SNCA and MAPT genes: Independent and joint effects in Parkinson disease in the Italian population

Luca Trotta [a,b,1], Ilaria Guella [a,1], Giulia Soldà [a], Francesca Sironi [b], Silvana Tesei [c], Margherita Canesi [c], Gianni Pezzoli [c], Stefano Goldwurm [c], Stefano Duga [a], Rosanna Asselta [a,*]

[a] Dipartimento di Biologia e Genetica per le Scienze Mediche, Università degli Studi di Milano, Milan, Italy
[b] Medical Genetics Laboratory, Foundation IRCCS Cà Granda Ospedale Maggiore Policlinico, Milan, Italy
[c] Parkinson Institute, Istituti Clinici di Perfezionamento, Milan, Italy

ARTICLE INFO

ABSTRACT

Background: Significant efforts have been focused on investigating the contribution of common variants to Parkinson disease (PD) risk. Several independent GWAS and metanalysis studies have shown a genome-wide significant association of single nucleotide polymorphisms (SNPs) in the α-synuclein (SNCA) and microtubule-associated protein tau (MAPT) regions. Here we investigated the role of SNCA and MAPT as PD susceptibility genes in a large Italian population of 904 patients and 891 controls. An evaluation of gene−gene and gene-environment interactions in association with PD was also attempted.
Methods: The SNCA Rep1 microsatellite was genotyped by a fluorescent PCR assay, whereas the SNPlex genotyping system was used to genotype 12 additional markers across the SNCA gene, and 2 SNPs tagging the risk MAPT H1 haplotype.
Results: Single-marker analysis demonstrated nominal evidence of association for: i) the 261-bp-long allele of Rep1; ii) 7 SNPs in the SNCA region (top SNP: rs356186, $P = 3.08 \times 10^{-04}$, intron 4); iii) both SNPs identifying the MAPT H1 haplotype ($P = 4.63 \times 10^{-04}$ and $P = 4.23 \times 10^{-04}$ for rs1800547 and rs9468, respectively). Moreover, we found a highly significant protective haplotype spanning ∼83 kb from intron 4 to the 3′ end of SNCA ($P = 1.29 \times 10^{-05}$).
Conclusions: Our findings strongly confirm SNCA and MAPT as major PD susceptibility genes for idiopathic PD in the Italian population. Interaction analyses did not evidence either epistatic effects between the two loci or gene-environment interactions.

© 2011 Elsevier Ltd.

1. Introduction

Parkinson disease (PD; OMIM#168600) is the second most common neurodegenerative disorder after Alzheimer disease, with a lifetime risk of developing the disease of 1−5% [1]. Up to now, five major causative genes have been identified in Mendelian forms (SNCA, PRKN, DJ1, PINK1, LRRK2), although mutations in these genes collectively explain less than 10% of familial PD cases [2]. In the vast majority of sporadic cases, the etiology is probably more complex, resulting from a combination of multiple environmental exposures (e.g.: pesticides, coffee consumption, cigarette smoking) and genetic risk factors acting on a background of aging. Indeed, a large amount of research has focused on identifying genetic variability that confers PD susceptibility [2].

Among putative PD susceptibility genes, both the α-synuclein gene (SNCA) and the microtubule-associated protein tau gene (MAPT) have been repeatedly investigated [3].

SNCA was the first causal gene identified in familial PD [3], and encodes α-synuclein, a major component of Lewy bodies (LB), a pathologic hallmark of PD [2]. The overproduction of α-synuclein is the direct cause of LB accumulation in neural cells of familial PD cases, whereas a more subtle overproduction of otherwise normal α-synuclein may contribute to PD pathogenesis in sporadic patients [2].

Rep1 (GenBank D4S3481) is a complex polymorphic microsatellite repeat located ∼10 kb upstream of the translation start site of SNCA. While several studies suggested that specific SNCA Rep1 alleles are associated with late-onset idiopathic PD, others have observed no association or reported an inverse association between the risk allele and PD ([4,5] and references therein). After many small studies, Maraganore and colleagues finally confirmed the association between the Rep1 polymorphism and PD in a large

* Corresponding author. Department of Biology and Genetics for Medical Sciences, University of Milan, Via Viotti, 3/5 - 20133 Milano, Italy. Tel.: +39 02 50315853; fax: +39 02 50315864.
E-mail address: rosanna.asselta@unimi.it (R. Asselta).
[1] These authors equally contributed to this work.

meta-analysis using more than 5000 samples from 11 different sites [5]. Moreover, functional analyses on the two most common Rep1 alleles (261 bp and 259 bp) suggested that the 261 bp-long risk allele is associated with an up-regulation of *SNCA* expression, whereas the 259 bp-long protective variant shows reduced gene expression [6–8].

In addition to variability in the promoter, genetic variation in the 3′ untranslated region of *SNCA* has been proposed to modulate predisposition to PD [9]. Association of haplotypes within linkage disequilibrium (LD) blocks located in the 3′ end of *SNCA* with sporadic PD has been reported [10,11]. Interestingly, a recent paper by Mata and colleagues showed a possible association between a single nucleotide polymorphism (SNP), rs356219, in the 3′ region of the gene and *SNCA* plasma levels [12].

