



OPEN ACCESS

ORIGINAL ARTICLE

Prevalence of rare mitochondrial DNA mutations in mitochondrial disorders

Sylvie Bannwarth,^{1,2} Vincent Procaccio,³ Anne Sophie Lebre,⁴ Claude Jardel,⁵ Annabelle Chaussenot,^{1,2} Claire Hoarau,^{1,2} Hassani Maoulida,⁶ Nathanaël Charrier,⁶ Xiaowu Gai,⁷ Hongbo M Xie,⁸ Marc Ferre,³ Konstantina Fragaki,^{1,2} Gaëlle Hardy,⁹ Bénédicte Mousson de Camaret,¹⁰ Sandrine Marlin,¹¹ Claire Marie Dhaenens,¹² Abdelhamid Slama,¹³ Christophe Rocher,¹⁴ Jean Paul Bonnefont,⁴ Agnès Rötig,⁴ Nadia Aoutil,⁵ Mylène Gilleron,⁵ Valérie Desquiret-Dumas,³ Pascal Reynier,³ Jennifer Ceresuela,¹⁰ Laurence Jonard,¹¹ Aurore Devos,¹² Caroline Espil-Taris,¹⁴ Delphine Martinez,⁹ Pauline Gaignard,¹³ Kim-Hanh Le Quan Sang,⁴ Patrizia Amati-Bonneau,³ Marni J Falk,¹⁵ Catherine Florentz,¹⁶ Brigitte Chabrol,¹⁷ Isabelle Durand-Zaleski,⁶ Véronique Paquis-Flucklinger^{1,2}

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/jmedgenet-2013-101604>).

For numbered affiliations see end of article.

Correspondence to

Professor V Paquis-Flucklinger, IRCAN UMR7284/INSERM U1081/UNS, School of Medicine, 28 av de Valombrose, 06107, Nice cedex 2, France; paquis@hermes.unice.fr

Received 13 February 2013

Revised 5 June 2013

Accepted 6 June 2013

Published Online First

11 July 2013

ABSTRACT

Background Mitochondrial DNA (mtDNA) diseases are rare disorders whose prevalence is estimated around 1 in 5000. Patients are usually tested only for deletions and for common mutations of mtDNA which account for 5–40% of cases, depending on the study. However, the prevalence of rare mtDNA mutations is not known.

Methods We analysed the whole mtDNA in a cohort of 743 patients suspected of manifesting a mitochondrial disease, after excluding deletions and common mutations. Both heteroplasmic and homoplasmic variants were identified using two complementary strategies (Surveyor and MitoChip). Multiple correspondence analyses followed by hierarchical ascendant cluster process were used to explore relationships between clinical spectrum, age at onset and localisation of mutations.

Results 7.4% of deleterious mutations and 22.4% of novel putative mutations were identified. Pathogenic heteroplasmic mutations were more frequent than homoplasmic mutations (4.6% vs 2.8%). Patients carrying deleterious mutations showed symptoms before 16 years of age in 67% of cases. Early onset disease (<1 year) was significantly associated with mutations in protein coding genes (mainly in complex I) while late onset disorders (>16 years) were associated with mutations in tRNA genes. *MTND5* and *MTND6* genes were identified as 'hotspots' of mutations, with Leigh syndrome accounting for the large majority of associated phenotypes.

Conclusions Rare mitochondrial DNA mutations probably account for more than 7.4% of patients with respiratory chain deficiency. This study shows that a comprehensive analysis of mtDNA is essential, and should include young children, for an accurate diagnosis that is now accessible with the development of next generation sequencing technology.

INTRODUCTION

Since the first identification of mitochondrial DNA (mtDNA) defects responsible for human diseases in

1988, it has been increasingly clear that mutations in the mtDNA represent an important cause of neuromuscular disorders.^{1 2} Nevertheless, today, the prevalence of mtDNA disease is still difficult to assess for two main reasons.³ First, large series are extremely rare because of clinical heterogeneity and diagnosis complexity.⁴ Second, molecular screening is usually restricted to detection of mtDNA deletions and a few common mutations without having accurate data on the prevalence of rare mutations.

Improvements in diagnosis methods led us to study a large cohort of patients from the French Mitochondrial Disease Network, including both children and adults suspected of manifesting a mitochondrial disorder. In our cohort, deletions and common mutations of mtDNA had already been excluded and we performed a comprehensive analysis of the mitochondrial genome in order to answer a number of key questions.

- What is the prevalence of rare mtDNA mutations in mitochondrial disorders?
- What are the phenotypes associated with these mutations?
- Is a comprehensive screening of mtDNA essential for optimal diagnostic approach?

PATIENTS AND METHODS

Study population

Over a 3 year period, 743 individuals suspected of having a mitochondrial disorder were included in the study. All were studied in French referral centres by hospital specialists. A standardised chart including clinical symptoms, imaging and extensive laboratory work was completed for all patients (ie, familial and clinical history, clinical presentation, brain MRI, metabolic screening, mitochondrial enzymes studies, histological and molecular analyses). Age of onset of clinical symptoms ranged from the neonatal period to 74 years of age. The patient population was based on the following inclusion criteria: (1) clinical features suggesting respiratory chain (RC) dysfunction and/or



Open Access
Scan to access more
free content

To cite: Bannwarth S, Procaccio V, Lebre AS, et al. *J Med Genet* 2013;**50**:704–714.

(2) metabolic screening suggesting RC dysfunction and/or (3) isolated or multiple RC complex deficiency and/or (4) histochemical evidence of mitochondrial abnormality in the muscle biopsy and (5) exclusion of mtDNA deletion(s) and m.3243A>G, m.8344A>G and m.8993T>C/G point mutations. Affected individuals presenting with mtDNA depletion or clear mendelian family history were also excluded. Among the 743 patients, 73 satisfied one, 301 two, 275 three and 94 four inclusion criteria. Patients were subdivided into three groups according to their age at presentation: <12 months of age, 1–16 years of age and >16 years of age.

