UNIVERSITÉ DE GENÈVE

Archive ouverte UNIGE
https://archive-ouverte.unige.ch

Article scientifique | Article | 2010

Published version | Open Access

This is the published version of the publication, made available in accordance with the publisher's policy.

Parkinson's disease-related LRRK2 G2019S mutation results from independent mutational events in humans

Lesage, Suzanne; Patin, Etienne; Condroyer, Christel; Leutenegger, Anne-Louise; Lohmann, Ebba; Giladi, Nir; Bar-Shira, Anat; Belarbi, Soraya; Hecham, Nassima; Pollak, Pierre; Ouvrard-Hernandez, Anne-Marie; Bardien, Soraya; Carr, Jonathan; Benhassine, Traki [**and 11 more**]

How to cite

LESAGE, Suzanne et al. Parkinson's disease-related LRRK2 G2019S mutation results from independent mutational events in humans. In: Human molecular genetics, 2010, vol. 19, n° 10, p. 1998–2004. doi: 10.1093/hmg/ddq081

This publication URL: https://archive-ouverte.unige.ch/unige:32884

Publication DOI: 10.1093/hmg/ddq081

© This document is protected by copyright. Please refer to copyright holder(s) for terms of use.

Human Molecular Genetics, 2010, Vol. 19, No. 10 **1998–2004**
doi:10.1093/hmg/ddq081
Advance Access published on March 2, 2010

Parkinson's disease-related LRRK2 G2019S mutation results from independent mutational events in humans

Suzanne Lesage[1,2,†], Etienne Patin[3,†], Christel Condroyer[1,2], Anne-Louise Leutenegger[4,5], Ebba Lohmann[1,2], Nir Giladi[6,8], Anat Bar-Shira[7], Soraya Belarbi[9], Nassima Hecham[9], Pierre Pollak[10], Anne-Marie Ouvrard-Hernandez[10], Soraya Bardien[11], Jonathan Carr[11], Traki Benhassine[12], Hiroyuki Tomiyama[13], Caroline Pirkevi[14], Tarik Hamadouche[15], Cécile Cazeneuve[16], A. Nazli Basak[14], Nobutaka Hattori[13], Alexandra Dürr[1,2,16], Meriem Tazir[9], Avi Orr-Urtreger[7,8], Lluis Quintana-Murci[3] and Alexis Brice[1,2,16,*] for the French Parkinson's Disease Genetics Study Group

[1]INSERM, UMR_S975 (Formerly UMR_S679), Paris, France, [2]Pierre et Marie Curie-Paris 6 University, UMR_S975 CRicm, Pitié-Salpêtrière, Paris, France, [3]Institut Pasteur, Human Evolutionary Genetics, CNRS URA3012, Paris, France, [4]INSERM, U946, Fondation Jean Dausset-CEPH, Paris, France, [5]Paris-Diderot University, Institut Universistaire d'Hématologie, paris, France, [6]Movement Disorders Unit Parkinson Center Department of Neurology and [7]Genetic Institute, Tel-Aviv Sourasky Medical Center, Tel-Aviv, Israel, [8]The Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel, [9]Service de Neurologie, CHU Mustapha, Alger, Algeria, [10]Service de Neurologie, CHU de Grenoble, Grenoble, France, [11]Division of Neurology, University of Stellenbosch, Cape Town, South Africa, [12]Laboratoire de biologie moléculaire, Université de Bab Ezzouar, Alger, Algérie, [13]Department of Neurology, Juntendo University School of Medicine, Tokyo, Japan, [14]Neurodegeneration Research Laboratory, Molecular Biology and Genetics Department, Boğaziçi University, Istanbul, Turkey, [15]Laboratoire de biologie moléculaire, Université M'hamed Bougara, Boumerdes, Algeria and [16]Department of Genetics and Cytogenetics, AP-HP, Pitié-Salpêtrière Hospital, Paris, France

Received December 8, 2009; Revised and Accepted February 17, 2010

Mutations in the *leucine-rich-repeat kinase 2* (*LRRK2*) gene have been identified in families with autosomal dominant Parkinson's disease (PD) and in sporadic cases; the G2019S mutation is the single most frequent. Intriguingly, the frequency of this mutation in PD patients varies greatly among ethnic groups and geographic origins: it is present at <0.1% in East Asia, ∼2% in European-descent patients and can reach frequencies of up to 15–40% in PD Ashkenazi Jews and North African Arabs. To ascertain the evolutionary dynamics of the G2019S mutation in different populations, we genotyped 74 markers spanning a 16 Mb genomic region around G2019S, in 191 individuals carrying the mutation from 126 families of different origins. Sixty-seven families were of North-African Arab origin, 18 were of North/Western European descent, 37 were of Jewish origin, mostly from Eastern Europe, one was from Japan, one from Turkey and two were of mixed origins. We found the G2019S mutation on three different haplotypes. Network analyses of the three carrier haplotypes showed that G2019S arose independently at least twice in humans. In addition, the population distribution of the intra-allelic diversity of the most widespread carrier haplotype, together with estimations of the age of G2019S determined by two different methods, suggests that one of the founding G2019S mutational events occurred in the Near East at least 4000 years ago.

*To whom correspondence should be addressed at: INSERM UMR_S975 (former UMR_S679), Hôpital de la Salpêtrière, 47, Boulevard de l'Hôpital, 75651 Paris cedex 13, France. Tel: +33 142162182; Fax: +33 144243658; Email: alexis.brice@upmc.fr
†The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

© *The Author 2010. Published by Oxford University Press. All rights reserved.*
For Permissions, please email: journals.permissions@oxfordjournals.org

Downloaded from http://hmg.oxfordjournals.org/ at Institut universitaire de hautes etudesinternationales - Bibliotheque on January 7, 2014

INTRODUCTION

Late-onset Parkinson's disease (PD [MIM# 168600]) is a severe, progressive, age-related neurodegenerative disorder with an estimated prevalence of ~2% among persons over 65 years of age (1). It is characterized clinically by bradykinesia, resting tremor, rigidity and postural instability, and neuropathologically by the loss of dopaminergic neurons in the brainstem and the presence of intracytoplasmic inclusions called Lewy bodies in surviving neurons.

