Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 12382–12387, October 1996
Genetics

Human brain contains high levels of heteroplasmy in the noncoding regions of mitochondrial DNA

(displacement loop/intraindividual variation/polymorphism/oxidative phosphorylation/aging)

ELENA E. JAZIN*†, LUCIA CAVELIER*†, INGER ERIKSSON*, LARS ORELAND‡, AND ULF GYLLENSTEN*§

Departments of *Medical Genetics and ‡Medical Pharmacology, Uppsala University, S-751 23 Uppsala, Sweden

Communicated by Rolf Luft, Karolinska Hospital, Stockholm, Sweden, August 13, 1996 (received for review October 27, 1995)

ABSTRACT **We have analyzed the level of intraindividual sequence variability (heteroplasmy) of mtDNA in human brain by denaturing gradient gel electrophoresis and sequencing. Single base substitutions, as well as insertions or deletions of single bases, were numerous in the noncoding control region (D-loop), and 35–45% of the molecules from a single tissue showed sequence differences. By contrast, heteroplasmy in coding regions was not detected. The lower level of heteroplasmy in the coding regions is indicative of selection against deleterious mutations. Similar levels of heteroplasmy were found in two brain regions from the same individual, while no heteroplasmy was detected in blood. Thus, heteroplasmy seems to be more frequent in nonmitotic tissues. We observed a 7.7-fold increase in the frequency of deletions/insertions and a 2.2-fold increase in the overall frequency of heteroplasmic mutations in two individuals aged 96 and 99, relative to an individual aged 28. Our results show that intraindividual sequence variability occurs at a high frequency in the noncoding regions of normal human brain and indicate that small insertions and deletions might accumulate with age at a lower rate than large rearrangements.**

mtDNA heteroplasmy has frequently been associated with human chronic degenerative diseases, and low levels have also been found in asymptomatic individuals (1, 2). However, most of the reported cases of heteroplasmy involve a single mutation and the complete spectrum of intraindividual variation has remained unknown. Early studies revealed a high degree of mitochondrial nucleotide sequence homogeneity for the cytochrome oxidase III gene in somatic cells (3), as well as for the D-loop region in retinal tissue (4). On the other hand, studies of human tissue culture cells have shown length heteroplasmy at a single position in the D-loop region (5) and recently at two positions in the first hypervariable segment of the D-loop region (6). In nonhuman species, heteroplasmy at the D-loop has been frequently described (7–12), mainly due to varying numbers of repeat units.

To understand the importance of heteroplasmy to overall mitochondrial function, as well as to segregation and inheritance of mitochondrial mutations, an assessment of the total amount of heteroplasmy in different tissues and age groups is necessary. For example, the frequency of variants at a single site displaying length heteroplasmy in the human mtDNA D-loop was recently shown to be stably inherited (13). Whether the total amount of heteroplasmy also remains similar between generations is, however, not known.

The progressive decrease in mitochondrial oxidative phosphorylation frequently seen with age has been hypothesized to be due to sustained damage to mitochondrial macromolecules caused by endogenously produced oxidants (14). Further, it has been suggested that the accumulation of somatic mutations in the mtDNA is at least partially responsible for the neurodegeneration associated with aging (15–20). The main experimental support for this hypothesis has been the demonstration that the level of a common 4977-bp deletion of the mtDNA (mtDNA^{4977}) in the brain is positively correlated with age (17, 18). Since the proportion of this particular deletion has been below the threshold for physiological significance (0.001–3%), it has been assumed that this mutation is indicative of a spectrum of somatic mutations produced by oxidative damage, which, together, exert an effect on mitochondrial function (17).

To determine the importance of heteroplasmy to aging and neurodegenerative disease, we have estimated the spectrum of mtDNA sequence heteroplasmy in the brain of normal individuals. The level of heteroplasmy was first assessed in five regions of the mtDNA in human brain by denaturing gradient gel electrophoresis (DGGE). This assay showed that the level of heteroplasmy is higher in the noncoding control region (D-loop) than in coding regions. Subsequently, multiple clones of the amplified fragments were sequenced to determine the nature and position of substitutions in coding and noncoding regions.

MATERIALS AND METHODS

Brain and Blood Samples. Brain samples for cloning experiments were obtained at autopsy from three individuals (INDs). IND 1 and IND 3 died without neurologic or neuromuscular disease; IND 3 committed suicide; and IND 2 was a patient with schizophrenia, diagnosed according to DSM-III-R (21). Samples from seven individuals were analyzed by DGGE. These included the three individuals above (INDs 1, 2, and 3 correspond to samples 2, 5, and 7 in Fig. 2) and an additional four individuals: one died without neurologic or neuromuscular disease (sample 6, age 56), two died after early-onset Alzheimer disease (samples 3 and 4, ages 65 and 78), and one died after late-onset Alzheimer disease (sample 1, age 95). Postmortem times before autopsy were on average 45 h. Brain samples (cortex gyrus frontalis and nucleus caudatus) were prepared with the addition of 0.75% deoxycholate as a solubilizing agent before homogenization (2 × 15 strokes) and sonication (1 min pulsed with a Branson sonicator), and the samples were stored at −70°C until DNA extraction. The blood sample from IND 4 was obtained from an asymptomatic individual and DNA was prepared as described below.

