European Journal of Human Genetics (1999) 7, 140–146

http://www.stockton-press.co.uk/ejhg

ARTICLE

Analysis of the trinucleotide CAG repeat from the human mitochondrial DNA polymerase gene in healthy and diseased individuals

Anja Rovio[1], Valeria Tiranti[2], Amy L Bednarz[3], Anu Suomalainen[4], Johannes N Spelbrink[1], Nicolas Lecrenier[5], Atle Melberg[6], Massimo Zeviani[2], Joanna Poulton[3], Françoise Foury[5] and Howard T Jacobs[1,7]

[1]*Institute of Medical Technology and Tampere University Hospital, University of Tampere, Finland*
[2]*Istituto Nazionale Neurologico 'C Besta', Milano, Italy*
[3]*Department of Pediatrics, University of Oxford, UK*
[4]*Department of Human Molecular Genetics, National Public Health Institute, Helsinki, Finland*
[5]*Unité de Biochimie Physiologique, Université Catholique de Louvain, Belgium*
[6]*Department of Neuroscience, University of Uppsala, Sweden*
[7]*Institute of Biomedical and Life Sciences, University of Glasgow, Scotland, UK*

The human nuclear gene (*POLG*) for the catalytic subunit of mitochondrial DNA polymerase (DNA polymerase γ) contains a trinucleotide CAG microsatellite repeat within the coding sequence. We have investigated the frequency of different repeat-length alleles in populations of diseased and healthy individuals. The predominant allele of 10 CAG repeats was found at a very similar frequency (approximately 88%) in both Finnish and ethnically mixed population samples, with homozygosity close to the equilibrium prediction. Other alleles of between 5 and 13 repeat units were detected, but no larger, expanded alleles were found. A series of 51 British myotonic dystrophy patients showed no significant variation from controls, indicating an absence of generalised CAG repeat instability. Patients with a variety of molecular lesions in mtDNA, including sporadic, clonal deletions, maternally inherited point mutations, autosomally transmitted mtDNA depletion and autosomal dominant multiple deletions showed no differences in *POLG* trinucleotide repeat-length distribution from controls. These findings rule out *POLG* repeat expansion as a common pathogenic mechanism in disorders characterised by mitochondrial genome instability.

Keywords: mitochondrial DNA; mitochondrial myopathy; microsatellite; trinucleotide repeat; polyglutamine tract; DNA polymerase; deletions; progressive external ophthalmoplegia; myotonic dystrophy

Correspondence: Howard T Jacobs, Institute of Medical Technology, University of Tampere, PO Box 607, 33101 Tampere, Finland. Tel. or Fax: 358 3 215 7731; E-mail: howy.jacobs@uta.fi
Received 4 March 1998; revised 5 June 1998; accepted 19 June 1998

Introduction

Mitochondrial DNA (mtDNA) mutations are associated with a wide spectrum of human diseases[1,2] and exhibit a variety of different types of inheritance. Clonal deletions or partial duplications of mtDNA typically manifest sporadically, whereas point mutations are usually maternally inherited. Two other kinds of mtDNA lesion appear to depend upon nuclear rather than mitochondrial genotype: mtDNA depletion[3–5] and multiple mtDNA deletions with autosomal dominant or recessive inheritance.[6–9] Autosomal dominant progressive external ophthalmoplegia (ad-PEO) associated with multiple deletions of mtDNA is genetically heterogeneous, with mapped loci on chromosomes 3p14.1–21.2 and 10q24, and other affected families unlinked to either of these.[7,8]

