Naumova AK, Leppert M, Barker DF, Morgan K, Sapienza C (1998) Parental origin-dependent, male offspring-specific transmission-ratio distortion at loci on the human X chromosome. Am J Hum Genet 62:1493–1499

Address for correspondence and reprints: Francesco Cucca, Istituto di Clinica e Biologia dell'Eta' Evolutiva, Via Jenner, University of Cagliari, Cagliari, Sardinia, Italy. E-mail: fcucca@mcweb.unica.it

0002-9297/2000/6601-0035$02.00

Am. J. Hum. Genet. 66:332–335, 2000

Predominance of the T14484C Mutation in French-Canadian Families with Leber Hereditary Optic Neuropathy Is Due to a Founder Effect

To the Editor:

The Leber hereditary optic neuropathy (LHON) phenotype was first defined, more than 125 years ago, as a maternally inherited optic neuropathy that primarily affects young adult men (Leber 1871). Loss of central vision may be acute or subacute, and peripheral vision is preserved. Slow delayed recovery can occur, but relative central scotomata usually remain (Johns et al. 1993). The chronic stage of LHON is characterized by optic atrophy (MIM 535000).

LHON is associated with three primary mtDNA mutations, all of which occur in genes coding for subunits of complex I of the mitochondrial respiratory chain: G3460A in *ND1,* G11778A in *ND4,* and T14484C in *ND6* (MIM 516006.0001) (Howell et al. 1995). All three of these mutations alter evolutionarily conserved amino acids and are not found in control individuals. The relative frequency of the three primary LHON mtDNA mutations varies considerably in different populations, although G11778A is the most common worldwide. We have recently found that T14484C is by far the most common mutation in French-Canadian families with LHON (Macmillan et al. 1998). The results of previous studies of mutation profiles of several metabolic disorders have demonstrated founder effects, including phenylketonuria (Rozen et al. 1994) and familial hyperchylomicronemia (De Braekeleer et al. 1991), in the French-Canadian population.

To test the hypothesis that the predominance of the T14484C mutation in French-Canadian families with LHON is caused by a founder effect, we sequenced a segment of the mtDNA displacement (D) loop and a segment of the control region from French-Canadian families with the T14484C mutation. Variation in these noncoding regions has been used extensively to study the evolution of modern populations. The regions sequenced included hypervariable regions (HVR) I and II (Vigilant et al. 1989). The D-loop region extends from nucleotide (nt) 189 backward through 0 to nt 16024 of the 16,569-bp mtDNA. The control region extends from nt 189 forward, toward the tRNAs. These two regions contain the fastest-evolving regions of mtDNA (Upholt and Dawid 1977), which have an estimated rate of evolution that is 2.8–5 times that of the remainder of the mitochondrial genome (Aquadro and Greenberg 1983; Cann et al. 1984). The average nucleotide-sequence variation in this region has been calculated to be 1.7% (Aquadro and Greenberg 1983). We reviewed the records of patients with suspected LHON who were independently referred, for molecular diagnosis, to the Montreal Neurological Hospital DNA Diagnostic Laboratory. By cycle sequencing with the use of a New England Biolabs kit (manual) or dye-labeled dideoxy terminators on an ABI system (automated), we sequenced regions that included HVR I and HVR II in the following individuals: one member of each French-Canadian family with the T14484C mutation, one member of a French-Canadian family with the G11778A mutation, and one member of a French-Canadian family without a family history of LHON and with none of the three primary LHON mutations. Sequencing was done in two reactions, by use of the following primer pairs: L15996 and H16401 (which includes HVR I) and L29 and H408 (which includes HVR II; Vigilant et al. 1989). "L" refers to "light" strand and "H" refers to "heavy" strand, and the numbers refer to the Cambridge sequence (Anderson et al. 1981). The institutional review boards of the Montreal Neurological Hospital and the University of Illinois at Chicago approved this study.

