Population genetics and disease susceptibility: characterization of central European haplogroups by mtDNA gene mutations, correlation with D loop variants and association with disease

Sabine Hofmann*, Michaela Jaksch+, Reimar Bezdol, Sabine Mertens, Simone Aholt, Armin Paprotta and Klaus-Dieter Gerbitz*

Institutes of Clinical Chemistry and Diabetes Research, Academic Hospital Schwabing, Koelner Platz 1, 80804 Munich, Germany

Received April 11, 1997; Revised and Accepted July 10, 1997

Mitochondrial (mt)DNA haplogroups in a German control group \((n = 67)\) were characterized by screening mitochondrial coding regions encompassing most of the ND, tRNA and cyt b genes. We used a PCR-SSCP screening approach followed by direct sequencing of polymorphic mtDNA fragments. Five major mtDNA lineages, diverging in at least nine different haplogroups, could be defined by characteristic polymorphic sites in mitochondrial genes. Additional sequencing of two hypervariable segments (HVS-I and II) of the non-coding displacement (D) loop in all control subjects revealed that certain D loop variants were strongly correlated with lineages and haplogroups, while others represented hotspots occurring frequently in different haplogroups. The existence of identified lineages and haplogroups received support from data in the literature, obtained by use of different approaches. Subsequently, we investigated four disease groups for association with these haplogroups: (i) LHON patients \((n = 55)\) carrying at least one of the primary/intermediate LHON mutations at nt 3460, 11778, 14484 and/or 15257; (ii) patients suffering from Wolfram or DIDMOAD syndrome \((n = 8)\); (iii) MELAS patients \((n = 9)\); (iv) a group of children, who died from ‘sudden infant death syndrome’ (SIDS) \((n = 9)\). The distribution patterns among the haplogroups of the disease groups (LHON, DIDMOAD and SIDS) differed considerably from the control population. LHON and DIDMOAD were significantly under-represented in the most frequent German haplogroup DC, but were concentrated in a mtDNA lineage defined by polymorphisms at nt 4216+11251+16126. As this lineage diverged into two precisely defined haplogroups, LHON and DIDMOAD could be assigned to the two haplogroups separately.

Strikingly, SIDS was often found in association with two rare German haplogroups. MELAS patients were equally distributed among German haplogroups and, moreover, did not reveal any accumulation of specific D loop variants. We conclude that certain European mtDNA haplogroups define a genetic susceptibility basis for various disorders.

INTRODUCTION

Attempts to identify genes responsible for penetrant single gene disorders have been extraordinarily successful during the last decade and, presumably, will lead to the detection of monogenic defects in most of the corresponding common disorders within the next few years. The major challenge for medical genetics in the near future, however, appears to be the identification of genes contributing to heredity in disorders which are not transmitted in a simple Mendelian fashion (1). Common disorders like hypertension, Alzheimer’s and Parkinson’s disease or diabetes mellitus may be oligo- or even polygenic. While in chromosomal DNA problems in defining specific genes and gene mutations that play an important role in this concerted action are multiple and complex, the other human genome, i.e. mtDNA, offers the possibility of approaching the problems from another, better defined perspective.

Mitochondria are the power stations of the cell, providing energy for every specialized cell function in the form of ATP or heat. Essential parts of the oxidative phosphorylation apparatus are encoded by the mitochondrial genome. Therefore, severe defects of the mitochondrial genome lead to disease (2). mtDNA has already been fully sequenced (3). It exhibits a few interesting features which make it a suitable tool for phylogenetic and pathogenetic studies. The mitochondrial genome is exclusively maternally inherited and demonstrates no recombination events, but does exhibit high mutation and mutation fixation rates. The study of mtDNA sequence variation from around the world has led to defined population-specific mtDNA lineages which can be...
Figure 1. Schematic representation of major mtDNA lineages and haplogroups in a German control population. The tree was constructed without use of any software packages by setting the most frequent variant on top (as a putative common ancestor) and separating other haplogroups by one or more variants. Nucleotide positions defining the clusters correspond to variants listed in Table 1. D loop variants are given in italic letters. D loop variants defining mitochondrial haplotypes are not depicted in this figure. The different levels of mutation clusters are not correlated with specific genetic distances.

traced back to an African origin of the mitochondrial gene pool of modern Homo sapiens (4–6).

Besides being useful for phylogenetic studies, there is growing evidence that certain mtDNA clusters are associated with distinct disorders (7–13). Since most of these studies have been carried out on Caucasians, the exact definition of Caucasoid-specific mtDNA lineages and haplogroups could have important implications for an understanding of the relationship between mutability of the mitochondrial genome and disease. However, most mtDNA haplotype studies have employed RFLP analysis with a limited number of restriction endonucleases. In other studies, dealing primarily with phylogenetic aspects, only the most polymorphic region of the genome, the displacement (D) loop, has been sequenced, without any relation to mitochondrial gene polymorphisms. We used another more differentiating approach. Employing SSCP optimized for mutation detection in mtDNA fragments (14) and direct sequencing of polymorphic fragments, we analyzed wide regions of the mtDNA in a group of healthy German subjects. Here we report that the bulk of the German population can be described by certain well-defined mtDNA lineages which are characterized by homoplasmic gene polymorphisms while, even within these haplogroups, sequencing of the D loop resulted in many different haplotypes. Distribution of various disease groups among these haplogroups revealed that some of the established mitochondrial gene haplogroups predispose to disease.

RESULTS

Mitochondrial gene polymorphisms in the German population

mtDNA haplogroups of 67 healthy German control subjects were characterized by PCR-SSCP followed by direct sequencing of polymorphic PCR fragments. In this ongoing study we have screened, so far, ~60% of ND, tRNA and cyt b genes and, additionally, small parts of COX, ATPase and rRNA genes. Using this approach, a total of 43 polymorphic sites compared with the Cambridge reference sequence (3) could be identified in our control group (Table 1). These were 27 polymorphic sites within ND genes, five in the cyt b gene, eight tRNA variants, one rRNA variant and two polymorphisms in non-coding regions (variants at nt 574 and 5580). The variants at nt 3423, 11335, 11447, 14199, 14272 and 15326 were detected in all 67 controls and were,
Table 1. Polymorphic sites in mitochondrial genes and two hypervariable segments (HVS-I and HVS-II) of the D loop from 67 German control subjects