MAPT encodes for the microtubule-associated protein tau, which regulates microtubule dynamics and assembles microtubules into parallel arrays within axons. The H1 and H2 haplotypes represent two distinct clades of subhaplotypes ensued from an inversion of ~900 kb on chromosome 17q21, spanning the entire *MAPT* coding region [13], and are tagged, among others, by genotypes at two SNPs: rs9468 and rs1800547 [14]. The H1 haplotype has been significantly associated with an increased risk for a number of neurodegenerative diseases [13]. Several studies proposed the *MAPT* H1 haplotype as susceptibility factor also for PD [15–17].

Associations of *SNCA* and *MAPT* with PD have been replicated not only in candidate gene studies, but also in several independent genome-wide association studies (GWAS), again suggesting the two genes as strong susceptibility factors for idiopathic PD in different populations [18–22]. These data were further confirmed by a very recent meta-analysis of PD GWAS [23]. However, no extensive investigations, including *SNCA-MAPT* interaction or gene-environment analyses, have been performed on the Italian population.

The present study was designed to investigate the role of *SNCA* and *MAPT* as susceptibility genes for PD in a large and well-characterized Italian population, and to evaluate possible interactions between the two loci.

2. Methods

2.1. Subjects

We studied 904 subjects with a diagnosis of PD and 891 controls who had contributed to the DNA Bank of the Parkinson Institute, Istituti Clinici di Perfezionamento, Milan, Italy ("Human genetic bank of patients affected by Parkinson disease and parkinsonisms"; http://www.parkinson.it/dnabank.html).

All patients were examined by movement disorder neurologists. The clinical diagnosis of PD was established according to the UK Parkinson Disease Society Brain Bank criteria ([24], and references therein) and required the presence of bradykinesia and at least one symptom among the followings: resting tremor, rigidity or postural instability, a positive response to dopaminergic therapy, and absence of atypical features or other causes of parkinsonism. The age at which the patient first noticed a PD-related symptom was considered the age of onset of the disease. Patients were classified as "familial" if at least one among their 1st or 2nd degree relatives had a formal diagnosis of PD. The remaining patients were classified as "sporadic".

The *LRRK2* G2019S missense substitution was screened in all subjects [25], whereas the *PRKN* gene was analyzed in all 62 PD patients with onset before the age of 40 [26]. *LRRK2* and *PRKN* mutation carriers were excluded from this study.

Control subjects, unrelated to the patients and with negative family history for movement disorders, were recruited among spouses and care-givers. With the exception of 4 patients, originating from Argentina, Albania, and Greece, all the affected subjects and controls were of Italian origin.

At the time of blood sampling, a standardized questionnaire, documenting medical diagnoses, lifestyle, and medications was filled in for all cases and controls. Smoking status was defined as current, former, or never smoker on the basis of self-reports (current and former smokers were grouped together in the single category of smokers in the statistical analyses). Coffee consumption was quantified on the basis of self-reports, with no consumption being defined as the intake of up to 2 coffee/day, and consumption as more than 2 coffee/day.

The study was approved by the local Ethics Committee. All participants signed an informed consent.

2.2. Genotyping assays

Genomic DNA extraction was performed from peripheral blood using a semi-automated extractor (QuickGene DNA Whole Blood Kit; FUJIFILM Europe GmbH Life Science, Düsseldorf, Germany). DNA samples were quantified using a Bio-Photometer (Eppendorf AG, Hamburg, Germany), standardized for concentration, and arrayed into 96-deep-well plates.

The *SNCA* Rep1 region was PCR amplified from genomic DNA using the primer couple: 5′-GACTGGCCCAAGATTAACCA-3′ (fluorescently labeled with 6-FAM) and 5′-CCTGGCATATTTGATTGCAA-3′ (Sigma–Aldrich Co., St Louis, MO, USA). PCR products were resolved by capillary electrophoresis on an ABI-3130XL Genetic Analyzer (Applied Biosystem, Foster City, CA, USA). Allelic sizes were assessed using the GeneMapper v4.0 software (Applied Biosytems). On the basis of the length of PCR products, Rep1 alleles were coded as follows: allele 1 = 259 bp, allele 2 = 261 bp, allele 3 = 263 bp, according to one of the previously described nomenclatures [4].

Fourteen biallelic markers were selected within the *SNCA* (12 SNPs) and *MAPT* (2 SNPs) gene regions. The 12 *SNCA* markers were chosen either on the basis of previous association data in Caucasian populations [10,27], or from dbSNP (http://www.ncbi.nlm.nih.gov/projects/SNP/); in all cases, SNPs showed a minor allele frequency of at least ~20% in the general population, and were selected in order to both assure an adequate coverage of the *SNCA* locus and meet the strict criteria for probe design of the genotyping assay. The 2 *MAPT* SNPs were chosen because both tagged the H1/H2 haplotypes.

Genotypes were determined using the Applied Biosystems SNPlex™ assay [28]. Detection was performed on an ABI-3130XL Genetic Analyzer; data interpretation was performed with the GeneMapper software. Reproducibility was assessed by comparing replicated samples.

2.3. Statistical analyses

2.3.1. Phenotypic data

The dependence of disease risk (the response variable) on potential explanatory factors [age, sex, smoking status, and coffee consumption] was analyzed with a conditional logistic regression model and statistical significance was tested using the likelihood ratio test and Wald's test [29]. In the regression, the response variable was set to 1 in cases and to 0 in controls. The effect of each explanatory factor was expressed as the odds ratio (OR) and 95% confidence interval (CI) computed from the corresponding estimated regression coefficient in the model. Unadjusted ORs were obtained by using a model that included only the factor of interest; adjusted ORs were obtained by using a model that included the factor of interest plus all of the remaining factors. All procedures were performed using the R program, version 2.10.1 (http://www.r-project.org/).

