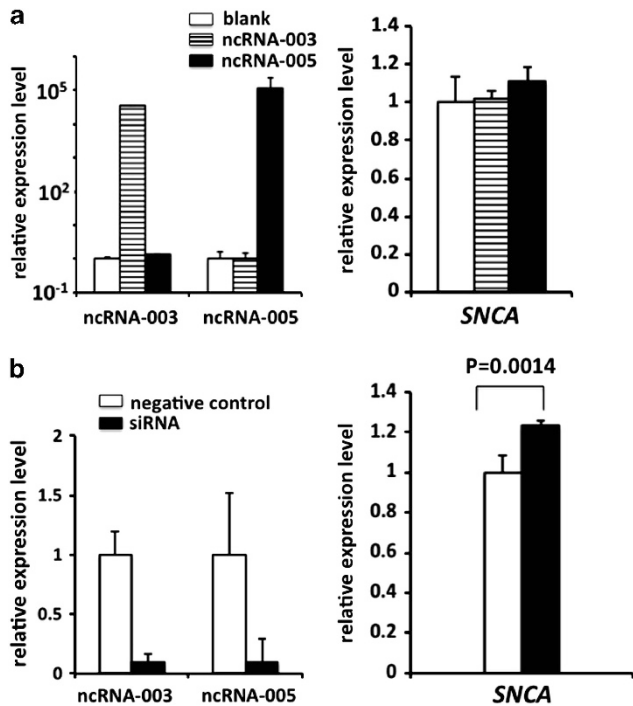
Figure 5 SNCA expression in SH-SY5Y cells after overexpression/knockdown of the ncRNA. (**a**) Quantification of SNCA mRNA in SH-SY5Y cells transfected by RP11-115D19.1 ncRNA-003 and -005 expression vectors or blank vector. Mean values of three ncRNA-003 plates (striped bar) or three ncRNA-005 (black bar) plates relative to the mean of three blank plates (white bar) are shown with s.d. Overexpression (over 10000-fold) of ncRNA-003/-005 was confirmed (left). SNCA expression did not change significantly by transfection of ncRNA-003 (P = 0.83, versus blank, t-test) or ncRNA-005 (P = 0.28, versus blank, t-test) (right). (**b**) Quantification of SNCA mRNA in SH-SY5Y cells transfected by siRNA targeted to RP11-115D19.1 or negative control RNA. Mean value of four siRNA plates (black bar) relative to the mean of four control plates (white bar) are shown with s.d. Repression of the ncRNA-003/005 (~0.1-fold) was confirmed (left). SNCA expression increased significantly (~1.2-fold, P = 0.0014, t-test) by transfection of siRNA (right).

approximately 20% in Caucasians but almost absent in Asians.[24] On the contrary, SNCA, LRRK2, BST1 and PARK16 susceptibilities are consistent in both the populations. To examine how SNPs might affect disease susceptibility, we focused on the SNCA SNPs that associate with PD across different ethnicities.

Of the four SNPs analyzed, only rs356219 showed distinct allele-specific features. The protective A allele showed greater activity in lucifersae assays than did the disease allele G. Moreover, specific binding of a nuclear protein to the protective A allele was observed using gel shift assays. When purified, this protein was identified as the transcription factor YY1, a ubiquitous zinc-finger transcription factor belonging to the Drosophila-Polycomb group protein family. The name Yin Yang 1 reflects its dual effects on transcriptional regulation, as YY1 can stimulate or repress gene expression depending on the cellular context. YY1 is associated with multiple biological functions such as proliferation, differentiation, apoptosis and tumorigenesis.[25,26] Comparison with known YY1 consensus binding motifs[27] supports the idea that the sequence of rs356219 would alter YY1 affinity in an allele-specific manner (Supplementary Figure 1). Based on the hypothesis that increased SNCA leads to LBD, it is reasonable to presume that YY1 could diminish SNCA gene expression levels by binding to the