| pos. | locus | base exchange | exchange | amino acid exchange | cons. |
|------|-------|--------------|----------|---------------------|-------|
| 574  | NC    | n.d.         |          |                     |       |
| 3197 | 16SrRNA T C | ~  | M         | ~                   |       |
| 3338 | ND1 T C  | syn         | H        | ~                   |       |
| 3341 | ND1 C T  | syn         | H        | ~                   |       |
| 3365 | ND1 T C  | Y/H         | H        | ~                   |       |
| 3423 | ND1 G/T  | syn         | L        | ~                   |       |
| 3446 | ND1 C/G  | P/A         | H        | ~                   |       |
| 4136 | ND1 A/G  | Y/C         | M        | ~                   |       |
| 4216 | ND1 T/C  | Y/H         | L        | ~                   |       |
| 4306 | IRNA-Glu T C | ~  | M         | ~                   |       |
| 4529 | ND2 A/T  | syn         | L        |                    | ~     |
| 4646 | ND2 T/C  | syn         | M        | ~                   |       |
| 4917 | ND2 A/G  | D/N         | H        | ~                   |       |
| 5580 | ND T/C  | ~           | n.d.     | ~                   |       |
| 5633 | IRNA-Ala C T | ~  | L         |                    | ~     |
| 7028 | COI C T  | syn         | L        | ~                   |       |
| 10034| IRNA-Gly T C | ~  | L         |                    | ~     |
| 10238| ND3 T C  | syn         | M        | ~                   |       |
| 10398| ND2 A/G  | T/A         | M        | ~                   |       |
| 10410| IRNA-Arg T C | ~  | L         |                    | ~     |
| 10463| IRNA-Arg T C | ~  | L         |                    | ~     |
| 10550| ND4L A/G | syn         | M        | ~                   | ~     |
| 11251| ND4 A/G  | syn         | M        | ~                   |       |
| 11253| ND4 T/C  | I/T         | L        | ~                   |       |
| 11299| ND4 T/C  | syn         | L        | ~                   |       |
| 11332| ND4 C/T  | syn         | L        | ~                   |       |
| 11335| ND4 T/C  | syn         | L        | ~                   |       |
| 11337| ND4 A/G  | N/S         | L        | ~                   |       |
| 11447| CG C/G   | syn         | L        | ~                   |       |
| 11467| ND4 A/G  | syn         | L        | ~                   |       |
| 11812| ND4 A/G  | syn         | L        | ~                   |       |
| 12308| IRNA-Leu2 A/G | ~  | H         |                    | ~     |
| 12372| ND5 G/A  | syn         | L        | ~                   |       |
| 12612| ND5 A/G  | syn         | L        | ~                   |       |
| 12624| ND5 A/T  | U/V         | H        | ~                   |       |
| 12705| ND5 C/T  | syn         | L        | ~                   |       |
| 13368| ND5 G/A  | syn         | L        | ~                   |       |
| 13708| ND5 A/G  | A/T         | M        | ~                   |       |
| 14199| ND5 G/T  | syn         | L        | ~                   |       |
| 14233| ND5 A/G  | syn         | M        | ~                   |       |
| 14405| cytb G/A | syn         | L        | ~                   |       |
| 15043| cytb A/G  | syn         | L        | ~                   |       |
| 15257| cytb C/T  | syn         | L        | ~                   |       |
| 15326| cytb A/G  | T/A         | L        | ~                   |       |
| 15905| IRNA-Thr C/T | ~  | L         |                    | ~     |
Variants in coding regions were detected either by PCR-SSCP and subsequent direct sequencing or by RFLP methods; D loop polymorphisms were identified by sequencing of HVS-I and HVS-II segments. The insertion variant at nt 574 indicates heteroplasmic C stretches. Control subjects were grouped according to their lineage association. The conservation of amino acids or nucleotides is indicated as high (H), medium (M) or low (L). In the case of an amino acid exchange, the conservation of the specific amino acid, in the case of a silent or a tRNA/rRNA mutation, the conservation of the nucleotide is given.

+ base exchange; -, no base exchange; n.d., not detected; NC, non-coding region

| Variant | Location | Conserved |
|---------|----------|-----------|
| 16362   | HVS-I    | T/C       |
| 16366   | HVS-I    | C/T       |
| 16369   | HVS-I    | G/A       |
| 16391   | HVS-I    | G/A       |
| 55      | HVS-II   | T/C       |
| 72      | HVS-II   | T/C       |
| 73      | HVS-II   | A/G       |
| 143     | HVS-II   | G/A       |
| 146     | HVS-II   | T/C       |
| 150     | HVS-II   | C/T       |
| 151     | HVS-II   | C/T       |
| 152     | HVS-II   | T/C       |
| 153     | HVS-II   | A/G       |
| 155     | HVS-II   | A/G       |
| 185     | HVS-II   | G/A       |
| 188     | HVS-II   | A/G       |
| 189     | HVS-II   | A/G       |
| 194     | HVS-II   | C/T       |
| 195     | HVS-II   | T/C       |
| 198     | HVS-II   | C/T       |
| 200     | HVS-II   | A/G       |
| 204     | HVS-II   | T/C       |
| 207     | HVS-II   | G/A       |
| 225     | HVS-II   | G/A       |
| 228     | HVS-II   | G/A       |
| 229     | HVS-II   | T/C       |
| 247     | HVS-II   | G/A       |
| 250     | HVS-II   | T/C       |
| 257     | HVS-II   | A/G       |
| 295     | HVS-II   | C/T       |
| 303     | HVS-II   | 1.C       |
| 303     | HVS-II   | 2.C       |
| 311     | HVS-II   | 1.C       |
| 311     | HVS-II   | 2.C       |
therefore, likely to represent errors in the reference sequence or very frequent polymorphisms. Only a few variants lead to an amino acid change at a highly conserved position (variants at nt 3448, 3394 and 12634) or are localized at highly conserved sites in tRNAs (variants at nt 12308 and 10463). In addition to our SSCP screening approach, the frequencies of seven further polymorphisms (nt 4136, 4216, 4336, 4917, 7028, 13708 and 15257) were investigated directly by RFLP methods (Table 1). Analysis of the whole sequence data set (a total of 50 polymorphic sites) revealed that our German control population could be described and characterized by five major mtDNA lineages (Fig. 1). Several branching points indicate important ancestral sites of the German tree. Thus a transition variant at nt 7028 separates the German population in one 7028(−) and four 7028(+) lineages. Among the 7028(+) subjects, the 4216+11251 lineage is observed in 19% (13/67) of our control group and defines the deepest node of two haplogroups consisting of the additional variants 10398+12612+13708 (haplogroup D12705, 7%) and 4917+10463+13368+14905+15607+15928 (haplogroup D4216, 12%) respectively. The second lineage 12308+12372 (16%) branches into three haplogroups, defined by nt 11467 (10%), subdiverging into haplogroups D12308 (11467+3397 6%) and D12308+16224 (4%), and by nt 4646 (haplogroup D4216 16224, 6%). The third lineage 12705 is found in 10% of our healthy subjects; it branches into haplogroups D12705 (10398, 4%) and D12705+10463 (6%). A fourth branch, haplogroup D12705 (4%), was defined by a C→T exchange at nt 15904 in the tRNA\(^{\text{Tyr}}\) gene. Two 7028(+) German controls were not further classified since no coding region variants have been identified so far. Among the 7028(−) subjects (46%, 31/67), representing the most frequent lineage D\(^{\text{C}}\), only six rare sub-branches, D\(^{\text{C}}\)3341, D\(^{\text{C}}\)3338, D\(^{\text{C}}\)5580, D\(^{\text{C}}\)10463, D\(^{\text{C}}\)14905 and D\(^{\text{C}}\)4336, could be defined using our sequence data set. Twenty three subjects were identical to the Cambridge reference sequence, at least with respect to the gene fragments investigated here. It should be noted, however, that this latter branch D\(^{\text{C}}\) was defined per exclusionem, i.e. by subtracting the other haplogroups from the Cambridge reference sequence.