Mutations in the *leucine-rich repeat kinase 2* (*LRRK2*) gene [MIM# 609007] at the PARK8 locus [MIM# 607060] are the most common known cause of autosomal dominant PD (2,3). It has been shown that mutations in the *LRRK2* gene can contribute significantly to the etiology of both familial and sporadic PD (4). *LRRK2* is located on chromosome 12q12 and consists of 51 exons. The product of *LRRK2*, also termed dardarin, is a highly conserved 2527 amino acid protein that has been suggested to be a cytoplasmic kinase involved in neuronal signaling (5). To date, at least 50 different possibly pathogenic *LRRK2* sequence variants have been reported in families presenting a wide variety of clinical and pathological features (6). The most common and most studied *LRRK2* mutation, c.6055 G > A in exon 41, encodes a G2019S change in the highly conserved activation loop of its MAP kinase domain (7–10). Haplotype analyses showed that the mutation is found on two distinct haplotypes: a common founder haplotype that is shared by most G2019S carriers and a second haplotype carried by three European-American families, suggesting that the G2019S mutation arose twice in these populations (11–16). The first haplotype was reported to derive from a common ancestor estimated to date back 2250–2600 years, whereas the second appeared more recently (15). A third haplotype, which differs from the European and North African haplotypes, was identified in three Japanese PD families with the mutation (17,18).

Despite the importance of the G2019S mutation in patients with PD and the considerable interest in the evolutionary history of this mutation in the natural setting, it has never been formally tested whether the three G2019S-carrying haplotypes derive from a single founder mutation or from three independent mutational events. In addition, the size of the sample in previous studies (Table 1), which is critical for a robust estimate of the origin and age of alleles (19), has been too small to allow for the evaluation of the age of G2019S in different ethnic groups. Indeed, it is now clear that the frequency of G2019S in PD patients depends greatly on their geographic origin: it accounts for 0 to 18% of familial and 0 to 6% of apparently sporadic PD in Europe with a North-South gradient of distribution (20). It is very rare in East Asia, but accounts for approximately 15 and 40% in Ashkenazi Jews and North African Arabs, respectively (14,21–28). The marked ethnic differences in the risk conferred by G2019S might result from the frequency of G2019S in non-affected populations (i.e. the odds ratio depends on the frequency of the disease-causing variant in controls) (29), or from different ethnic-related genetic backgrounds, which might include mutations in modifier genes.

Here we sought to assess the evolutionary history of the G2019S mutation in the natural setting, i.e. in a large multiethnic population panel.

RESULTS

To assess the evolutionary history of the G2019S mutation, we first determined the haplotypic phase of the 191 genotyped subjects (382 chromosomes) carrying or not the mutation, for the 74 markers genotyped (see Materials and Methods). Our haplotype data clearly indicated that the G2019S mutation was found on three different haplotypes: haplotype 1, which was the most frequent and widespread (11–16,24); haplotype 2, which was initially identified in three European-American families only (15), and was found in two French families in the present study; haplotype 3, which was primarily observed in Japanese PD patients (17,18), but also in a Turkish family (30). Previous studies have shown that these three haplotypes could be distinguished by markers only ~5 kb upstream and downstream of G2019S (15,18), suggesting that the mutation arose independently in each case.

To formally test whether the G2019S mutation has a single or multiple independent origins, we used our resequencing data from 226 subjects to identify the single nucleotide polymorphisms (SNPs) closest to G2019S that distinguished the three haplotypes (6). Haplotype 2 differed from haplotype 1 at markers 2.6 kb upstream (rs6581668) and 1.7 kb downstream (rs10506155) of G2019S, and haplotype 3 differed from haplotype 1 at markers 2.6 kb upstream (rs6581668) and 6.2 kb downstream (rs10784522) of G2019S (Supplementary Material, Table S1). It is thus highly improbable, given the very short distance separating these SNPs, that haplotypes 2 and 3 were generated by recombination between haplotype 1 and other non-carrier haplotypes. This observation suggests that the G2019S appeared more than once in humans.

To evaluate the phylogenetic relationships between G2019S carrier and non-carrier haplotypes, we reconstructed the network of our haplotype data (see Materials and Methods), restricted to unrelated individuals and to the 32 SNPs contained in the ~250 kb interval between D12S2514 and D12S2519 to avoid reticulations in the network due to recombination events or recurrent mutations; *Phase* estimated that the local recombination rate increased by a factor of 130 at microsatellite D12S2519 (data not shown). Interestingly, the network showed that a non-carrier form of each of the three G2019S-carrying haplotypes was also found at high frequency in the general population (Fig. 1). In addition, haplotype 1 and haplotypes 2 and 3 were located at the opposite tips of the network, illustrating the high level of divergence among the three haplotypes. Altogether, our results demonstrate that the G2019S mutation occurred independently at least twice during recent human evolution; once on a haplotype 1 background and separately on the evolutionary-related haplotype 2 or 3.

To gain further insight into the history of the G2019S mutation, we estimated the age of the mutation event occurring on haplotype 1; haplotypes 2 and 3 were not studied further because of the small number of individuals carrying these haplotypes. We used two different maximum-likelihood methods to estimate allele age, both based on the distribution of recombination events and the proportion of recombinant haplotypes around the mutation of interest (see Materials and Methods). G2019S was estimated to have appeared on haplotype 1 between 2925 and 6250 years ago, depending

Downloaded from http://hmg.oxfordjournals.org/ at Institut universitaire de hautes etudesinternationales - Bibliotheque on January 7, 2014

Table 1. Previous and present studies of the evolutionary history of the *LRRK2* haplotype 1 carrying the G2019S mutation