One candidate gene for involvement in mtDNA disorders showing nuclear inheritance is *POLG*, encoding the catalytic subunit of mitochondrial DNA polymerase (DNA polymerase γ). The gene has been mapped to the region of chromosome 15q24–15q26,[10–12] and exhibits an unusual feature not shared with its orthologues in mouse, *Drosophila* or yeast, namely the presence of a trinucleotide (CAG) microsatellite repeat within the N-terminal region of the coding sequence.[11,13] The sequenced gene (Figure 1) contains 10 consecutive, glutamine-encoding CAG codons, followed by a single CAA and two further CAGs. Instability at trinucleotide repeats is associated with various human disorders, including Huntington's disease (HD), myotonic dystrophy (DM), and several forms of spinocerebellar atrophy (eg SCA1). For review see La Spada[14] and Koshy *et al*.[15] Such disorders may be conveniently sub-classified according to where the repeat is found in relation to the coding sequence of the affected gene. Where the repeat is found in coding DNA, as here, such disorders are usually dominant, reflecting a gain of function associated with expanded repeat number. Expansions to 35 or more repeats are usually sufficient to create a pathological phenotype in such cases, eg in HD (see Wellington *et al*[16] for review). Where the repeat is found outside coding DNA (eg in Friedreich ataxia), inheritance is usually recessive, reflecting loss of function, and repeat expansions can be, and usually are, much larger. Different repeat-length alleles of a single gene can be associated with distinct pathological phenotypes, for example in the case of the CAG repeat in the androgen receptor gene, where short 'normal' alleles are associated with enhanced predisposition to prostate cancer,[17] long 'normal' alleles with androgen insensitivity, and expanded alleles ($\geq$ 40 repeats) with Kennedy's disease. Dominant disorders characterised by repeat expansion also exhibit the phenomenon of anticipation: the progressive increase in repeat length and in disease severity from generation to generation.

Expansion of the *POLG* CAG repeat would be a plausible disease mechanism in disorders characterised by mtDNA instability inherited as an autosomal trait, since expansion of the polyglutamine tract might impair the function of the mitochondrial DNA polymerase in a dominant-negative manner. Interestingly, anticipation has been reported in one family with ad-PEO unlinked to the mapped ad-PEO loci on chromosomes 3p or 10q.[18] A modest polyglutamine expansion might also be associated with apparently sporadic cases of mtDNA rearrangement, or might even promote the mutational events that manifest subsequently as diseases associated with inherited point mutations of mtDNA. In addition, dysfunction of the mitochondrial DNA polymerase not associated with repeat size might be of pathological significance. Repeat-length variation in the *POLG* gene might therefore provide a useful linkage marker for evaluating the involvement of the gene in diverse human disorders.

In order to investigate the possible role of the *POLG* gene in human disease, we have studied repeat-length variation at the locus, both in healthy populations and in various groups of patients. Our aim was to look both for gross expansions associated with disease, as well as to examine possibly more subtle relationships between *POLG* repeat-length and pathological phenotype. The results indicate that *POLG* repeat expansion cannot be a frequent cause of mitochondrial disease. Diversity at the *POLG* locus is also lower than optimal for use generally as a linkage marker, but sufficient that the gene can be excluded in specific cases.

Results and Discussion

Nested primers flanking the *POLG* CAG repeat were designed, based on a region of the coding sequence[11,13] contained within a single exon (see Figure 1). The primers amplified the predicted fragments from genomic DNA (Figure 1), as verified by direct sequencing. Testing against the monochromosomal inter-species hybrid and Genebridge 4 radiation hybrid panels[19] confirmed that they detected only the expected gene sequence from chromosome 15q, and not a pseudogene