Blood and tissue samples were obtained after parents of affected children and adult patients had given informed consent. Authorisations for individual data protection were obtained.

OXPPOS spectrophotometric measurements

Spectrophotometric studies of oxidative phosphorylation (OXPPOS) complexes and citrate synthase were performed on tissue homogenates and fibroblasts, as previously described.⁵

Surveyor and MitoChip analysis

The entire human mtDNA was first analysed by the Surveyor method, as previously described.^{6–7} Then, screening was performed with a resequencing chip (MitoChip, GeneChip Mitochondrial Resequencing Array 2.0) according to the manufacturer's instructions (Affymetrix, Inc).⁸ A custom bioinformatics pipeline was then used for MitoChip analysis, as previously described.⁹ Sequences were compared with the human mtDNA consensus sequence, Genbank No J01415.2.¹⁰

Variants reported in the MITOMAP database (<http://www.mitomap.org/MITOMAP>) as polymorphisms and present four times or more in the mtDB database (<http://www.mtodb.igp.uu.se/>) were considered as non-deleterious and excluded from the study.

Statistical analysis

Variables

The variables used refer to clinical data of patients, age at onset (<1 year, 1–16 years, >16 years) and mutation localisation (tRNA genes vs protein coding genes).

Descriptive statistics

Descriptive statistics were used to characterise the population. The unit of analysis was the patient and the analyses were done on the entire population. All variables were categorical.

Multiple correspondence analysis and hierarchical ascendant classification

Analyses were performed on the patient population carrying pathogenic mutations. The unit of analysis was the patient. To explore the genotype–phenotype relationships, we first used multiple correspondence analysis (MCA)¹¹ and next considered the coordinates of the observations on the retained factorial axes as new variables used for hierarchical ascendant classification (HAC).¹² MCA allows a graphical representation of the strength of the association between categories of selected variables. The points corresponding to the categories of each variable (eg, age of onset of disease in three categories) are plotted at distances from the origin that are inversely related to the number of patients in a given category. Benzecri's method was used to select the number of axes considered in the analyses.¹³ MCA was built using age at onset and mutation localisation as supplementary (non-active) variables as they were the variables to explain. The variables with a low frequency (<15%) were

excluded from the analysis and Cramer's V test was performed to identify the variables strongly associated with others, which were analysed as non-active variables. To build homogeneous clusters of patients, we performed HAC based on the Ward method. The clustering process was plotted as a dendrogram, with horizontal branches representing the combination of two clusters and vertical branches the degree of dissimilarity between combined clusters. Long distances of the vertical segments indicated large differences between the clusters. Associations between profiles of patients and variables used in the model were tested with Fisher's exact test. Test values were calculated to measure the association between supplementary variables and profiles of patients. We did not use other clustering methods, such as K means or latent class analysis, because of the small size of the population. We chose hierarchical clustering because of the small datasets and the absence of previous assumptions about the number of clusters. All statistical analyses were performed with SAS V.9.3 software (SAS Institute, Cary, North Carolina, USA). A two tailed $p \leq 0.05$ was considered statistically significant.

RESULTS

The 743 patients were subdivided into 395 men and 348 women (ratio 1.13 : 1) with an average age of 25 years at the time of the study. Clinical data are shown in the online supplementary table S1 for the entire population, and depending on age of onset. Early clinical symptoms were seen during the first year of life for 356 (48%) patients, between 1 year and 16 years for 192 (25.8%) patients and in adulthood for the remaining 180 (24.2%) patients. In 15 (2%) of 743 cases, age of onset was unknown. Neurological presentations were the most frequent (72.7% of total population). Before 1 year, the most common symptom was psychomotor retardation (55.6%) whereas cerebellar ataxia was commonly seen after 1 year of age or during adulthood (37% and 25%, respectively). Epilepsy and myoclonics were also very common (32.2%) whatever the age of presentation. Muscular involvement was among the second rank of clinical disorders, with chronic progressive external ophthalmoplegia (CPEO) mainly seen during childhood and adulthood (27.1% and 32.2%, respectively). Postnatal growth failure was highly represented before 1 year (29.5%). Hypertrophic cardiomyopathy, and liver and kidney deficiencies, were also mainly associated with onset before 1 year (11%, 27.8% and 11.5%, respectively). Optic atrophy was most frequent when the disorder started during childhood (11.5%) whereas peripheral neuropathy was present in 24.4% of patients with adult onset. Paraclinical investigations are presented in the online supplementary table S2. Brain MRI was abnormal in 82.1% of cases before 1 year of age, with a diagnosis of Leigh or Leigh-like syndrome in 43.1% of patients. In patients with adult onset, white matter involvement was the most common pattern (43.8%) whereas posterior fossa atrophy/hypoplasia was the most frequent when onset was during childhood (29.5%). A lactate peak was detected in 60.6% of cases by MRI spectroscopy, mainly in onset before 1 year (68.1%). Hyperlactacidaemia was present in about 50% of patients but hyperlactatorachia was the most common metabolic abnormality (75.3%), mainly during childhood (83.3%) (see online supplementary table S2). Results obtained from tissue biopsies, mainly skeletal tissue, are presented in the online supplementary table S3. The percentage of ragged red fibres (RRF) and COX negative fibres increased with age of onset, reaching 48.9% and 56.9%, respectively, after 16 years of age. In muscle, the biochemical activity of RC complexes was significantly reduced in

59.4% of total biopsies. In liver tissue, a RC deficiency was seen in 53.9% of cases before 1 year of age.

