REPORT

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

Cyrus P. Zabetian, Carolyn M. Hutter, Dora Yearout, Alexis N. Lopez, Stewart A. Factor,* Alida Griffith, Berta C. Leis, Thomas D. Bird, John G. Nutt, Donald S. Higgins, John W. Roberts, Denise M. Kay, Karen L. Edwards, Ali Samii, and Haydeh Payami

The leucine-rich repeat kinase 2 (*LRRK2*) G2019S mutation is the most common genetic determinant of Parkinson disease (PD) identified to date. It accounts for 1%–7% of PD in patients of European origin and 20%–40% in Ashkenazi Jews and North African Arabs with PD. Previous studies concluded that patients from these populations all shared a common Middle Eastern founder who lived in the 13th century. We tested this hypothesis by genotyping 25 microsatellite and single-nucleotide–polymorphism markers in 22 families with G2019S and observed two distinct haplotypes. Haplotype 1 was present in 19 families of Ashkenazi Jewish and European ancestry, whereas haplotype 2 occurred in three European American families. Using a maximum-likelihood method, we estimated that the families with haplotype 1 shared a common ancestor 2,250 (95% confidence interval 1,650–3,120) years ago, whereas those with haplotype 2 appeared to share a more recent founder. Our data suggest two separate founding events for G2019S in these populations, beginning at a time that coincides with the Jewish Diasporas.

Parkinson disease (PD [MIM 168600]) is the second-most common neurodegenerative disorder and is characterized by bradykinesia, resting tremor, rigidity, and postural instability.[1,2] Although approximately one in five patients with PD reports a family history of the disease, pedigrees that demonstrate clear Mendelian inheritance are rare.[3,4] Linkage analysis of extended pedigrees has implicated five genes in autosomal recessive (*PARK2* [MIM 600116 and 602544], *PINK1* [MIM 605909], and *PARK7* [MIM 606324]) or dominant (*SNCA* [MIM 163890] and *LRRK2* [MIM 609007]) parkinsonism that closely resembles PD.[5–10] However, with the exception of *LRRK2*, mutations in these genes rarely result in typical late-onset PD. A single *LRRK2* mutation (G2019S) accounts for 0.7%–1.6% of sporadic and 2.8%–6.6% of familial PD cases of European origin.[11–16] More-recent studies indicate a remarkably high prevalence of G2019S in Ashkenazi Jews (13.3% sporadic; 29.7% familial) and North African Arabs (40.8% sporadic; 37.0% familial) with PD.[17,18] Patients heterozygous for G2019S display clinical characteristics that are indistinguishable from those of patients with PD in the general population.[11–18]

Data from five studies of 90 unrelated G2019S-bearing subjects of European or Middle Eastern–North African (MENA) origin revealed that all shared the same haplotype, consistent with a common founder.[11,16–19] The high prevalence of the mutation in Ashkenazi Jews and North African Arabs has led to the hypothesis that the mutation originated in the Middle East.[18] Lesage and colleagues[19] used a likelihood-based haplotype approach[20] in the study of six families of North African or European origin carrying G2019S and estimated that these individuals shared a common founder ~725 years ago.[19] Given the fact that G2019S is widely distributed across Europe and occurs at high frequency among Ashkenazi Jews and North African Arabs, this relatively recent estimated date is difficult to reconcile with established patterns of human migration.[21,22] Therefore, we collected genotype data on a larger sample of subjects, to examine the hypotheses that (1) patients with PD who are of European and MENA origin share a single ancestral haplotype and (2) the most recent common ancestor lived in the 13th century.

The study population was derived from 1,611 unrelated patients with PD and 1,647 control individuals enrolled at seven movement-disorder clinics in three U.S. cities (Albany, NY; Portland, OR; and Seattle). All patients met standardized diagnostic criteria for PD, as determined by a movement-disorder specialist.[23] Controls had no history of parkinsonism, and most were spouses of patients with PD. The institutional review boards at each participating

From the Geriatric Research Education and Clinical Center (C.P.Z.; D.Y.; A.N.L.; T.D.B.) and Parkinson's Disease Research Education and Clinical Center (A.S.), Veterans Affairs Puget Sound Health Care System, Department of Neurology, University of Washington School of Medicine (C.P.Z.; D.Y.; A.N.L.; T.D.B.; A.S.), Department of Epidemiology, University of Washington School of Public Health and Community Medicine (C.M.H.; K.L.E.), and Virginia Mason Medical Center (J.W.R.), Seattle; Parkinson's Disease and Movement Disorder Clinic, Albany Medical Center (S.A.F.; D.S.H.), and Genomics Institute, Wadsworth Center, New York State Department of Health (D.M.K.; H.P.), Albany; Booth Gardner Parkinson's Care Center, Evergreen Hospital Medical Center, Kirkland, WA (A.G.; B.C.L.); and Department of Neurology, Oregon Health and Science University, Portland (J.G.N.)

Received June 23, 2006; accepted for publication July 28, 2006; electronically published August 17, 2006.