2.3.2. Genetic data

Allele and genotype frequencies of the Rep1 marker were calculated by using the Genepop program, option 5 (http://genepop.curtin.edu.au/). The same software, option 1, was used to test the deviation from the Hardy–Weinberg equilibrium (HWE). Allelic and genotypic frequencies of Rep1 were compared by means of a χ^2 statistic, generated through the Monte Carlo approach using the CLUMP program [30]. CLUMP uses the Monte Carlo procedure by measuring the differences in frequencies between cases and controls, and then produces multiple simulated datasets to calculate how many times the observed differences might be generated by chance if the frequencies were the same. Significance was assessed by carrying out 10,000 simulations. The use of the Monte Carlo simulation avoids the need for the Bonferroni (or others) correction.

Analysis of deviation from HWE, genotyping success rate, allelic and genotypic frequencies, and association with PD of SNPs in *SNCA* and *MAPT* were performed using the PLINK software (http://pngu.mgh.harvard.edu/~purcell/plink/). For each SNP, standard case–control analyses on allele and genotype frequency data were performed with χ^2 statistics (Fisher exact test). In the text, all *P*-values for individual SNPs, OR, and 95% CI are presented as non-corrected for multiple testing and refer to the minor allele. Associations were also retested adjusting the logistic regression models for sex, smoking, and coffee consumption, previously tested for association with the disease, using the R software.

For *SNCA*, the LD structure was determined from our SNP data using the Haploview v4.0 program (http://www.broadinstitute.org/haploview). LD blocks were defined using the "spine of LD" setting. Haplotype-based association analysis was performed considering only those SNPs mapping within the same LD block, using the PLINK software.

Power estimates indicated that, if each analyzed polymorphism (disease allele frequency of 10%) were to directly confer a 1.5-fold increase in the relative risk of PD, the case and control groups used in this research would be of sufficient size to have 99% power to detect a significant association at the 0.05 level.

2.3.3. Secondary analyses

All secondary analyses were accomplished using the PLINK software. In particular, a pair-wise epistasis analysis of the 15 genotyped polymorphisms was performed using the "–epistasis" option, which provides a logistic regression test for

interaction that assumes an allelic model for both the main effects and the interactions.

A case-only analysis was also performed. For both analyses, only pairs of SNPs on different chromosomes (i.e. in linkage equilibrium) were considered.

As for gene-environment interaction, we focused on smoking and coffee consumption, because their individual inverse associations with PD are well established [31]. A Breslow-Day test of homogeneity of OR for the identification of possible interactions between *SNCA/MAPT* with smoking and coffee consumption was performed. In this case, clusters were stratified according to the presence/absence of the habit for each SNP, and their disease/gene ORs compared.

3. Results

3.1. PD patients' characteristics

Demographics of the subjects included in this study are showed in Table 1.

Among the 904 PD patients, 543 (60.1%) were male, the mean age at onset was 56.1 years (range 14—82), and the mean disease duration was 10.6 years (range 4—42). Two hundred and eighty nine patients were classified as "familial", the remaining subjects as "sporadic".

Coffee consumption and cigarette smoking habits were both more represented in controls than in cases, and showed a strong protective effect [OR = 0.54 (95% CI = 0.42—0.69) and OR = 0.68 (95% CI = 0.53—0.88), respectively]; also sex difference was very marked, but this is ascribable to a recruitment bias, as most controls are partners of PD patients.

3.2. Single-marker association analysis with SNCA and MAPT

An association analysis on the entire PD cohort was performed by genotyping 15 polymorphic markers: the Rep1 microsatellite located in the 5′ regulatory region of *SNCA*, 12 SNPs covering the whole *SNCA* gene, and 2 SNPs, located ∼50 kb apart from each other, both tagging the *MAPT* H1/H2 haplotypes.

The mean genotyping success rate was 90.2% and the accuracy was >99%, according to random duplicated genotyping of 5% of samples. Only one SNP, rs356221, did not result in HWE and was excluded from further analysis.

3.2.1. The Rep1 polymorphism

Genotyping of the Rep1 polymorphism revealed three predominant alleles (length: 259 bp, 261 bp, and 263 bp) (Table 2). Before the association test, rare alleles with frequencies less than 0.5% were excluded from the analysis, as previously described [10].

Association was first tested using CLUMP. Rep1 did not demonstrate association in the allelic analysis, but showed a robust

Table 2
Analysis of allele and genotype frequency differences between PD cases and controls for the Rep1 polymorphism.