Study	Number of genotyped markers	Number of genotyped unrelated subjects and their origin	Number of haplotype 1 analyzed[a]	Shortest shared haplotype 1 (kb)	Age estimation (95% CI)[b]
Kachergus *et al*. (12)	17 microsatellites, 4 SNPs	7 Norwegians, 3 Americans, 2 Irish, 1 Polish	13	145.14	ND
Goldwurm *et al*. (11)	13 microsatellites, 6 SNPs	14 Italians, 1 Brazilian, 1 Portuguese, 1 Greek, 1 Moroccan	18	161.00	ND
Lesage *et al*. (13)	17 microsatellites, 4 SNPs	3 Algerians, 3 Moroccans, 1 Tunisian, 3 French, 1 Dutch, 1 Portuguese, 1 American, 1 Belgian	14	59.96	725 (375–1375)
Zabetian *et al*. (15)	13 microsatellites, 12 SNPs	13 Europeans, 9 Jewish ancestry	19	242.96	1875 (1375–2600)
Zabetian *et al*. (18)	7 microsatellites, 9 SNPs	3 Europeans, 1 Ashkenazi Jew, 2 Japanese	4	171.00	ND
Tomiyama *et al*. (17)	9 microsatellites, 3 SNPs	3 Israelis, 3 Tunisians, 1 Moroccan, 1 Japanese	7	158.40	ND
Warren *et al*. (16)	2 microsatellites, 39 SNPs	17 Tunisians	17	116.00	2600 (1950–3850)
Bar-Shira *et al*. (32)	15 microsatellites, 1 SNP	127 Ashkenazi Jews	78	243.00	1525 (1300–1800)
This study	20 microsatellites, 54 SNPs	67 North-African Arabs, 18 Europeans, 32 Ashkenazi Jews, 5 Sephardic Jews, 1 Japanese, 1 Turkish, and 2 families with mixed ancestries	99[c]	6.28	3525–4800 (2925–6250) (see Fig. 2)

[a]The number of independent (i.e. identified in unrelated individuals) G2019S-carrying haplotypes analyzed in the study and defined as haplotype 1.
[b]Ages are given in years. The intergenerational time interval was set at 25 years.
[c]Twenty-six of the 191 individuals genotyped were excluded because their phase reconstructions were of poor quality, 60 were excluded because they were related to at least one other individual, 2 were excluded because they were of mixed origins and 4 carried G2019S haplotype 2 or 3.

on the method and the data used [Fig. 2, grey dots; method a: 4037 years (3492–4875), method b: 4800 years (3825–6250), method c: 3525 years (2925–4275)]. Because the most recent common ancestor of the populations studied probably predates this time period (31), our result is consistent with a scenario in which the G2019S mutation appeared in a particular population from the Near East, North Africa or Europe and then spread by gene flow to other populations.

To further test this hypothesis, we next studied the population distribution of the intra-allelic diversity of haplotype 1 in more detail. Subjects were grouped according to their ethnic origin. Two subjects were excluded because they were of mixed ancestries (one from South Africa and one from the USA; Supplementary Material, Table S2). We built the network of all haplotypes defined as haplotype 1, using all 54 genotyped SNPs. Interestingly, the intra-allelic diversity was not evenly distributed among the different populations studied (Fig. 3). Haplotypes of North-African Arabs were principally observed in the innermost part of the network, which was the most frequent haplotype and was defined as the ancestral haplotype on which the G2019S mutation appeared. In contrast, the tips of the network corresponded to haplotypes of Ashkenazi Jews (Fig. 3), which underwent recombination events in the vicinity of the G2019S mutation.

The ethnic-specific distribution of recombining G2019S-carrying haplotypes prompted us to estimate the age of G2019S in each population separately, but using the same definition of the ancestral non-recombining haplotype determined in the whole sample. The two methods showed the age of G2019S to be greater in the Ashkenazi Jewish population [Fig. 2, green dots; method a: 9102 years (7430–11 767), method b: 5775 years (3875–9500), method c: 4550 years (3250–6425)] than in the other two populations.

DISCUSSION

In this study, we analyzed in detail the haplotypic diversity of the G2019S mutation responsible for PD in humans, in a large multi-ethnic sample of subjects. Resequencing of the G2019S genomic region as well as phylogenetic analyses of carrier and non-carrier haplotypes clearly support that G2019S occurred independently in humans at least twice, once on the haplotype 1, a widely distributed carrier haplotype, and once on the phylogenetically related haplotypes 2 and 3. Haplotype 2 has been only found in European-descent individuals, whereas haplotype 3 has been observed in few Japanese and Turkish families. However, a complete picture of the geographic distribution of these haplotypes is missing. A larger sample of individuals with haplotypes 2 and 3 would be helpful to determine whether the G2019S mutation also appeared independently on haplotypes 2 and 3.

Our age estimation of the G2019S mutation on haplotype 1 is older than previously reported, especially in Ashkenazi Jews. Generally, this difference with previous studies is likely due to the multi-ethnic origins and the 5-fold greater size of our entire sample (see Table 1, except for the Ashkenazi sample) (32), rather than the higher marker density in our study, because our estimates were still higher than those previously reported when we restricted our analyses to four SNPs (data not shown). One of the studies, however, estimated the age of G2019S in a large sample of 77 subjects, mostly Ashkenazi Jews, at 1525 years with 95% confidence intervals between 1300 and 1800 (32). Although our Ashkenazi sample includes 28 of these subjects, our estimation of the age of G2019S in Ashkenazi Jews is much earlier, using the same method and microsatellite data. This difference partially results from the fact that our sample also includes Ashkenazi

Downloaded from http://hmg.oxfordjournals.org/ at Institut universitaire de hautes etudesinternationales - Bibliotheque on January 7, 2014