Address for correspondence and reprints: Dr. Cyrus P. Zabetian, Geriatric Research Education and Clinical Center S-182, Veterans Affairs Puget Sound Health Care System, 1660 South Columbian Way, Seattle, WA 98108. E-mail: zabetian@u.washington.edu

* Present affiliation: Department of Neurology, Emory University School of Medicine, Atlanta.

Am. J. Hum. Genet. 2006;79:752–758. 0002-9297/2006/7904-0019$15.00

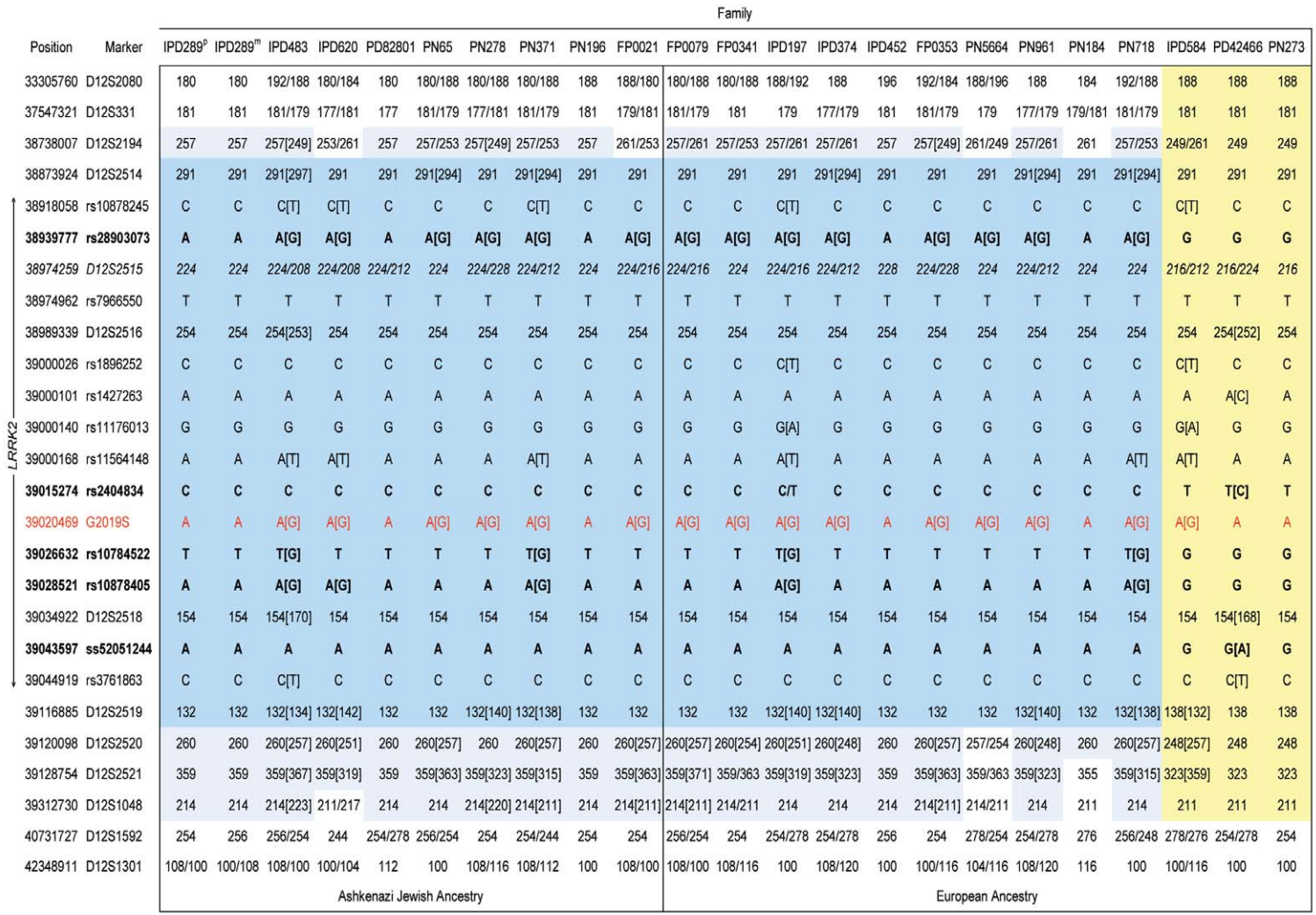
site approved the study, and all subjects gave informed consent. Among these subjects, we identified 21 patients (1 homozygous, 20 heterozygous) and 1 control individual with G2019S (19 of these subjects have been reported elsewhere).[12,24] For the present study, we included these 22 subjects, 23 of their family members (6 affected and 17 unaffected), and 30 unrelated G2019S-negative European American controls. Of the G2019S-positive families, 13 were of (non-Jewish) European origin, 3 defined themselves as Ashkenazi Jews, and 6 were of unspecified Jewish ancestry. However, the probands from those six families all indicated that their ancestors emigrated from central/eastern Europe or Russia and were thus likely descendants of the Ashkenazim.