SNCA Rep1	PD cases N (%)	Controls N (%)	Allele distribution
Alleles			
259	468 (27.3%)	496 (30.01%)	T1: $p = 0.18$
261	1157 (67.5%)	1064 (64.56%)	T2: $p = 0.18$
263	89 (5.2%)	88 (5.43%)	T3: $p = 0.072$
Total	1714 (100%)	1648 (100%)	T4: $p = 0.16$
SNCA Rep1	PD cases N (%)	Controls N (%)	Genotype distribution
Genotypes			
259/259	83 (9.68%)	74 (9%)	
259/261	282 (32.91%)	331 (40.17%)	T1: $p = 0.05$
259/263	20 (2.33%)	17 (2.06%)	T2: $\boldsymbol{p = 0.026}$
261/261	406 (47.38%)	334 (40.53%)	T3: $\boldsymbol{p = 0.002}$
261/263	63 (7.35%)	65 (7.88%)	T4: $\boldsymbol{p = 0.014}$
263/263	3 (0.35%)	3 (0.36%)	
Total	857 (100%)	824 (100%)	

Alleles having frequencies less than 0.5% were excluded from analysis.
N, number of PD cases or controls.
T1, Pearson's χ^2 statistic of the "raw" contingency table.
T2, the χ^2 statistic of a table with rare alleles/genotypes grouped together to prevent small expected cell counts.
T3, the largest of the χ^2 statistics of 2×2 tables, each of which compares one allele/genotype with the rest grouped together.
T4, the largest of the χ^2 statistics of all possible 2×2 tables, comparing any combination of alleles/genotypes with the rest.
Significant p values are indicated in bold.

association in the genotypic analysis, being the three χ^2 statistics T2, T3, and T4 significant, and the forth one (T1) close to the significance level (Table 2). In particular, frequency differences in cases vs. controls were striking for the 259/261 and 261/261 genotypes (∼33% vs. 40%, and 47.4% vs. 40.5%, respectively). Hence, association of Rep1 with PD was also analyzed assuming a biallelic model, only considering individuals carrying the 259 and/or 261 alleles (∼90% of cases and controls). This approach again revealed no association in the allelic analysis (adjusted *P*-value = 0.398) (Table 3), but evidenced significant associations both in the genotypic analysis (*P* = 0.005) and in the analysis performed assuming a dominant model of inheritance (*P* = 0.004; the 261 bp-long being the risk allele) (Supplementary Table 1).

3.2.2. SNPs in SNCA

Seven SNPs mapping in the *SNCA* region showed a significant association with the disease; three retained their significance even after correction for covariates (Table 3). One (rs2737029) of these resulted associated with an increased risk of PD (OR = 1.22, 95% CI = 1.03—1.44), whereas the remaining two showed a similar strength of association and a protective effect (top SNP rs356186, *P* = 0.005, OR = 0.74, 95% CI = 0.59—0.91) (Table 3). Interestingly,

Table 1
Characteristics of PD patients and control subjects.

	Cases ($n = 904$)	Controls ($n = 891$)	Unadjusted OR (95% CI)	Adjusted OR (95% CI)
Age (y ± SD)[a]	66.6 ± 10.9	62.4 ± 14.6	n.s.	−
Mean onset age (y ± SD)[b]	56.1 ± 11.0	n.a.	−	−
Mean disease duration (y ± SD)	10.6 ± 5.9	n.a.	−	−
Familial cases (%)	32	n.a.	−	−
Female (%)	39.9	65.5	0.35 (0.29—0.42)	0.31 (0.25—0.40)
Coffee consumption (%)[c]				
No (up to 2 coffee/day)	71.2	56.6	0.53 (0.42—0.67)	0.54 (0.42—0.69)
Yes (more than 2 coffee/day)	28.8	43.4		
Smoking (%)[c]				
No	63.8	59.2	0.82 (0.67—1.00)	0.68 (0.53—0.88)
Yes + former	36.2	40.8		

Data are shown either as mean ± standard deviation (SD) or as %; y, years; n.s., not significant; n.a., not applicable.
[a] Age at blood draw.
[b] Age at which the patient noticed the first PD symptoms.
[c] Data are based on patients and controls self report. Current and former smokers were aggregated in the single category of smokers.

Table 3
Allelic frequencies and association test of Rep1 and selected SNPs in the *SNCA* and *MAPT* regions.

Chr	bp[a]	Locus	SNP	Major/minor Allele	MAF cases	MAF controls	Unadjusted analysis		Adjusted analysis[b]	
							P value	OR (95% CI)	P value	OR (95%CI)
4	90847150	*SNCA*	rs356180	T/C	0.34	0.31	0.051	1.16 (1.00–1.36)	0.227	1.12 (0.93–1.36)
4	90856624	*SNCA*	rs356219	G/A	0.41	0.38	**0.027**	1.17 (1.02–1.35)	0.077	1.17 (0.98–1.39)
4	90860363	*SNCA*	rs356220	T/C	0.41	0.37	**0.023**	1.18 (1.02–1.36)	0.060	1.19 (0.99–1.42)
4	90876514	*SNCA*	rs3775423	T/C	0.08	0.07	0.575	1.08 (0.83–1.40)	0.210	1.23 (0.89–1.69)
4	90885064	*SNCA*	rs356203	G/A	0.42	0.38	0.033	1.16 (1.01–1.34)	0.085	1.16 (0.98–1.38)
4	90906950	*SNCA*	rs356192	C/T	0.28	0.26	0.320	1.08 (0.92–1.27)	0.355	1.09 (0.90–1.33)
4	90924387	*SNCA*	rs356186	A/G	0.17	0.22	**3.08×10^{-04c}**	0.73 (0.61–0.87)	**0.005**	0.74 (0.59–0.91)
4	90930793	*SNCA*	rs2737029	C/T	0.44	0.40	0.015	1.19 (1.03–1.36)	**0.020**	1.22 (1.03–1.44)
4	90948625	*SNCA*	rs2197120	A/G	0.18	0.23	**7.20×10^{-04c}**	0.75 (0.63–0.89)	**0.007**	0.75 (0.61–0.92)
4	90959901	*SNCA*	rs2737020	C/T	0.26	0.28	0.151	0.89 (0.77–1.04)	0.391	0.92 (0.76–1.11)
4	90979851	*SNCA*	rs2583988	T/C	0.31	0.27	0.029	1.19 (1.02–1.38)	0.118	1.17 (0.96–1.41)
4	90986232	*SNCA*	Rep1	259/261	0.29	0.32	**0.046**	0.85 (0.73–1.00)	0.398	0.92 (0.76–1.11)
17	41407682	*MAPT*	rs1800547	G(H2)/A(H1)	0.19	0.24	**4.63×10^{-04c}**	0.74 (0.62–0.87)	**0.002**[c]	0.72 (0.58–0.89)
17	41457408	*MAPT*	rs9468	C(H2)/T(H1)	0.19	0.24	**4.23×10^{-04c}**	0.71 (0.60–0.83)	**4.62×10^{-04c}**	0.69 (0.56–0.85)