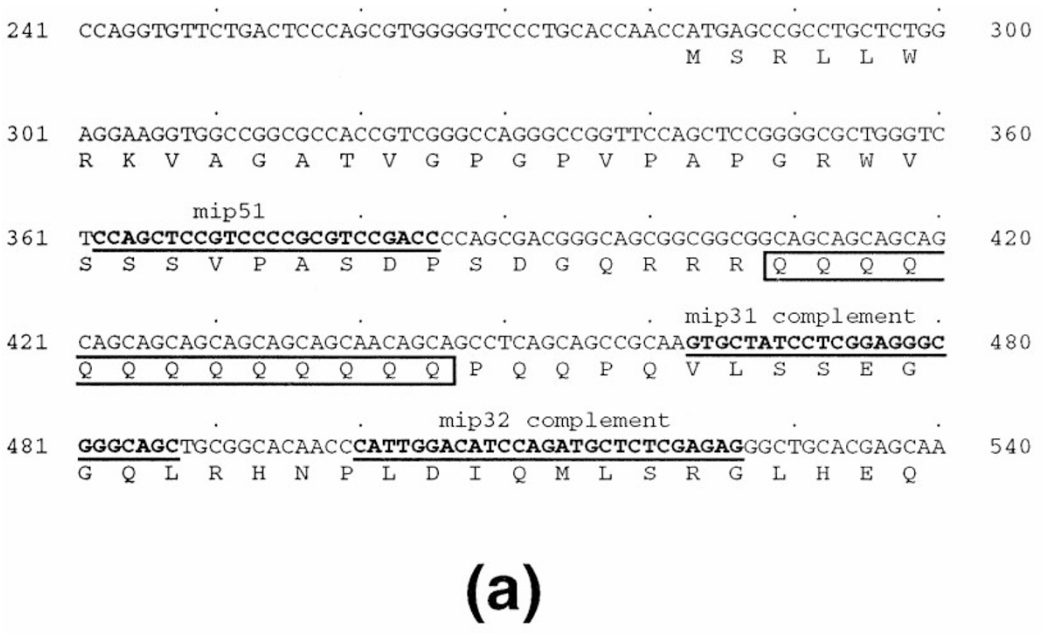

Figure 1 *(a) Partial sequence of the human gene for the catalytic subunit of the mitochondrial DNA polymerase (POLG), from Genbank data entry MIHSDNAPL (accession number X98093), numbered as in the database. The N-terminus of the coding region is shown in the one-letter amino acid code, with the polyglutamine tract boxed. The sequences corresponding to the PCR primers used here (or their complement) are bold, underlined. (b) 2.5% agarose gel of PCR products made using primer pairs mip51/mip31 and mip51/mip32, on human genomic DNA template.*

(data not shown). The 5' primer (mip51) was also tested with other, more downstream 3' primers from the same exon of the *POLG* gene, successfully and routinely amplifying genomic DNA fragments up to 0.6 kb in size (data not shown). This indicates that even grossly expanded CAG repeats (eg of the order of 150) would be detectable.

POLG trinucleotide repeat-length was then evaluated in two groups of healthy controls, plus a number of patient groups, as shown in Table 1, using a fluorescence-based PCR method.[20] Genotype was also confirmed in many cases by direct sequencing of PCR products or by cloning. The previously reported variant of 10 repeats[11,13] was found much more frequently than any other allele. In an ethnically diverse sample of 73 unrelated, healthy persons, it was found at a frequency of 88%. In a sample of 61 unrelated, healthy Finns, a genetically isolated and bottlenecked population that has been much exploited in disease linkage studies, the frequency of the common allele was almost the same (89%). Several groups of patients were also studied, including a series of 51 unrelated DM patients from the UK, and a group of 12 unrelated Chinese patients with acute aminoglycoside ototoxicity (AAO), one of whom had the np 1555 mtDNA mutation. The frequency of the common (10-repeat) allele was not significantly different (z-test, $P > 0.05$) in any group studied, nor was any evidence found for an association between a specific, unusual allele and any of the patient groups investigated. Amongst DM patients, there was furthermore no correlation between disease severity and *POLG* genotype. The overall frequency of the 10-repeat allele in 214 ethnically diverse, unrelated individuals included in the study was 88%. The next commonest allele was of 11 repeats (7%), with other, rarer alleles ranging from 5 to 13 repeats comprising the remainder. Observed homozygosity values for the common allele are close to equilibrium predictions.

