Am. J. Hum. Genet. 67:528, 2000

Age and Origin of the *PRNP* E200K Mutation Causing Familial Creutzfeldt-Jacob Disease in Libyan Jews

To the Editor:

Creutzfeldt-Jacob disease (CJD [MIM 123400]), the most prevalent of the human spongiform encephalopathies (HSEs), is a rapidly progressive neurodegenerative disease that manifests itself as a sporadic, transmissible, or familial disorder (Johnson and Gibbs 1998). Patients with CJD generally develop neurological dysfunction in midlife and die within 6–24 mo of onset. The largest cluster of CJD occurs among Libyan Jews, where the incidence (1/10,000) is ~100 times higher than incidence worldwide. The origin of the higher incidence of CJD in this population is an intriguing problem that has not yet been resolved.

Although a few studies (reviewed in Meiner et al. 1997) had pointed out the familial predisposition to CJD, it was first speculated that its frequency in Libyan Jews could be explained by their habit of consuming lightly grilled sheep's brains or eyeballs (Herzberg et al. 1974; Alter and Kahana 1976), reflecting a shared environmental risk (exposure to scrapie-infected meat) rather than any genetic factor. This hypothesis, based on the unrealistic assumption that scrapie was widespread in Libya and that a marked culinary difference between Jews of Libyan and other North African origins existed, suffered a strong blow when two cases of CJD among young Jews born in Israel to immigrants from Libya were discovered (Nisipeanu et al. 1990; Hsiao et al. 1991). They could not have been exposed to scrapie, since it does not exist in Israel. The most likely explanation soon became that CJD could be inherited genetically in a pattern similar to that of another HSE, Gerstmann-Sträussler-Scheinker disease (MIM 137440). Genetic linkage between a missense mutation (E200K: 598A→G; 200Glu→Lys) in the prion protein (*PRNP*) gene and CJD was established in Libyan Jews (LOD score >4.85 [Gabizon et al. 1993]). The E200K mutation, which accounts for >70% of cases of familial CJD, was first identified in a Polish family and subsequently in patients living in England, France, Austria, Slovakia, Chile, the United States, and Japan. The analysis of its geographic distribution (Goldfarb et al. 1991) suggested that the E200K mutation originated in Spain >5 centuries ago, possibly in a Jewish person, and spread to Mediterranean and continental countries after the expulsion of Sephardic Jews from Spain. The hypothesis of propagation through a limited number of successful migrants was supported by the discovery of a higher frequency, in patients with CJD and their unaffected relatives, compared with the general Libyan Jewish population, of *PRNP*'s 129M polymorphism in the normal allele. However, the same hypothesis was disputed (Gabizon et al. 1993; Korczyn 1994) on the basis of lack of evidence for the presence of the E200K mutation in Spain and among Jews living in other countries to which Sephardim have emigrated and because it seems to disregard the intermarriage rules followed by the local Jewish communities.

Recently, the E200K mutation has been discovered in other European countries—that is, in Italy, as well as in Spain itself. More decisively, Lee et al. (1999) reported in the *Journal* that Libyan, Tunisian, Italian (continental), Chilean, and Spanish families with CJD share a major haplotype on chromosome 20p12-pter, to which the *PRNP* gene has been mapped, whereas families with CJD from Germany, Sicily, Austria, and Japan bear different corresponding haplotypes, suggesting independent mutational events. The prominence of this study for the reinstatement of the "founder effect" hypothesis to explain the Libyan focus of familial CJD cannot be dismissed. However, the probational strength of haplotype data presented by Lee et al. (1999) can be even more convincing if they are quantitatively analyzed for linkage disequilibrium (LD) decay over time and the results compared with the Libyan Jewish population's history. To perform this, two different methods were used, both of which are based on the genetic clock equation $\ln P = -\theta g$, relating the time (in generations, g) tracing back to the most recent common ancestor of mutant chromosomes, the frequency of recombination between the disease locus and the marker (θ), and the probability that a marker's allele on a disease chromosome is the ancestral one (P). Taking the frequencies on E200K-bearing (p_d) and normal (p_n) Libyan Jewish chromosomes for the alleles of three microsatellite markers flanking the *PRNP* locus and defining the putative common ancestral haplotype (D20S116–20, D20S482–14, and D20S895–18/19; table 2 of Lee et al. 1999), the LD measure, $\delta = (p_d - p_n)/(1 - p_n)$, was calculated according to Bengtsson and Thomson (1981). Recombination fraction values (θ) were conveniently estimated from chromosome-20 physical mapping (LDB and the Sanger Centre) under the assumption of a genetic to physical distance ratio of 3.7 obtained from regression of centimorgan versus megabyte values for 20p12-pter loci. Applying the algorithm of Risch et al. (1995), $g = \log\delta/\log(1 - \theta)$, the estimated age (in generations, g) of the E200K mutation is 16.8 ± 2.0 SD (95% confidence interval [CI] 11.7–21.8 g) (table 1). Using the iterative method of Reich and Goldstein (1999), which models the regeneration of the ancestral haplotype by the recombination and mutation process through a Markov transition matrix that gives the probabilities that any one haplotype will be transformed into any other one in a single generation, a somewhat lower age was ob-