We analyzed 27 independently referred French-Canadian families with LHON and the T14484C mutation, all of which were homoplasmic for the T14484C mutation. We found eight homoplasmic transition mutations (C16069T, T16126C, G16213A, A73G, G185A, G228A, A263G, and C295T; table 1), compared with the Cambridge sequence, in the families with the T14484C mutation. Of the 27 families analyzed, 26 shared identical substitutions at all eight sites that were different from the Cambridge sequence. Six of these mutations were found only in families with the T14484C mutation and not in either the family with the G11778A family or the family without LHON (table 1). The mutated sites were distributed throughout the D-loop and control region: one was located in the large central conserved sequence block (CSB; A73G), one (G228A) was in CSB 1, two (T16126C and G16213A) were in HVR I, three (G185A, A263G, and C295T) were in HVR II, and one (C16069T) was just outside HVR I.

In addition, 22/27 families with LHON and the T14484C mutation had a C insertion in the homopolymeric stretch of C's before the T at position 310. All 27

Table 1

Variable Sites in the D-Loop and Control Regions of mtDNA from French-Canadian Families with LHON and the T14484C Mutation (27 Pedigrees), with LHON and the G11778A Mutation, and without LHON

	Presence of Mutation[a]							
Pedigree	C16069T	T16126C	G16213A	A73G	G185A	G228A	A263G	C295T
1	+	+	+	+	+	+	+	+
2	+	+	+	+	+	+	+	+
3	+	+	+	+	+	+	+	+
4	+	+	*	+	+	+	+	A
5	+	+	+	+	+	+	+	+
6	+	+	+	+	+	=	+	+
7	+	+	+	+	+	+	+	+
8	+	+	+	+	+	+	+	+
9	+	+	+	+	+	+	+	+
10	+	+	+	+	+	+	+	+
11	+	+	+	+	+	+	+	+
12	+	+	+	+	+	+	+	+
13	+	+	+	+	+	+	+	+
14	+	+	+	+	+	+	+	+
15	+	+	+	+	+	+	+	+
16	+	+	+	+	+	+	+	+
17	+	+	+	+	+	+	+	+
18	+	+	+	+	+	+	+	+
19	+	+	+	+	+	+	+	+
20	+	+	+	+	+	+	+	+
21	+	+	+	+	+	+	+	+
22	+	+	+	+	+	+	+	+
23	+	+	+	+	+	+	+	+
24	+	+	+	+	+	+	+	+
25	+	+	+	+	+	+	+	+
26	+	+	+	+	+	+	+	+
27	+	+	+	+	+	+	+	+
G11778A	+	*	*	*	*	*	+	*
Without LHON	*	*	*	*	*	*	+	*

[a] A plus sign (+) indicates that the transition mutation indicated was present at that site. An asterisk (*) indicates the Cambridge consensus sequence. "A" indicates adenine.

families with the T14484C mutation, the family with the G11778A mutation, and the family without LHON had a C insertion in the homopolymeric stretch of C's after the T at position 310. Finally, there were some sites at which the sequence appeared to be heteroplasmic; however, these results require independent confirmation.

These data demonstrate that French-Canadian families with LHON and the T14484C mutation likely share the same maternal lineage and suggest that they may all have been derived from a single founder woman. All the observed sequence variants—with the exception of the insertions in the homopolymeric tract of C—were transitions, as compared with the Cambridge sequence; the transition-to-transversion ratio of nucleotide substitutions has been reported to be very high in this region (Aquadro and Greenberg 1983). All our patients were positive for mutations at positions 3394, 4216, and 13708, relative to the standard sequence, and it is likely that they belong to the same haplogroup as does the patient described by Brown et al. (1992). This haplogroup is clearly related to white haplogroup J, which is characterized by mutations at positions 4216 and 13708 and which is where T14484C mutations cluster in the European population (Brown et al. 1997; Torroni et al. 1997). The most-similar published haplotype is that of the haplogroup J T14484C "TAS2" individual, in which seven of the mutations (C16069T, T16126C, A73G, G185A, G228A, A263G, and C295T) were common (Howell et al. 1995). There were four additional TAS2 mutations that were not found and three additional TAS2 mutations that were not evaluated in the French-Canadian families. The French-Canadian mutation array is even less similar to those reported for five other haplogroup J T14484C individuals, seven G11778A individuals, and three G3460A individuals (Howell et al. 1995; Hofmann et al. 1997). Furthermore, the French-Canadian array is different from that reported for haplogroup J individuals without LHON (Hofmann et al. 1997).