The presence of deletions and common point mutations (m.3243A>G, m.8344A>G and m.8993T>C/G) of mtDNA had previously been ruled out. The first part of the study aimed to search for heteroplasmic mutations by Surveyor analysis. After eliminating the known polymorphisms (list available on request) and, among new sequence variations, the 52 synonymous variants and the 28 nucleotide substitutions in the D loop region (list available on request), a total of 82 different heteroplasmic sequence variations corresponding to putative or known pathogenic mutations were retained (see online supplementary table S4).^{14–16} The 56 different variants corresponding to putative mutations were not reported in the MITOMAP database and were present less than four times in the mtDB database. Two variants only were found twice. The m.1938A>T variant in *MTRNR2* was found in a child who presented with psychomotor delay, epilepsy, ptosis, deafness and white matter lesions on brain MRI. The second patient was an adult presenting with epilepsy, cerebellar ataxia and stroke-like episodes. Neuroimaging revealed patterns of Leigh/MELAS (mitochondrial encephalomyopathy, lactic acidosis and stroke-like episodes) overlap syndrome. In both cases, biochemical analysis revealed no RC deficiency on muscle biopsy. The second variant, m.13604G>C, was in the *MTDN5* gene. The first child died at birth and presented a complex I (CI) deficiency in the liver. The second child presented with myopathy, CPEO and deafness. Muscle biopsy showed RRF and COX– fibres but biochemical analysis was not done.

Twenty-six different sequence variations were already known as pathogenic mutations. The characteristics of the 34 corresponding patients (34/743, 4.6%) are presented in table 1. Four mutations (m.9185T>C, m.10191T>C, m.12706T>C and m.13514A>G) were found in two patients, m.14487T>C in three patients and m.13513G>A in four patients. Recurrent mutations were responsible for Leigh syndrome in these patients, except for one, and encoded subunits of CI, except the m.9185T>C variant located in the *MTATP6* gene (table 1). This work expands the clinical spectrum of mtDNA disorders, with the example of m.636A>G and m.12236G>A mutations that had been previously reported in non-syndromic hearing loss and that are associated with progressive multisystemic disorders in our cohort.

In 74% of cases (25 of 34 patients), symptoms started before 16 years of age (figure 1). When onset was before 1 year, all patients presented a multisystemic disorder with abnormal brain MRI, hyperlactataemia and elevation of the lactate to pyruvate ratio. Leigh syndrome was found in 85% of cases (six of seven patients), and usually associated with a CI deficiency in muscle (four of six patients with muscle biopsy, 66.7%). When the onset was between 1 and 16 years of age, ocular symptoms were present in 44.4% of patients (8/18). In adult onset, cerebellar ataxia was found in 44.4% of cases (four of nine patients) and biochemical analysis identified a multiple RC deficiency in muscle in 42.9% (three of seven patients with muscle biopsy). In agreement with the mitochondrial enzyme data, mutations were mainly located in CI genes before 1 year (five of seven patients, 71.4%) and in tRNA genes in adult onset (six of nine patients, 66.7%) (figure 1). Statistical analysis confirmed these results. MCA, used as a step of data reduction, retained 12 active variables: affected maternal relatives, psychomotor retardation, regression, hypotonia, stroke-like episodes, epilepsy, movement disorders, CPEO, optic atrophy, deafness, postnatal growth retardation and Leigh syndrome. The following variables

were strongly correlated with others and were added as non-active: cerebellar ataxia, white matter involvement and posterior fossa atrophy or hypoplasia. HAC results showed three distinct profiles (figure 1). The first profile of patients was associated with early onset disease (<1 year) and mutations in protein coding genes. The following variables were overrepresented in this group of patients ($p<0.05$): psychomotor retardation, hypotonia, optic atrophy, postnatal growth retardation and Leigh syndrome. The second profile of patients was associated with late onset (>16 years) and mutations in tRNA genes. In this profile, affected maternal relatives were common (36%) ($p=0.13$). We did not find any association for the last profile of patients with a disease onset during childhood (1–16 years). A dendrogram for cluster model is illustrated in the online supplementary figure S1.

The second part of the study was to identify homoplasmic mutations using MitoChip analysis in patients without pathogenic heteroplasmic variants. Known polymorphisms and 180 novel variants (140 synonymous and 40 in the D loop) were excluded (list available on request). The online supplementary table S5 shows the 120 different homoplasmic variants found in 131 patients, which correspond to potential and deleterious mutations. Among the 104 variants corresponding to putative pathogenic mutations, four were present more than once (see online supplementary table S5). The m.3228_3229 insA variant in *MTRNR2* was found in two adults with a mitochondrial myopathy and in one child with Leigh syndrome. The m.8654T>C (p.Ile43Thr) variant in *MTATP6* was associated with exercise intolerance and multiple RC deficiency in muscle of an adult patient. The second affected individual was a child with Leigh syndrome. None showed complex V deficiency. The last two variants were in *MTCYB*. Two adult patients, presenting with exercise intolerance, carried the m.15122A>G (p.Thr126Ala) variant with classical mitochondrial features on muscle biopsy and CIII deficiency in only one of two. The m.15774T>C (p.Val343Ala) variant was associated with two early onset presentations (<1 year). The first child presented with Toni Debré Fanconi syndrome and the second one had a neurological presentation (psychomotor delay, hypotonia and epilepsy). In both cases, RRF fibres were observed on muscle biopsy but we found no RC deficiency.

Sixteen different sequence variations were already known as pathogenic mutations and the characteristics of the 21 corresponding patients (2.8%) are presented in table 2. The m.1555A>G mutation, only seen in a homoplasmic state, was present in three patients with a multisystemic disorder; only one was deaf. Two other mutations were present more than once, and both had been previously found in our study in the heteroplasmic state (m.4317delA, m.13513G>A).

Ocular involvement was observed in two-thirds of patients and deafness in half of affected individuals during childhood (figure 2). However, the population size was too small to perform statistical studies.