### D loop variants and correlation with mitochondrial gene mutations

The sequences of two hypervariable segments of the control region, HVS-I and HVS-II, were determined in all control subjects. Sequencing of a total of 674 bp revealed 90 variants (Table 1) corresponding to 13% of variable sites. As can be seen in Figure 1, many D loop variants were associated with defined lineages or haplogroups, which means that they were haplogroup specific. An A→G exchange at nt 263 and an insertion of one C at nt 311 were found in all subjects, indicating either frequent polymorphisms in the German population or errors in the reference sequence. All 7028(+) lineages (except lineage 15904) had a G→A exchange at nt 073. Lineage 15904 was additionally characterized by the D loop variant at nt 072. Variant 16126 was significantly associated with lineage 4216+11251, variant 295 was solely found in haplogroup D\(^{15904}\) and variants 16294 and 16296 were exclusively observed in haplogroup D\(^{2119}\). Also, several D loop polymorphisms clearly correlated with coding region variants within the lineages 12308+12372 and 12705 and corresponding sub-clusters. On the other hand, a number of D loop polymorphisms represented hypervariable sites, as they have arisen multiple times during evolution in different lineages/haplogroups. Examples for such hotspots are the polymorphisms at nt 16189, 16304, 16311, 16362, 152, 195, 200 and 303.

With respect to the entire data set (coding and D loop regions), we observed 65 different haplotypes in 67 German controls. All control subjects characterized by at least one coding region polymorphism could be distinguished from one another by additional D loop polymorphisms. Only two pairs of control subjects belonging to haplogroup D\(^{\text{C}}\) (subjects 731/765 and 742/764) shared identical D loop sequences.

### Association with disease

Having established the characterization of our healthy central European (German) control group by several major mtDNA lineages segregated into carefully defined genetic sub-branches (haplogroups) and haplotypes (defined by additional D loop variants) we investigated whether distinct disorders cluster in certain lineages, haplogroups or haplotypes. For this purpose we collected our different disease cohorts. These were patients suffering from LHON, MELAS, DIDMOAD and children who had died from SIDS. Table 2 summarizes the distribution of these disease groups among German mitochondrial lineages and haplogroups.

### Table 2. Association of four disease groups with mtDNA lineages and sub-clusters

| Lineage/haplogroup | Frequencies | Controls | LHON | DIDMOAD | SIDS | MELAS |
|--------------------|-------------|----------|------|----------|------|-------|
| haplogroup D\(^{\text{C}}\) | 46% | 23% | 0% | 33% | 55% |
| 4216+11251 lineage | 19% | 59% | 75% | 33% | 33% |
| 4917 haplogroup | 12% | 0% | 63% | 0% | 22% |
| 13708 haplogroup | 7% | 59% | 0% | 0% | 11% |
| 13708(−) side branch | 0% | 0% | 12% | 33% | 0% |
| 12308+12372 lineage | 16% | 12% | 12% | 11% | 0% |
| 574.insC subcluster | 0% | 0% | 0% | 11% | n.d. |
| 12705 lineage | 10% | 6% | 0% | 22% | 11% |
| 574.insC sub-cluster | 3% | 0% | 0% | 22% | n.d. |
| 15904 lineage | 4% | 0% | 0% | 0% | 0% |

Frequencies are given in percentage occurrence. All values were rounded off. The frequencies within different groups do not reach 100% since minor sub-branches have not been considered. With respect to the LHON group, only a subgroup of 17 patients has been included in this table.
LHON. The LHON group (n = 55), clinically defined by bilateral optic neuropathy, was sub-divided with respect to the occurrence of a primary (nt 11778, 14484 and 3460) and/or intermediate (nt 15257) LHON mutation and consisted of twenty four 11778 carriers, twenty one 14484 carriers and five 3460 carriers. The 15257 mutation, classified as an intermediate mutation with respect to its pathogenic power, was associated with the 11778 mutation in three cases, with the 14484 mutation in six cases and was found five times without any of the established primary mutations. Primarily the mutations 11778, 14484 and 15257 were found with increased frequencies in association with haplogroup D\textsuperscript{13708} (15): 86% of 15257 carriers, 66% of 14484 carriers and 46% of 11778 carriers belonged to this haplogroup. These frequencies of LHON in haplogroup D\textsuperscript{13708} were significantly higher than would be expected from the haplogroup distribution found in our control population (P < 0.001).

To answer the question as to whether the primary LHON mutations were preferentially associated with a distinct haplotype defined by D loop variants we sequenced the two hypervariable D loop regions from 10 representative LHON patients belonging to haplogroup D\textsuperscript{13708}. Figure 2 shows that these LHON patients, although belonging to a precisely defined haplogroup, separated into two major branches and eight sub-branches mainly by HSV-1 variants. Only two (of five) 11778 carriers and two (of two) 3460 carriers showed identical D loop sequences, while all other LHON patients represented different haplotypes. This pattern strongly suggests three important points. First, LHON is associated with haplogroup D\textsuperscript{13708} independent of the nature of the underlying primary LHON mutation. Second, most of the primary LHON mutations must have arisen several times within this haplogroup; for instance, the 11778 mutation at least four times and the 15257 intermediate LHON mutation at least three times; only the association of the 3460 mutation within haplogroup D\textsuperscript{13708} could be traced back to one mutational event. Thirdly, disease susceptibility is associated with mtDNA gene variants and not with any specific D loop haplotype, making founder effects unlikely.