Control Samples. Plasmid I and plasmid II were individual pUC19 clones, obtained from amplification and cloning of the D-loop fragment spanning nucleotides 8–429, from blood (plasmid I) and hair (plasmid II), from two different asymptomatic individuals.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: mtDNA^{4977}, common 4977-bp deletion of the mtDNA; DGGE, denaturing gradient gel electrophoresis.

†E.E.J. and L.C. contributed equally to this work.
§To whom reprint requests should be addressed at: Department of Medical Genetics, Biomedical Center, Box 589, S-751 23 Uppsala, Sweden.

DNA Extraction. Genomic DNA was prepared by proteinase K digestion followed by phenol extraction and ethanol precipitation. DNA amounts were estimated by DNA fluorescence (Hoechst 33258; Hoechst Pharmaceuticals).

DGGE. For DGGE, DNA was PCR-amplified using 29- to 33-mer oligonucleotide primers starting at the positions indicated in Fig. 1. The primers used were as follows: for region A, the primers used were light 15761 (5′-ACGAATTCGAGGACAACCAGTAAGCTAC) and heavy 16430 (5′-CCGGATCCTGCGGGATATTGATTTCACGG); for region B, light 8 (5′-GTAGGATCCGGTCTATCACCCTATTAACCAC) and heavy 429 (5′-GTAAAGCTTCTGTTAAAAGTGCATACCGCCA); for region C, light 3148 (5′-ACGAATTCCCTACTTCACAAAGCGCCTTCC) and heavy 3440 (5-CCGGATCCAGCCCGTAGGGGCCTACCACGTT); for region D, light 6903 (5′-ACGAATTCTGATCTGCTGCAGTGCTCTGA) and heavy 7230 (5′-CCGGATCCAGTCCGAGTAACGTCGGGGCA); and for region E, light 11938 (5′-ACGAATTCCCTACTTACAGGACTCAACA) and heavy 12472 (5′-CCGGATCCTAATAAAGGTGGATGCGACA). The 50-μl amplification reaction contained 2.5 μM of each primer, 10% glycerol, 0.5% Nonidet P-40/Tween 20, 0.8 mM total dNTP, PCR buffer, 50 ng of total genomic DNA, and 1.5 units of *Taq* polymerase (AmpliTaq; Perkin–Elmer). The PCR protocol consisted of 25 cycles, each with 1 min each at 94°C, 55°C, and 72°C. Denaturing gradient gels were prepared and run as described (22). Briefly, two gel solutions containing the desired end concentrations of denaturant were prepared by mixing 0% and 100% denaturant stock solutions. The 0% stock consisted of 7% acrylamide in Tris·acetate EDTA (TAE) electrophoretic buffer, and the 100% stock consisted of the same plus 40% formamide and 7 M urea. Twelve milliliters of each solution, usually 30 or 40% in the upstream chamber and 60 or 80% in the lower one, were mixed in a gradient maker together with 70 μl of 10% ammonium persulfate and 7 μl of *N*,*N*,*N*′,*N*′-tetramethylethylenediamine (TEMED) and poured to form a 1-mm thick gel with a linear gradient. Gels were electrophoresed at 60°C and 80 V for 20 h in a unit of slab gel electrophoresis (model SE600; Hoeffer) and stained with ethidium bromide.

PCR Amplification and DNA Sequencing. Two mtDNA regions were analyzed by sequencing of clones to determine the level of heteroplasmy. The D-loop region (region B), was amplified with the primers light 8 and heavy 429, and region C was amplified with light 148 and heavy 440, described above. The amplification before cloning and sequencing was initiated from genomic DNA containing 10^7–10^8 mtDNA copies per reaction. The initial mtDNA copy number of the samples was calculated using a competitive PCR assay (23). From 75 to 170 ng of genomic DNA were used in a 50-μl amplification reaction together with 25 pmol of each primer, 0.5 μl of BSA stock (New England Biolabs), 1 unit of Vent R polymerase (New England Biolabs), 1.6 mM of dNTPs, and the PCR buffer recommended by the suppliers (New England Biolabs). The PCR cycle consisted of 1 min each at 94°C, 55°C, and 72°C and was repeated 35 times. In the control experiments, 3×10^8 and 1×10^{11} copies were used per reaction of plasmid I and plasmid II, respectively. All amplified DNA fragments were treated with 4 units of T4 DNA polymerase at 37°C for 15 min to create blunt ends, the polymerase was heat-inactivated at 85°C for 5 min, and the fragments were ligated to the *Sma*I site of pUC18 overnight at 37°C. Transformations were performed using standard procedures, and cloned inserts were amplified by PCR and sequenced using solid phase T7 polymerase dye primer sequencing using an automated fluorescent dye DNA sequencer (model 373A; Applied Biosystems). The nucleotides at positions with potential variability were confirmed either by sequencing of the opposite strand or by repeating the sequencing reaction.