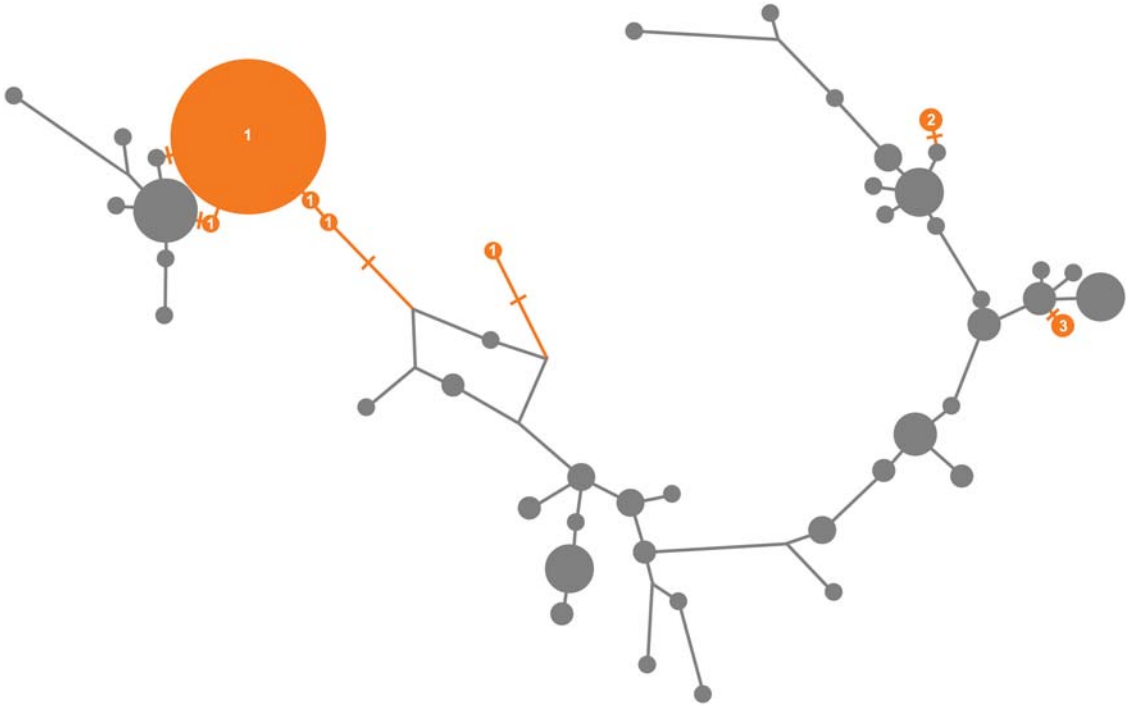

Figure 1. Haplotype network of the *LRRK2* G2019S core region in our multi-ethnic panel. Each circle represents a haplotype, the diameter of which is proportional to the haplotype frequency in the sample. Haplotypes in orange carry the G2019S mutation. The numbers in the orange circles indicate the class of the G2019S-carrying haplotypes, as defined in previous studies (12,13,15,18). Each line connecting two haplotypes represents a mutational event. The length of lines is proportional to the number of mutational events. Lines in orange correspond to the appearance of G2019S. We restricted this analysis to unrelated individuals and to the 32 SNPs located between the D12S2514 and D12S2519 microsatellites to avoid reticulations in the network due to recombination events or recurrent mutations. Interestingly, haplotype 1 and haplotypes 2 and 3 were located on each edge of the network, suggesting multiple origins of the G2019S mutation.

Jews from France, in whom the recombination event closest to G2019S was found. This difference also results from our use of a different definition and size of the ancestral non-recombining haplotype 1. Because of the differences in the ethnic origin of the samples used in our study and in Bar-Shira *et al*.'s work, we defined the ancestral haplotype from our whole multi-ethnic sample, while Bar-Shira *et al*.'s group defined it according to their sample of Ashkenazi Jews alone (32). Consequently, our estimation reflects the age of occurrence of the mutation, while the age estimated by Bar-Shira *et al*. reflects the founder effect of the mutation in the Ashkenazim. Indeed, when using the definition of the ancestral haplotype used in Bar-Shira *et al*. study (C.P. Zabetian and C.M. Hutter, personal communication), our age estimation in Ashkenazi Jews was more recent [3250 years (2300–4725)] than that using our multi-ethnic definition of the ancestral haplotype [3525 years (2925–4275)].

Aside from methodological issues, our estimations of the age of G2019S in ancestors of Ashkenazi Jews—ranging from 4500 to 9100 years depending on the method used—do not coincide with current historical records, because the history of the Ashkenazim is probably at most 2000 years old (33,34) and because other populations of Jewish ancestry do not present the mutation (e.g. Iraqi Jews) (26,32). One possible scenario that may explain our results assumes that the *LRRK2* mutation initially arose in the Near East—as previously proposed (24)– at least ∼4000 years ago. Because of a founder effect, the ancestors of present-day Ashkenazi Jews may have kept the low-frequency G2019S mutation through the different diasporas, whereas Near Eastern daughter populations lost the mutation. The mutation might then have been reintroduced by recurrent gene flow from Ashkenazi populations to other Jewish, European and North-African populations. The existence of such gene flows has been demonstrated by phylogeographic studies of mtDNA (33,34). The present-day frequency of the mutation in control populations (∼0.05% in Europeans, ∼0.5% in North-African Arabs and ∼1% in Ashkenazi Jews) (20) may support this scenario. However, other scenarios might be considered, taking into account the absence of the *LRRK2* mutation in other populations of Jewish ancestry (26,32).

In conclusion, we have shown using a large multi-ethnic population panel, that the PD-related G2019S mutation has appeared independently in humans at least twice, and that one of these events occurred most likely in the near-East at least 4000 years ago. The late onset of PD and the incomplete penetrance of the G2019S mutation (24) may have prevented elimination of the mutation by negative selection, and it has thus been continuously transmitted since its ancient origin. The higher prevalence of haplotype 1 with respect to haplotypes 2 and 3 is likely due to the earlier origin of haplotype 1, rather than different strengths of negative selection. Indeed, no major differences in the clinical features were noted between patients with haplotypes 1, 2 and 3

Downloaded from http://hmg.oxfordjournals.org/ at Institut universitaire de hautes etudesinternationales - Bibliotheque on January 7, 2014