We systematically selected markers, to determine whether our subjects carrying G2019S shared a common haplotype, using data from an ongoing study in which we are screening the entire *LRRK2* coding region and flanking intronic sequence for new mutations. Complete sequence data on all 51 exons were available for one subject homozygous and for three subjects heterozygous for G2019S. We compared these four individuals at 90 polymorphic sites and found five (*rs28903073, rs2404834, rs10784522, rs10878405,* and *ss52051244*) at which one subject (IPD584) differed from one or more of the others (IPD289, IPD452, and/or PD82801) at both alleles. We genotyped these five SNPs, 18 markers included in previous studies,[11,16,17,19,24] and two additional microsatellites (*D12S331* and *D12S1592*) in all 75 subjects in our study population. The markers (12 SNPs and 13 microsatellites) spanned 9 Mb across the *PARK8* region. SNP genotyping was performed by sequencing with the Applied Biosystems BigDye Terminator v3.1 Cycle Sequencing Kit. Microsatellites were amplified by PCR with use of fluorescently labeled F primers, were run on an ABI PRISM 3130 Genetic Analyzer, and were analyzed using GeneMapper 4.0 software (Applied Biosystems).

We estimated the age of the most recent common founder using the program Estiage.[20] This maximum-likelihood method uses information on the recombination fractions between the mutation and each marker, the frequencies of the shared allele at each marker, and the position of the first marker in each direction that is no longer shared, to calculate the number of generations (with 95% CI) elapsed since the most recent common ancestor introduced the mutation into the population. We defined a marker as shared among families if a single (best-call) allele was included in all disease haplotypes or in at least half and at significantly greater frequency (Fisher's exact test $\alpha = .05$) than in 60 inferred control haplotypes (collapsing all other alleles into a single bin). Genetic-map positions for each marker were derived from the linkage-mapping server MAP-O-MAT, and physical positions were taken from the National Center for Biotechnology Information (NCBI) human genome assembly Build 35.[25] Because many of the markers used in the present study were tightly linked, genetic distances between them could not be accurately estimated from available genetic maps. Instead, we used the genetic length and physical distance between the extreme markers (*D12S2080* and *D12S1301*) to calculate an average of 0.30 cM per Mb across the entire analyzed region and then computed recombination fractions using the Kosambi mapping function.

In instances where marker phase could not be unambiguously resolved by pedigree data, we used PHASE v2.1.1[26] (Matthew Stephens' Web site) to infer haplotypes. Since the use of inferred haplotypes can introduce additional uncertainty into age estimates, we first divided PHASE calls for each allele at a given marker into three categories by probability: best ($P > .5$), plausible ($.05 < P < .5$), and unlikely ($P < .05$). We then performed four separate runs, in Estiage, to estimate the age of the most recent founder by using haplotypes from only the families in which phase could be unambiguously resolved across shared markers or with the most likely (best call only), longest plausible, or shortest plausible haplotypes inferred by PHASE (Matthew Stephens' Web site) in all families.

Two clearly distinct disease haplotypes (referred to hereafter as "haplotype 1" and "haplotype 2") became evident on inspection of the data (fig. 1). Haplotype 1 included a core region of 17 consecutive markers spanning 243 kb (*D12S2514* to *D12S2519*) that was shared by 19 families, with the exception of one marker, *D12S2515*. Our group and others have argued that *D12S2515* is unstable and that the allelic differences previously observed among disease haplotypes at this marker were likely due to recurrent mutation rather than recombination.[11,24] Data from the current study further support this idea; all 19 families with haplotype 1 shared alleles at three markers upstream from *D12S2515,* including the A allele of *rs28903073,* which was observed elsewhere in only 1 of 906 European American control chromosomes (fig. 1).[24] This suggests that the A allele is identical by descent for haplotype 1 carriers and that any discordant alleles at intervening markers (between *rs28903073* and G2019S) in these individuals arose via new mutations. Thus, *D12S2515* was excluded from the data set used to reconstruct haplotypes and was not used in subsequent common founder-age estimates. Haplotype 2 was present in three families of European descent (fig. 1) and could be differentiated from the core region of haplotype 1 by five intragenic SNPs and one extragenic microsatellite (*D12S2519*). These included the two closest flanking markers, which were located 5 kb upstream (*rs2404834*) and 6 kb downstream (*rs10784522*) from G2019S. Haplotype 2 extended a minimum of 6 Mb across the *PARK8* region, from the 5′-most marker assayed (*D12S2080*) to *D12S1592*. Haplotypes 1 and 2 appear to be rare in populations of European ancestry, since the frequency of haplotype 1 can be no greater than that of the A allele at *rs28903073* (0.1%),[24] and haplotype 2 was not observed in our sample of 60 control chromosomes. Taken together, these data strongly suggest that the disease chromosomes in our study population originated from two separate founders.