RESULTS

Screening for Heteroplasmy in the mtDNA Using DGGE. To determine the extent of heteroplasmy in different regions of the mtDNA, we used DGGE to examine the PCR products derived from two brain tissues from each of seven individuals. Five regions of the mtDNA, encompassing almost 15% of the genome, were chosen to screen for variability (Fig. 1, regions A–E); the two segments of the D-loop (A and B) were selected due to their high level of interindividual polymorphism (25, 26), and the other segments were selected because they represent coding regions conserved between hominoids (27). Before DGGE analysis, the five regions were analyzed for their denaturation profiles (Fig. 2 *Right*). The fragments were chosen so as to include at least one high melting domain, to allow for detection of substitutions. DGGE analysis revealed interindividual variability in all five regions. However, the gels did not show any heteroplasmy for regions C–E (Fig. 2). Region B revealed smears with additional faint sharp bands, while region A resulted in single strong sharp bands and additional bands with a low staining intensity (Fig. 2), suggesting the presence of a mixture of sequences in these two regions of the mtDNA. Reconstruction experiments with cloned mtDNA sequences from the same regions were used to verify that the complex banding pattern on the DGGE gel reflected sequence heterogeneity. For instance, DGGE analysis of individual clones from region B resulted in sharp bands (results not shown), while pools of clones resulted in the complex pattern seen in Fig. 2 (region B), supporting that heteroplasmic mutations are responsible for the complex DGGE banding pattern.

Nature and Frequency of the Heteroplasmic Substitutions in the mtDNA. To determine the nature, distribution, and amount of heteroplasmic substitutions in the D-loop, a mtDNA fragment encompassing nucleotides 8–429 (Fig. 1, region B) was amplified from four individuals by PCR using Vent polymerase, and ≈30 clones from each sample were sequenced. The frequency of errors introduced by the Vent DNA polymerase or cloning steps was estimated by reamplification and cloning of the same D-loop region present in two clones (control in Tables 1 and 2), that were processed in

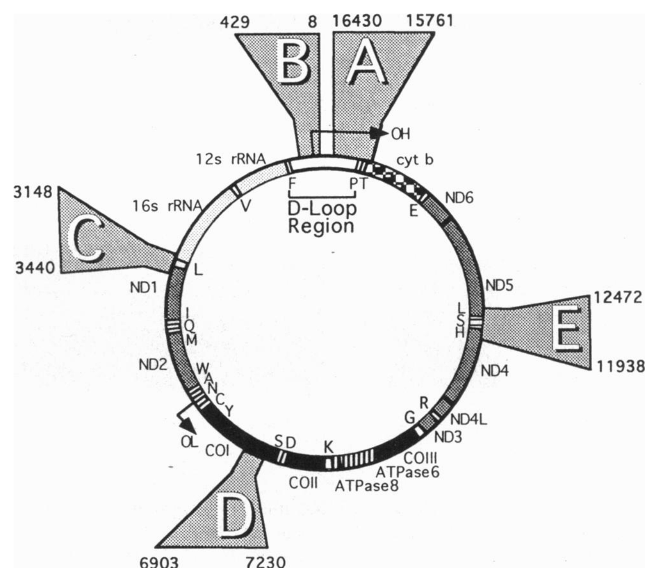

FIG. 1. Map of human mtDNA showing the localization of the five regions (A–E) analyzed by PCR and DGGE. Numbering of base pairs is according to Anderson *et al.* (24). The position of the mitochondrial genes is shown on the outside of the circle, while the positions of tRNAs are indicated by a letter, according to the carried amino acid, on the inside of the circle. The light (O_L) and heavy (O_H) origins of replication are marked.