Although the T16213A mutation has been reported in several populations (Horai and Hayasaka 1990; Lum et al. 1994; Mountain et al. 1995), this is the first report

of its association with a family with LHON. The single family (pedigree 4) that did not share this mutation may have undergone a reversion mutation to the original sequence; the representative of this family also had a C→A transversion at site 295. The small changes in the D-loop–sequence homopolymeric tracts of C starting at position 303 are variants of the published sequence first reported by Greenberg et al. (1983). The C insertion after the T at position 310 was present in all French-Canadian families, and it was also found in the family with TAS2. However, the C insertion before the T at position 310 was not universally present in the French-Canadian families with the T14484C mutation. These insertions are very unstable and may have arisen after the migration of the common female ancestor to New France. These French-Canadian families with LHON and the T14484C mutation represent a unique genetic resource in which to evaluate the rate of accumulation of sequence variants and the resolution of heteroplasmy in mtDNA on a time scale of a few hundred years.

Acknowledgments

We thank the families for their participation in this study. This study was supported by a University of Illinois at Chicago Campus Research Board grant (to C.M.), grants from the Medical Research Council of Canada, and an FRSQ-Hydro-Quebec technology transfer grant (to E.A.S.). E.A.S. is a Montreal Neurological Institute Killam Scholar.

C. Macmillan,[1] T. A. Johns,[2] K. Fu,[2,3] and E. A. Shoubridge[2,3,4]

[1]*Departments of Neurology and Pediatrics, University of Illinois at Chicago, Chicago; and* [2]*Montreal Neurological Institute and Departments of* [3]*Human Genetics and* [4]*Neurology and Neurosurgery, McGill University, Montreal*

Electronic-Database Information

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

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/ (for Leber optic atrophy [MIM 535000] and MTND6*LHON14484A [MIM 516006.0001]).

References

Anderson S, Bankier AT, Barrell BG, de Bruijn MHL, Coulson AR, Drouin J, Eperon IC, et al (1981) Sequence and organization of the human mitochondrial genome. Nature 290: 457–465

Aquadro CF, Greenberg BD (1983) Human mitochondrial DNA variation and evolution: analysis of nucleotide sequences from seven individuals. Genetics 103:287–312

Brown MD, Sun F, Wallace DC (1997) Clustering of Caucasian Leber hereditary optic neuropathy patients containing the 11778 or 14484 mutations on an mtDNA lineage. Am J Hum Genet 60:381–387

Brown MD, Voljavec AS, Lott MT, MacDonald I, Wallace DC (1992) Leber's hereditary optic neuropathy: a model for mitochondrial neurogenetic diseases. FASEB J 6:2791–2799

Cann RL, Brown WM, Wilson AC (1984) Polymorphic sites and the mechanism of evolution in human mitochondrial DNA. Genetics 106:479–499

De Braekeleer M, Dionne C, Gagne C, Julien P, Brun D, Ven Murthy MR, Lupien PJ, et al (1991) Founder effect in familial hyperchylomicronemia among French Canadians of Quebec. Hum Hered 41:168–173

Greenberg BD, Newbold JE, Sugino A (1983) Intraspecific nucleotide sequence variability surrounding the origin of replication in human mitochondrial DNA. Gene 21:33–49