No expanded alleles were observed in healthy controls or in any of the patient groups studied, including the 51 DM subjects. This rules out *POLG* repeat expansion as a contributory factor in DM, and is consistent with repeat instability being locus-specific in this disorder. Expanded alleles were also absent from diverse patients with mtDNA disorders, as summarised in Table 2. These results do not exclude the possibility of very large expansions undetectable by conventional PCR, but coding region expansions of the type found in other disorders (i.e. up to several hundred copies) would have been easily observed. Expansions in the CAG repeat of the DMPK gene of over 500 repeat units were easily detected in DM patient DNAs, using primers that amplify a similarly sized (165 bp) fragment spanning the repeat, and otherwise identical materials and methods (data not shown). Most patients we studied were homozygous for the common *POLG* allele, (except where stated in Table 2), but heterozygotes were found amongst several patient groups, in which cases it is possible to exclude even very large expansions undetectable by PCR.

Table 1 Allele frequencies[a] at the POLG trinucleotide repeat in different populations

Allele (no. of repeats)	*Healthy controls[b] (non-Finns)*	*Healthy controls[b] (Finns)*	*Healthy controls (all)*	*DM patients (UK)*	*AAO patients[c] (China)*	*Total[d]*
5	0	0	0	0.01	0	<0.01
7	0	0	0	0	0.04	<0.01
8	0	0	0	0.01	0	<0.01
9	0.01	0.02	0.01	0.01	0.04	0.01
10	0.88	0.89	0.89	0.80	0.92	0.88
11	0.09	0.06	0.07	0.15	0	0.07
12	0.04	0.02	0.03	0.02	0	0.02
13	0	0.01	<0.01	0	0	<0.01
No. of persons	73	61	134	51	12	214
10/10 homozygotes	57 (78%)	48 (79%)	105 (78%)	33 (65%)	10 (83%)	164 (77%)

[a]All allele frequencies quoted to 2 decimal places, except where rounding would give 0.00, which is quoted as <0.01. Therefore, not all columns total 1.00.
[b]Healthy controls had no hearing impairment and reported no manifestations of systemic disease.
[c]Only 12 Chinese patients with acute aminoglycoside ototoxicity (AAO) were included, hence the reported allele frequencies of 0.04 represent single alleles in two heterozygotes.
[d]The final column includes 17 individuals not listed under any other category. Hence the total of 214 persons is greater than the sum of controls, UK DM patients and chinese AAO patients. The additional persons were one hearing-impaired individual, plus 16 English patients with various mitochondrial and non-mitochondrial disorders.

Table 2 Exclusion of *POLG* CAG repeat expansion in cases of mtDNA disorders

Clinical phenotype	*Molecular phenotype*	*Literature reference*	*No. of patients studied*	*Comment*
ad-PEO	multiple mtDNA deletions	8	3	no linkage to ad-PEO loci on chromosomes 3p or 10q
ad-PEO	multiple mtDNA deletions	7	2	Finnish family linked to chromosome 10q24
ad-PEO with anticipation	multiple mtDNA deletions	18	2	Swedish family, no linkage to ad-PEO loci on chromosomes 3p or 10q
PEO	single mtDNA macrodeletion	24	10	sporadic cases
infantile mt myopathy with COX deficiency	mtDNA depletion	25, 26	3, plus unaffected relatives	mtDNA depletion verified in muscle and/or liver: one Italian, two English patients
infantile mt myopathy with COX deficiency	suspected mtDNA depletion	26	3	English patients described in Holt *et al*[24]
fatal neonatal hepatic failure	mtDNA depletion in liver	5	2, plus unaffected relatives	consanguineous family described in Bakker *et al*,[5] heterozygous
MELAS	np 3243 point mutation	27, 28	2	plus 36 unaffected maternal relatives from Finnish and English pedigrees with the mutation
MERRF	np 8344 point mutation	29	2	
Non-syndromic deafness	np 1555 mutation	30	16	Hispanic pedigrees with the mutation; total includes 8 unaffected maternal relatives; 7 heterozygotes
NARP	np 8993	31	1	from UK

Genotyping of the *POLG* repeat furthermore enabled us to exclude the locus from further consideration in the case of the family with hepatic mtDNA depletion, assuming autosomal recessive inheritance.[21] Linkage studies have also excluded the *POLG* region of chromosome 15q from involvement in ad-PEO in Finnish, Italian and Swedish families in which it is unlinked to the previously mapped loci on chromosomes 3 or 10 (A Suomalainen 1998, unpublished data).