Analysis of all (hetero- and homoplasmic) deleterious mutations confirmed the role of rare mtDNA mutations in paediatric presentations (onset <16 years for 37 of 55 patients, 67.3%) (figure 3). Neuromuscular symptoms were clearly predominant in the 55 patients carrying pathogenic mutations. Liver failure was associated with onset before 1 year, ocular involvement was present in 50% of patients (12/24) in childhood presentations and diabetes mellitus and deafness were found in 20.8% of cases in adult onset disorders (5 of 24 patients). Statistical analysis confirmed previous findings. The same active and non-active variables were retained by MCA. Results of HAC on the

Table 1 Characteristics of patients carrying pathogenic mutations identified by Surveyor analysis

Gene	Mutation	Gender/age at onset (years)	Clinical symptoms	Brain MRI	Lactate	L/P	OH-but/ Ac-Ac	Muscle histology	RC
<i>MTTF</i>	m.636A>G	M<1	CNS (En, R, ME), Mu (H, ptosis), GI, anaemia, neutropenia	Leigh, WM, thalamic lesions, lactate peak	↑ (B, CSF)	↑	N	N	↓CI (Mu)
<i>MTTL1</i>	m.3258T>C	1<M<16	CNS (A, R, ME), HCM, Mu (SM)	N	↑ (B, CSF, U)	↑	↑	N	↓CI+IV (Mu)
<i>MTND1</i>	m.3460G>A	F>16	CNS (E, A, MD)	Cerebral atrophy	ND	ND	ND	N	ND
<i>MTT1</i>	m.4317A>G	M>16	CNS (A, My, PS)	Leigh-like, WM, cerebellar atrophy	ND	ND	ND	N	↓CI (Mu)
<i>MTTN</i>	m.4317 delA	F>16	CNS (S-I), Mu (ptosis), D	WM, stroke-like	ND	ND	ND	RRF, COX-	↓CI (Mu)
	m.5667G>A	1<M<16	Mu (ptosis)	ND	ND	ND	ND	RRF, COX-	N (Mu)
	m.5728T>C	1<M<16	Mu (SM, ptosis)	ND	ND	ND	ND	Lipidosis, RRF, COX-	N (Mu)
<i>MTTS1</i>	m.7472insC	F>16	CNS (PaS), Mu (SM, myalgias), DM	ND	ND	ND	ND	RRF, COX-	N (Mu)
<i>MTTK</i>	m.8362T>G	F>16	CNS (A), D, DM, lipomas	ND	ND	ND	ND	ND	ND
<i>MTATP6</i>	m.8609_8610insC	M<1	CNS (PMR, E, R, A, Dys), Mu (H), PN-G	Cerebellar atrophy	↑ (B)	↑	ND	ND	↓CV (Mu) N (F)
<i>MTATP6</i>	m.9185T>C	1<F<16	CNS (PMR, R, A), Re, IU-G, PN-G, anaemia	Leigh	↑ (B)	↑	N	Lipidosis, RRF	↓CV (Mu)
		1<M<16	CNS (PMR, A), PN	Leigh, cerebellar atrophy	↑ (CSF)	↑	N	N	N (Mu) N (F)
<i>MTND3</i>	m.10191T>C	M<1	CNS (PMR, A, E, R, S-I, Dys), Mu (PEO), OA, IU-G	Leigh, stroke-like, thalamus lesions, WM, lactate peak	↑ (B, CSF)	↑	ND	Lipidosis	↓CI (Mu)
<i>MTND4</i>	m.11778G>A	1<F<16	CNS (Dys, dementia, PS), OA, Mu (SM, ptosis)	WM, cerebral atrophy	↑ (U)	ND	ND	N	↓CI+III (Mu)
<i>MTTS2</i>	m.12236G>A	F>16	CNS (A), PN, Mu (myalgias)	ND	ND	ND	ND	N	↓CI+III+IV (Mu) ↓CI+III+IV (F)
<i>MTTL2</i>	m.12276G>A m.12316G>A	1<F<16	Mu (SM, PEO), PR	ND	↑ (B)	ND	ND	RRF	ND
		M>16	Mu (PEO), PN	N	ND	ND	ND	Lipidosis, RRF, COX-	↓CI+IV (Mu)
<i>MTND5</i>	m.12414delT m.12706T>C	1<F<16	Mu (exercise intolerance)	ND	ND	ND	ND	Lipidosis	↓CI (Mu)
		1<M<16	CNS (PMR, E, A), OA, D, PN-G	Leigh, stroke-like, cerebellar atrophy	ND	ND	ND	ND	↓CI (Mu) N (F)
<i>MTND5</i>	m.13513G>A m.13514A>G	M>16	CNS (S-I, En)	ND	ND	ND	ND	N	↓CI+III (Mu) ↓CI (F)
		F<1	CNS (PMR, S-I), Mu (H, ptosis), HCM, OA, D	Leigh, stroke-like,	↑ (B)	↑	↑	RRF	↓CI (Mu)
		M<1	CNS (PMR, R, S-I, M-I H, B) PN, Mu (PEO)	Leigh, stroke-like, WM	↑ (B, CSF)	↑	↑	N	ND
		1<F<16	CNS (A, ME), cataract Re, D, PN-G	Leigh cerebral atrophy	↑ (B)	↑	ND	RRF	N (Mu)
		1<F<16	CNS (A), OA, PN-G	Leigh, cerebellar atrophy	↑ (B)	ND	ND	Lipidosis, RRF	↓CI (Mu)
<i>MTND6</i>	m.14487T>C	1<M<16	Mu (ptosis), VA, leucopenia	Leigh	N	↑	N	Lipidosis RRF	N (Mu) N (F)
		F<1	CNS (PMR, A), Mu (H), L, PN-G	Leigh	↑ (B, CSF)	↑	↑	Lipidosis	N (Mu) N (L)
<i>MTND6</i>	m.14487T>C	1<F<16	CNS (PMR, S-I), Mu (PEO), PN-G	Leigh	↑ (CSF)	↑	↑	N	N (Mu)
		1<M<16	CNS (R, Dys, PS)	Leigh, stroke-like	↑ (B)	↑	ND	ND	N (Mu) N (F)

Continued

Table 1 Continued

Gene	Mutation	Gender/age at onset (years)	Clinical symptoms	Brain MRI	Lactate	L/P	OH-but/ Ac-Ac	Muscle histology	RC
MTTE	m.14535_14536insC	M<1	CNS (En, E), Mu (H), L	Leigh, thalamic lesions	↑ (B, CSF)	↑	↑	N	↓CI (Mu)
	m.14710G>A	M>16	DM, D	ND	ND	ND	ND	ND	↓CI (F)
	m.14710G>A	1<F<16	CNS (M-H), Mu (PEO), PR	ND	ND	ND	ND	Lipidosis RRF, COX-	N (Mu)
MTCYB	m.14724G>A	1<M<16	CNS (A, My), Mu (exercise intolerance), Re, PR, D, Gi, DM	Leukodystrophy, cerebral atrophy	↑ (B)	ND	ND	RRF	ND
	m.15699G>C	1<M<16	CNS (PMR, A), Mu (H), PN, OA	WM, cerebellar atrophy	↑ (B, CSF)	↑	↑	RRF	↓CIII (F) N (Mu)
MTTG MTND3	m.10010T>C m.10191T>C	1<M<16	CNS (Dys, PMR)	Leigh, cerebellar atrophy	N	N	↑	N	↓CI (Mu) ↓CI (F)

The last line on the table corresponds to a patient who carried two mutations.