Rep1 was evaluated assuming a biallelic model, by considering only the two most frequent alleles (259 and 261 bp).
Chr, chromosome; MAF, minor allele frequency; L, 95% CI lower limit; U, 95% CI upper limit.
Significant p-values are indicated in bold.
[a] Position is given according to UCSC Genome Browser [http://genome.ucsc.edu/, Mar. 2006 (NCBI36/hg18) assembly].
[b] Adjusted for sex, smoke, and coffee consumption.
[c] Significant even after the conservative Bonferroni correction for 14 SNP tests.

these two SNPs were the only ones showing a significant association in the genotypic analysis (Supplementary Table 1).

3.2.3. SNPs in MAPT

Both selected *MAPT* SNPs showed a strong signal of association (top SNP: rs9468, *P* = 0.00046, OR = 0.69, 95% CI = 0.56–0.85) (Table 3). The tagged H1 haplotype was over-represented in patients with respect to controls (81% vs. 76%; OR = 1.39 and 1.45 for rs1800547 and rs9468, respectively). The association was also confirmed by genotype distribution analysis (Supplementary Table 1).

3.3. LD blocks and haplotype analyses of the SNCA locus

LD within *SNCA* was evaluated using Haploview. The whole region showed an overall strong degree of LD; nonetheless, 2 distinct haplotype blocks could be defined, one covering a 37 kb-long region and corresponding to the 5′ end of the gene, the other spanning from intron 4 to the 3′ region of *SNCA* (∼83 kb) (Fig. 1). Haplotype analysis was accomplished considering only SNPs mapping within the same LD block. The CACCATAA haplotype, composed of SNPs rs356180-rs356219-rs356220-rs3775423-rs356203-rs356192-rs356186-rs2737029 (block 1), was less represented in PD cases (13.9%) respect to controls (19.6%), thus resulting a protective haplotype (*P* = 1.29 × 10^{-05}). In block 2, a second protective haplotype was found (16.1 and 21.1% in patients and controls; *P* = 2 × 10^{-04}), together with two predisposing haplotypes, whose frequency is however below 2% (Supplementary Table 2).

3.4. Gene–gene and gene-environment interaction analyses

Considering that some evidence suggests that genetic variants in *SNCA* and *MAPT* could interact in modulating risk for PD [14], an attempt to evaluate gene–gene interactions was performed. All pair-wise combinations of SNPs in linkage equilibrium (i.e. lying on different chromosomes) were tested, but no allelic epistasis was observed (Supplementary Table 3). A similar analysis, performed only on cases, did not show any significant interaction (Supplementary Table 3).

Because both coffee consumption and smoking status showed a significant protective effect against PD, a preliminary test for homogeneity was used to identify their possible interactions with genotyped polymorphisms. However, no tendency to heterogeneity was detected between groups stratified according to the presence/absence of the habit for any SNP (lowest *P*-value = 0.18) (data not shown). No further analysis was hence attempted.

4. Discussion

Many efforts have been focused on the identification of common variants conferring susceptibility to PD: replication studies on candidate genes, GWAS, and large meta-analyses provide a strong support for a role of *SNCA* and *MAPT* in the genetic predisposition to the disease [4,5,9–23]. Considering that no extensive data on the Italian population are currently available, we investigated the involvement of *SNCA* and *MAPT* genetic variants in PD susceptibility in our cohort of 904 Italian PD patients and 891 controls.

Concerning the *SNCA* gene, we first analyzed the Rep1 polymorphism, confirming its association with PD and emphasizing the role of the 261 bp-long allele as a risk factor for the disease, especially under the dominant model of inheritance. Indeed, the large meta-analysis performed by Maraganore and colleagues [5] showed a robust association of both the Rep1 259 bp-long allele (protective effect) and the 263 bp-long allele (predisposing effect), but not of the 261 bp-long allele. The lack of association of the Rep1 261 bp-long allele in the meta-analysis could be due to the large fluctuation in its frequencies observed in the analyzed populations (range: 62–71% and 61–72% in cases and controls, respectively). In any case, our results well correlate with those reported by Cronin and colleagues, who studied the effects of three distinct Rep1 variants in the brains of human *SNCA* transgenic mice. In animals homozygous for the risk allele (261 bp), they observed a 1.7-fold increase of the *SNCA* mRNA and a 1.25-fold increase in the protein level, suggesting that this allele may mimic *SNCA* locus multiplication [7].