Hofmann S, Jaksch M, Bezold R, Mertens S, Aholt S, Paprotta A, Gerbitz KD (1997) Population genetics and disease susceptibility: characterization of central European haplogroups by mtDNA gene mutations, correlation with D loop variants and association with disease. Hum Mol Genet 6: 1835–1846

Horai S, Hayasaka K (1990) Intraspecific nucleotide sequence differences in the major noncoding region of human mitochondrial DNA. Am J Hum Genet 46:828–842

Howell N, Kubacka I, Halvorson S, Howell B, McCullough DA, Mackey D (1995) Phylogenetic analysis of the mitochondrial genomes from Leber hereditary optic neuropathy pedigrees. Genetics 140:285–302

Johns DR, Heher KL, Miller NR, Smith KH (1993) Leber's hereditary optic neuropathy: clinical manifestations of the 14484 mutation. Arch Ophthalmol 111:495–498

Leber T (1871) Ueber hereditaere und congenital-angelegte Sehnervenleiden. Arch Ophthalmol 17:249–291

Lum JK, Rickards O, Ching C, Cann RL (1994) Polynesian mitochondrial DNAs reveal three deep maternal lineage clusters. Hum Biol 66:567–590

Macmillan C, Kirkham T, Fu K, Allison V, Andermann E, Chitayat D, Fortier D, et al (1998) Pedigree analysis of French Canadian families with T14484C Leber's hereditary optic neuropathy. Neurology 50:417–422

Mountain JL, Hebert JM, Bhattacharyya S, Underhill PA, Ottolenghi C, Gadgil M, Cavalli-Sforza LL (1995) Demographic history of India and mtDNA-sequence diversity. Am J Hum Genet 56:979–992

Rozen R, Mascisch A, Lambert M, Laframboise R, Scriver CR (1994) Mutation profiles of phenylketonuria in Quebec populations: evidence of stratification and novel mutations. Am J Hum Genet 55:321–326

Torroni A, Petrozzi M, D'Urbano L, Sellitto D, Zeviani M, Carrara F, Carducci C, et al (1997) Haplotype and phylogenetic analyses suggest that one European-specific mtDNA background plays a role in the expression of Leber hereditary optic neuropathy by increasing the penetrance of the primary mutations 11778 and 14484. Am J Hum Genet 60: 1107–1121

Upholt WB, Dawid WB (1977) Mapping of mitochondrial DNA of individual sheep and goats: rapid evolution in the D loop region. Cell 11:571–583

Vigilant L, Pennington R, Harpending H, Kocher TD, Wilson AC (1989) Mitochondrial DNA sequences in single hairs from a southern African population. Proc Natl Acad Sci USA 86:9350–9354

Address for correspondence and reprints: Dr. Carol Macmillan, Department of Neurology, M/C 796, University of Illinois at Chicago, 912 South Wood Street, Room 855N, Chicago, IL 60612-7330. E-mail: cmacmill@uic.edu

0002-9297/2000/6601-0036$02.00

Am. J. Hum. Genet. 66:335–338, 2000

Testing for Linkage Disequilibrium, Maternal Effects, and Imprinting with (In)complete Case-Parent Triads, by Use of the Computer Program LEM

To the Editor:

The traditional transmission/disequilibrium test (TDT) and related tests (see Thomson 1995) require complete triads of genotyped cases plus both parents, in order to test for linkage disequilibrium in the presence of population admixture. A problem in empirical research is that some of the genotype measurements will usually be missing. These incomplete triads must be discarded to ensure the validity of the TDT (Curtis and Sham 1995). Recently, Weinberg (1999*a*) developed likelihood-ratio tests (LRTs) that used the expectation-maximization (EM) algorithm (Dempster et al. 1977), to use incomplete triads as well. Weinberg's tests capitalize on the fact that parent-child dyads may be informative about the genotype of the missing parent. For instance, if a child and a parent are both homozygous for the variant allele, the genotype of the missing parent should comprise at least one copy. Simulations showed that the EM-LRTs were more powerful than the traditional tests that exclude incomplete triads and that they recaptured much of the loss in information caused by missing parental genotypes.