The amount of heterozygosity at the *POLG* repeat is probably too low to enable it to be employed as a useful marker in disease linkage studies generally, although it will have utility in specific cases. The fact that a similar pattern of variation is evident in widely scattered populations suggests the possibility of selection, acting to drive the repeat length down (or up) to the modal value of ten CAG repeats. However, all other alleles found amongst patients were also found in controls, hence there is no evidence for association of particular repeat-length variants with any given pathological phenotype.

Methods

DNA Samples

DNA was obtained from blood or muscle biopsy of healthy controls and patients, as described previously,[22] but with proteinase K inactivated by incubation at 92°C for 10 min. All the samples were taken with informed consent. DNA from Chinese patients with acute aminoglycoside ototoxicity was kindly donated by Dr W-Q Qiu (Tiedao Medical College, Shanghai, China), from UK myotonic dystrophy patients by Dr Helen Harley (University of Wales, Cardiff), from Spanish subjects with the np 1555 mtDNA mutation by Dr Ignacio del Castillo (Hospital Ramón y Cajal, Madrid, Spain), and from members of Finnish pedigrees carrying the

np 3243 mtDNA mutation by Dr Kari Majamaa (University of Oulu, Finland). DNA samples for the Genebridge 4 radiation hybrid panel[19] and the monochromosomal human-rodent hybrid panel were supplied by the UK MRC HGMP Resource Centre, Cambridge.

Oligonucleotide primers

Nested oligonucleotides corresponding to regions of the N-terminal coding sequence of *POLG*, as indicated in Figure 1, were purchased from Life Technologies (Paisley, Scotland) or DNA Technology (Aarhus, Denmark). For fluorescent PCR[20] one of the primers (mip51) was 5' pre-labelled with ROX (Perkin-Elmer dye) by the manufacturer.

PCR

Unlabelled PCR reactions used primer pairs mip51/mip31 (126 bp product from reference sequence) or mip51/mip32 (165 bp product) at 0.8 μM each, 200 μM dNTPs (Amersham Pharmacia Biotech, Uppsala, Sweden), plus 0.25 units of Dynazyme thermostable DNA polymerase (Finnzymes, Espoo, Finland) in the manufacturer's buffer. Reactions comprised 30 cycles of denaturation for 1 min at 95°C, annealing for 45 s at 62°C, and extension for 2 min at 72°C (5 min extension in final cycle). Products were analysed by agarose or polyacrylamide gel electrophoresis. For fluorescent PCR, 12.5 μl reactions contained 0.2 μM each primer, 200 μM dNTPs (Amersham Pharmacia Biotech, Uppsala, Sweden), 0.5 μl of DNA template[22] and 0.16 units of Dynazyme thermostable DNA polymerase (Finnzymes, Espoo, Finland) in the manufacturer's buffer, and used the same cycle conditions, except the extension step which was for 1 min only. Reaction products were diluted 1:10 in water, and samples containing 1 μl of diluted PCR product, 12 μl of deionised formanide and 0.15 μl of Tamra Genescan-350 DNA size standards (Perkin-Elmer, Norwalk, CT, USA) were analysed on the ABI 310 Genetic Analyzer (Perkin-Elmer, Norwalk, CT, USA), using the manufacturer's data collection and Genescan analysis software.[20,23]

Cloning and cycle sequencing

Unlabelled PCR products were purified on QIAquick spin columns (Qiagen). Cloning used the TA-cloning kit (Invitrogen), and insert-containing plasmids were identified by miniprepping and *Eco*RI digestion. Cycle sequencing employing the ABI 310 Genetic Analyzer used the Terminator Ready Reaction Kit (Perkin-Elmer), together with unlabelled mip51 primer (direct sequencing of PCR products) or M13 reverse primer (plasmid sequencing).