Besides Rep1, we performed an association analysis between 12 SNPs covering the whole *SNCA* gene region and PD. We found strong signals of association clustered in a region of ∼24 kb located in the middle of *SNCA* intron 4 (flanking SNPs: rs356186 and rs2197120; Table 3 and Fig. 1), where only few putative enhancers
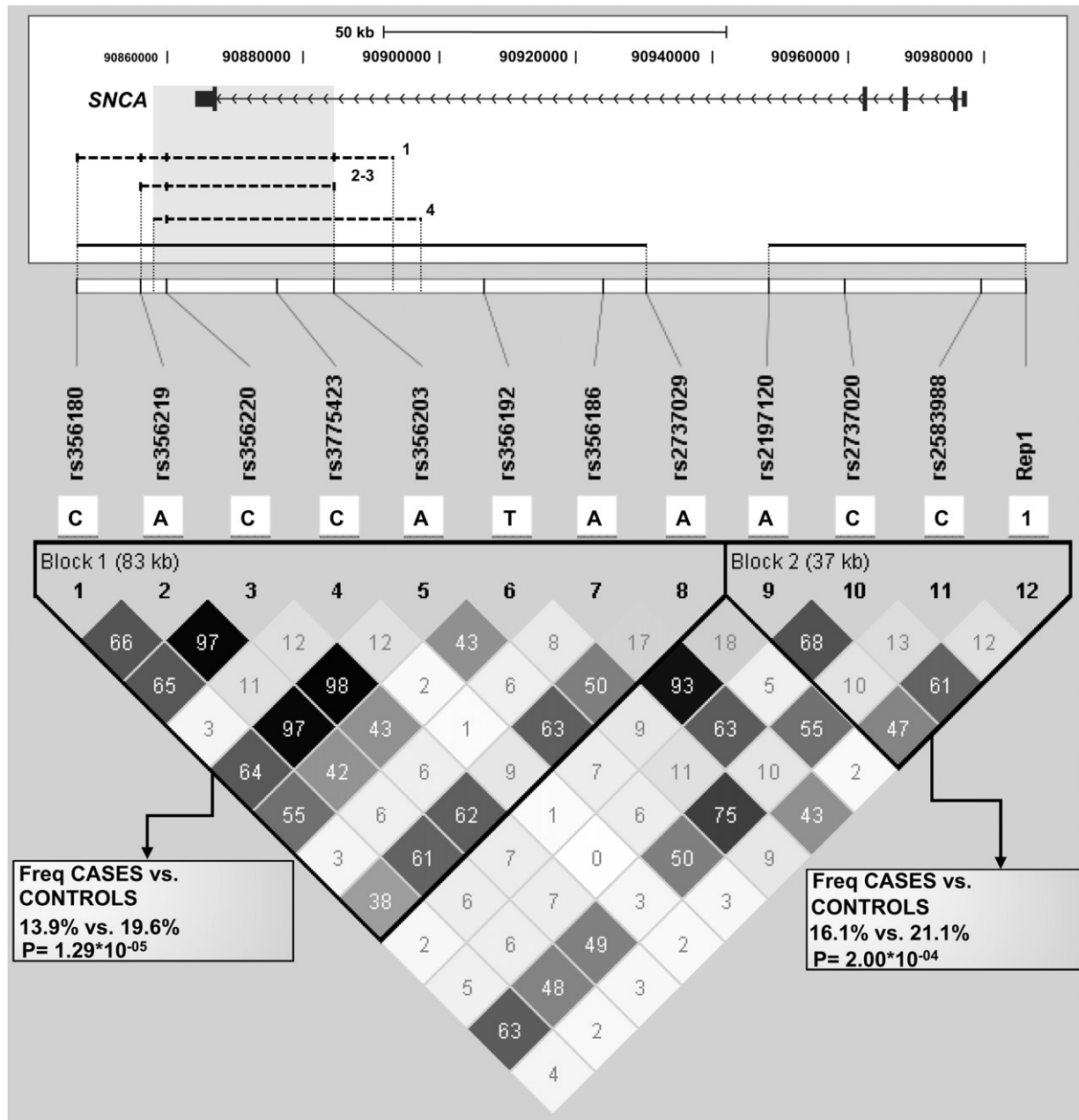
Fig. 1. LD structure of the *SNCA* locus. The structure of the *SNCA* gene is shown (drawn to scale; exons are represented by boxes, introns by horizontal lines, which also indicate the transcriptional direction of the gene). At the top of the gene, the ruler indicates the gene size, whereas numbers below the ruler depict the position within chromosome 4 (UCSC Genome Browser, Mar. 2006 release, NCBI36/hg18). Genotyped polymorphisms are listed, and their locations are shown by lines relative to the *SNCA* gene. Below the gene scheme, associated haplotypes identified in this (continuous) or previous (dashed) works are indicated as horizontal lines; the overlapping region is shaded in gray; numbers beside previously-reported haplotypes correspond to: 1) [9]; 2) [10]; 3) [19]; 4) [18]. Only SNPs shared among our and other haplotypes are indicated with a vertical hyphen in haplotypes 1—4; capital letters indicate alleles exerting a protective effect, small letters indicate alleles exerting a predisposing effect. For the Rep1 microsatellite, only the two most frequent alleles were considered: the 259 bp-long allele was named 1, the 261 bp-long allele was named 2. In the lower part of the figure, the LD structure of the *SNCA* locus is shown. Pairwise LD values, estimated for the genotyped SNPs, are represented by boxes. The standard color scheme of Haploview was used to display the strength of LD: black indicates strong LD, grey intermediate, whereas white denotes no LD. r2 values are shown within the boxes.