tained (14 ± 2 SD; 95% CI 9–19 *g*; table 2). When supplemented by setting the genetic clock according to the Luria-Delbrück approach (Labuda et al. 1996), under the assumption of a mean population growth rate of .5, the estimate rises to 22.7 ± 2.1 SD (95% CI 17.4–27.9 *g*). Given an intergenerational interval of 30 years (Tremblay and Vézina 2000) and an average year of birth of 1950 for all affected and carrier subjects, the present results would date the most recent common ancestor bearing the E200K mutation back to 1450–1530 or to the second half of the 13th century. This dating points to the origin of CJD in Libyan Jews at the time, or before, that Jewish families of Iberian origin settled in Libya after their expulsion from Spain in 1492 and from Portugal in 1497 (Barnavi 1992). Intrafamily marriages were a common practice among them (Gabizon et al. 1993), and a constant growth of this isolated population in the centuries after immigration has been documented (see, e.g., Goldberg 1971). However, other events, such as immigration of Jews from the neighboring island of Jerba at approximately the same time (Udovitch and Valensi 1984) or bottleneck effects, should be taken into consideration. Despite the methodological limits associated with LD-based allele age estimation, persuasive further evidence for the hypothesis of a "Spanish founder effect" in Libyan Jewish CJD genetic epidemiology can be drawn from the analysis of the haplotype data reported by Lee et al. (1999).

ROBERTO COLOMBO

Human
Biology and Genetics Research Unit
Department of Psychology
Catholic University of the Sacred Heart
Milan

Electronic-Database Information

Accession numbers and URLs for data in this article are as follows:

LDB, the Genetic Location Database, Department of Human Genetics, University of Southampton, UK, http://cedar.genetics.soton.ac.uk/public_html/ldb.html

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim (for CJD [MIM 123400] and Gerstmann-Sträussler-Scheinker disease [MIM 137440])

Sanger Centre, The, http://webace.sanger.ac.uk/cgi-bin/ace/simple/

References

Alter M, Kahana E (1976) Creutzfeldt-Jacob disease among Libyan Jews in Israel. Science 192:428

Barnavi E (ed) (1992) A historical atlas of the Jewish people. Schocken Books, New York

Bengtsson BO, Thomson G (1981) Measuring the strength of associations between HLA antigens and diseases. Tissue Antigens 18:356–363

Gabizon R, Rosenmann H, Meiner Z, Kahana E, Shugart Y, Ott J, Prusiner SB (1993) Mutation and polymorphism of the prion protein gene in Libyan Jews with Creutzfeldt-Jacob disease (CJD). Am J Hum Genet 53:828–835

Goldberg H (1971) Ecologic and demographic aspects of rural Tripolitanian Jewry. Int J Middle East Stud 2:245–265

Goldfarb LG, Brown P, Mitrova E, Cervenakova L, Goldin L, Korczyn AD, Chapman J, et al (1991) Creutzfeldt-Jacob disease associated with the PRNP codon 200Lys mutation: an analysis of 45 families. Eur J Epidemiol 7:477–486