↓, decreased; ↑, increased; B, blood; CI, II, III, IV, V, respiratory chain complexes; CNS, central nervous system (A, cerebellar ataxia; B, bulbar involvement; CSF, cerebro spinal fluid; Dys, dystonia; E, epilepsy; En, encephalopathy; F, female; M, male; MD, movement disorders; ME, myoclonic epilepsy; M-H, migraine-like headaches; MRI, magnetic resonance imaging; My, myoclonies; PMR, parkinsonian syndrome; PS, pyramidal syndrome; R, regression; S-I, stroke-like episode); COX-, COX negative fibres; D, deafness; DM, diabetes mellitus; F, fibroblasts; Gi, gastrointestinal disorders; HCM, hypertrophic cardiomyopathy; IU-G, intrauterine growth failure; L, liver involvement; L/P, lactate/pyruvate; Mu, muscle (H, hypotonia; PEO, progressive external ophthalmoplegia; SM, skeletal myopathy); N, normal; ND, not determined; OA, optic atrophy; OH-but/Ac-Ac, hydroxy-butyrate/acetate; PN, peripheral neuropathy; PN-G, postnatal growth failure; PR, pigmentary retinopathy; RC, respiratory chain; Re, renal involvement; RRF, ragged red fibres; U, urine; VA, decreased visual acuity; WM, white matter involvement.

55 patients showed three distinct profiles (figure 3). Early onset disease and mutations in protein coding genes were associated with a profile of patients characterised by psychomotor retardation, regression, stroke-like episodes, movement disorders, Leigh syndrome and white matter involvement. Late onset was associated with mutations in the tRNA genes. Affected maternal relatives were also found predominantly in this second profile (33%), despite being non-significantly associated ($p = 0.18$). Again, we found no pattern associated with an onset during childhood. A dendrogram for cluster model is illustrated in the online supplementary figure S2.

DISCUSSION

In OXPHOS disorders, identification of a causative mutation(s) is critical to confirm the diagnosis and to propose accurate genetic counselling and prenatal diagnosis, but molecular analysis is problematic due to the large genetic heterogeneity of mtDNA related disorders. Primary mtDNA defects include large rearrangements and point mutations. In clinical practice, most suspected patients do not have their entire mitochondrial genome sequenced but are usually tested only for deletions and a small number of common mutations. Several groups have reported that testing for common mtDNA defects in adults with suspected mitochondrial disease can have quite a high diagnostic yield, ranging from 10% to 40% of patients tested.^{3 4 17} Testing for common mtDNA mutations has a much lower diagnostic yield in children, ranging from 2–4% to 25%.^{18 19} Recently, looking for 10 mtDNA point mutations in 3168 neonatal cord blood samples revealed that at least 1 in 200 healthy humans harbours a pathogenic mtDNA mutation, although at low mutant heteroplasmy.²⁰ However, to date, the rate of rare mutations is totally unknown. To answer this question, we analysed the whole mitochondrial genome of 743 patients suspected of manifesting a mitochondrial disease after excluding deletions and the most common mutations.

A total of 55 patients (7.4%) harboured pathogenic mutations responsible for an onset disorder before adulthood in 67% of cases and before 1 year of age in 24% of cases. Previous studies have suggested that 50% or more of all patients have adult onset. However, in these studies, the most common mutations (mainly deletions and m.3243A>G) that were mainly found in adult presentations (CPEO, maternally inherited deafness and diabetes (MIDD), etc) were not excluded. The French Mitochondrial Disease Network includes approximately 90% of patients investigated for mitochondrial disease in France. The 743 patients included in the study were doubly selected because they were highly ‘suspect’ of RC disorder, and they did not carry common mtDNA mutations. Our results suggest that in the French population, the clinical presentations that are not associated with common mutations begin mainly before adulthood. Further studies using next generation sequencing technology will be necessary to confirm this point.

Assuming that heteroplasmic and homoplasmic mutations may have different effects, we also analysed both populations separately. As expected, heteroplasmic mutations were more frequent than homoplasmic mutations (4.6% vs 2.8%). Symptoms appeared before the age of 16 years in 75% of patients carrying heteroplasmic pathogenic mutations and in 58% of patients carrying homoplasmic pathogenic mutations. If a number of factors have led to an exaggerated perception that the mtDNA genome is rarely involved in children affected by mitochondrial disorders, our results show that rare mtDNA mutations are a non-negligible cause of paediatric RC deficiency because 6.75%