are reported (see the ENCODE track, UCSC Genome browser, http://genome.ucsc.edu/). These signals were further confirmed by haplotype analysis, which revealed the existence of a strong protective allele spanning from the *SNCA* 3′ flanking region up to intron 4 (Fig. 1). Our results corroborate the hypothesis of a major susceptibility block located in the 3′ region of *SNCA* in Caucasian populations, as reported in other studies [10,11], but extend the putative susceptibility region toward intron 4, as previously hypothesized [9,10]. Marker saturation of this overlooked region could provide more accuracy in defining LD blocks and haplotypes,

thus hopefully facilitating the identification of functional variants underlying PD susceptibility.

Since the discovery of the involvement of *MAPT* in neurological disorders [13], several studies investigated its possible involvement in PD, especially in the light of in-vitro studies suggesting an interaction between the *MAPT*-encoded tau protein and α-synuclein [32]. Although initially controversial, mounting evidence now suggests the H1 haplotype as a strong PD susceptibility factor [18—23]. The association study performed in our population further confirms the role of the H1-clade as risk allele: strong signals of

associations were observed both in the allelic and in the genotypic analyses. In particular, the association we observed between the H1 haplotype (tagged by rs9468) and PD was highly significant ($P = 4.62 \times 10^{-4}$), and the effect size was even stronger than that observed for the best associated SNP in *SNCA* (OR = 1.45, 95% CI = 1.18−1.78 vs. OR = 1.35, 95% CI = 1.1−1.69 for *MAPT* and *SNCA*).

Given our results both on *SNCA* and *MAPT*, we attempted to evaluate potential gene−gene and gene-environment interactions. No pair-wise interaction was observed between polymorphisms of the two genes, with no evidence suggesting cumulative, synergistic, or antagonistic effects, as for the majority of other Caucasian populations ([33] and references therein). Only a few association studies have examined gene-environment interactions in PD, the most studied environmental exposures being coffee drinking and cigarette smoking. Cumulative and interactive effects of cigarette smoking and gene polymorphisms in *NOS1* (neuronal nitric oxide synthase), *NOS2A* (inducible nitric oxide synthase), and *SNCA* were reported to modulate the risk of PD [31,34], whereas a significant interaction was observed between the *APOE* gene and coffee intake [34]. Our analysis, though performed on a relatively large number of PD cases and controls, failed to observe any interaction, again underscoring the need for larger association studies to define the role of gene-environment interactions in the development of PD.

In conclusion, our results strongly confirm the susceptibility role of *SNCA* and *MAPT* genes in idiopathic PD in the Italian population, and add a further piece of information toward the future identification of the actual functional variants underlying these associations and the construction of a cumulative score of genetic variants to predict individual disease risk.

Acknowledgments

Appendix. Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.parkreldis.2011.10.014.