Acknowledgements

We thank Kari Majamaa, Ignacio del Castillo, Wei-Qin Qiu, Helen Harley and Peter Harper for supplying DNA samples, Erkki Mäkelä for statistical analysis, plus the many patients and healthy controls who took part in the study. Funding for this work was obtained from the Royal Society (University Research Fellowship to JP), Wellcome Trust, British Diabetic Association, Finnish Academy, Muscular Dystrophy Group of Great Britain, Juselius Foundation, Tampere University Hospital Medical Research Fund, Italian Telethon Foundation (grant no. 767 to MZ) and the European Union Human Capital and Mobility programme.

References

1 Larsson NG, Clayton DA: Molecular genetic aspects of human mitochondrial disorders. *Annu Rev Genet* 1995; **29**: 151–178.
2 Fadic R, Johns DR: Clinical spectrum of mitochondrial diseases. *Seminars Neurol* 1996; **6**: 1–20.
3 Moraes CT, Shanske S, Tritschler HJ *et al*: MtDNA depletion with variable tissue expression: a novel genetic abnormality in mitochondrial diseases. *Am J Hum Genet* 1991; **48**: 492–501.
4 Bodnar AG, Cooper JM, Holt IJ, Leonard JV, Schapira AH: Nuclear complementation restores mtDNA levels in cultured cells from a patient with mtDNA depletion. *Am J Hum Genet* 1993; **53**: 663–669.
5 Bakker HD, Scholte HR, Dingemans KP, Spelbrink JN, Wijburg FN, Van den Bogert C: Depletion of mitochondrial deoxyribonucleic acid in a family with fatal neonatal liver disease. *J Pediatr* 1996; **128**: 683–687.
6 Zeviani M, Bresolin N, Gellera C *et al*: Nucleus-driven multiple large-scale deletions of the human mitochondrial genome: a new autosomal dominant disease. *Am J Hum Genet* 1990; **47**: 904–914.
7 Suomalainen A, Kaukonen J, Amati P *et al*: An autosomal locus predisposing to deletions of mitochondrial DNA. *Nat Genet* 1995; **9**: 146–151.
8 Kaukonen JA, Amati P, Suomalainen A *et al*: An autosomal locus predisposing to multiple deletions of mtDNA on chromosome 3p. *Am J Hum Genet* 1996; **58**: 763–769.
9 Bohlega S, Tanji K, Santorelli FM, Hirano M, al-Jishi A, Di Mauro S: Multiple mitochondrial DNA deletions associated with autosomal recessive ophthalmoplegia and severe cardiomyopathy. *Neurology* 1996; **46**: 1329–1334.
10 Walker RL, Anziano P, Meltzer PS: A PAC containing the human mitochondrial DNA polymerase gamma gene (*POLG*) maps to chromosome 15q25. *Genomics* 1997; **40**: 376–378.
11 Ropp PA, Copeland WC: Cloning and characterization of the human mitochondrial DNA polymerase, DNA polymerase gamma. *Genomics* 1996; **36**: 449–458.