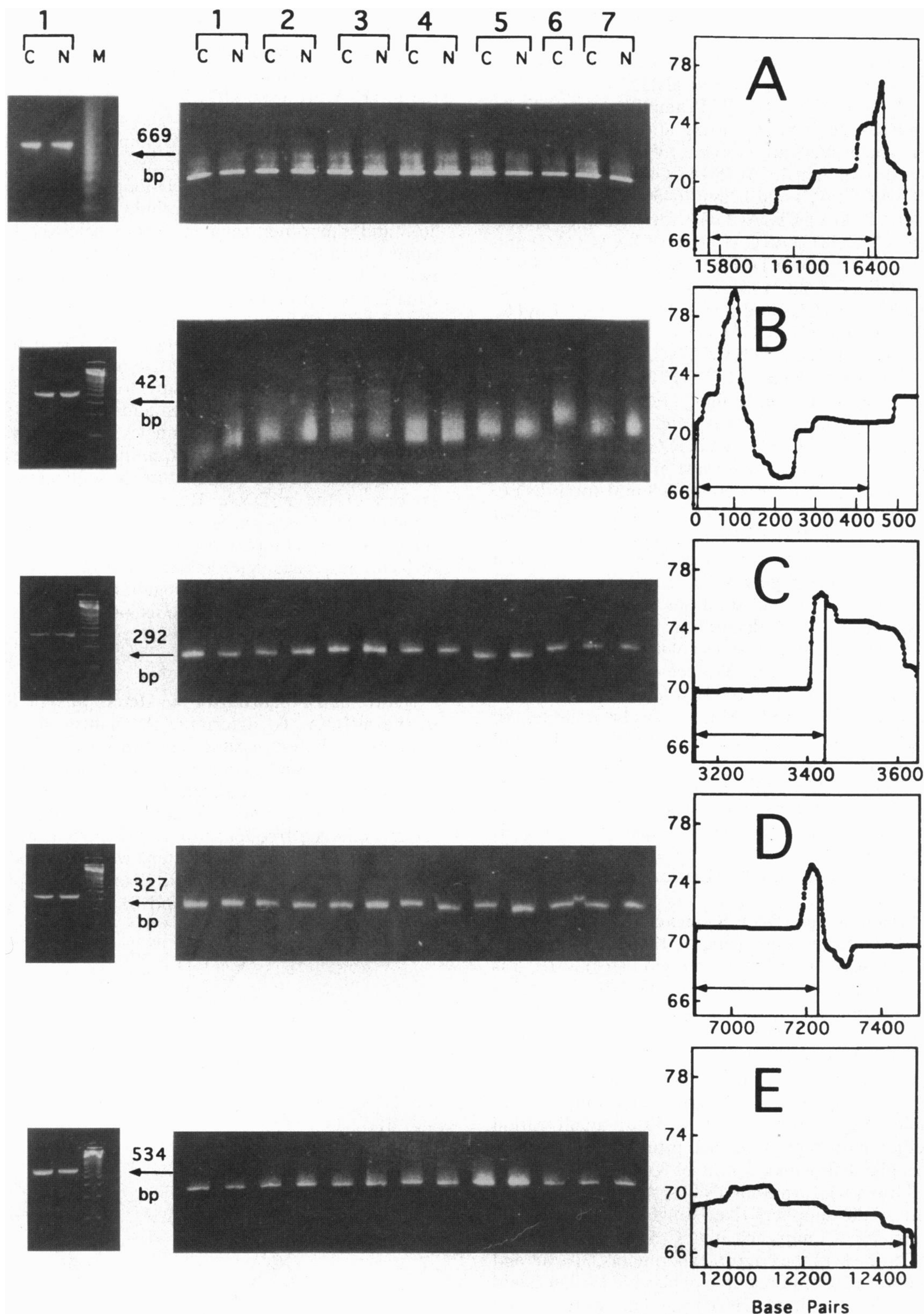

FIG. 2. Analysis of PCR-amplified mtDNA fragments by DGGE. (*Left*) Ethidium bromide-stained agarose gels with PCR-amplified mtDNA fragments [only sample 1 and a 100-bp ladder as a size marker (lane labeled M)]. (*Middle*) Ethidium bromide-stained denaturing gradient gels with fragments amplified from seven brains. Sample 1, age 95; sample 2, age 94; sample 3, age 65; sample 4, age 78; sample 5, age 99; sample 6, age 56; and sample 7, age 28. Samples were obtained from two regions of the brain, cortex gyrus frontalis (lanes labeled C) and nucleus caudatus (lanes labeled N). (*Right*) Melting maps of the five analyzed mtDNA fragments obtained using EZMELT software (28). Numbering of base pairs is according to Anderson *et al.* (24). The midpoint melting temperature at which there is a 50% probability of the DNA sequence melting is shown on the *y* axis. The double-headed arrows indicate the beginning and end of each selected fragment.

parallel with the brain samples. Artifactual single base substitutions were found to occur at a rate of 2×10^{-4} per base (Table 2, maximum of two experiments), corresponding to an approximate rate of 8×10^{-6} per nucleotide per cycle, while no deletions/insertions were found in the control experiments (Table 1).

Sequence analysis of clones from two brain regions (cortex gyrus frontalis and nucleus caudatus) revealed a large number of single base substitutions and deletions/insertions (Table 1). Length variation occurred in at least one position for each individual, with deletions/insertions appearing in ≈30% of the mtDNA molecules. The two most common length variants involved the removal of a G from position 71 and a C from position 309 (Table 1), the latter corresponding to a common polymorphism. In addition, a number of single base substitutions, predominantly transitions, were found between the brain sequences. The frequency of heteroplasmic substitutions in the brain samples was significantly higher than that for controls (e.g., IND 1, gyrus frontalis/Exp I versus plasmid/Exp I, χ^2 = 16.4, df = 1, $P < 0.00007$; IND 2, gyrus frontalis versus plasmid/Exp I, χ^2 = 14.4, df = 1, P = 0.0002), indicating that most of the sequence variants found in the brain are of mitochondrial genomic origin. Also, the frequency of substitutions in the brain samples was higher than that for blood (IND 1, gyrus frontalis/Exp I versus IND 4, χ^2 = 10.57, df = 1, P = 0.0011). Since the frequency of substitutions in the blood was not significantly higher than for the control samples, we are unable to detect heteroplasmy in this tissue. The accuracy of the estimate of heteroplasmy was assessed by repeating the complete experiment with cortex gyrus frontalis twice for IND 1. Both the total amount of heteroplasmy, and the particular substitutions detected were very similar between these experiments (Tables 1 and 2). The two brain regions were found to contain about the same amount of heteroplasmy, both for IND 1 and IND 2 (Table 2). The frequency of artifactual single base substitutions is one order of magnitude lower than the calculated level of heteroplasmy. Therefore the heteroplasmy level might be overestimated by no more than 10%.

To determine the presence of heteroplasmy in coding regions, sequencing of clones for region C (encompassing the tRNA *leu* gene and parts of the 16S rRNA and the *ND1* gene) was performed for gyrus frontalis of IND 1. No heteroplasmic substitutions were detected among 23 different clones (Table 2). This is in accordance with the sharp bands obtained on the DGGE gels for this coding region (Fig. 2, region C).