	Position	Marker	IPD289[P]	IPD289[M]	IPD483	IPD620	PD82801	PN65	PN278	PN371	PN196	FP0021	FP0079	FP0341	IPD197	IPD374	IPD452	FP0353	PN5664	PN961	PN184	PN718	IPD584	PD42466	PN273
			Family																						
	33305760	D12S2080	180	180	192/188	180/184	180	180/188	180/188	180/188	188	188/180	180/188	180/188	188/192	188	196	192/184	188/196	188	184	192/188	188	188	188
	37547321	D12S331	181	181	181/179	177/181	177	181/179	177/181	181/179	181	179/181	181/179	181	179	177/179	181	181/179	179	177/179	179/181	181/179	181	181	181
	38738007	D12S2194	257	257	257[249]	253/261	257	257/253	257[249]	257/253	257	261/253	257/261	257/253	257/261	257/261	257	257[249]	261/249	257/261	261	257/253	249/261	249	249
	38873924	D12S2514	291	291	291[297]	291	291	291[294]	291	291[294]	291	291	291	291	291	291[294]	291	291	291	291[294]	291	291[294]	291	291	291
	38918058	rs10878245	C	C	C[T]	C[T]	C	C	C	C[T]	C	C	C	C	C[T]	C	C	C	C	C	C	C	C[T]	C	C
	38939777	**rs28903073**	**A**	**A**	**A[G]**	**A[G]**	**A**	**A[G]**	**A[G]**	**A[G]**	**A**	**A[G]**	**A[G]**	**A[G]**	**A[G]**	**A[G]**	**A**	**A[G]**	**A[G]**	**A[G]**	**A**	**A[G]**	**G**	**G**	**G**
	38974259	*D12S2515*	*224*	*224*	*224/208*	*224/208*	*224/212*	*224*	*224/228*	*224/212*	*224*	*224/216*	*224/216*	*224*	*224/216*	*224/212*	*228*	*224/228*	*224*	*224/212*	*224*	*224*	*216/212*	*216/224*	*216*
	38974962	rs7966550	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
	38989339	D12S2516	254	254	254[253]	254	254	254	254	254	254	254	254	254	254	254	254	254	254	254	254	254	254	254[252]	254
	39000026	rs1896252	C	C	C	C	C	C	C	C	C	C	C	C	C[T]	C	C	C	C	C	C	C	C[T]	C	C
	39000101	rs1427263	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A[C]	A
LRRK2	39000140	rs11176013	G	G	G	G	G	G	G	G	G	G	G	G	G[A]	G	G	G	G	G	G	G	G[A]	G	G
	39000168	rs11564148	A	A	A[T]	A[T]	A	A	A	A[T]	A	A	A	A	A[T]	A	A	A	A	A	A	A[T]	A[T]	A	A
	39015274	**rs2404834**	**C**	**C**	**C**	**C**	**C**	**C**	**C**	**C**	**C**	**C**	**C**	**C**	**C/T**	**C**	**C**	**C**	**C**	**C**	**C**	**C**	**T**	**T[C]**	**T**
	39020469	G2019S	A	A	A[G]	A[G]	A	A[G]	A[G]	A[G]	A	A[G]	A[G]	A[G]	A[G]	A[G]	A	A[G]	A[G]	A[G]	A	A[G]	A[G]	A	A
	39026632	**rs10784522**	**T**	**T**	**T[G]**	**T**	**T**	**T**	**T**	**T[G]**	**T**	**T**	**T**	**T**	**T[G]**	**T**	**T**	**T**	**T**	**T**	**T**	**T[G]**	**G**	**G**	**G**
	39028521	**rs10878405**	**A**	**A**	**A[G]**	**A[G]**	**A**	**A**	**A**	**A[G]**	**A**	**A**	**A**	**A**	**A[G]**	**A**	**A**	**A**	**A**	**A**	**A**	**A[G]**	**G**	**G**	**G**
	39034922	D12S2518	154	154	154[170]	154	154	154	154	154	154	154	154	154	154	154	154	154	154	154	154	154	154	154[168]	154
	39043597	**ss52051244**	**A**	**A**	**A**	**A**	**A**	**A**	**A**	**A**	**A**	**A**	**A**	**A**	**A**	**A**	**A**	**A**	**A**	**A**	**A**	**A**	**G**	**G[A]**	**G**
	39044919	rs3761863	C	C	C[T]	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C[T]	C
	39116885	D12S2519	132	132	132[134]	132[142]	132	132	132[140]	132[138]	132	132	132	132	132[140]	132[140]	132	132	132	132[140]	132	132[138]	138[132]	138	138
	39120098	D12S2520	260	260	260[257]	260[251]	260	260[257]	260	260[257]	260	260[257]	260[257]	260[254]	260[251]	260[248]	260	260[257]	257/254	260[248]	260	260[257]	248[257]	248	248
	39128754	D12S2521	359	359	359[367]	359[319]	359	359[363]	359[323]	359[315]	359	359[363]	359[371]	359/363	359[319]	359[323]	359	359[363]	359/363	359[323]	355	359[315]	323[359]	323	323
	39312730	D12S1048	214	214	214[223]	211/217	214	214	214[220]	214[211]	214	214[211]	214[211]	214/211	214	214	214	214[211]	214/211	214	211	214	211	211	211
	40731727	D12S1592	254	256	256/254	244	254/278	256/254	254	254/244	254	254	256/254	254	254/278	254/278	256	254	278/254	254/278	276	256/248	278/276	254/278	254
	42348911	D12S1301	108/100	100/108	108/100	100/104	112	100	108/116	108/112	100	108/100	108/100	108/116	100	108/120	100	100/116	104/116	108/120	116	100	100/116	100	100
			Ashkenazi Jewish Ancestry										European Ancestry												