protective allele. However, elevated luciferase activity of the construct containing the rs356219-A protective allele seems contradictory. To address this issue, we transfected YY1 expression vector to human neuroblastoma cell line SH-SY5Y of which rs356219 genotype is GA and compared allele-specific expression level using 3'-UTR SNP rs356165 as a marker. YY1 was successfully overexpressed in SH-SY5Y and little difference of SNCA expression levels between the alleles was observed (Figure 3c), suggesting little possibility of allele-specific transcription regulation by YY1–rs356219 interaction, at least in SH-SY5Y cells. We raised another possibility that rs356219 might influence transcription of a certain gene regulating SNCA expression. Recent Ensemble Genome Browser (http://www.ensembl.org/) showed novel antisense ncRNA RP11-115D19.1 in SNCA 3'-flanking region. We also investigated the expression of RP11-115D19.1, because rs356219 is located in introns of the two spliced isoforms, RP11-115D19.1-003 and -005 (Figure 3a). Surprisingly, RP11-115D19.1 expression was prominently stimulated by YY1 overexpression, in contrast to the little change in SNCA expression (Figure 3b). From these findings, it would be informative to analyze RP11-115D19.1 expression levels stimulated by YY1 in relation to rs356219 genotypes. Because no SNP was found in SH-SY5Y RP11-115D19.1 exons, we designed an experiment to transfect YY1 expression vector to human lymphoblasts G or A in rs356219 and compare SNCA mRNA expression levels between G and A. However, because of low transfection efficiency (~10% by electroporation), we could not induce YY1 in lymphoblasts (data not shown).

Our in vitro and cellular experiments showed the possibility that rs356219 might influence YY1-stimulated transcription of RP11-115D19.1. To extend this, we investigated steady state SNCA and RP11-115D19.1 expression levels in autopsied frontal cortices in relation to rs356219 genotypes. SNCA expression levels in cortices tended to be positively correlated with the number of the disease allele of rs356219 (Supplementary Figure 2a), replicating our previous result.[13] Our result was not consistent with other groups' previous reports. Fuchs et al.[28] reported that higher SNCA mRNA correlated with rs356219 disease allele in the substantia nigra of the midbrain, however, with the protective allele in the cerebellum. On the other hand, Linnertz et al.[29] reported that higher SNCA mRNA correlated with rs356219 protective allele in the temporal cortices and midbrain and unchanged among the genotypes in the frontal cortices. The discrepancies among these studies might be from the number and variation of samples.

We also analyzed steady state in vivo RP11-115D19.1 expression levels and found little difference among genotypes (Supplementary Figure 2b). However, it was of note that RP11-115D19.1 expression levels were strongly and positively correlated with those of SNCA (Figure 4a). Recent reports focused on SNP-coexpression associations.[30,31] We found a positive correlation between SNCA and RP11-115D19.1 in either of the rs356219 genotype. The significance of correlation tended to increase according to the number of the disease allele G (Supplementary Figure 3). To confirm this tendency, large number of samples must be analyzed.

We also quantified RP11-115D19-003 and SNCA mRNA levels in lymphoblasts originated from PD patients harboring GG and AA (Supplementary Figure 4). Similarly to the brain, there were no significant differences in RP11-115D19-003 and SNCA expression levels between GG and AA lymphoblasts. In contrast to the brain, however, there was no significant correlation between RP11-115D19-003 and SNCA levels in lymphoblasts. Although the numbers examined were limited, these results suggest interaction between RP11-115D19 and SNCA expressions may be brain-specific.

The coexpression suggests functional interaction between the two genes. *RP11-115D19.1* transcripts span 0.4–1.8 kb, classified by size to long noncoding RNAs (lncRNAs).[32] Although the function of majority of lncRNAs remains uncovered, some are supposed to participate in regulating transcription of coding genes. For example, it was reported that mRNA of *BACE1, β-site APP (amyloid precursor protein) cleaving enzyme 1*, was stabilized by *BACE1* antisense transcript, a lncRNA including an exon complementary to a *BACE-1* exon.[33] On the contrary, recent work reported that *brain-derived neurotrophic factor (BDNF)* was repressed by its antisense RNA transcript.[34] These reports showed that some antisense ncRNAs downregulated and others upregulated transcription of the sense genes. *RP11-115D19.1-005* transcript overlaps *SNCA* 3′-UTR partly in tail-to-tail manner (Figure 3a), suggesting putative regulatory activity on *SNCA* expression. *RP11-115D19.1-005* transcript was hardly detected in steady state brain mRNA; however, it was prominently stimulated by YY1 in SH-SY5Y cells, as well as in *-003* transcript.