The widespread use of this valuable approach, however, seems hampered by a lack of accessible software. Weinberg, for instance, used the commercial package GLIM, which is good and flexible software but not very user friendly (see remarks on their Internet site), and it requires programming in order to perform the EM-LRTs. To suggest an alternative, we discuss the script to perform Weinberg's tests (1999*b*) for linkage disequilibrium, maternal effects, or parent-of-origin effects in LEM, which is a program for log-linear analysis with missing data that uses the EM algorithm (Vermunt 1997*a*, 1997*b*). An important advantage of LEM is that, with this script, all the tests discussed by Weinberg (1999*b*) can readily be performed in the presence of all possible patterns of missing data, without programming work or the need to learn more LEM syntax. Furthermore, the program is optimized for rapid convergence with EM algorithm, and standard errors of the estimates, fit indices, and a number of appropriate tests are automatically reported in the output so that they do not have to be programmed separately. A final advantage is that the program (which has a DOS and a Windows version) and the manual can be downloaded free of charge on the Internet at the Web site for Methoden en Technieken van Onderzoek (mto).

With a biallelic locus assumed, the genotypes of the mother (M), father (P), and child (C) contain no copy, one copy, or two copies of the variant allele. If the D's are dummy variables (e.g., $D_{(C=1)}$ means that the variable is 1 in all triads in which C = 1 and is 0 otherwise), then the log of the expected cell counts $E(n_{MPC})$ of Weinberg's (1999*b*, see table 1) full model can be written as

$$\ln[E(n_{MPC})] = \gamma_j + \beta_p D_{(C=1)} + \beta_2 D_{(C=2)} + \alpha_1 D_{(M=1)} + \alpha_2 D_{(M=2)} + \ln(w_{MPC}) ,$$

where $e^{\gamma_j} = \mu_j$ are the mating-type–stratum effects (e is the natural exponent), $e^{\beta_p} = R_p$ is the ratio of the risk of disease for genotypes with one copy versus no copies of the variant allele, $e^{\beta_2} = R_2$ is the risk ratio when the genotype comprises two versus no copies of the variant allele, $e^{\alpha_1} = S_1$ is the risk ratio or maternal effect when the mother has one copy versus no copies of the variant allele, and $e^{\alpha_2} = S_2$ is the risk ratio when the mother has two copies versus no copies of the variant allele. The w_{MPC} are cell weights (this becomes clearer when the component is moved to the left-hand side of the equation, so that we obtain $\ln[E(n_{MPC})] - \ln[E(w_{MPC})] = \ln[E(n_{MPC}/w_{MPC})]$), or, in GLIM terminology, $\ln(w_{MPC})$ is called the "offset." The weights can have four different values. First, they can be 0. Because the expected counts in these cells have to be multiplied with $e^{\ln(0)} = 0$, the implication is that the cell frequencies are fixed at 0. This weight is therefore assigned to combinations—such as M = 2, P = 2, and C < 2—that, for theoretical reasons, cannot occur. They are also useful in the context of recovery of information from incomplete triads. For example, if, in the situation described above, the genotype of the child is missing, the 0 weights for C < 2 imply that the missing genotype must comprise two copies of the variant allele. Second, the weights can be 1, so that the expected cell counts are multiplied with $e^{\ln(1)} = 1$, implying that the frequencies as predicted by R_p, R_2, S_1, and S_2 remain unaltered. Third, in the triads M = 2, P = 1, C = 1; M = 2, P = 0, C = 1; and M = 1, P = 0, C = 1 (M > F), where the child receives the copy of the variant allele from the mother, the weights equal the "parent of origin" or "imprinting" effect I_m. Because