Figure 1. Comparison of G2019S-containing haplotypes across the *LRRK2* region. Five intragenic markers (*bold type*) delineate two distinct haplotypes in which complete allele sharing is indicated by blue (haplotype 1) or yellow (haplotype 2) shading. For haplotype 1, significant but incomplete allele sharing is denoted by light-blue shading. For markers for which phase could not be unambiguously determined, both alleles are shown with the best call, as inferred by PHASE (Matthew Stephens' Web site) listed first (*alleles of probability <.05 are in brackets*). The proband of family IPD289 was homozygous for G2019S; thus, haplotypes for the paternal ([P]) and maternal ([M]) lineages are provided separately. *D12S2515* (*shown in italics*) is unstable and was not included in haplotype analyses.

We then estimated the age of the most recent common founder among the 19 families with haplotype 1, designating *D12S2194* and *D12S1048* as the boundaries of the shared haplotype as determined by the aforementioned criteria (table 1). Our first approach, which was the one most comparable to that used by Lesage and colleagues,[19] restricted the analysis to families with phase-known data within the shared region. This reduced the data set to five families (IPD289, IPD452, PD82801, PN184, and PN196) representing six disease chromosomes and produced an age estimate of 65 (95% CI 36–121) generations. After incorporating the remaining 14 families for which phase was inferred, the best-call and longest haplotypes were identical and yielded an estimate of 75 (95% CI 55–104) generations, whereas the shortest haplotypes estimate was 84 (95% CI 61–117) generations. With use of the most widely published intergenerational interval, 25 years,[27] our best-call data indicate that these families shared a common ancestor 1,875 (95% CI 1,375–2,600) years ago. However, several recent studies have concluded that 30 years is a better approximation of generation time for the period of human history in question.[27–29] A 30-year interval with best-call haplotypes increases the age of the founding event to 2,250 (95% CI 1,650–3,120) years ago. We did not perform age estimates for the three families with haplotype 2, because the sample size was small and we were not able to define the 5′ limit of the shared region. However, since the genetic length of the shared region is inversely proportional to the age of the founding event, these three families likely shared a common ancestor more recently than did the haplotype 1 families.

Our study is the first to provide evidence that G2019S-positive patients of European ancestry originated from more than one founder. In our sample, haplotypes 1 and 2 could be adequately distinguished only by inclusion of five intragenic SNPs added on the basis of extensive sequence comparisons performed on a subset of probands. Only one of these five SNPs (*rs28903073*) has been included in similar analyses by other groups,[11] so it is possible that both haplotypes were present among the 90 G2019S-bearing families of European or MENA origin reported elsewhere.[11,16–19] However, we suspect that the majority of these families will prove to carry haplotype 1, on the basis of allele sharing at four extragenic microsatellites (*D12S2519, D12S2520, D12S2521,* and *D12S1048*) assayed in four of five studies[11,16,17,19] and at *rs28903073,* which was genotyped in a single study of 18 (largely Italian) families.[11] This is consistent with our data and suggests that haplotype 1 is predominant in European Americans and Ashkenazi Jews—and likely in North African Arabs as well—carrying G2019S, which would be explained by a common Middle Eastern founder. The geographic origin of haplotype 2 is less certain, but all three families in which it occurred originated primarily from Western Europe.

Our age estimates for the most-recent common founder for G2019S (haplotype 1) were two- to threefold larger than the one calculated by Lesage and colleagues.[19] There were three potential differences of relevance between the two studies: the physical distance shared among disease chromosomes, our use of inferred haplotypes, and the estimation of recombination fraction between markers. We observed a minimum shared region of 243 kb flanking G2019S in our sample, whereas Lesage et al.[19] reported a much shorter segment, 60 kb. We suspect that this was due largely to their using an unstable marker (*D12S2515*) to delineate the 5′ boundary of the shared segment and possibly unknowingly combining subjects with haplotypes 1 and 2 in the analysis. Slight differences in the marker sets genotyped in the two studies might also have had an effect. However, this would not explain the difference in age estimates, since our inclusion of longer disease haplotypes would have the opposite effect (i.e., predict a more recent ancestor). Likewise, our use of inferred haplotypes in some families does not explain the discrepancy, because, even when the analysis was restricted to chromosomes of known phase, as was done in the study of Lesage et al.,[19] our estimate was still more than twice as large (65 vs. 29 generations). Furthermore, our inclusion of reconstructed haplotypes in 14 of 19 families allowed more-efficient use of the data set, with minimal added uncertainty in the estimate, as evidenced by the narrow range of ages calculated using the longest or shortest plausible haplotypes (75 vs. 84 generations). The only remaining explanation to account for the discrepancy between the two studies is that we must have used lower recombination fractions in the maximum-likelihood calculations (table 1). Lower recombination fractions correspond to a decreased genetic length of the segment shared among disease chromosomes and an increased estimated time since the last common ancestor. This general relationship is apparent when our estimate of 75 generations, based on a shared segment of <0.2 cM, is compared with two previous studies that used Estiage to estimate a founder age of 46 generations (0.7–12.6 cM shared) for the *AAAS* (MIM 605378) IVS14+1G→A mutation and 23 generations (7 cM shared) for the *VAPB* P56S mutation (MIM 605704 and 608627).[20,30] In contrast, Lesage and colleagues[19] calculated an age of only 29 generations, using a shared physical distance less than what we used. We cannot determine the degree to which our estimated recombination fractions differed from theirs, because they have not published these data, although the genetic distances used in both studies were derived from the same source (MAP-O-MAT).