To investigate how *RP11-115D19.1* may influence *SNCA* expression, we performed cellular overexpression and knockdown experiments. We found that siRNA-mediated knockdown of the ncRNA increased *SNCA* expression (∼1.2-fold) in SH-SY5Y cells (Figure 5b). The effect seems small; however, it could be sufficient to influence susceptibility to late-onset sporadic PD and DLB, caused by multiple genetic and environmental factors. The reason why overexpression of the ncRNA did not suppress *SNCA* expression remains unknown. We hypothesize that ncRNA may influence *SNCA* expression in a locus-dependent manner and that repressive effect may be saturated at endogeneous (low) expression level of ncRNA.

We performed similar analysis using human embryonic kidney-derived HEK293 cells. YY1-induced expression of ncRNA was replicated in HEK293 cells (Supplementary Figure 5). The degree of ncRNA stimulation in HEK293 cells, however, was smaller than those in SH-SY5Y cells. *SNCA* expression levels in HEK293 cells did not change significantly after YY1 overexpression, ncRNA overexpression or ncRNA knockdown. These suggest that ncRNA's repressive effect on *SNCA* expression may be specific to neuroblastoma cells.

Based on our demonstration that ncRNA has directly repressive effect on *SNCA* expression *in vitro*, positive correlation between *SNCA* and its antisense ncRNA *in vivo* may be explained as follows: One possibility is that expression of antisense ncRNA may coordinate with that of sense gene *SNCA* under locus-specific transcriptional regulation. Another possibility is that certain neurotoxic factors, including oxidative stress, may cause stimulation of *SNCA* and ncRNA expression levels simultaneously. Simultaneous stimulation of *SNCA* and *YY1* may be also possible, because of correlation of *SNCA* and *YY1* expression levels in the brain (Figure 4c) and YY1-induced expression of ncRNA in cells (Figure 3). In both the hypotheses (Supplementary Figure 6), ncRNA may be beneficial to maintain *SNCA* expression levels within normal range.

Accumulation of SNCA is thought to be neurotoxic. On the other hand, the protective effects of SNCA are also reported. Chandra *et al.*[35] reported that transgenic expression of *SNCA* rescued neurodegeneration in *CSPα* knockout mice. Musgrove *et al.*[36] reported that endogenous SNCA upregulation in response to weak oxidative stress was neuroprotective against additional acute oxidative stress in primary cultured neurons. These suggest that SNCA exerts neuroprotective effects under disease or stress. However, repeated stresses or lasting disease state may cause SNCA accumulation. Extraordinary upregulation of SNCA may be also harmful. Therefore, *SNCA* expression levels must be fine-tuned to maintain proper range *in vivo*.

To clarify the regulatory mechanism of the ncRNA, with or without neurotoxic stresses, further intensive approach should be necessary.

In conclusion, our findings of the protective-allele specific YY1 and antisense ncRNA raised a novel possible mechanism to regulate *SNCA* expression.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