DIDMOAD. We were able to demonstrate a statistically significant association between DIDMOAD and a distinct haplogroup (12). Six out of eight patients (75%) were associated with the 4216+11251+16126 lineage, of which five (63%) clustered in haplogroup D\textsuperscript{4216}. Figure 2 shows that these LHON patients, although belonging to a precisely defined haplogroup, separated into two major branches and eight sub-branches mainly by HSV-1 variants. Only two (of five) 11778 carriers and two (of two) 3460 carriers showed identical D loop sequences, while all other LHON patients represented different haplotypes. This pattern strongly suggests three important points. First, LHON is associated with haplogroup D\textsuperscript{13708} independent of the nature of the underlying primary LHON mutation. Second, most of the primary LHON mutations must have arisen several times within this haplogroup; for instance, the 11778 mutation at least four times and the 15257 intermediate LHON mutation at least three times; only the association of the 3460 mutation within haplogroup D\textsuperscript{13708} could be traced back to one mutational event. Thirdly, disease susceptibility is associated with mtDNA gene variants and not with any specific D loop haplotype, making founder effects unlikely.

SIDS. Nine children who had obviously died from SIDS were investigated for their association with haplogroups and haplotypes (Table 2). Three of the SIDS infants (33%) belonged to D\textsuperscript{3208}, one to D\textsuperscript{13708} (11%), another to lineage 12705 (11%) and the remaining five patients were members of haplogroup D\textsuperscript{C} (55%). D loop sequencing of all MELAS patients (data not shown) did not reveal any concentration in a distinct haplotype and showed that the MELAS mutations must have arisen multiple and independent times. Furthermore, we could not detect a higher frequency of the 16189 mutation, as reported by others (16), since only one MELAS patient (11%) harbored this polymorphism (controls 16%).

Figure 2. Schematic representation of haplotypes observed in 10 selected LHON subjects belonging to haplogroup D\textsuperscript{4216+11251+10398+12612+13708}. D loop polymorphisms are given in italics.
### DISCUSSION

The mitochondrial genome had been fully sequenced 16 years ago (3) and many subsequent publications worldwide have revealed legions of polymorphic sites, as well as population-specific lineages, haplogroups and haplotypes. Nevertheless, a valid systematic classification remains elusive, although various authors have created their own nomenclature (12,18–20). The rich variability within the D loop region compared with the relatively constant constellation within the gene regions provides useful criteria for phylogenetic and pathogenetic studies. Furthermore, the fact that distinct nucleotide positions have different tendencies to mutate has important implications for phylogenetic and pathogenetic studies. We therefore suggest a classification of populations based on mtDNA gene variants. Our study demonstrates that the German population can be exactly characterized by a few haplogroups defined by common clusters of homoplasmic mtDNA mutations within the coding regions of the mitochondrial genome. Due to the lower mutability within the gene regions compared with the non-coding D loop, our haplogroup definition is affected with a much lower degree of uncertainty compared with the use of D loop sequences. For example, most of the identified coding region variants (~70%) occurred at least twice but showed strict haplogroup specificity, while ~30% of variants occurred only once in our control group. The only exception was the variant at nt 10398, which must have arisen multiple times during evolution as we found it in different haplogroups. In contrast, only ~25% of D loop polymorphisms could be classified as haplogroup-specific variants and the majority represented highly variable sites or very rare variants. Furthermore, the use of gene mutations leads to a more precise definition, allowing characterization of common extant sub-branches. Nevertheless, as we investigated our control population also with respect to the two hypervariable D loop segments we are able to compare our data with those described in the literature.

Comparing our results with those recently obtained by a combined RFLP/sequencing approach (19,21) and a HVS-I sequencing approach (20) lends support to the existence of the major German lineages and haplogroups (Table 3). In our control population a transition at nt 7028, corresponding to an AluI site at nt 7025 in Torroni et al. (19,21), separates the lineage D^C, which is 7028(−) in our nomenclature, from all other lineages. Furthermore, among 7028(+) subjects the major lineages 4216+11251, 12308+12372 and 12705 are all characterized by the transitional variant 073 in our study, a polymorphism also mentioned by Richards et al. (20) as a major separator between the consensus haplogroup (Cambridge Reference Sequence) and four longer branches. The main lineage 4216+11251, which is strongly correlated with the HVS-I variant 16126 (Fig. 1), diverges into haplogroups D^4216, I^11251 and D^12308. These well-defined haplogroups were characterized by a number of homoplasmic gene mutations as well as D loop variants. They seem to be consistent with the haplogroups J and T in Torroni’s grouping and 2A and 2B in Richard’s nomenclature. Both haplogroups were reported to occur only in European populations. However, as mentioned by Richards et al. (20), the ancestral state of haplogroups 2A and 2B, namely the 16126 mutation alone, has only been found in the Middle East. Our second major lineage, 12308+12372, splits into three haplogroups designated D^12308, I^12308 and D^12308, each characterized by several homoplasmic mitochondrial gene mutations. D^12308 is probably consistent with Richard’s group 4 and 12705 branched into a 10398(−) and a 10398(+) haplogroup. The lineage 12308+12372, which is associated with nt 3197, indicates that D^12308 is consistent with groups 5 and U in Richard’s (20) and Torroni’s (21) studies respectively. Our third sub-cluster of lineage 12308+12372, haplogroup D^12308, although not equivalent to any specific haplogroup described in other studies, might also belong to Torroni’s haplogroup U. Lineage 12705 branched into a 10398(−) and a 10398(+) haplogroup. The latter (D^12705) was defined by several further mitochondrial gene and D loop variants; it is likely to be consistent with Richard’s group 3A. A sub-branch of D^12705 is defined by a heteroplasmic C insertion at nt 574, described by Torroni et al. (19) as a characteristic part of their haplogroup I. The lineage D^C in our study, called haplogroup H in Torroni’s nomenclature, is equivalent to group 1 of Richards et al. (20). It is by far the most frequent group and comprises between 40 and 70% of the

### Table 3. Comparison of major German lineages and haplogroups with Caucasian data from the literature

| Lineage | This study (SSCP/sequencing) | Torroni et al. (21*) (RFLP/D loop sequencing) | Richards et al. (20*) (D loop sequencing) |
|---------|-------------------------------|--------------------------------------------|------------------------------------------|
|         | Frequency                     | Frequency                                  | Divergence (years)                       | Divergence (years) |
| 7028(−) | D^7                           | H                                          | 40%                                      | 22 500–45 000     |
| 4216+11251 | D^4216 +11251                | J                                          | 3–15%                                   | 10 500–21 000     |
| 12308+12372 | D^12308 +12372                | T                                          | 6–22%                                   | 8276–10 909       |
| 12705+16233 | D^12705 +16233               | U                                          | 10–16%                                  | 51 034–67 273     |
| 15904  | D^15904                       | W,X?                                       | 4%                                      | 10 500–21 000     |

*Data are from Torroni et al. (21) and Richards et al. (20).
European population. However, as this group is defined by a substraction of the other group-specific polymorphisms, it remains the most poorly defined group which, obviously, will be subdivided into further groups as our study continues. Examples of this are the rare haplogroups D\textsuperscript{C}1341, D\textsuperscript{C}3338, D\textsuperscript{C}5580, D\textsuperscript{C}10463, D\textsuperscript{C}14905 and D\textsuperscript{C}4336, which were 7028(–) and were found infrequently in our control population. D loop sequences of all 31 members of lineage D\textsuperscript{(C)} (Table 1) revealed 29 different haplotypes. Thus the existence of major and frequent haplogroups within this lineage seems unlikely.