Our data are consistent with the hypothesis of Ozelius and colleagues[18] that Europeans, Ashkenazi Jews, and North African Arabs with G2019S arose from a common Middle Eastern founder, if only those individuals with haplotype 1 are considered. Our best estimate suggests that the common ancestor lived 2,250 years ago, during the period of the Jewish Diasporas (586 BC to 70 AD), at a time when the ancestral Jewish population and some Arab communities existed in close proximity. Similar age estimates have been calculated for other disease alleles found at high fre-

Table 1. Data Used in Estimating the Age of the G2019S Haplotype 1 Common Founder

	Data Used for Family																						
Marker	IPD289(p)[a]	IPD289(m)[a]	IPD483	PN65	PN371	PN196	FP0079	FP0341	IPD452	FP0353	PN718	PD82801	PN278	IPD374	PN961	IPD197	FP0021	PN184	PN5664	IPD620	Distance[b]	θ[c]	Frequency[d]
D12S331	181	181	181	181	181	181	181	181	181	181	181	177	177	177	177	179					1.473	.0045	
D12S2194	257	257	257	257	257	257	257	257	257	257	257	257	257	257	257	257	261	261	261	253	.282	.0009	.20
D12S2514	291	291	291	291	291	291	291	291	291	291	291	291	291	291	291	291	291	291	291	291	.147	.0004	.53
rs10878245	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	.102	.0003	.60
rs28903073	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	.081	.0002	.001
rs7966550	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	.046	.0001	.85
D12S2516	254	254	254	254	254	254	254	254	254	254	254	254	254	254	254	254	254	254	254	254	.031	.00009	.63
rs1896252	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	.020	.00006	.62
rs1427263	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	.020	.00006	.68
rs11176013	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	.020	.00006	.62
rs11564148	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	.020	.00006	.39
rs2404834	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	.005	.00002	.91
G2019S	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	0	0	
rs10784522	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	.006	.00002	.40
rs10878405	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	.008	.00002	.40
D12S2518	154	154	154	154	154	154	154	154	154	154	154	154	154	154	154	154	154	154	154	154	.014	.00004	.80
ss52051244	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	.023	.00007	.88
rs3761863	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	.024	.00007	.67
D12S2519	132	132	132	132	132	132	132	132	132	132	132	132	132	132	132	132	132	132	132	132	.096	.0003	.30
D12S2520	260	260	260	260	260	260	260	260	260	260	260	260	260	260	260	260	260	260	257	260	.100	.0003	.20
D12S2521	359	359	359	359	359	359	359	359	359	359	359	359	359	359	359	359	359	355		359	.108	.0003	.08
D12S1048	214	214	214	214	214	214	214	214	214	214	214	214	214	214	214	214	214			211	.292	.0009	.21
D12S1592	254	256	256	256	254	254	256	254	256	254	256	254	254	254	254	254	254				1.711	.0052	

NOTE.—The interval between markers *D12S2194* and *D12S1048* was defined as shared and, within this segment phase, was inferred for 14 of 20 disease chromosomes. Marker *D12S2515* is unstable and was excluded from the analysis.

[a] The proband of family IPD289 was homozygous for G2019S; thus, haplotypes from the paternal (p) and maternal (m) lineages are provided separately.

[b] Physical distance (Mb) between G2019S and each marker, from NCBI Build 35.

[c] Recombination fraction between G2019S and each marker, calculated using the Kosambi mapping function and an average of 0.30 cM per Mb across the entire region.

[d] Frequencies of the shared allele were taken from Kachergus et al.[16] for microsatellites and 30 International HapMap Project CEPH trios for SNP markers, except for *rs28903073* (which was from a previous report[24]) and *ss52051244* (which was derived from 60 European American control chromosomes genotyped in this study).

quency among Ashkenazi Jews, non-Ashkenazi Jews, and Arabs, including the factor XI deficiency type II *F11* (MIM 264900) E117X mutation and the *APC* (MIM 175100) I1307K variant, which increases risk for colorectal cancer.[22,31] Thus, Lesage and colleagues' conclusion that the G2019S founding event occurred in the 13th century[19] appears implausible, since, by that time, Jews destined to become Ashkenazim had been separated from Arab and other Jewish communities for many centuries.