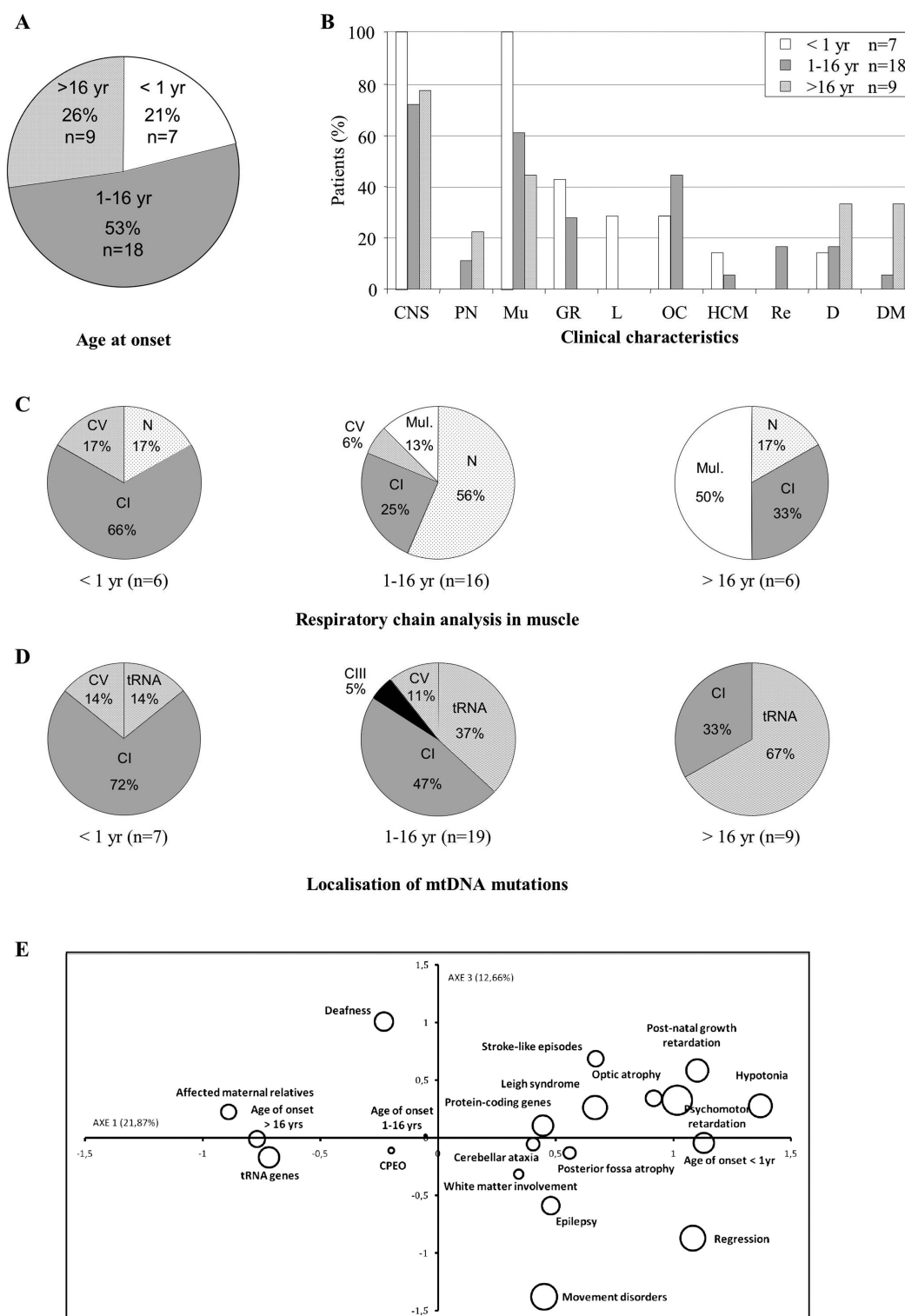


Figure 1 Analysis of the population harbouring heteroplasmic pathogenic mitochondrial DNA (mtDNA) mutations according to age at onset. n, number of patients. (A) Age at onset (in years). (B) Clinical characteristics. CNS, central nervous system; D, deafness; DM, diabetes mellitus. GR, growth retardation (intrauterine or postnatal); HCM, hypertrophic cardiomyopathy; L, liver involvement; Mu, muscle; OC, ocular involvement (optic atrophy or pigmentary retinopathy); PN, peripheral neuropathy; Re, renal involvement. (C) Respiratory chain analysis by spectrophotometry. CI, CV, respiratory chain deficiency in complex I, in complex V; mul, multiple respiratory chain deficiency; N, normal. (D) Localisation of mtDNA mutations. CI, CIII, CV, genes encoding subunits of complexes I, III, V; tRNA, genes encoding tRNA. (E) Statistical analyses with hierarchical ascendant classification of pathogenic heteroplasmic mutations found in the 34 patients. The upper right quadrant is complementary to the bottom left quadrant. The numbers on the axes are the scores that represent the contribution of each feature to the overall inertia. Our model (axes 1–3) explains 48% of total inertia. The circles represent the variables. Their size is proportional to the quality of representation of the variables on the plane (axes 1 and 3). Patients with disease onset <1 year are clustered with mutations in protein coding genes, psychomotor retardation, hypotonia, optic atrophy, postnatal growth retardation and Leigh syndrome. Patients with disease onset >16 years are clustered with mutations in tRNA genes and affected maternal relatives.

Table 2 Characteristics of patients with pathogenic mutations identified by MitoChip analysis