The relative frequencies of the haplogroups described differ slightly within the central European area, as shown in Table 3. Haplogroups D\textsuperscript{C}1216, D\textsuperscript{C}13708 and D\textsuperscript{C}4917 were found with a frequency of 7 and 12% respectively in our German control group. The distribution of both haplogroups, however, differs considerably across the Middle East and Europe (20,21,23,24). For instance, Torroni’s group T (our haplogroup D\textsuperscript{C}1216 and Richards’ group 2B) was not detectable in Turkey or Iceland (20), but was found with higher frequencies in the Middle East (12%), in Sardinia, Spain, Portugal (all 13%) (20,22) and, to the remarkably high degree of 75%, in a Ladin speaking isolate in Val Gardena in the northern part of Italy (24). It should be noted here that two of our DIDMOAD index patients belonging to haplogroup D\textsuperscript{C}1216 were of Turkish origin (12). Richard’s haplogroup 2A (our haplogroup D\textsuperscript{C}1216) also varies widely within European populations, ranging from 2% in the Basques to 22% in Cornwall (20). Sub-clusters of lineage 12308+12372, averaging 7% in European countries (19,20), could not be found among Asians or Native Americans (20). Lineage 12705 is also found to a large degree in the Middle East (24%) and in Turkey (18%), but only to a much lesser extent in several European populations (20). We presume that part of the reported differences in the central European mitochondrial haplogroup distribution is due to uncertainties produced by hypervariable hotspots in the D loop sequences.

Recently Zischler et al. (6) described an integration of a mitochondrial D loop fragment into human chromosome 11, where it presumably evolves as a chromosomal pseudogene (‘fossil’) and, therefore, more slowly than its mitochondrial counterparts. The common ancestor of the insertion and the contemporary mitochondrial gene pool probably originated shortly before the genesis of a common ancestor of the current mitochondrial gene pool. When compared with this nuclear ‘fossil’, the Cambridge reference sequence is found to contain 25 base exchanges. Of the mitochondrial lineages and haplogroups described here, only the 12308+12372 and 12705 lineages share several additional positions with the fossil sequence, namely an A at nt 16129, a C at nt 16249, a T at nt 16278, a C at nt 16311, a C at nt 16263 and a C at 16356. As pointed out by Torroni et al. (21), haplogroup U (our haplogroup D\textsuperscript{C}1216) is the only one that Europeans share with Africans. Thus the 12308+12372 and, probably, 12705 lineages seem to be the oldest ones, originating possibly in Africa and expanding into the Middle East and Europe (21).

There also exists some uncertainty about the tree position and the value of the A→G exchange at nt 10398. While Torroni et al. (19,21) suggested that the presence or absence of this DdeI site at nt 10394 separates all Caucasian mtDNAs into two major groups, our data suggest this polymorphism as being a hotspot occurring in different haplogroups. Otherwise, the haplogroups D\textsuperscript{C}1216 and D\textsuperscript{C}4917 would not share variants 4216+11251+16126 as an ancestral lineage and, furthermore, the major lineages 11251+4216+16126, 12308+12372 and 12705 must have arisen at least twice during evolution. A further possibility is that 10398 is indeed very ancient, but underwent back-mutation in certain haplogroups.

Most of the D loop variants characteristic of certain haplogroups have also been described in the skeleton network recently published by Richards et al. (20), summarizing the Caucasian HVS-I sequence data from DiRienzo and Wilson (22), Pierry et al. (23) and Pult et al. (25). Thus even at the level of the HVS-I fragment the Caucasian population can be identified by a limited number of variants. On the other hand, several positions within the D loop sequences represent mutational hotspots. Furthermore, most of the D loop positions investigated in the present and other studies (>80%) remain unaffected. Thus D loop positions seem not to be randomly affected by mutational events. Rather, selection of mutational sites seems to be taking place. If so, this would have important implications for phylogenetic studies, because widely used methods such as maximum parsimony analysis (22,26–28) and pairwise difference analysis (22,29) usually assign an equal weight to all nucleotide exchanges, as recently stated by Torroni et al. (21). Hence, for further phylogenetic studies the single D loop positions would have to be weighted individually.

The disease groups were chosen for several reasons. The LHON and MELAS groups consisted of patients routinely diagnosed in our laboratory. They represented typical maternally transmitted disorders with specific homoplasmic (most LHON patients) or heteroplasmic (all MELAS patients) mtDNA point mutations. Diagnostic studies 3 years ago on the LHON group had revealed that the bulk of the LHON patients were associated with a cluster of so-called secondary LHON mutations (9). By extending these studies to a group of healthy German subjects we found the frequency of haplogroup D\textsuperscript{C}1216, D\textsuperscript{C}13708 to be significantly higher in our LHON cases compared with controls (12,15), irrespective of the nature of the primary/intermediate LHON mutations. D loop sequencing of 10 LHON patients belonging to this haplogroup revealed that the 11778, 14484 and 15257 LHON mutations have arisen several times independently within haplogroup D\textsuperscript{C}13708 (Fig. 2). The accumulation of LHON on this haplogroup background is unlikely to be the result of founder mutations, thus haplogroup D\textsuperscript{C}13708 seems prone to formation or expression of LHON mutations. The observation that another disorder exhibiting optic atrophy in combination with diabetes mellitus is concentrated in haplogroup D\textsuperscript{C}4917 (12) makes disease susceptibility at the level of lineage 4216+11251+16126 likely. DIDMOAD has been proposed to be a mitochondrial-mediated disease (30). However, it lacks maternal transmission and is thus unlikely to be commonly caused by severe mtDNA mutations (12,31,32). It has recently been mapped to chromosome 4p (33).