Our data provide an important framework for further investigation of the widespread expansion of G2019S over the past two millennia. Mapping the frequency of disease haplotypes 1 and 2 in European and MENA populations, with use of the marker set described above, will be critical in this endeavor, particularly in non–Ashkenazi Jewish and non-Jewish communities in the Middle East. Defining the age and origin of haplotype 2 might help to explain the variable prevalence of the mutation across Europe. Finally, G2019S was recently identified in an Asian population[24] on a haplotype background clearly distinct from the two observed here, which indicates that at least three founding events have occurred. It is becoming clear that the genetic and demographic processes that have shaped the current distribution of G2019S across the world are more complex than previously recognized.

Acknowledgments

We thank the individuals who participated in the study, and we thank Stephen Ayres, Sharon Evans, Erica Martinez, and Galen Richards, for technical support and assistance with subject recruitment. This work was supported by National Institutes of Health grants K08 NS044138, P30 AG008017, and R01 NS036960; the Michael J. Fox Foundation, Edmond J. Safra Global Genetics Consortia Initiative; a Department of Veterans Affairs Merit Review Award; American Parkinson Disease Association; and the Veterans Integrated Service Network 20 Geriatric Research Education and Clinical Center, Mental Illness Research Education and Clinical Center, and Parkinson's Disease Research Education and Clinical Centers.

Web Resources

The URLs for data presented herein are as follows:

International HapMap Project, http://www.hapmap.org/
MAP-O-MAT, http://compgen.rutgers.edu/mapomat/
Matthew Stephens' Web site, http://www.stat.washington.edu/stephens/software.html (for PHASE software)
NCBI, http://www.ncbi.nlm.nih.gov/ (for the human genome assembly)
Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/ (for PD, *PARK2, PINK1, PARK7, SNCA, LRRK2, AAAS, VAPB, F11,* and *APC*)

References

1. Nussbaum RL, Ellis CE (2003) Alzheimer's disease and Parkinson's disease. N Engl J Med 348:1356–1364
2. Samii A, Nutt JG, Ransom BR (2004) Parkinson's disease. Lancet 363:1783–1793
3. Payami H, Larsen K, Bernard S, Nutt J (1994) Increased risk of Parkinson's disease in parents and siblings of patients. Ann Neurol 36:659–661
4. Bonifati V, Fabrizio E, Vanacore N, De Mari M, Meco G (1995) Familial Parkinson's disease: a clinical genetic analysis. Can J Neurol Sci 22:272–279
5. Kitada T, Asakawa S, Hattori N, Matsumine H, Yamamura Y, Minoshima S, Yokochi M, Mizuno Y, Shimizu N (1998) Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism. Nature 392:605–608
6. Bonifati V, Rizzu P, van Baren MJ, Schaap O, Breedveld GJ, Krieger E, Dekker MC, Squitieri F, Ibanez P, Joosse M, van Dongen JW, Vanacore N, van Swieten JC, Brice A, Meco G, van Duijn CM, Oostra BA, Heutink P (2003) Mutations in the *DJ-1* gene associated with autosomal recessive early-onset parkinsonism. Science 299:256–259
7. Valente EM, Abou-Sleiman PM, Caputo V, Muqit MM, Harvey K, Gispert S, Ali Z, Del Turco D, Bentivoglio AR, Healy DG, Albanese A, Nussbaum R, Gonzalez-Maldonado R, Deller T, Salvi S, Cortelli P, Gilks WP, Latchman DS, Harvey RJ, Dallapiccola B, Auburger G, Wood NW (2004) Hereditary early-onset Parkinson's disease caused by mutations in *PINK1*. Science 304:1158–1160
8. Polymeropoulos MH, Lavedan C, Leroy E, Ide SE, Dehejia A, Dutra A, Pike B, Root H, Rubenstein J, Boyer R, Stenroos ES, Chandrasekharappa S, Athanassiadou A, Papapetropoulos T, Johnson WG, Lazzarini AM, Duvoisin RC, Di Iorio G, Golbe LI, Nussbaum RL (1997) Mutation in the α-synuclein gene identified in families with Parkinson's disease. Science 276: 2045–2047
9. Paisan-Ruiz C, Jain S, Evans EW, Gilks WP, Simon J, van der Brug M, de Munain AL, Aparicio S, Gil AM, Khan N, Johnson J, Martinez JR, Nicholl D, Carrera IM, Pena AS, de Silva R, Lees A, Marti-Masso JF, Perez-Tur J, Wood NW, Singleton AB (2004) Cloning of the gene containing mutations that cause *PARK8*-linked Parkinson's disease. Neuron 44:595–600
10. Zimprich A, Biskup S, Leitner P, Lichtner P, Farrer M, Lincoln S, Kachergus J, Hulihan M, Uitti RJ, Calne DB, Stoessl AJ, Pfeiffer RF, Patenge N, Carbajal IC, Vieregge P, Asmus F, Muller-Myhsok B, Dickson DW, Meitinger T, Strom TM, Wszolek ZK, Gasser T (2004) Mutations in *LRRK2* cause autosomal-dominant parkinsonism with pleomorphic pathology. Neuron 44:601–607
11. Goldwurm S, Di Fonzo A, Simons EJ, Rohe CF, Zini M, Canesi M, Tesei S, 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. Kay DM, Zabetian CP, Factor SA, Nutt JG, Samii A, Griffith A, Bird TD, Kramer P, Higgins DS, Payami H (2006) Parkinson's disease and *LRRK2*: frequency of a common mutation in U.S. movement disorder clinics. Mov Disord 21:519–523
13. Gilks WP, Abou-Sleiman PM, Gandhi S, Jain S, Singleton A, Lees AJ, Shaw K, Bhatia KP, Bonifati V, Quinn NP, Lynch J, Healy DG, Holton JL, Revesz T, Wood NW (2005) A common *LRRK2* mutation in idiopathic Parkinson's disease. Lancet 365:415–416
14. Nichols WC, Pankratz N, Hernandez D, Paisan-Ruiz C, Jain S, Halter CA, Michaels VE, Reed T, Rudolph A, Shults CW, Singleton A, Foroud T (2005) Genetic screening for a single common *LRRK2* mutation in familial Parkinson's disease. Lancet 365:410–412