Gene	Mutation	Gender/age at onset	Clinical symptoms	Brain MRI	Lactate	L/P	OH-but/ Ac-Ac	Muscle histology	RC	
<i>MTRNR1</i>	m.1555A>G	1<F<16	CNS (dementia), Mu (SM), HCM, D	WM	ND	ND	ND	RRF, COX–, lipidosis	N (Mu)	
		M<1	CNS (PMR), Mu (SM, H), IU-G, PN-G	N	ND	↑	ND	ND	↓CI, CIII, CIV, CV (Mu, F)	
		1<M<16	OA, PN, PN-G	Cerebellar atrophy	ND	N	ND	ND	N (Mu)	
<i>MTTV</i>	m.1644G>A	M>16	CNS (E, S-I, dementia)	Cerebral atrophy	ND	ND	ND	COX–	↓CIV (Mu)	
<i>MTTL1</i>	m.3303C>T	M<1	Mu (H), HCM, L	N	↑(B)	↑	N	Lipidosis	↓CI+III (Mu) N (F)	
<i>MTND1</i>	m.3395A>G	F>16	Mu (SM, myalgia, ptosis), HCM, D, DM	ND	ND	ND	ND	COX–	N (F)	
		m.3460A>G	M>16	CNS (A), PN, Mu (ptosis), HCM, L, OA	Leigh	ND	N	ND	RRF, COX–, lipidosis	↓CI (Mu)
		m.3890G>A	M<1	CNS (PMR, E, A), Mu (PEO), L, OA	Leigh, thalamus lesions	↑ (B, CSF)	↑	ND	Lipidosis	N (Mu) ↓CI, CIII, CIV (F)
<i>MTTI</i>	m.4300A>G	F>16	HCM	ND	ND	ND	ND	COX–, lipidosis	ND	
		m.4316A>G	F>16	Mu (SM, PEO)	ND	ND	ND	ND	RRF, COX–	↓CIII+IV (Mu)
		m.4317 delA	M<1	CNS (PMR, En, R, E)	N	ND	ND	ND	Lipidosis	↓CIV (Mu)
<i>MTTW</i>	m.5521G>A	M>16	Mu (ptosis)	ND	ND	ND	ND	RRF, COX–	↓CIII (Mu)	
		F>16	CNS (PMR, A, E, MD, S-I), Mu (SM), eating disorder	Leukodystrophy, cerebellar atrophy	↑ (B)	↑	N	RRF, COX–, lipidosis	↓CI, CIV, ↓CII+III, CIII (Mu)	
		M>16	CNS (A, dementia), Mu (SM), D	N	ND	ND	ND	RRF, COX–, lipidosis	↓CIV (Mu)	
<i>MTATP6</i>	m.9185T>C	F>16	CNS (A)	ND	ND	ND	ND	N	N (Mu)	
<i>MTND5</i>	m.13513G>A	1<M<16	CNS (En, E, M-IH), PN, Mu (SM), HCM, PR, OA, D	Cerebellar atrophy, stroke-like	ND	↑	ND	Mitochondrial proliferation	↓CI (Mu)	
		1<M<16	CNS (PMR, A), Mu (PEO), OA, IU-G, PN-G	ND	ND	ND	ND	ND	ND	
<i>MTND6</i>	m.14459G>A	F<1	CNS (R)	Leigh	N (B)	ND	ND	ND	N (Mu), ↓CI (L), ↓CI+IV (F)	
		M<1	CNS (PMR, Dys), Mu (ptosis)	Leigh	↑ (B, CSF)	ND	ND	ND	↓CI (Mu, F)	
<i>MTCYB</i>	m.15234G>A	1<M<16	CNS (PMR, S-I, PS), Mu (ptosis)	Leigh	↑ (CSF)	N	N	N	↓CI (Mu)	
		1<M<16	CNS (R, PMR, E, S-I, M-IH), Mu (PEO, SM), D, cataract	Leigh, Stroke-like, leukodystrophy,	↑ (B, CSF)	↑	↑	RRF	↓CI, CIII, CIV (M), N (F)	

↓, decreased; ↑, increased; B, blood; CI, II, III, IV, V, respiratory chain complexes; CNS, central nervous system (A, cerebellar ataxia; Dys, dystonia; E, epilepsy; En, encephalopathy; MD, movement disorders; M-I H, migraine-like headaches; PMR, psychomotor retardation; PS, pyramidal syndrome; R, regression; S-I, stroke-like episode); COX–, COX negative fibres; D, deafness; DM, diabetes mellitus; F, fibroblasts; HCM, hypertrophic cardiomyopathy; IU-G, intrauterine growth failure; L, liver involvement; L/P, lactate/pyruvate; Mu, muscle (H, hypotonia; PEO, progressive external ophthalmoplegia; SM, skeletal myopathy); N, normal; ND, not determined; OA, optic atrophy; OH-but/Ac-Ac, hydroxy-butylrate/aceto-acetate; PN, peripheral neuropathy; PN-G, postnatal growth failure; PR, pigmentary retinopathy; RC, respiratory chain; RRF, ragged red fibres; WM, white matter involvement.

of patients with onset before adulthood (37 of 548) carried rare pathogenic mtDNA mutations.

One of the main results of this study was that the location of mtDNA mutations influenced the age of onset of OXPHOS diseases. We found a significant correlation between mutations in protein coding (mainly CI) genes and early onset disorder (<1 year) on the one hand, and between mutations in tRNA genes and adult onset on the other hand; the hierarchical clustering allowing unambiguous class assignments of patients with early and late onset disease. These results suggest that mutations in tRNA genes are much better tolerated than those that directly affect the functioning of CI whose defect is the most frequently observed. It was previously thought that a mtDNA origin accounted for only 5–10% of paediatric cases.²¹ However, a recent study of 109 children with isolated CI deficiency identified pathogenic mtDNA mutations in 29% of cases.²² In our study, mutations in *MTND* genes were found in 11 of 75 patients (14.7%) with an isolated CI deficiency before the age of 16 years but this percentage is likely underestimated because other deleterious mutations were likely present among the 44 variants corresponding to putative mutations that we identified

in *MTND* genes. Among CI genes, *MTND5* and *MTND6* corresponded to 'hotspots' for disease causing mutations, being, respectively, found mutated in 11 (20%) and seven (12.7%) of 55 patients, with Leigh syndrome accounting for the large majority of associated phenotypes.