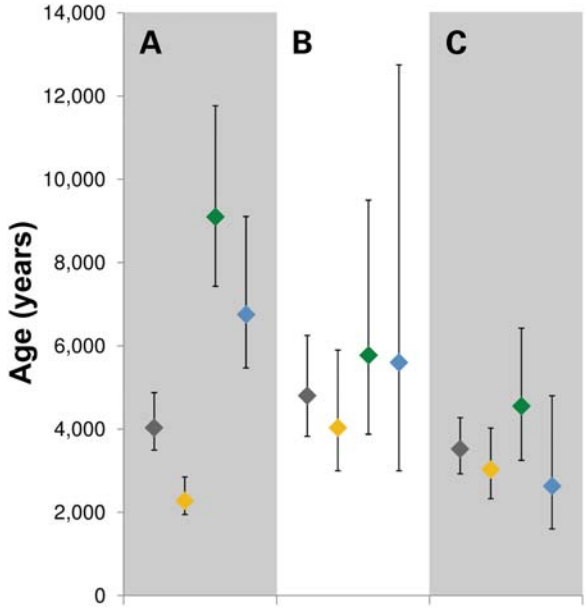

Figure 2. Age of the G2019S mutation in our whole sample and in three ethnic groups, estimated by different methods. The age of G2019S on haplotype 1 was estimated using (**A**) the SNP-based method (42) with 11 informative SNPs, (**B**) *Estiage* (43) with the same 11 informative SNPs, (**C**) *Estiage* with eight informative microsatellites (D12S2080, D12S2194, D12S2514, D12S2516, D12S2519, D12S2520, D12S2521 and D12S1301). The intergenerational time interval was set at 25 years. Vertical bars represent 95% CI of estimations. Grey, overall samples; light orange, North-African Arabs; dark green, Ashkenazi Jews; blue, Europeans.

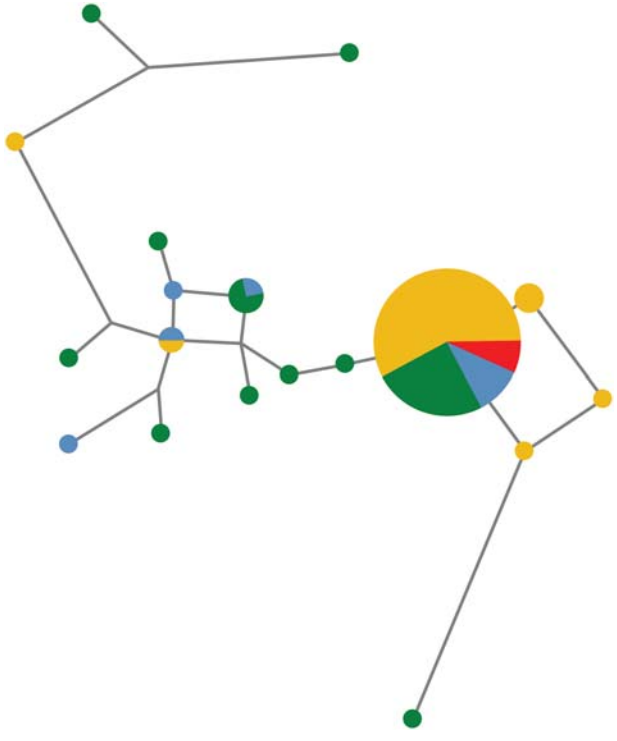

Figure 3. Haplotype network of the G2019S-carrying haplotype 1 in our multi-ethnic panel. This analysis was performed using all 54 genotyped SNPs in unrelated individuals only. Light orange, North-African Arabs; dark green, Ashkenazi Jews; blue, Europeans; salmon, Sephardic Jews.

(17,26,28,35,36). However, the small number of families sharing haplotypes 2 and 3 does not allow detecting significant differences between these three groups. Further studies of large samples of affected and unaffected individuals from different human populations will help understand the strong ethnic variation in the risk of PD conferred by this mutation.

MATERIALS AND METHODS

Study populations

This population data set was mostly derived from 1230 unrelated patients with PD and 391 control individuals recruited from the French PD genetic study group and diverse collaborations within the Mediterranean basin. All patients met standardized diagnostic criteria for PD and were examined by a movement disorder specialist. Blood samples were obtained from all participants with their informed consent. In all, we assembled a series of 125 index cases with the G2019S mutation (120 heterozygous and 5 homozygous) and one control carrier (the spouse of a PD patient who did not carry the *LRRK2* G2019S mutation). Most of these G2019S-positive families have been reported elsewhere (13,14,17,26,28,30,35,37,38). Additional members of 16 of the families (22 affected and 43 unaffected) were included in the present study. Of the 126 families, 67 were of North-African Arab origin (54 from Algeria, 10 from Morocco and 3 from Tunisia), 18 were of North/Western European descent (12 from France, 2 from Portugal, 2 from Italy, 1 from Belgium and 1 from the Netherlands), 37 were of Jewish origin, including 32 Ashkenazi Jews (10 from Poland, 4 from Romania, 4 from Russia, 3 from Byelorussia, 3 from the Ukraine, 2 from Germany, 2 from Hungary, 3 from France and 1 expatriate family from South Africa) and 5 Sephardic Jews (3 from Algeria, 1 from Tunisia and 1 from Morocco), 1 family was from Japan, 1 from Turkey and 2 families were of mixed ancestries (1 from South Africa and 1 from the USA).

Marker selection and genotyping

In addition to the 25 markers previously described (12,15), we designed four additional short tandem repeat (STR) and 45 SNP markers within the 16 Mb *LRRK2* region delimited by the D12S2194 and D12S2520 STR markers (Supplementary Material, Table S1) (primer sequences and PCR conditions upon request). The 4 STRs—D12S$(CAAA)_n$, $(AG)_n$, $(AT)_n$ and $(TA)_n$—were identified using the program Tandem Repeats Finder (39). The 45 polymorphic SNPs were identified by sequencing the 51 exons and flanking regions of the whole *LRRK2* gene in 226 patients with PD (6). All 74 markers (20 STRs and 54 SNPs) within the PARK8 region were genotyped in all individuals, except for 28 index cases of Ashkenazi origin from Northern/Eastern Europe, who were analyzed for a subset of 41 markers (26,32). SNP genotyping was performed by sequencing with the Big Dye Terminator Cycle Sequencing Ready Reaction kit and analysis with DNA Sequencing Analysis (version 5.1) and Seqscape (version 2.1.1) software (Applied Biosystems, Foster City, CA, USA). The STR markers were genotyped by multiplexing appropriate fluorescently-labeled primers. The fluorescent

Downloaded from http://hmg.oxfordjournals.org/ at Institut universitaire de hautes etudesinternationales - Bibliotheque on January 7, 2014

PCR products were then pooled for analysis on an ABI 3730 and 310 automated analyzers with Genescan 3.7 and Genotyper 3.7 software (Applied Biosystems).