1 de Rijk, M. C., Tzourio, C., Breteler, M. M., Dartigues, J. F., Amaducci, L., Lopez-Pousa, S. *et al.* Prevalence of parkinsonism and Parkinson's disease in Europe: the EUROPARKINSON Collaborative Study. European Community Concerted Action on the Epidemiology of Parkinson's disease. *J. Neurol. Neurosurg. Psychiatry* **62,** 10–15 (1997).
2 Shults, C. W. Lewy bodies. *Proc. Natl Acad. Sci. USA* **103,** 1661–1668 (2006).
3 Farrer, M. J. Genetics of Parkinson disease: paradigm shifts and future prospects. *Nat. Rev. Genet.* **7,** 306–318 (2006).
4 Lesage, S. & Brice, A. Parkinson's disease: from monogenic forms to genetic susceptibility factors. *Hum. Mol. Genet.* **18,** R48–R59 (2009).
5 Warner, T. T. & Schapira, A. H. Genetic and environmental factors in the cause of Parkinson's disease. *Ann. Neurol.* **53** (Suppl 3), S16–S23 (2003).
6 Polymeropoulos, M. H., Lavedan, C., Leroy, E., Ide, S. E., Dehejia, A., Dutra, A. *et al.* Mutation in the α-synuclein gene identified in families with Parkinson's disease. *Science* **276,** 2045–2047 (1997).
7 Spillantini, M. G., Schmidt, M. L., Lee, V. M., Trojanowski, J. Q., Jakes, R. & Goedert, M. α-Synuclein in Lewy bodies. *Nature* **388,** 839–840 (1997).
8 Singleton, A. B., Farrer, M., Johnson, J., Singleton, A., Hague, S., Kachergus, J. *et al.* α-Synuclein locus triplication causes Parkinson's disease. *Science* **302,** 841 (2003).
9 Miller, D. W., Hague, S. M., Clarimon, J., Baptista, M., Gwinn-Hardy, K., Cookson, M. R. *et al.* α-Synuclein in blood and brain from familial Parkinson disease with *SNCA* locus triplication. *Neurology* **62,** 1835–1838 (2004).
10 Chartier-Harlin, M. C., Kachergus, J., Roumier, C., Mouroux, V., Douay, X., Lincoln, S. *et al.* α-Synuclein locus duplication as a cause of familial Parkinson's disease. *Lancet* **364,** 1167–1169 (2004).
11 Ibanez, P., Bonnet, A. M., Debarges, B., Lohmann, E., Tison, F., Pollak, P. *et al.* Causal relation between α-synuclein gene duplication and familial Parkinson's disease. *Lancet* **364,** 1169–1171 (2004).
12 Mueller, J. C., Fuchs, J., Hofer, A., Zimprich, A., Lichtner, P., Illig, T. *et al.* Multiple regions of α-synuclein are associated with Parkinson's disease. *Ann. Neurol.* **57,** 535–541 (2005).
13 Mizuta, I., Satake, W., Nakabayashi, Y., Ito, C., Suzuki, S., Momose, Y. *et al.* Multiple candidate gene analysis identifies α-*synuclein* as a susceptibility gene for sporadic Parkinson's disease. *Hum. Mol. Genet.* **15,** 1151–1158 (2006).
14 Satake, W., Nakabayashi, Y., Mizuta, I., Hirota, Y., Ito, C., Kubo, M. *et al.* Genome-wide association study identifies common variants at four loci as genetic risk factors for Parkinson's disease. *Nat. Genet.* **41,** 1303–1307 (2009).
15 Simon-Sanchez, J., Schulte, C., Bras, J. M., Sharma, M., Gibbs, J. R., Berg, D. *et al.* Genome-wide association study reveals genetic risk underlying Parkinson's disease. *Nat. Genet.* **41,** 1308–1312 (2009).
16 Bower, J. H., Maraganore, D. M., McDonnell, S. K. & Rocca, W. A. Incidence and distribution of parkinsonism in Olmsted County, Minnesota, 1976-1990. *Neurology* **52,** 1214–1220 (1999).
17 Adachi, O., Kawai, T., Takeda, K., Matsumoto, M., Tsutsui, H., Sakagami, M. *et al.* Targeted disruption of the *MyD88* gene results in loss of IL-1- and IL-18-mediated function. *Immunity* **9,** 143–150 (1998).

18 Ozaki, K., Ohnishi, Y., Iida, A., Sekine, A., Yamada, R., Tsunoda, T. *et al.* Functional SNPs in the lymphotoxin-α gene that are associated with susceptibility to myocardial infarction. *Nat. Genet.* **32,** 650–654 (2002).

19 Kuruma, H., Egawa, S., Oh-Ishi, M., Kodera, Y., Satoh, M., Chen, W. *et al.* High molecular mass proteome of androgen-independent prostate cancer. *Proteomics* **5,** 1097–1112 (2005).

20 Rappsilber, J., Ishihama, Y. & Mann, M. Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Anal. Chem.* **75,** 663–670 (2003).