Goldstein DB, Pollock DD (1997) Launching microsatellites: a review of mutation processes and methods in phylogenetic interference. J Hered 88:335–342

Herzberg L, Herzberg BN, Gibbs CJ Jr, Sullivan W, Amyx H, Gajdusek DC (1974) Creutzfeldt-Jacob disease: hypothesis for the high incidence in Libyan Jews in Israel. Science 186: 848

Hsiao K, Meiner Z, Kahana E, Cass C, Kahana I, Avrahami D, Scarlato G, et al (1991) Mutation of the prion protein in Libyan Jews with Creutzfeldt-Jacob disease. N Engl J Med 324:1091–1097

Johnson RT, Gibbs CJ (1998) Creutzfeldt-Jacob disease and related transmissible spongiform encephalopathies. N Engl J Med 339:1994–2004

Korczyn AD (1994) Neurologic genetic diseases of Jewish people. Biomed Pharmacother 48:391–397

Labuda M, Labuda D, Korab-Laskowska M, Cole DEC, Zietkiewicz E, Weissenbach J, Popowska E, et al (1996) Linkage disequilibrium analysis in young populations: pseudo-vitamin D–deficiency rickets and the founder effect in French Canadians. Am J Hum Genet 59:633–643

Lee HS, Sambuughin N, Cervenakova L, Chapman J, Pocchiari M, Litvak S, Qi HY, et al (1999) Ancestral origins and worldwide distribution of the *PRNP* 200K mutation causing familial Creutzfeldt-Jacob disease. Am J Hum Genet 64: 1063–1070

Meiner Z, Gabizon R, Prusiner SB (1997) Familial Creutzfeldt-Jacob disease: codon 200 prion disease in Libyan Jews. Medicine 76:227–237

Nisipeanu P, El Ad B, Korczyn AD (1990) Spongiform encephalopathy in an Israeli born to immigrants from Libya. Lancet 336:686

Reich DE, Goldstein DB (1999) Estimating the age of mutations using variation at linked markers. In: Goldstein DB, Schlötterer C (eds) Microsatellites: evolution and applications. Oxford University Press, Oxford, pp 129–138

Risch N, de Leon D, Ozelius L, Kramer P, Almasy L, Singer B, Fahn S, et al (1995). Genetic analysis of idiopathic torsion dystonia in Ashkenazi Jews and their recent descent from a small founder population. Nat Genet 9:152–159

Tremblay M, Vézina H (2000) New estimates of intergenerational time intervals for the calculation of age of origins of mutations. Am J Hum Genet 66:651–658

Udovitch AL, Valensi L (1984) The last Arab Jews: the communities of Jerba, Tunisia. Harwood Academic Press, New York/London, p 178

Weber JL, Wong C (1003) Mutation in human short tandem repeats. Hum Mol Genet 2:1123–1128

Address for correspondence and reprints: Dr. Roberto Colombo, Human Biology and Genetics Research Unit, Department of Psychology, Catholic University of the Sacred Heart, Largo A. Gemelli 1, I-20123 Milan, Italy. E-mail: rcolombo@aixmiced.mi.unicatt.it

0002-9297/2000/6702-0035$02.00

Table 1

Estimation of the Age of the E200K Mutation in the *PRNP* Gene, Causing CJD in Libyan Jews (Risch et al.'s [1995] Method, without and with the Luria-Delbrück Correction for Population Growth Rate [Labuda et al. 1996])

	Distance from *PRNP* locus[a]				Haplotype data[b]		LD[c]	Estimated age[d]		
Marker	Mb	cM	θ	Allele	p_d	p_n	(δ)	g	g_0	g_c
D20S116	.57	2.11	.0211	20	.81	.29	.73	14.6	6.3	20.9
D20S482	.27	1	.001	14	.93	.59	.83	18.6	12.4	31
D20S895	.35	1.3	.013	18+19	.87	.35	.8	17.1	7.3	24.4

[a] Genetic distances (in centimorgans) were calculated using a conversion factor of 3.7 cM/Mb obtained from regression of centimorgan vs. megabyte values for 20p12-pter loci (mapping data are from LDB).