Statistical analyses

The haplotypic phase of all 191 genotyped subjects (382 chromosomes, with or without the G2019S mutation) was obtained using the Bayesian method implemented in *Phase* v.2.1.1 (40). We divided the phasing procedure into three steps, because some of the individuals sampled were related and some data were missing (28 Ashkenazi patients from Northern/Eastern Europe were genotyped for a subset of only 41 markers). First, we phased markers without missing data in related individuals (with family data), specifying in *Phase* the parent/child relationships. The haplotypes of this first step were reconstructed with 100% confidence. Second, we phased markers without missing data in all unrelated individuals (without family data), together with one child per family whose haplotypes were retrieved from the first step and set as 'known haplotypes' in the second step. This second step was run 10 times with different seeds and the most probable run was kept. Third, we phased all markers in all unrelated individuals, retrieving the phase at markers without missing data from the second step and specifying that their phase is known. This third step was run 10 times with different seeds and the most probable run was kept. Missing genotypes were imputed by *Phase*. Sixty related individuals and 26 unrelated individuals with a low maximal probability of haplotype pairs ($<10\%$) were discarded from subsequent analyses. Our final data set was composed of 105 individuals, including 105 G2019S-carrier haplotypes and 105 non-carrier haplotypes. Network reconstruction of all carrier and non-carrier haplotypes was obtained using the median-joining algorithm implemented in NETWORK v.4.5.1 (41). G2019S allele age was calculated using two different methods. The first method, designed for SNP data, assumes a null mutation rate and uses the recombination rate only as a molecular clock (42). The second approach, which was implemented in *Estiage* and designed for microsatellite data, uses both the mutation and recombination rates as molecular clocks (43). Local recombination rates were obtained from the HapMap II recombination map for SNPs (44), and from the Map-O-Mat database for microsatellites, when available. The two maps were congruent with each other. Rates that were not available were extrapolated from physical distances using a scale of 0.55 cM per Mb (estimated from the physical and genetic distances between D12S2194 and D12S1301). A stepwise mutation rate at 10^{-3} mutations per site per generation (45) was used in *Estiage* to model mutations of microsatellites. We used a 25-year intergeneration interval for both methods.

WEB RESOURCES

Online Mendelian Inheritance in Man (OMIM) (for PD, *PARK8* and *LRRK2*), http://www.ncbi.nlm.nih.gov/Omim/.

Genbank [for Homo sapiens *LRRK2* (accession number AY792511)], http://www.ncbi.nlm.nih.gov/Genbank/.

dbSNP, www.ncbi.nlm.nih.gov/projects/SNP/.

The International HapMap Project, http://www.hapmap.org/.

Map-O-Mat, http://compgen.rutgers.edu/mapomat/.

Phase program, http://depts.washington.edu/ventures/UW_Technology/Express_Licenses/PHASEv2.php.

Network program, http://www.fluxus-engineering.com/sharenet.htm.

Tandem Repeats Finder program, http://tandem.bu.edu/trf/trf.html.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at *HMG* online.

ACKNOWLEDGEMENTS

The authors are grateful to the patients and their families. We thank Merle Ruberg for critical reading of the manuscript, Frédéric Austerlitz for help in the analyses, Doron Behar for useful suggestions on Jewish genetics and history, Cyrus Zabetian and Carolyn Hutter for helpful discussions, and the DNA and Cell Bank for sample preparation.

Conflict of Interest statement. None declared.

FUNDING

This work was supported by the Agence Nationale de la Recherche (ANR-05-NEUR-019).

APPENDIX

The participants in the French Parkinson's Disease Genetics Study Group are: Y. Agid, A.-M. Bonnet, M. Borg, A. Brice, E. Broussolle, Ph. Damier, A. Destée, A. Dürr, F. Durif, E. Lohmann, M. Martinez, C. Penet, P. Pollak, O. Rascol, F. Tison, C. Tranchant, A. Troiano, M. Vérin, F. Viallet, M. Vidailhet.

REFERENCES

1. de Rijk, M.C., Breteler, M.M., Graveland, G.A., Ott, A., Grobbee, D.E., van der Meche, F.G. and Hofman, A. (1995) Prevalence of Parkinson's disease in the elderly: the Rotterdam Study. *Neurology*, **45**, 2143–2146.
2. Paisan-Ruiz, C., Jain, S., Evans, E.W., Gilks, W.P., Simon, J., van der Brug, M., Lopez de Munain, A., Aparicio, S., Gil, A.M., Khan, N. *et al.* (2004) Cloning of the gene containing mutations that cause PARK8-linked Parkinson's disease. *Neuron*, **44**, 595–600.
3. Zimprich, A., Biskup, S., Leitner, P., Lichtner, P., Farrer, M., Lincoln, S., Kachergus, J., Hulihan, M., Uitti, R.J., Calne, D.B. *et al.* (2004) Mutations in LRRK2 cause autosomal-dominant parkinsonism with pleomorphic pathology. *Neuron*, **44**, 601–607.
4. Giasson, B.I. and Van Deerlin, V.M. (2008) Mutations in LRRK2 as a cause of Parkinson's disease. *Neurosignals*, **16**, 99–105.
5. Cookson, M.R., Dauer, W., Dawson, T., Fon, E.A., Guo, M. and Shen, J. (2007) The roles of kinases in familial Parkinson's disease. *J. Neurosci.*, **27**, 11865–11868.
6. Lesage, S., Condroyer, C., Lannuzel, A., Lohmann, E., Troiano, A., Tison, F., Damier, P., Thobois, S., Ouvrard-Hernandez, A.M., Rivaud-Pechoux, S. *et al.* (2009) Molecular analyses of the LRRK2 gene in European and North African autosomal dominant Parkinson's disease. *J. Med. Genet.*, **46**, 458–464.
7. Di Fonzo, A., Rohe, C.F., Ferreira, J., Chien, H.F., Vacca, L., Stocchi, F., Guedes, L., Fabrizio, E., Manfredi, M., Vanacore, N. *et al.* (2005) A