Age-Related Accumulation of Mutations in the D-Loop. The frequency of heteroplasmic substitutions was not significantly different between the two older individuals (aged 96 and 99) for any of the two brain regions (Table 2). By contrast, the brain of the younger individual (age 28) contained a significantly lower frequency of mutations (IND 1, gyrus frontalis/Exp II versus IND 3, χ^2 = 4.35, df = 1, P = 0.037). Deletions/insertions were 7.7 times more frequent in the elderly individuals (average for IND 1, gyrus frontalis/Exp I and II versus IND 3, χ^2 = 5.79, df = 1, P = 0.016; IND 2, gyrus frontalis versus IND 3, χ^2 = 7.62, df = 1, P = 0.006), while the difference with respect to single base substitutions was not statistically significant.

Table 1. Heteroplasmic substitutions in the D-loop

		IND 1 Age = 96			IND 2 Age= 99		IND 3 Age= 28	IND 4	Control	
		G		N	G	N	F	Blood		
		Exp I	Exp II						Exp I	Exp II
	No. of Clones	1 1 1 3 4 1 18	1 1 1 1 5 1 1 1 4 2 29	2 1 1 1 1 1 1 27	1 1 1 1 8 1 17	1 1 2 1 1 1 4 19	1 1 1 1 23	1 1 24	1 29	1 15
Position										
41	C	. . . .	. . . .	. . . .	. . . .	T	. . . . .	. . .	. .	. .
61	C	. . . .	. . . .	. . . .	. . . .	. T	. . . . .	. . .	. .	. .
64	C	T	T	. . . .	. . . .	. . . .	. . . . .	. . .	. .	. .
66	G	. . . .	. . . .	. . . .	T	. . . .	. . . . .	. . .	. .	. .
69	G	. . . .	. . . .	. . . .	. . . .	. . . .	C	. . .	. .	T .
71	G	. - - - - . .	. - - - -	-	. . . .	. . -	. . . . .	. . .	. .	. .
73	A	G G G G G G G	G G G G G G G G G G G	G G G G G G G G	G G G G G G G	G G G G G G G G	. G . . .	. . .	. .	G G
91	C	. . . .	. . . .	. . . .	. . . .	. . . T	. . . . .	. . .	. .	. .
111.1	-	. . . .	. . . . . C	. . . .	. . . .	. . . .	. . . . .	. . .	. .	. .
114	C	. . . .	. . . .	. . . .	. . . .	. . . .	. . . . .	A . .	. .	. .
115	T	. . . .	. . . .	. . . .	. C	. . . .	. . . . .	. . .	. .	. .
141	C	. . . .	. . . .	. . . .	. . . .	. . . .	. . . . .	. T .	. .	. .
144	C	. . . .	. . . . . . T	. . . .	. . . .	. . . .	. . . . .	. . .	. .	. .
146	T	C C C C C C C	C C C C C C C . C C C	C C C C C C C C	. . . .	. . . .	. . . . .	. . .	. .	. .
147	C	. . . .	. A	. . . .	. . . .	. . . .	. . . . .	. . .	. .	. .
183	A	. . . .	. . . .	. . . .	G G G G G G G	G G G G . G G G	. . . . .	. . .	. .	. .
185	G	. . . .	. . . . A . . .	. . . .	. . . .	. . . .	. . . . .	. . .	. .	. .
189	A	. G	. . G . . . . . G . .	. G . . . G . .	. . . .	. . . .	. . . . .	. . .	. .	G G
195	T	C C C C C C C	C C C C C C C . C C C	C C A C C C C C	. . C	. . . .	. . . . .	. . .	. .	C C
198	C	. . . .	. . . .	. . . T	. . . .	. . . .	. . . . .	. . .	. .	. .
204	T	. . . .	. . . .	. . . .	. . . .	. . . . . C . .	. . . . .	. . .	. .	C C
211	A	. . T	. . . .	. . . .	. . . .	. . . .	. . . . .	. . .	. .	. .
215	A	. . . .	. . . .	. . . . G . . .	. . . .	. . . .	. . . . .	. . .	. .	. .
228	G	. . . .	. . . . A . . .	. . . .	. . . .	. . . .	. . . . .	. . .	. .	. .
252	T	. . . .	. . . .	. . . .	. . . .	. . . .	. . C . .	. . .	. .	. .
263	A	G G G . G G G	G G G G G G G G G G G	G G G G G G G G	G G G G G G G	G G G G G G G G	G G G G G	G G G	G G	G G
295	C	. . . .	. . . . T . . .	. . . .	. . . .	. . . .	. . . . .	. . .	. .	. .
304	C	. . . . .b . .c	. . . .	. . . . a . a .d	. . . . .a . .	a . . a . a .b .	. . A . .c	. . .e	. .	. .
308	C	. . . . . . .a	. . . A	.a	. . . .	. . . .	. . . . .	. . .a	. .	. .
309	C	. . . .	. . . .	. . . . . - . .	. . . .	. . . .	. . . - .	. . .	. .	. .
309.1	-	. . . .	. . . .	. . . .	C C C . . C C	. C C . C . . C	. . . . .	. . .	C C	. .
309.2	-	. . . .	. . . .	. . . .	. . . . . C .	. . . .	. . . . .	. . .	. .	. .
314	C	. . . . . A .a	. . . .	. . . .	. . . .	. . . .	. . . . .	. . .a	. .	. .
315	C	. . . .	. . . .	. . . .	. . . A . . .	. . . .	. . T . .	. . .b	. .	. .
316	G	. . . .	. . . .	. . . . . . A .	. . . .	. . . .	. . . . .	. . .a	. .	. .
361	A	. . . .	. . . . G .	. . . .	. . . .	. . . .	. . . . .	. . .c	. .	. .
378	C	. . . .	. . . .	. . . .	. . . .	. . . .	. . . . .	. . .c	T .	. .