Children who have died from SIDS have been included in a mtDNA screening program and haplotype analysis for the following reasons. SIDS is the most common cause of death in infants from the first week to the first year of age, however, despite extensive global research the reasons and the pathomechanisms triggering the fatal event remain generally unknown. In a relatively high proportion SIDS is preceded by one or more so-called ‘near miss for SIDS’ or ‘apparent life threatening events (ALTE)’ which are characterized by respiratory arrest, cyanosis, increased temperature with massive sweating, etc. Obviously
ALTE symptoms can be related to alterations in coupling of the respiratory chain and the oxidative phosphorylation apparatus. Furthermore, certain mtDNA mutations might affect coupling (2) and this could have pronounced effects, especially in the mitochondria-rich brown adipose tissue, which regulates thermogenesis in the newborn during the first year of life. Although a detailed mtDNA screening for disease mutation by our group (unpublished results) did not reveal any clear-cut evidence for SIDS-causing mtDNA mutations, our haplogroup analysis showed a different distribution compared with our control population. First, the SIDS infants accumulated in a side-branch of 4216+11251, characterized by additional variants at nt 10398 and 12612, a haplogroup not found in 67 German controls. Second, 3/9 SIDS children carried a heteroplasmic C insertion at nt 574 while only two control subjects (3%) carried this polymorphism, in both cases on a haplogroup D12905 background. These unstable C stretches in the non-coding region have been previously discussed as predisposing to somatic mutations (19), as also suggested for unstable stretches of C residues around nt 16189 (34). Whether this variant plays a role in susceptibility for SIDS needs further investigation.

The association of different disorders with certain mitochondrial haplogroups seems to us the most important result of our study. If this can be confirmed by further studies it will have unforeseen consequences and implications for genetic research in general. If the idea of disease-susceptible haplogroups holds true, populations with, for example, higher frequencies of haplogroups D12905 and D21697 should show an increase in the penetrance of distinct disorders. On the other hand, populations with higher frequency of haplogroup D could be relatively protected from such diseases.

Is there any rational explanation for the association of certain mtDNA haplogroups and diseases? mtDNA encodes several peptide subunits necessary for optimal functioning of the respiratory chain and the oxidative phosphorylation apparatus in the mitochondrion, responsible for the essential production of ATP in all organs and cells, especially in those with high energy demand such as brain, heart and skeletal muscle. It is therefore not surprising that single mtDNA gene mutations cause functional deficiencies in the corresponding gene products, thus leading to insufficient function of the energy-providing machinery and, ultimately, to monogenic mitochondrial diseases of the brain, heart, skeletal muscle and neuroendocrine system (2). To our knowledge there are only a few publications dealing with the role of polymorphic mitochondrial genetic backgrounds in cellular function and individual cellular fitness. In 1991 Dionne et al. (35) correlated the mtDNA RFLP pattern of athletes with maximal oxygen uptake and response to endurance training and found that distinct sequence variation clusters were associated with different levels of physical fitness. Recently Vergani et al. (36) showed that the LHON 11778 mutation leads to a deficiency in ATP production, especially when associated with the polymorphic site at nt 4216. This study supplied, for the first time, a functional basis for the assumption that lineages 4216+11251 might increase the risk of expression and penetrance of LHON and, possibly, other disorders. Thus the disease susceptibility of certain haplogroups might be caused by relative differences in energy production due to synergistically working point mutations, as already proposed by others (10,11,13). Furthermore, certain haplogroups (or specific mtDNA variants) might be more prone to the formation of disease-causing mutations; this possibility has been discussed for the heteroplasmic C stretches at nt 574 (19) which were detected with high frequency in our SIDS group. A third possibility is that different mtDNA mutation clusters might generate different amounts of oxygen radicals, with their respective consequences on DNA, protein and lipid structures. However, up to now evidence for any such explanation is weak; obviously, further studies are needed.

**Materials and Methods**

**Glossary**

As several mitochondrial lineages and sub-clusters are population specific, we used more informative and distinctive terms for the haplogroups characterized in our study. Thus the ‘D’ in, for example, our term D21697 defines a German population; furthermore, the upper mtDNA position indicates the branching site defining the mtDNA lineage, while the lower mtDNA position characterizes the sub-cluster of mitochondrial haplogroup. Mitochondrial haplotypes are defined by a cluster of recurrent D loop variants belonging to one defined haplogroup.

**Subjects**

Our control group consisted of 67 healthy German blood donors who were collected randomly with respect to age and sex. All eight DIDMOAD patients have been described previously (12). The LHON group consisted of 55 patients clinically defined by bilateral optic neuropathy. All nine MELAS patients presented with typical clinical features of this syndrome. LHON and MELAS patients were analyzed for the respective disease-causing mtDNA mutations during routine diagnosis in our laboratory. Cardiac muscle probes from each SIDS children were provided by the Department of Forensic Medicine, Munich.

**DNA extraction**

Total DNA was prepared from tissue (SIDS disease group) or blood cells by standard proteinase K digestion at 65°C for 2–4 h and NaCl extraction. DNA was precipitated using isopropanol at −20°C, washed in 70% ethanol, vacuum dried and resuspended in 10 mM Tris–HCl, 1 mM EDTA. DNA concentration was assayed spectrophotometrically at 260 nm.

**PCR**

PCR amplification was performed on 0.2 μg DNA template using Taq DNA polymerase (Perkin-Elmer). The cycle profile was as follows: 94°C 5 min; 30 cycles of 60°C 1 min, 72°C 2 min, 94°C 1 min; a final cycle of 60°C 1 min, 72°C 4 min. The primers specific for the light (L) and heavy (H) strand of the mtDNA (3) were designed using the OLIGO programme for DOS (v.3.3). Upstream primers: L516 (516–536), L3153 (3153–3172), L4237 (4237–4259), L4430 (4430–4452), L5342 (5424–5447), L6325 (6325–6351), L7458 (7458–7487), L9964 (9964–9995), L10289 (10289–10314), L11158 (11158–11177), L11692 (11692–11714), L12210 (12210–12237), L12495 (12495–12525), L13324 (13324–13346), L14108 (14108–14131), L14415 (14415–14433), L14815 (14815–14842), L15015 (15015–15042), L15149 (15149–15174) and L15806 (15806–15833). Downstream primers: H768 (768–748), H3426 (3426–3404), H4505 (4505–4482), H4678 (4678–4656), H5920 (5920–5900), H6667 (6667–6639), H7817 (7817–7788), H10289 (10288–10264), H10628 (10628–10601),
HI1502 (11502–11478), H12017 (12017–11995), H12549 (12549–12527), H12788 (12788–12763), H13735 (13735–13713), H14347 (14347–14328), H14862 (14862–14842), H15161 (15161–15138), H15315 (15315–15293), H15522 (15522–15496) and H16075 (16075–16052). Amplifications with these primers were carried out at an annealing temperature of 60°C. The primers for amplification of hypervariable segments in the non-coding region were L16064 (16064–16084)/H16433 (16433–16403) and L16538 (16538–16561)/H363 (363–346), with calculated annealing temperatures of 61°C (L16064/H16433) and 54°C (L16538/H363). All oligonucleotides were chemically synthesized by MWG-Biotech (Ebersberg, Germany).