21 Rylski, M., Amborska, R., Zybura, K., Konopacki, F. A., Wilczynski, G. M. & Kaczmarek, L. Yin Yang 1 expression in the adult rodent brain. *Neurochem. Res.* **33,** 2556–2564 (2008).

22 International Parkinson Disease Genomics Consortium, Nalls, M. A., Plagnol, V., Hernandez, D. G., Sharma, M., Sheerin, U. M. *et al.* Imputation of sequence variants for identification of genetic risks for Parkinson's disease: a meta-analysis of genome-wide association studies. *Lancet* **377,** 641–649 (2011).

23 Zabetian, C. P., Hutter, C. M., Factor, S. A., Nutt, J. G., Higgins, D. S., Griffith, A. *et al.* Association analysis of *MAPT* H1 haplotype and subhaplotypes in Parkinson's disease. *Ann. Neurol.* **62,** 137–144 (2007).

24 Stefansson, H., Helgason, A., Thorleifsson, G., Steinthorsdottir, V., Masson, G., Barnard, J. *et al.* A common inversion under selection in Europeans. *Nat. Genet.* **37,** 129–137 (2005).

25 Shi, Y., Lee, J. S. & Galvin, K. M. Everything you have ever wanted to know about Yin Yang 1...... *Biochim. Biophys. Acta* **1332,** F49–F66 (1997).

26 Gordon, S., Akopyan, G., Garban, H. & Bonavida, B. Transcription factor YY1: structure, function, and therapeutic implications in cancer biology. *Oncogene* **25,** 1125–1142 (2006).

27 Shrivastava, A. & Calame, K. An analysis of genes regulated by the multi-functional transcriptional regulator Yin Yang-1. *Nucleic Acids Res.* **22,** 5151–5155 (1994).

28 Fuchs, J., Tichopad, A., Golub, Y., Munz, M., Schweitzer, K. J., Wolf, B. *et al.* Genetic variability in the *SNCA* gene influences α-synuclein levels in the blood and brain. *FASEB J.* **22,** 1327–1334 (2008).

29 Linnertz, C., Saucier, L., Ge, D., Cronin, K. D., Burke, J. R., Browndyke, J. N. *et al.* Genetic regulation of α-synuclein mRNA expression in various human brain tissues. *PLoS One* **4,** e7480 (2009).

30 Kayano, M., Takigawa, I., Shiga, M., Tsuda, K. & Mamitsuka, H. Efficiently finding genome-wide three-way gene interactions from transcript- and genotype-data. *Bioinformatics* **25,** 2735–2743 (2009).

31 Wang, Y., Joseph, S. J., Liu, X., Kelley, M. & Rekaya, R. SNPxGE^2: a database for human SNP-coexpression associations. *Bioinformatics* **28,** 403–410 (2012).

32 Costa, F. F. Non-coding RNAs: meet thy masters. *Bioessays* **32,** 599–608 (2010).

33 Faghihi, M. A., Modarresi, F., Khalil, A. M., Wood, D. E., Sahagan, B. G., Morgan, T. E. *et al.* Expression of a noncoding RNA is elevated in Alzheimer's disease and drives rapid feed-forward regulation of β-secretase. *Nat. Med.* **14,** 723–730 (2008).

34 Modarresi, F., Faghihi, M. A., Lopez-Toledano, M. A., Fatemi, R. P., Magistri, M., Brothers, S. P. *et al.* Inhibition of natural antisense transcripts in vivo results in gene-specific transcriptional upregulation. *Nat. Biotechnol.* **30,** 453–459 (2012).

35 Chandra, S., Gallardo, G., Fernández-Chacón, R., Schlüter, O. M. & Südhof, T. C. α-synuclein cooperates with CSPα in preventing neurodegeneration. *Cell* **123,** 383–396 (2005).

36 Musgrove, R. E. J., King, A. E. & Dickson, T. C. Neuroprotective upregulation of endogeneous alpha-synuclein precedes ubiquitination in cultured dopaminergic neurons. *Neurotox. Res.* **19,** 592–602 (2011).

Supplementary Information accompanies the paper on Journal of Human Genetics website (http://www.nature.com/jhg)