12 Zullo SJ, Butler L, Zahorchak RJ, MacVille M, Wilkes C, Merril CR: Localization by fluorescence *in situ* hybridization (FISH) of human mitochondrial polymerase γ (*POLG*) to human chromosome band 15q24 15q26, and of mouse mitochondrial polymerase γ (*Polg*) to mouse chromosome band 7E, with confirmation by direct sequence analysis of bacterial artificial chromosomes (BACs). *Cytogenet Cell Genet* 1997; **78**: 281–284.
13 Lecrenier N, Van Der Bruggen P, Foury F: Mitochondrial DNA polymerases from yeast to man: a new family of polymerases. *Gene* 1997; **185**: 147–152.
14 La Spada AR: Trinucleotide repeat instability: genetic features and molecular mechanisms. *Brain Pathol* 1997; **7**: 943–963.
15 Koshy BT, Zoghbi HY: The CAG/polyglutamine tract diseases: gene products and molecular pathogenesis. *Brain Pathol* 1997; **7**: 927–942.
16 Wellington CL, Brinkman RR, O'Kusky JR, Hayden MR: Toward understanding the molecular pathology of Huntington's Disease. *Brain Pathol* 1997; **7**: 979–1002.
17 Giovannucci E, Stampfer MJ, Krithivas K *et al*: The CAG repeat within the androgen receptor gene and its relationship to prostate cancer. *Proc Natl Acad Sci USA* 1997; **94**: 3320–3323.
18 Melberg A, Arnell H, Dahl N *et al*: Anticipation of autosomal dominant progressive external ophthalmoplegia with hypogonadism. *Muscle Nerve* 1996; **19**: 1561–1569.
19 Gyapay G, Schmitt K, Fizames C *et al*: A radiation hybrid map of the human genome. *Hum Mol Genet* 1996; **5**: 339–346.
20 *ABI PRISM™ 310 Genetic Analyser, GeneScan™ Chemistry Guide*. Perkin Elmer Corporation: July 1995.
21 Spelbrink JN, Van Galen MJM, Zwart R *et al*: Familial mitochondrial DNA depletion in liver: haplotype analysis of candidate genes. *Hum Genet* 1998; **102**: 327–331.
22 Reid FM, Vernham GA, Jacobs HT: A novel mitochondrial point mutation in a maternal pedigree with sensorineural deafness. *Hum Mutat* 1994; **3**: 243–247.
23 Wenz H, Robertson JM, Menchen S *et al*: High-precision genotyping by denaturing capillary electrophoresis. *Genome Res* 1998; **8**: 69–80.
24 Holt IJ, Harding AE, Morgan-Hughes JA: Deletions of muscle mitochondrial DNA in patients with mitochondrial myopathies. *Nature* 1988; **331**: 717–719.
25 Mariotti C, Uziel G, Carrara F *et al*: Early onset encephalomyopathy associated with tissue-specific mitochondrial DNA depletion – a morphological, biochemical and molecular genetic study. *J Neurol* 1995; **242**: 547–556.
26 Poulton J, Sewry C, Potter CG *et al*: Variation in mitochondrial DNA levels in muscle from normal controls. Is depletion of mitochondrial DNA in patients with mitochondrial myopathy a distinct clinical syndrome? *J Inherit Metab Dis* 1995; **18**: 4–20.
27 Goto Y, Nonaka I, Horai S: A mutation in the tRNALeu (UUR) gene associated with the MELAS subgroup of mitochondrial encephalomyopathies. *Nature* 1990; **348**: 651–653.
28 Kobayashi Y, Momoi MY, Tominaga K *et al*: A point mutation in the mitochondrial $tRNA^{Leu}$(UUR) gene in MELAS (mitochondrial myopathy, encephalopathy, lactic acidosis and stroke-like episodes). *Biochem Biophys Res Comm* 1990; **173**: 816–822.
29 Shoffner JM, Lott MT, Lezza AM, Seibel P, Ballinger SW, Wallace DC: Myoclonic epilepsy and ragged-red fiber disease (MERRF) is associated with a mitochondrial DNA tRNALys mutation. *Cell* 1990; **61**: 931–937.
30 Prezant TR, Agapian JV, Bohlman MC *et al*: Mitochondrial ribosomal RNA mutation associated with both antibiotic-induced and non-syndromic deafness. *Nat Genet* 1993; **4**: 289–294.
31 Holt IJ, Harding AE, Petty RKH, Morgan-Hughes JA: A new mitochondrial disease associated with mitochondrial DNA heteroplasmy. *Am J Hum Genet* 1990; **46**: 428–433.