15. Di Fonzo A, Rohe CF, Ferreira J, Chien HF, Vacca L, Stocchi F, Guedes L, Fabrizio E, Manfredi M, Vanacore N, Goldwurm S, Breedveld G, Sampaio C, Meco G, Barbosa E, Oostra BA, Bonifati V (2005) A frequent *LRRK2* gene mutation associated with autosomal dominant Parkinson's disease. Lancet 365: 412–415
16. Kachergus J, Mata IF, Hulihan M, Taylor JP, Lincoln S, Aasly J, Gibson JM, Ross OA, Lynch T, Wiley J, Payami H, Nutt J, Maraganore DM, Czyzewski K, Styczynska M, Wszolek ZK, Farrer MJ, Toft M (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
17. Lesage S, Durr A, Tazir M, Lohmann E, Leutenegger AL, Janin S, Pollak P, Brice A (2006) *LRRK2* G2019S as a cause of Parkinson's disease in North African Arabs. N Engl J Med 354: 422–423
18. Ozelius LJ, Senthil G, Saunders-Pullman R, Ohmann E, Deligtisch A, Tagliati M, Hunt AL, Klein C, Henick B, Hailpern SM, Lipton RB, Soto-Valencia J, Risch N, Bressman SB (2006) *LRRK2* G2019S as a cause of Parkinson's disease in Ashkenazi Jews. N Engl J Med 354:424–425
19. Lesage S, Leutenegger A-L, Ibanez P, Janin S, Lohmann E, Dürr A, Brice A, French Parkinson's Disease Genetics Study Group (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
20. Genin E, Tullio-Pelet A, Begeot F, Lyonnet S, Abel L (2004) Estimating the age of rare disease mutations: the example of triple-A syndrome. J Med Genet 41:445–449
21. Colombo R (2000) Age estimate of the N370S mutation causing Gaucher disease in Ashkenazi Jews and European populations: a reappraisal of haplotype data. Am J Hum Genet 66: 692–697
22. Niell BL, Long JC, Rennert G, Gruber SB (2003) Genetic anthropology of the colorectal cancer-susceptibility allele *APC* I1307K: evidence of genetic drift within the Ashkenazim. Am J Hum Genet 73:1250–1260
23. Gibb WR, Lees AJ (1988) The relevance of the Lewy body to the pathogenesis of idiopathic Parkinson's disease. J Neurol Neurosurg Psychiatry 51:745–752
24. Zabetian CP, Morino H, Ujike H, Yamamoto M, Oda M, Maruyama H, Izumi Y, Kaji R, Griffith A, Leis BC, Roberts JW, Yearout D, Samii A, Kawakami H (2006) Identification and haplotype analysis of *LRRK2* G2019S in Japanese patients with Parkinson disease. Neurology (http://www.neurology.org/cgi/content/abstract/01.wnl.0000227732.37801.d4v1) (electronically published May 25, 2006) (accessed August 11, 2006)
25. Kong X, Matise TC (2005) MAP-O-MAT: internet-based linkage mapping. Bioinformatics 21:557–559
26. Stephens M, Scheet P (2005) Accounting for decay of linkage disequilibrium in haplotype inference and missing-data imputation. Am J Hum Genet 76:449–462
27. Fenner JN (2005) Cross-cultural estimation of the human generation interval for use in genetics-based population divergence studies. Am J Phys Anthropol 128:415–423
28. Tremblay M, Vézina H (2000) New estimates of intergenerational time intervals for the calculation of age and origins of mutations. Am J Hum Genet 66:651–658
29. Sigurðardóttir S, Helgason A, Gulcher JR, Stefansson K, Donnelly P (2000) The mutation rate in the human mtDNA control region. Am J Hum Genet 66:1599–1609
30. Nishimura AL, Al-Chalabi A, Zatz M (2005) A common founder for amyotrophic lateral sclerosis type 8 (ALS8) in the Brazilian population. Hum Genet 118:499–500
31. Goldstein DB, Reich DE, Bradman N, Usher S, Seligsohn U, Peretz H (1999) Age estimates of two common mutations causing factor XI deficiency: recent genetic drift is not necessary for elevated disease incidence among Ashkenazi Jews. Am J Hum Genet 64:1071–1075