[b] p_d and p_n are the frequencies for the marker allele on disease-mutation–bearing and normal chromosomes, respectively. Source of data: table 2 of Lee et al. (1999).

[c] Linkage disequilibrium index, calculated according to Bengtsson and Thompson (1981): $\delta = (p_d - p_n)/(1 - p_n)$.

[d] g and g_c are the estimated number of generations obtained by use of Risch et al.'s (1995) algorithm, without and with the Luria-Delbrück correction of the genetic clock (Labuda et al. 1996), respectively: $g = \log\delta/\log(-\theta)$ and $g_c = g + g_0$, where $g_0 = -(1/d)\ln(\theta \times f_d)$, assuming $d = 0.5$ and $f_d = 1/d$.

Table 2

Estimation of the Age of the E200K Mutation in the *PRNP* Gene, Causing CJD in Libyan Jews (Reich and Goldstein's [1999] Iterative Method, without and with the Luria-Delbrück Correction for Population Growth Rate [Labuda et al. 1996])

	Distance from *PRNP* locus[a]				Mutation data[b]		Haplotype data[c]		Estimated age[d]		
Marker	Mb	cM	θ	Allele	μ	f	p_d	p_n	g	g_0	g_c
D20S116	.57	2.11	.0211	20	.00056	.32	.81	.29	14	6.3	20.3
D20S482	.27	1	.001	14	.0021	.29	.93	.59	12	12.4	24.4
D20S895	.35	1.3	.013	18+19	.00056	.09	.87	.35	16	7.3	23.3

[a] See footnote a to table 1.

[b] μ is the assumed frequency of mutation at the marker locus (dinucleotide repeats [D20S116 and D20S482], $\mu \approx .00056$; tetranucleotide repeat [D20S482], $\mu \approx .0021$; Weber and Wong 1993) and f is the observed frequency of all one-mutant neighbors of the ancestral allele in the control population (source of data, table 1 of Lee et al. 1999).

[c] See footnote b to table 1.

[d] g and g_c are the estimated number of generations obtained by use of Reich and Goldstein's (1999) algorithm, without and with the Luria-Delbrück correction of the genetic clock (Labuda et al. 1996), respectively. To provide a complete model for the haplotype's evolutionary process, according to the method of Reich and Goldstein (1999), a Markov transition matrix (**K**) for each marker was generated, which gives the probabilities that any one haplotype will be transformed into any other one in a single generation. **K** was calculated as the weighted sum of matrices corresponding to recombination (**R**), mutation (**M**), and no event occurring (**I**): $\mathbf{K} = \theta\mathbf{R} + \mu\mathbf{M} + (1 - \theta - \mu)\mathbf{I}$, where μ is the frequency of mutation at the marker locus. The matrix **R** has the elements $R_{11} = p_n$, $R_{12} = p_n$, $R_{21} = 1 - p_n$, and $R_{22} = 1 - p_n$. Under the assumption of a stepwise mutation model for microsatellites (Goldstein and Pollock 1997) and a distribution of marker allele sizes on disease chromosomes that matches that seen in the control population, **M** has the elements $M_{11} = 0$, $M_{12} = f/2$, $M_{21} = 1$, and $M_{22} = 1 - f/2$, where f is the frequency of all one-mutant neighbors of the ancestral allele in the control population. With the parameters of **K** specified, the number of generations that have passed since the foundation event was estimated by multiplying the state vector (q, $1 - q$) by **K** iteratively, until the observed proportion of ancestral haplotypes was reached ($q \approx p_d$). Iteration began at a frequency vector of (1,0), corresponding to the archetypal condition of only ancestral haplotypes. The number of times that **K** had been multiplied yielded an estimate of g (Reich and Goldstein 1999). The iterative procedure was implemented through the facilities of the SPSSv.9 matrix language. Iterations stopped at an average distance of $\pm 5 \times 10^{-3}$ from the corresponding p_d value, which is the maximum attainable accuracy, being p_d values accurate to $\pm 1 \times 10^{-2}$. The Luria-Delbrück-corrected age (Labuda et al. 1996) is given as $g_c = g + g_0$, where $g_0 = -(1/d)\ln(\theta \times f_d)$, assuming $d = 0.5$ and $f_d = 1/d$.