Heteroplasmic substitutions in the D-loop. Distribution of point mutations and insertions/deletions in the D-loop region amplified from brain samples of three individuals and one blood sample, as well as two controls (plasmid). The individual clones are grouped according to their sequence, and the number of clones of each group is indicated. Mutants are shown as differences from a reference mtDNA sequence (24) shown on the left of the table. Positions indicated as numbers followed by a decimal point correspond to additions of nucleotides not present in the reference sequence. A dash indicate a deletion of a nucleotide. a, b, c, d, and e designate 1, 3, 4, 11, and 14 clones not determined at this position, respectively. The tissues analyzed were: G, cortex gyrus frontailis; N, nucleus caudatus; and F, frontal cortex. The D-loop region of IND 1 was analyzed in two separte experiments (Exp I and Exp II) and two different control experiments were performed (ExpI and Exp II).

Table 2. Mutation frequencies

Individual	Region	Experiment	Frequency of point mutations		Frequency of deletions/insertions		Frequency of total substitutions	
			Absolute	Relative	Absolute	Relative	Absolute	Relative
IND 1, age = 96	D-loop							
	Gyrus frontalis	Exp I	7/10991	6.4×10^{-4}	9/10991	8.2×10^{-4}	16/10991	1.5×10^{-3}
		Exp II	16/17780	9.0×10^{-4}	9/17780	5.1×10^{-4}	25/17780	1.4×10^{-3}
	Nucleus caudatus		6/13265	4.5×10^{-4}	3/13265	2.3×10^{-4}	9/13265	6.8×10^{-4}
	Region C (tRNAleu^{+})							
	Gyrus frontalis		0/5704	0	0/5704	0	0/5704	0
IND 2, age = 99	D-loop							
	Gyrus frontalis		4/11370	3.5×10^{-4}	10/11370	8.8×10^{-4}	14/11370	1.2×10^{-3}
	Nucleus caudatus		5/11370	4.4×10^{-4}	9/11370	7.9×10^{-4}	14/11370	1.2×10^{-3}
IND 3, age = 28	D-loop							
	Frontal cortex		5/10233	4.9×10^{-4}	1/10233	9.8×10^{-5}	6/10233	5.9×10^{-4}
IND 4	D-loop							
	Blood		2/9598	2.1×10^{-4}	0/9598	0	2/9598	2.1×10^{-4}
Plasmid	D-loop	Exp I	1/11370	8.8×10^{-5}	0/11370	0	1/11370	8.8×10^{-5}
		Exp II	1/6064	1.7×10^{-4}	0/6064	0	1/6064	1.7×10^{-4}

Frequency of heteroplasmic mutations in the region B (D-loop) and at region C (tRNAleu+, tRNA *leu* gene and parts of the 16S rRNA and the *ND1* gene). The absolute and relative frequencies of the mutations shown in Table 1 were calculated for four individuals and two plasmid controls. The frequency of mutations at coding regions (region C) was only estimated in cortex gyrus frontalis of IND 1. Absolute frequencies are expressed as the number of mutations over the total number of base pairs sequenced per sample, while relative frequencies are the number of mutations per base pair.

Distribution of Heteroplasmy and Polymorphism in the D-Loop. We investigated whether the heteroplasmic nucleotide positions coincided with sites for polymorphism (29). In general, regions in the D-loop with many heteroplasmic sites also showed high levels of polymorphism (Fig. 3). However, the most 5′ segment of the D-loop was depauperate of polymorphism relative to heteroplasmy. Transitions were more frequent than transversions among heteroplasmic single base substitutions (11 transversions versus 34 transitions), consistent with the mutational bias previously observed at polymorphic positions (30, 33) .

DISCUSSION

Heteroplasmy Is Highest in Noncoding Regions. Our analysis revealed high levels of heteroplasmy in the D-loop, while no variability was detected in the coding regions. This distinction between coding and noncoding regions may be attributable either to differences in mutation rate or selection pressure. Certain portions of the D-loop might have a higher mutation rate, due to errors introduced during initiation of replication. The other possibility is that lower levels of heteroplasmy are allowed for in the coding regions, presumably since these are under conservative selection pressures. Consequently, mitochondria with mtDNA molecules carrying detrimental mutations in coding regions are eliminated.