Downloaded from http://hmg.oxfordjournals.org/ at Institut universitaire de hautes etudesinternationales - Bibliotheque on January 7, 2014

frequent LRRK2 gene mutation associated with autosomal dominant Parkinson's disease. *Lancet*, **365**, 412–415.
8. Gilks, W.P., Abou-Sleiman, P.M., Gandhi, S., Jain, S., Singleton, A., Lees, A.J., Shaw, K., Bhatia, K.P., Bonifati, V., Quinn, N.P. *et al.* (2005) A common LRRK2 mutation in idiopathic Parkinson's disease. *Lancet*, **365**, 415–416.
9. Nichols, W.C., Pankratz, N., Hernandez, D., Paisan-Ruiz, C., Jain, S., Halter, C.A., Michaels, V.E., Reed, T., Rudolph, A., Shults, C.W. *et al.* (2005) Genetic screening for a single common LRRK2 mutation in familial Parkinson's disease. *Lancet*, **365**, 410–412.
10. Mata, I.F., Wedemeyer, W.J., Farrer, M.J., Taylor, J.P. and Gallo, K.A. (2006) LRRK2 in Parkinson's disease: protein domains and functional insights. *Trends Neurosci.*, **29**, 286–293.
11. Goldwurm, S., Di Fonzo, A., Simons, E.J., Rohe, C.F., Zini, M., Canesi, M., Tesei, S., Zecchinelli, A., Antonini, A., Mariani, C. *et al.* (2005) The G6055A (G2019S) mutation in LRRK2 is frequent in both early and late onset Parkinson's disease and originates from a common ancestor. *J. Med. Genet.*, **42**, e65.
12. Kachergus, J., Mata, I.F., Hulihan, M., Taylor, J.P., Lincoln, S., Aasly, J., Gibson, J.M., Ross, O.A., Lynch, T., Wiley, J. *et al.* (2005) Identification of a novel LRRK2 mutation linked to autosomal dominant parkinsonism: evidence of a common founder across European populations. *Am. J. Hum. Genet.*, **76**, 672–680.
13. Lesage, S., Leutenegger, A.L., Ibanez, P., Janin, S., Lohmann, E., Durr, A. and Brice, A. (2005) LRRK2 haplotype analyses in European and North African families with Parkinson disease: a common founder for the G2019S mutation dating from the 13th century. *Am. J. Hum. Genet.*, **77**, 330–332.
14. Lesage, S., Durr, A., Tazir, M., Lohmann, E., Leutenegger, A.L., Janin, S., Pollak, P. and Brice, A. (2006) LRRK2 G2019S as a cause of Parkinson's disease in North African Arabs. *N. Engl. J. Med.*, **354**, 422–423.
15. Zabetian, C.P., Hutter, C.M., Yearout, D., Lopez, A.N., Factor, S.A., Griffith, A., Leis, B.C., Bird, T.D., Nutt, J.G., Higgins, D.S. *et al.* (2006) LRRK2 G2019S in families with Parkinson disease who originated from Europe and the Middle East: evidence of two distinct founding events beginning two millennia ago. *Am. J. Hum. Genet.*, **79**, 752–758.
16. Warren, L., Gibson, R., Ishihara, L., Elango, R., Xue, Z., Akkari, A., Ragone, L., Pahwa, R., Jankovic, J., Nance, M. *et al.* (2008) A founding LRRK2 haplotype shared by Tunisian, US, European and Middle Eastern families with Parkinson's disease. *Parkinsonism Relat. Disord.*, **14**, 77–80.
17. Tomiyama, H., Li, Y., Funayama, M., Hasegawa, K., Yoshino, H., Kubo, S., Sato, K., Hattori, T., Lu, C.S., Inzelberg, R. *et al.* (2006) Clinicogenetic study of mutations in LRRK2 exon 41 in Parkinson's disease patients from 18 countries. *Mov. Disord.*, **21**, 1102–1108.
18. Zabetian, C.P., Morino, H., Ujike, H., Yamamoto, M., Oda, M., Maruyama, H., Izumi, Y., Kaji, R., Griffith, A., Leis, B.C. *et al.* (2006) Identification and haplotype analysis of LRRK2 G2019S in Japanese patients with Parkinson disease. *Neurology*, **67**, 697–699.
19. Slatkin, M. and Rannala, B. (2000) Estimating allele age. *Annu. Rev. Genomics Hum. Genet.*, **1**, 225–249.
20. Healy, D.G., Falchi, M., O'Sullivan, S.S., Bonifati, V., Durr, A., Bressman, S., Brice, A., Aasly, J., Zabetian, C.P., Goldwurm, S. *et al.* (2008) Phenotype, genotype, and worldwide genetic penetrance of LRRK2-associated Parkinson's disease: a case-control study. *Lancet Neurol.*, **7**, 583–590.
21. Lu, C.S., Simons, E.J., Wu-Chou, Y.H., Fonzo, A.D., Chang, H.C., Chen, R.S., Weng, Y.H., Rohe, C.F., Breedveld, G.J., Hattori, N. *et al.* (2005) The LRRK2 I2012T, G2019S, and I2020T mutations are rare in Taiwanese patients with sporadic Parkinson's disease. *Parkinsonism Relat. Disord.*, **11**, 521–522.
22. Tan, E.K., Shen, H., Tan, L.C., Farrer, M., Yew, K., Chua, E., Jamora, R.D., Puvan, K., Puong, K.Y., Zhao, Y. *et al.* (2005) The G2019S LRRK2 mutation is uncommon in an Asian cohort of Parkinson's disease patients. *Neurosci. Lett.*, **384**, 327–329.
23. Fung, H.C., Chen, C.M., Hardy, J., Hernandez, D., Singleton, A. and Wu, Y.R. (2006) Lack of G2019S LRRK2 mutation in a cohort of Taiwanese with sporadic Parkinson's disease. *Mov. Disord.*, **21**, 880–881.
24. Ozelius, L.J., Senthil, G., Saunders-Pullman, R., Ohmann, E., Deligtisch, A., Tagliati, M., Hunt, A.L., Klein, C., Henick, B., Hailpern, S.M. *et al.* (2006) LRRK2 G2019S as a cause of Parkinson's disease in Ashkenazi Jews. *N. Engl. J. Med.*, **354**, 424–425.