References

[1] Lees AJ, Hardy J, Revesz T. Parkinson's disease. Lancet 2009;373:2055−66.
[2] Lesage S, Brice A. Parkinson's disease: from monogenic forms to genetic susceptibility factors. Hum Mol Genet 2009;18:R48−59.
[3] Bekris LM, Mata IF, Zabetian CP. The genetics of Parkinson disease. J Geriatr Psychiatry Neurol 2010;23:228−42.
[4] Farrer M, Maraganore DM, Lockhart P, Singleton A, Lesnick TG, de Andrade M, et al. Alpha-synuclein gene haplotypes are associated with Parkinson's disease. Hum Mol Genet 2001;10:1847−51.
[5] Maraganore DM, de Andrade M, Elbaz A, Farrer MJ, Ioannidis JP, Krüger R, et al. Collaborative analysis of alpha-synuclein gene promoter variability and Parkinson disease. JAMA 2006;296:661−70.
[6] Chiba-Falek O, Nussbaum RL. Effect of allelic variation at the NACP-Rep1 repeat upstream of the alpha-synuclein gene (SNCA) on transcription in a cell culture luciferase reporter system. Hum Mol Genet 2001;10:3101−9.
[7] Cronin KD, Ge D, Manninger P, Linnertz C, Rossoshek A, Orrison BM, et al. Expansion of the Parkinson disease-associated SNCA-Rep1 allele upregulates human alpha-synuclein in transgenic mouse brain. Hum Mol Genet 2009;18:3274−85.
[8] Chiba-Falek O, Kowalak JA, Smulson ME, Nussbaum RL. Regulation of alpha-synuclein expression by poly (ADP ribose) polymerase-1 (PARP-1) binding to the NACP-Rep1 polymorphic site upstream of the SNCA gene. Am J Hum Genet 2005;76:478−92.
[9] Mizuta I, Satake W, Nakabayashi Y, Ito C, Suzuki S, Momose Y, et al. Multiple candidate gene analysis identifies alpha-synuclein as a susceptibility gene for sporadic Parkinson's disease. Hum Mol Genet 2006;15:1151−8.
[10] Mueller JC, Fuchs J, Hofer A, Zimprich A, Lichtner P, Illig T, et al. Multiple regions of alpha-synuclein are associated with Parkinson's disease. Ann Neurol 2005;57:535−41.
[11] Pankratz N, Nichols WC, Elsaesser VE, Pauciulo MW, Marek DK, Halter CA, et al. Alpha-synuclein and familial Parkinson's disease. Mov Disord 2009;24:1125−31.
[12] Mata IF, Shi M, Agarwal P, Chung KA, Edwards KL, Factor SA, et al. SNCA variant associated with Parkinson disease and plasma alpha-synuclein level. Arch Neurol 2010;67:1350−6.
[13] Kalinderi K, Fidani L, Bostantjopoulou S. From 1997 to 2007: a decade journey through the H1 haplotype on 17q21 chromosome. Parkinsonism Relat Disord 2009;15:2−5.
[14] Goris A, Williams-Gray CH, Clark GR, Foltynie T, Lewis SJ, Brown J, et al. Tau and alpha-synuclein in susceptibility to, and dementia in, Parkinson's disease. Ann Neurol 2007;62:145−53.
[15] Tobin JE, Latourelle JC, Lew MF, Klein C, Suchowersky O, Shill HA, et al. Haplotypes and gene expression implicate the MAPT region for Parkinson disease: the genepd study. Neurology 2008;71:28−34.
[16] Zabetian CP, Hutter CM, Factor SA, Nutt JG, Higgins DS, Griffith A, et al. Association analysis of MAPT H1 haplotype and subhaplotypes in Parkinson's disease. Ann Neurol 2007;62:137−44.
[17] Refenes N, Bolbrinker J, Tagaris G, Orlacchio A, Drakoulis N, Kreutz R. Role of the H1 haplotype of microtubule-associated protein tau (MAPT) gene in Greek patients with Parkinson's disease. BMC Neurol 2009;9:26.
[18] 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 2009;41:1303−7.
[19] Simón-Sánchez J, Schulte C, Bras JM, Sharma M, Gibbs JR, Berg D, et al. Genome-wide association study reveals genetic risk underlying Parkinson's disease. Nat Genet 2009;41:1308−12.
[20] Saad M, Lesage S, Saint-Pierre A, Corvol JC, Zelenika D, Lambert JC, et al. Genome-wide association study confirms BST1 and suggests a locus on 12q24 as the risk loci for Parkinson's disease in the European population. Hum Mol Genet 2011;20:615−27.
[21] Spencer CC, Plagnol V, Strange A, Gardner M, Paisan-Ruiz C, Band G, et al. Dissection of the genetics of Parkinson's disease identifies an additional association 5' of SNCA and multiple associated haplotypes at 17q21. Hum Mol Genet 2011;20:345−53.
[22] Edwards TL, Scott WK, Almonte C, Burt A, Powell EH, Beecham GW, et al. Genome-wide association study confirms SNPs in SNCA and the MAPT region as common risk factors for Parkinson disease. Ann Hum Genet 2010;74:97−109.
[23] Consortium IPDG. Imputation of sequence variants for identification of genetic risks for Parkinson's disease: a meta-analysis of genome-wide association studies. Lancet 2011;377:641−9.
[24] Hughes A, Daniel S, Lees A. Improved accuracy of clinical diagnosis of Lewy body Parkinson's disease. Neurology 2001;57:1497−9.
[25] Goldwurm S, Zini M, Di Fonzo A, De Gaspari D, Siri C, Simons E, et al. LRRK2 G2019S mutation and Parkinson's disease. A clinical, neuropsychological and neuropsychiatric study in a large italian sample. Parkinsonism Relat Disord 2006;12:410−9.
[26] Sironi F, Primignani P, Zini M, Tunesi S, Ruffmann C, Ricca S, et al. Parkin analysis in early onset Parkinson's disease. Parkinsonism Relat Disord 2008;14:326−33.
[27] Ross OA, Gosal D, Stone JT, Lincoln SJ, Heckman MG, Irvine GB, et al. Familial genes in sporadic disease: common variants of alpha-synuclein gene associate with Parkinson's disease. Mech Ageing Dev 2007;128:378−82.
[28] Tobler AR, Short S, Andersen MR, Paner TM, Briggs JC, Lambert SM, et al. The SNPlex genotyping system: a flexible and scalable platform for snp genotyping. J Biomol Tech 2005;16:398−406.
[29] Breslow NE, Day NE. Statistical methods in cancer research. In: The design and analysis of cohort studies, vol. ii. IARC Sci Publ; 1987. p. 1−406.
[30] Sham PC, Curtis D. Monte Carlo tests for associations between disease and alleles at highly polymorphic loci. Ann Hum Genet 1995;59:97−105.
[31] Gao HM, Hong JS. Gene-environment interactions: key to unraveling the mystery of Parkinson's disease. Prog Neurobiol 2011;94:1−19.
[32] Giasson BI, Forman MS, Higuchi M, Golbe LI, Graves CL, Kotzbauer PT, et al. Initiation and synergistic fibrillization of tau and alpha-synuclein. Science 2003;300:636−40.
[33] Elbaz A, Ross OA, Ioannidis JP, Soto-Ortolaza AI, Moisan F, Aasly J, et al. Independent and joint effects of the MAPT and SNCA genes in Parkinson disease. Ann Neurol 2011;69:778−92.
[34] McCulloch CC, Kay DM, Factor SA, Samii A, Nutt JG, Higgins DS, et al. Exploring gene-environment interactions in Parkinson's disease. Hum Genet 2008;123:257−65.