**Single-strand conformation polymorphism (SSCP) analysis**

SSCP was performed using non-denaturing 10% polyacrylamide gels (10% acrylamide, 10% bisacrylamide, 1× TBE, 10% glycerine) as previously described (14). Aliquots of 3.5 µl PCR product were mixed with 5.5 µl formamide, denatured for 5 min at 95°C, placed immediately on ice and loaded into the preformed slots. The electrophoretic separation was carried out using 1× TBE, 10% glycerine as electrophoresis buffer on a horizontal electrophoresis apparatus for 20 h at 180 V and 4°C. Single strands were visualized by sensitive silver staining according to the manufacturer's instructions (Pharmacia).

**Direct sequencing of PCR products**

PCR-amplified fragments were purified by electrophoretic separation on 1.5% agarose gels, excision of the DNA band of interest and extraction of DNA using quick-spin columns (Qiagen). Aliquots of 0.1 µg DNA were cycle sequenced according to the chain termination method using the Delta Taq Cycle Sequencing Kit (US Biochemical, Cleveland, OH) and [35S]dATP incorporation (0.5 µl, 1000 Ci/mmol; ICN). The reaction products were finally denatured by addition of formamide and heating and analyzed using denaturing polyacrylamide electrophoresis. For autoradiography we used Hyperfilm-beta max autoradiography films (Kodak).

**Statistical analysis**

Differences in haplogroup frequencies between disease and control groups were assessed by Fisher's two-tailed exact test.

**ACKNOWLEDGEMENTS**

We thank the Gesundheitsamt München for providing us with samples of a defined group of German blood donors, the Department of Forensic Medicine, University of Munich, for providing us with tissue samples of SIDS infants and several hospitals in Germany for allowing us to investigate their DIDMOAD and LHON patient samples. Correction of the manuscript by Raji Jayaraman is gratefully acknowledged. This study was partially supported by grants from the Deutsche Forschungsgemeinschaft (Ge 213/7-1 and Ja 802/1-1).

**ABBREVIATIONS**

(d)ATP, (deoxy)adenosine triphosphate; COX, cytochrome c oxidase; cyt b, cytochrome b; D loop, displacement loop; DIDMOAD, diabetes insipidus, diabetes mellitus, optic atrophy, deafness; HVS, hypervariable segment; LHON, Leber’s hereditary optic neuropathy; MELAS, mitochondrial encephalomyopathy, lactic acidosis and stroke-like episodes; mtDNA, mitochondrial DNA; ND, NADH dehydrogenase; RFLP, restriction fragment length polymorphism; SIDS, sudden infant death syndrome; SSCP, single-strand conformation polymorphism; TBE, Tris, borate, EDTA.

**REFERENCES**

1. Weissman, S.M. (1995) Genetic basis for common polygenic diseases. *Proc. Natl. Acad. Sci. USA*, 92, 8543–8544.
2. Wallace, D.C. (1994) Mitochondrial DNA sequence variation in human evolution and disease. *Proc. Natl. Acad. Sci. USA*, 91, 8739–8746.
3. Anderson, A., Bankier, A.T., Barrell, B.G., de Bruijn, M.H.L., Coulson, A.R., Drouin, J., Eperon, I.C., Nierlich, D.P., Roe, B.A., Sanger, F., Schuler, G.P., Smith, A.H.J., Staden, R. and Young, I.G. (1981) Sequence and organisation of the human mitochondrial genome. *Nature*, 290, 547–553.
4. Cann, R., Stoneking, M. and Wilson, A.C. (1987) Mitochondrial DNA and human evolution. *Nature*, 325, 311–315.
5. Vigilant, L., Stoneking, M., Harpending, H., Hawkes, K. and Wilson, A.C. (1991) African populations and the evolution of human mitochondrial DNA. *Science*, 253, 1503–1507.
6. Zischler, H., Geisert, H., von Haeseler, A. and Paabo, S. (1995) A nuclear ‘fossil’ of the mitochondrial D-loop and the origin of modern humans. *Nature*, 378, 489–492.
7. Ozawa, T., Tanaka, M., Ino, H., Ohno, K., Sano, T., Wada, Y., Yoneda, M., Tanno, Y., Miyatake, T., Tanaka, T., Iyomaya, S., Ikebe, S., Hattori, N. and Mizuno, Y. (1991) Distinct clustering of point mutations in mitochondrial DNA among patients with mitochondrial encephalomyopathies and with Parkinson’s disease. *Biochem. Biophys. Res. Commun.*, 176, 938–946.
8. Shoffner, J.M., Brown, M.D., Torroni, A., Lott, M.T., Cabell, M.F., Mirra, S.S., Beal, M.F., Yang, C.-C., Gearing, M., Salvo, R., Watts, R.L., Juncos, J.L., Hansen, L.A., Cram, B.J., Fayad, M., Reckord, C.L. and Wallace, D.C. (1993) Mitochondrial DNA variants observed in Alzheimer disease and Parkinson disease patients. *Genomics*, 17, 171–184.
9. Obermaier-Kusser, B., Lorenz, B., Schubring, S., Popratto, A., Zerres, K., Meitinger, T., Meire, F., Cochaux, P., Blankenagel, A., Komm娘娘, G., Jaksh, M. and Gerbitz, K.-D. (1994) Features of mtDNA mutation patterns in European pedigrees and sporadic cases with Leber hereditary optic neuropathy. *Am. J. Hum. Genet.*, 55, 1063–1066.
10. Lerttit, P., Kapsa, R., Jean-Francois, M.J.B., Thyagarajan, D., Noer, A.S., Marzuki, S. and Byrne, E. (1994) Mitochondrial DNA polymorphism in disease: a possible contributor to respiratory dysfunction. *Hum. Mol. Genet.*, 3, 1973–1981.
11. Brown, M.D., Torroni, A., Reckord, C.L. and Wallace, D.C. (1995) Phylogenetic analysis of Leber’s hereditary optic neuropathy mitochondrial DNA’s indicates multiple independent occurrences of the common mutations. *Hum. Mutat.*, 6, 311–325.
12. Hofmann, S., Bezdol, R., Jaksh, M., Obermaier-Kusser, B., Mertens, S., Kaufhold, P., Rahl, W., Hecker, W. and Gerbitz, K.-D. (1997) Wolfram (DIDMOAD) syndrome and Leber hereditary optic neuropathy (LHON) are associated with distinct mitochondrial DNA haplotypes. *Genomics*, 39, 8–18.
13. Torroni, A., Petrozzi, M., D’Urbanlo, L., Sollitto, D., Zeviani, M., Carrara, F., Carducci, C., Leuzzi, V., Carelli, V., Barbomio, P., D Neri, A. and Scozzari, R. (1997) Haplotype and phylogenetic analysis 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 11758 and 14484. *Am. J. Hum. Genet.*, 60, 1107–1121.
14. Jaksh, M., Gerbitz, K.-D. and Kilger, C. (1996) Screening for mitochondrial DNA (mtDNA) point mutations using nonradioactive single strand conformation polymorphisms (SSCP) analysis. *Clin. Biochem.*, 28, 503–509.
15. Hofmann, S., Bezdol, R., Jaksh, M., Kaufhold, P., Obermaier-Kusser, B. and Gerbitz, K.-D. (1997) Disease relevance of the so-called secondary LHON mutations. *Am. J. Hum. Genet.*, 60, 1539–1542.
16. Morton, K.J., Poulton, J. and Sykes, B. (1995) Multiple independent occurrence of the 3243 mutation in mitochondrial tRNAleu(UUR) in patients with the MELAS phenotype. *Hum. Mol. Genet.*, 4, 1689–1691.
17. Marchington, D.R., Poulton, J., Sellar, A. and Holt, I.J. (1996) Do sequence variants in the major non-coding region of the mitochondrial genome influence mitochondrial mutations associated with disease? *Hum. Mol. Genet.*, 5, 473–479.