25. Punia, S., Behari, M., Govindappa, S.T., Swaminath, P.V., Jayaram, S., Goyal, V., Muthane, U.B., Juyal, R.C. and Thelma, B.K. (2006) Absence/rarity of commonly reported LRRK2 mutations in Indian Parkinson's disease patients. *Neurosci. Lett.*, **409**, 83–88.
26. Orr-Urtreger, A., Shifrin, C., Rozovski, U., Rosner, S., Bercovich, D., Gurevich, T., Yagev-More, H., Bar-Shira, A. and Giladi, N. (2007) The LRRK2 G2019S mutation in Ashkenazi Jews with Parkinson disease: is there a gender effect? *Neurology*, **69**, 1595–1602.
27. Hulihan, M.M., Ishihara-Paul, L., Kachergus, J., Warren, L., Amouri, R., Elango, R., Prinjha, R.K., Upmanyu, R., Kefi, M., Zouari, M. *et al.* (2008) LRRK2 Gly2019Ser penetrance in Arab-Berber patients from Tunisia: a case-control genetic study. *Lancet Neurol.*, **7**, 591–594.
28. Lesage, S., Belarbi, S., Troiano, A., Condroyer, C., Hecham, N., Pollak, P., Lohman, E., Benhassine, T., Ysmail-Dahlouk, F., Durr, A. *et al.* (2008) Is the common LRRK2 G2019S mutation related to dyskinesias in North African Parkinson disease? *Neurology*, **71**, 1550–1552.
29. Altshuler, D., Daly, M.J. and Lander, E.S. (2008) Genetic mapping in human disease. *Science*, **322**, 881–888.
30. Pirkevi, C., Lesage, S., Condroyer, C., Tomiyama, H., Hattori, N., Ertan, S., Brice, A. and Basak, A.N. (2009) A LRRK2 G2019S mutation carrier from Turkey shares the Japanese haplotype. *Neurogenetics*, **10**, 271–273.
31. Adams, S.M., Bosch, E., Balaresque, P.L., Ballereau, S.J., Lee, A.C., Arroyo, E., Lopez-Parra, A.M., Aler, M., Grifo, M.S., Brion, M. *et al.* (2008) The genetic legacy of religious diversity and intolerance: paternal lineages of Christians, Jews, and Muslims in the Iberian Peninsula. *Am. J. Hum. Genet.*, **83**, 725–736.
32. Bar-Shira, A., Hutter, C.M., Giladi, N., Zabetian, C.P. and Orr-Urtreger, A. (2009) Ashkenazi Parkinson's disease patients with the LRRK2 G2019S mutation share a common founder dating from the second to fifth centuries. *Neurogenetics*, **10**, 355–358. Epub 2009 March 13.
33. Behar, D.M., Metspalu, E., Kivisild, T., Achilli, A., Hadid, Y., Tzur, S., Pereira, L., Amorim, A., Quintana-Murci, L., Majamaa, K. *et al.* (2006) The matrilineal ancestry of Ashkenazi Jewry: portrait of a recent founder event. *Am. J. Hum. Genet*, **78**, 487–497.
34. Behar, D.M., Metspalu, E., Kivisild, T., Rosset, S., Tzur, S., Hadid, Y., Yudkovsky, G., Rosengarten, D., Pereira, L., Amorim, A. *et al.* (2008) Counting the founders: the matrilineal genetic ancestry of the Jewish Diaspora. *PLoS One*, **3**, e2062.
35. Lesage, S., Ibanez, P., Lohmann, E., Pollak, P., Tison, F., Tazir, M., Leutenegger, A.L., Guimaraes, J., Bonnet, A.M., Agid, Y. *et al.* (2005) G2019S LRRK2 mutation in French and North African families with Parkinson's disease. *Ann. Neurol.*, **58**, 784–787.
36. Lohmann, E., Leclere, L., De Anna, F., Lesage, S., Dubois, B., Agid, Y., Dürr, A. and Brice, A. and French Parkinson's Disease Genetics Study Group (2009) A clinical, neuropsychological and olfactory evaluation of a large family with LRRK2 mutations. *Parkinsonism Relat. Disord.*, **15**, 273–276.
37. Lesage, S., Janin, S., Lohmann, E., Leutenegger, A.L., Leclere, L., Viallet, F., Pollak, P., Durif, F., Thobois, S., Layet, V. *et al.* (2007) LRRK2 exon 41 mutations in sporadic Parkinson disease in Europeans. *Arch. Neurol.*, **64**, 425–430.
38. Lesage, S., Leclere, L., Lohmann, E., Borg, M., Ruberg, M., Durr, A. and Brice, A. (2007) Frequency of the LRRK2 G2019S mutation in siblings with Parkinson's disease. *Neurodegener. Dis.*, **4**, 195–198.
39. Benson, G. (1999) Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res.*, **27**, 573–580.
40. Stephens, M. and Donnelly, P. (2003) A comparison of bayesian methods for haplotype reconstruction from population genotype data. *Am. J. Hum. Genet.*, **73**, 1162–1169.
41. Bandelt, H.J., Forster, P. and Rohl, A. (1999) Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.*, **16**, 37–48.
42. Austerlitz, F., Kalaydjieva, L. and Heyer, E. (2003) Detecting population growth, selection and inherited fertility from haplotypic data in humans. *Genetics*, **165**, 1579–1586.
43. Genin, E., Tullio-Pelet, A., Begeot, F., Lyonnet, S. and Abel, L. (2004) Estimating the age of rare disease mutations: the example of Triple-A syndrome. *J. Med. Genet.*, **41**, 445–449.
44. Frazer, K.A., Ballinger, D.G., Cox, D.R., Hinds, D.A., Stuve, L.L., Gibbs, R.A., Belmont, J.W., Boudreau, A., Hardenbol, P., Leal, S.M. *et al.* (2007) A second generation human haplotype map of over 3.1 million SNPs. *Nature*, **449**, 851–861.
45. Brinkmann, B., Klintschar, M., Neuhuber, F., Huhne, J. and Rolf, B. (1998) Mutation rate in human microsatellites: influence of the structure and length of the tandem repeat. *Am. J. Hum. Genet.*, **62**, 1408–1415.

Downloaded from http://hmg.oxfordjournals.org/ at Institut universitaire de hautes etudesinternationales - Bibliotheque on January 7, 2014