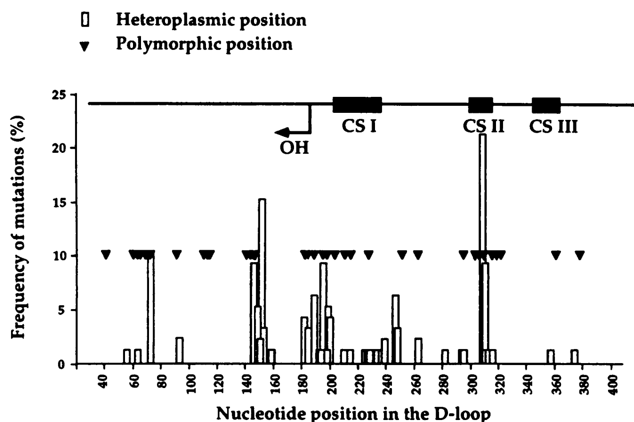

FIG. 3. Distribution of heteroplasmic substitutions and polymorphisms in the D-loop region. The frequency of interindividual polymorphism was calculated from 50 previously reported sequences (30). The heteroplasmic positions reported in Table 1, are indicated with arrowheads. Also shown at the top are the positions of the conserved sequence blocks (CS I, II, and III) previously described (31, 32) and the origin of heavy strand replication (O_H).

Heteroplasmy Is a Natural State for Brain Tissue. In addition to the variation between coding and noncoding regions with respect to the level of heteroplasmy, our study showed differences among tissue types. Heteroplasmy for a large number of single base substitutions and deletions/insertions is a natural state for the brains of asymptomatic individuals, while heteroplasmy appears undetectable in dividing tissue such as blood. Heteroplasmy might exist also in blood, but at levels similar to, or below, the intrinsic error rate of our method. The absence of heteroplasmic variation in leucocytes could be explained by their shorter life span, while in the brain, a tissue with nondividing cellular lineages and high energy requirement, the accumulation of heteroplasmic substitutions and mtDNA rearrangements can occur. Since we do not have any information regarding the distribution of the heteroplasmic substitutions on individual cells or their potential effect on mitochondrial function, it is difficult to evaluate their significance to processes such as aging and neurodegeneration.

Frequent Small Insertions and Deletions in the Control Region. In the brain, several positions showed heteroplasmy for single base deletions at higher frequencies than those reported for the mtDNA^{4977} deletion (17, 18). These positions might represent hotspots for substitutions that either occur early in the ontogeny and/or have expanded clonally in certain tissues. Different mechanisms can explain the high frequency of small insertion/deletions in the D-loop. Deletions in polycytosine or polyguanine tracks may occur at higher frequency and/or may not be as readily repaired as single base mismatches. On the other hand, a shorter replication time for deleted molecules that has been hypothesized to account for the faster rate of accumulation of large deletions (1) is unlikely to be of importance for deletions of a single base pair. Finally, some of the heteroplasmic deletions and insertions may not be selectively neutral. The positive or mild negative effect of certain natural polymorphisms in mtDNA coding regions has been hypothesized (2). Indeed, several of the heteroplasmic

substitutions occur in semiconserved regions that are important for the control of transcription and replication (31).

The excess of transitions among heteroplasmic substitutions that we found parallels the results of studies of polymorphism (30), but the bias in our study was less pronounced. This could reflect that many of the heteroplasmic mutations are recent, and therefore only a few, if any, positions have experienced multiple changes. Segments of the D-loop showing the highest levels of polymorphism also contain a large proportion of the heteroplasmic sites—for example, nucleotides 180-200, surrounding the origin of replication, and nucleotides 300-320, containing the RNase mitochondrial RNA processing (MRP) cleavage site. However, in the 5′ part of the D-loop, heteroplasmic sites are more frequent than polymorphisms. The similarity in the distribution and substitution pattern for heteroplasmic and polymorphic sites indicate that the same mutational mechanisms might be involved in the generation of heteroplasmy and polymorphism.

Increase of Mutant mtDNA with Age. The frequency of heteroplasmy was higher in two individuals aged 96 and 99 compared with the 28-year-old individual, consistent with results for the accumulation with age of the $mtDNA^{4977}$ deletion (17, 18), a duplication (34), and the myoclonic epilepsy and ragged-red fibers (MERRF) point mutation (35). The $mtDNA^{4977}$ deletion has been reported to increase 50- to 2000-fold over ≈70 years (18). The increase of small insertions and deletions with age in our study was much lower (≈7.7-fold), while point mutations did not show any significant increase. The less pronounced difference between age groups in our study could reflect a higher baseline level, due to the misincorporation rate of the Vent polymerase. Thus, it remains possible that our intrinsic error rate is higher than the level of heteroplasmy in the younger individual, resulting in a lower rate of increase than that actually existing. Also, since we did not study the same individual at different ages, it also remains a possibility that the small insertions and deletions substitutions are constitutive and do not accumulate with age.

In summary, we have shown that the level of heteroplasmy in noncoding regions of the mtDNA in the brain is high, while that for coding regions heteroplasmy is lower. Further experiments measuring the accumulated damage in the coding regions will be needed to clarify the functional importance of heteroplasmic mutations in aging human brain.

We are grateful to Dr A. Brookes for his comments on the manuscript. This work was supported by grants from the Beijer Foundation, the Swedish Medical Research Council, the Swedish Natural Sciences Research Council, the Marcus Borgström Foundation, and the Sörensen Foundation.