18. Stoneking, M., Hedgecock, D., Higuchi, R.G., Vigilant, L. and Erlich, H.A. (1991) Population variation of human mtDNA control region sequences detected by enzymatic amplification and sequence-specific oligonucleotide probes. *Am. J. Hum. Genet.*, 48, 370–382.

19. Torroni, A., Lott, M., Cabell, M., Chen, Y., Lavergne, L. and Wallace, D.C. (1994) MiDNA and the origin of Caucasians: identification of ancient Caucasian-specific haplogroups, one of which is prone to a recurrent somatic duplication in the D-loop region. *Am. J. Hum. Genet.*, 55, 780.

20. Richards, M., Corte-Real, H., Forster, P., Macaulay, V., Wilkinson-Herbots, H., Demaine, A., Papina, S., Hodges, R., Bandelt, H.-J. and Sykes, B. (1996) Paleolithic and neolithic lineages in the European mitochondrial gene pool. *Am. J. Hum. Genet.*, 59, 185–203.

21. Torroni, A., Huoponen, K., Fracalacci, P., Petrozzi, M., Morelli, L., Scozzari, R., Obinu, D., Savontaus, M.-L. and Wallace, D.C. (1996) Classification of European mtDNAs from an analysis of three European populations. *Genetics*, 144, 1835–1850.

22. DiRenzo, A. and Wilson, A.C. (1991) Branching pattern in the evolutionary tree for human mitochondrial DNA. *Proc. Natl. Acad. Sci. USA*, 88, 1597–1601.

23. Peercy, R., Sullivan, K.M., Benson, N. and Gill, P. (1993) The application of mitochondrial DNA typing to the study of white Caucasian genetic identification. *Int. J. Legal Med.*, 106, 85–90.

24. Stenico, M., Negro, L., Bertorelle, G., Calafell, F., Capitanio, M., Corrain, C. and Barbujani, G. (1996) High mitochondrial DNA sequence diversity in linguistic isolates of the Alps. *Am. J. Hum. Genet.*, 59, 1363–1375.

25. Pult, I., Sajantila, A., Simanainen, J., Georgiev, O., Schaffner, W. and Piäko, S. (1994) Mitochondrial DNA sequences from Switzerland reveal striking homogeneity of European populations. *Biol. Chem. Hoppe-Seyler*, 375, 837–849.

26. Vigilant, L., Stoneking, M., Harpending, H., Hawkes, K. and Wilson, A.C. (1991) African populations and the evolution of human mitochondrial DNA. *Science*, 253, 1503–1507.

27. Torroni, A., Schurr, T.G., Cabell, M.F., Brown, M.D., MeiJ.V., Larsen, M., Smith, D.G., Vullo, C.M. and Wallace, D.C. (1993) Asian affinities and continental radiation of the four founding Native American mtDNAs. *Am. J. Hum. Genet.*, 53, 563–590.

28. Torroni, A., Sukernik, R.I., Schurr, T.G., Starikovskaya, Y.B., Cabell, M.F., Crawford, M.H., Comuzzie, A.G. and Wallace, D.C. (1993b) MiDNA variation of aboriginal Siberians reveals distinct genetic affinities with Native Americans. *Am. J. Hum. Genet.*, 53, 591–608.

29. Rogers, A.R. and Harpending, H. (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Mol. Biol. Evol.*, 9, 552–569.

30. Bu, X. and Rotter, J.J. (1999) Hypothesis: Wolfram syndrome: a mitochondrial mediated disorder. *Lancet*, 342, 598–600.

31. Van den Ouweland, J.M.W., Brunning, G.J., Linnout, D., Wit, J.-M., Veldhuijzen, E.E. and Maassen, J.A. (1992) Mutations in mitochondrial rRNA genes: non-linkage with syndromes of Wolfram and chronic progressive external ophthalmoplegia. *Nucleic Acids Res.*, 20, 679–682.

32. Barrett, T.G., Bundey, S.E. and Macleod, A.E. (1995) Neurodegeneration and diabetes: UK nationwide study of Wolfram (DIDMOAD) syndrome. *Lancet*, 346, 1458–1463.

33. Polymeropoulos, M.H., Swift, R.F. and Swift, M. (1996) Linkage of the gene for Wolfram syndrome to markers on the short arm of chromosome 4. *Nature Genet.*, 8, 95–97.

34. Bendall, K. and Sykes, B. (1995) Length heteroplasmy in the first hypervariable segment of the human mitochondrial DNA control region. *Am. J. Hum. Genet.*, 57, 248–256.

35. Dionne, F.T., Turcotte, L., Thibault, M.-C., Boulay, R., Skinner, J.S. and Bouchard, C. (1991) Mitochondrial DNA sequence polymorphism, VO₂max, and response to endurance training. *Med. Sci. Sports Exercise*, 23, 177–185.

36. Vergani, L., Martinuzzi, A., Carelli, V., Cortelli, P., Montagna, P., Schievano, G., Carrozzo, R., Angelini, C. and Lugaresi, E. (1995) MiDNA mutations associated with Leber’s hereditary optic neuropathy: studies on cytoplasmic hybrid (cybrid) cells. *Biochem. Biophys. Res. Commun.*, 210, 880–888.