1. Wallace, D. C. (1992) *Annu. Rev. Biochem.* **61,** 1175–1212.
2. Wallace, D. C. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 8739–8746.
3. Monnat, R. J. J. & Loeb, L. A. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 2895–2899.
4. Bodenteich, A., Mitchell, L. G. & Merril, C. R. (1991) *Gene* **108,** 305–309.
5. Hauswirth, W. W. & Clayton, D. A. (1985) *Nucleic Acids Res.* **13,** 8093–8104.
6. Bendall, K. E. & Sykes, B. C. (1995) *Am. J. Hum. Genet.* **57,** 248–256.
7. Hale, L. R. & Singh, R. S. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 8813–8817.
8. Harrison, R. G., Rand, D. M. & Wheeler, W. C. (1985) *Science* **228,** 1446–1448.
9. Hoeh, W. R., Blakley, K. H. & Brown, W. M. (1991) *Science* **251,** 1488–1490.
10. Densmore, L. D., Wright, J. W. & Brown, W. M. (1985) *Genetics* **110,** 689–707.
11. Boursot, P., Yonekawa, H. & Bonhomme, F. (1987) *Mol. Biol. Evol.* **4,** 46–55.
12. Ghivizzani, S. C., Mackay, S. L., Madsen, C. S., Laipis, P. J. & Hauswirth, W. W. (1993) *J. Mol. Evol.* **37,** 36–37.
13. Kim, Y. L., Brown, M. D. & Wallace, D. C. (1995) *Anal. Biochem.* **224,** 608–611.
14. Shinenaga, M. K., Hagen, T. M. & Ames, B. N. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 10771–10778.
15. Cortopassi, G. A. & Arnheim, N. (1990) *Nucleic Acids Res.* **18,** 6927–6933.
16. Cortopassi, G. A., Shibata, D., Soong, N. W. & Arnheim, N. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 7370–7374.
17. Soong, N. W., Hinton, D. R., Cortopassi, G. & Arnheim, N. (1992) *Nat. Genet.* **2,** 318–323.
18. Corral-Debrinski, M., Horton, T., Lott, M. T., Shoffner, J. M., Beal, M. F. & Wallace, D. C. (1992) *Nat. Genet.* **2,** 324–329.
19. Lee, H. C., Pang, C. Y., Hsu, H. S. & Wei, Y. H. (1994) *Biochim. Biophys. Acta* **1226,** 37–43.
20. Baumer, A., Zhang, C., Linnane, A. W. & Nagley, P. (1994) *Am. J. Hum. Genet.* **54,** 618–630.
21. American Psychiatric Association (1987) *Diagnostic and Statistical Manual of Mental Disorders* (Am. Psychiatr. Assoc., Washington, DC), 3rd Ed., revised.
22. Myers, R. M., Sheffield, V. C. & Cox, D. R. (1988) in *Detection of Single Base Changes in DNA: Ribonuclease Cleavage and Denaturing Gradient Gel Electrophoresis*, ed. Davies, K. E. (IRL, Oxford), pp. 95–139.
23. Cavelier, L., Jazin, E. E., Eriksson, I., Prince, J., Bave, U., Oreland, L. & Gyllensten, U. (1995) *Genomics* **29,** 217–224.
24. Anderson, S., Bankier, A. T., Barrel, B. G., Bruijn, M. H. L. D., Coulson, A. R., Drouin, J., Eperon, I. C., Nierlich, D. P., Roe, B. A., Sanger, F., Schreier, P. H., Smith, A. J. H., Staden, R. & Young, I. G. (1981) *Nature (London)* **290,** 457–465.
25. Greenberg, B. D., Newbold, J. E. & Sugino, A. (1983) *Gene* **21,** 33–49.
26. Cann, R. L. & Wilson, A. C. (1983) *Genetics* **104,** 699–711.
27. Kondo, R., Horai, S., Satta, Y. & Takahata, N. (1993) *J. Mol. Evol.* **36,** 517–531.
28. Behrens, J., Warram, J. H., Lerman, L. S. & Krolewski, A. S. (1990) *Am. J. Hum. Genet.* **47,** Suppl. 1, A173 (abstr.).
29. Vigilant, L., Stoneking, M., Harpending, H., Hawkes, K. & Wilson, A. C. (1991) *Science* **253,** 1503–1507.
30. Stoneking, M., Hedgecock, D., Higuchi, R. G., Vigilant, L. & Erlich, H. A. (1991) *Am. J. Hum. Genet.* **48,** 370–382.
31. Bennett, J. L. & Clayton, D. A. (1990) *Mol. Cell. Biol.* **10,** 2191–2201.
32. Walberg, M. W. & Clayton, D. A. (1981) *Nucleic Acids Res.* **9,** 5411–5421.
33. Stoneking, M. (1990) *J. Mol. Evol.* **31,** 343–346.
34. Lee, H. C., Pang, C. Y., Hsu, H. S. & Wei, Y. H. (1994) *FEBS Lett.* **354,** 79–83.
35. Munscher, C., Rieger, T., Muller, H. J. & Kadenbach, B. (1993) *FEBS Lett.* **317,** 27–30.