A total of 167 patients (22.4%) harboured putative pathogenic mutations, with homoplasmic novel variants being more frequent than heteroplasmic variants (14.6% vs 7.8%). Interpretation of the clinical significance of rare variants is challenging due to the highly polymorphic nature of mtDNA. The existing scoring systems to evaluate the pathogenicity of a mtDNA variant consider several parameters, including single fibre PCR or cybrid studies, multiple family reports, segregation within the family, measurable biochemical defect, evolutionary conservation, etc.^{23–24} Collection of data and tools to perform functional studies is in progress but we currently do not have the information needed to precisely define the status of the 160 different variants identified. Nevertheless, 32 patients (4.3%) carried novel variants in protein coding genes with high or medium amino acid conservation and a damaging and not tolerated status according to our in silico analysis using PolyPhen 2

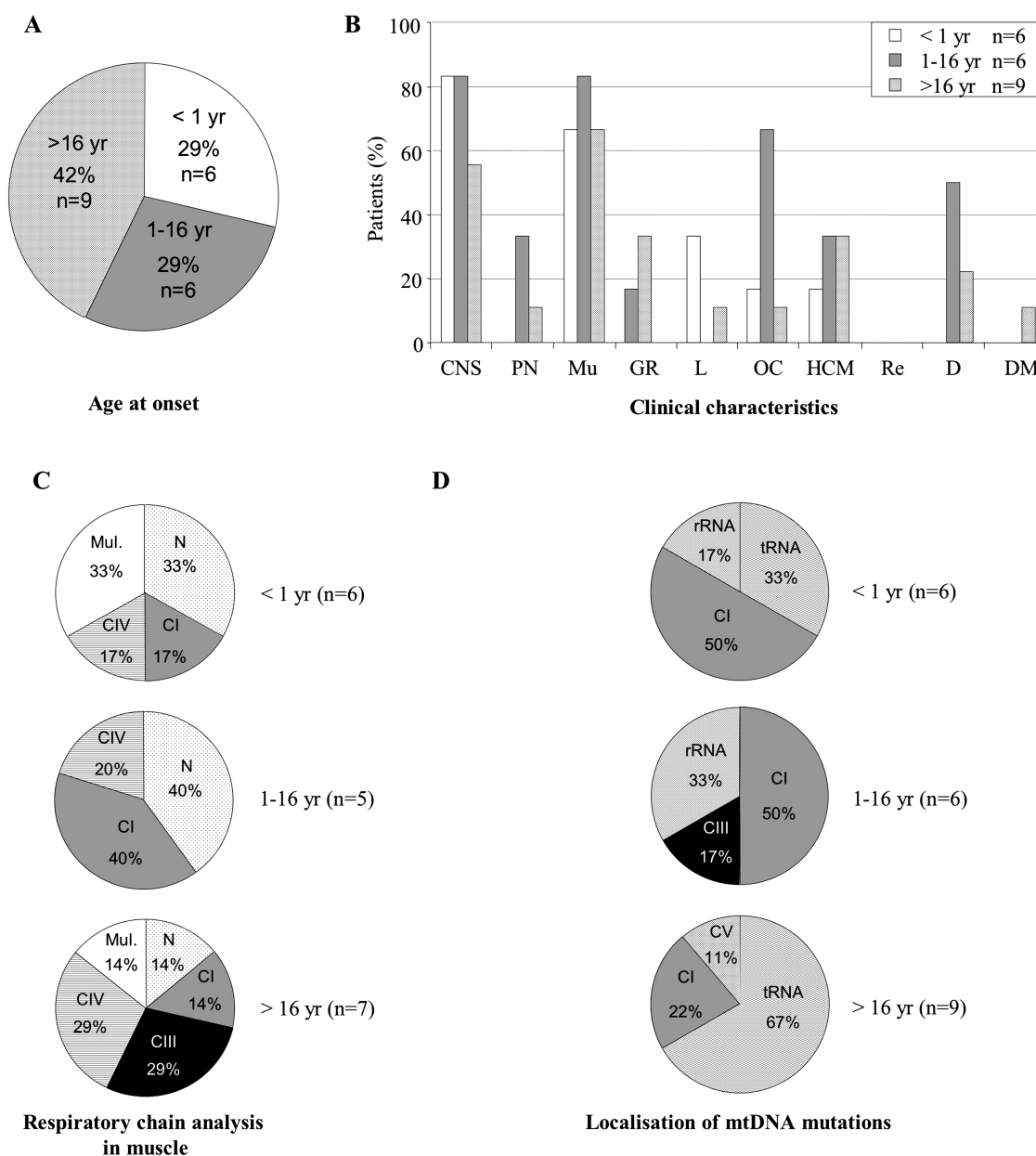


Figure 2 Analysis of the population harbouring homoplasmic pathogenic mitochondrial DNA (mtDNA) mutations according to age at onset. n, number of patients. (A) Age at onset (in years). (B) Clinical characteristics. CNS, central nervous system; D, deafness; DM, diabetes mellitus; GR, growth retardation (intrauterine or postnatal); HCM, hypertrophic cardiomyopathy; L, liver involvement; Mu, muscle; OC, ocular involvement (optic atrophy or pigmentary retinopathy); PN, peripheral neuropathy; Re, renal involvement. (C) Respiratory chain analysis by spectrophotometry. CI, CIII, CIV, respiratory chain deficiency in complex I, complex III, complex IV; mul, multiple respiratory chain deficiency; N, normal. (D) Localisation of mtDNA mutations. CI, CIII, CV, genes encoding subunits of complexes I, III, V; rRNA, genes encoding rRNA; tRNA, genes encoding tRNA.

and SIFT, respectively (see online supplementary tables S4 and S5).²⁵ In 15 of 32 patients, the identified variant was correlated with a corresponding biochemical defect (not shown). Regarding tRNA coding genes, mutations were detected in 18 of 22 genes and were found spread all over the structural domains of the corresponding tRNAs. Seventeen of the 36 newly detected variants affect highly conserved nucleotides (table 3). Despite the fact that conservation is often poor in picking pathogenicity,^{28 29} this is a criterion that may be taken into account to prioritise future functional studies allowing assessment of tRNA variants using scoring systems.

In conclusion, we have reported the first study to determine the prevalence of rare mtDNA mutations in RC disorders.

Although further analyses are needed to define precisely the pathogenic or polymorphic nature of all of the novel variants identified, our results show that rare mtDNA mutations account for more 7.4% of patients with mitochondrial disorders. Careful selection of the population studied and the choice of tissues concerned by molecular analysis, which was carried out mainly in muscle (88% of cases), may explain this mutation rate, probably higher if one takes into account some new variants are probably deleterious. In this cohort of patients, we excluded those carrying a deletion or common mtDNA mutations. It is interesting to note that within the French Mitochondrial Disease Network, we have a rate of 8% for these common abnormalities of the mitochondrial

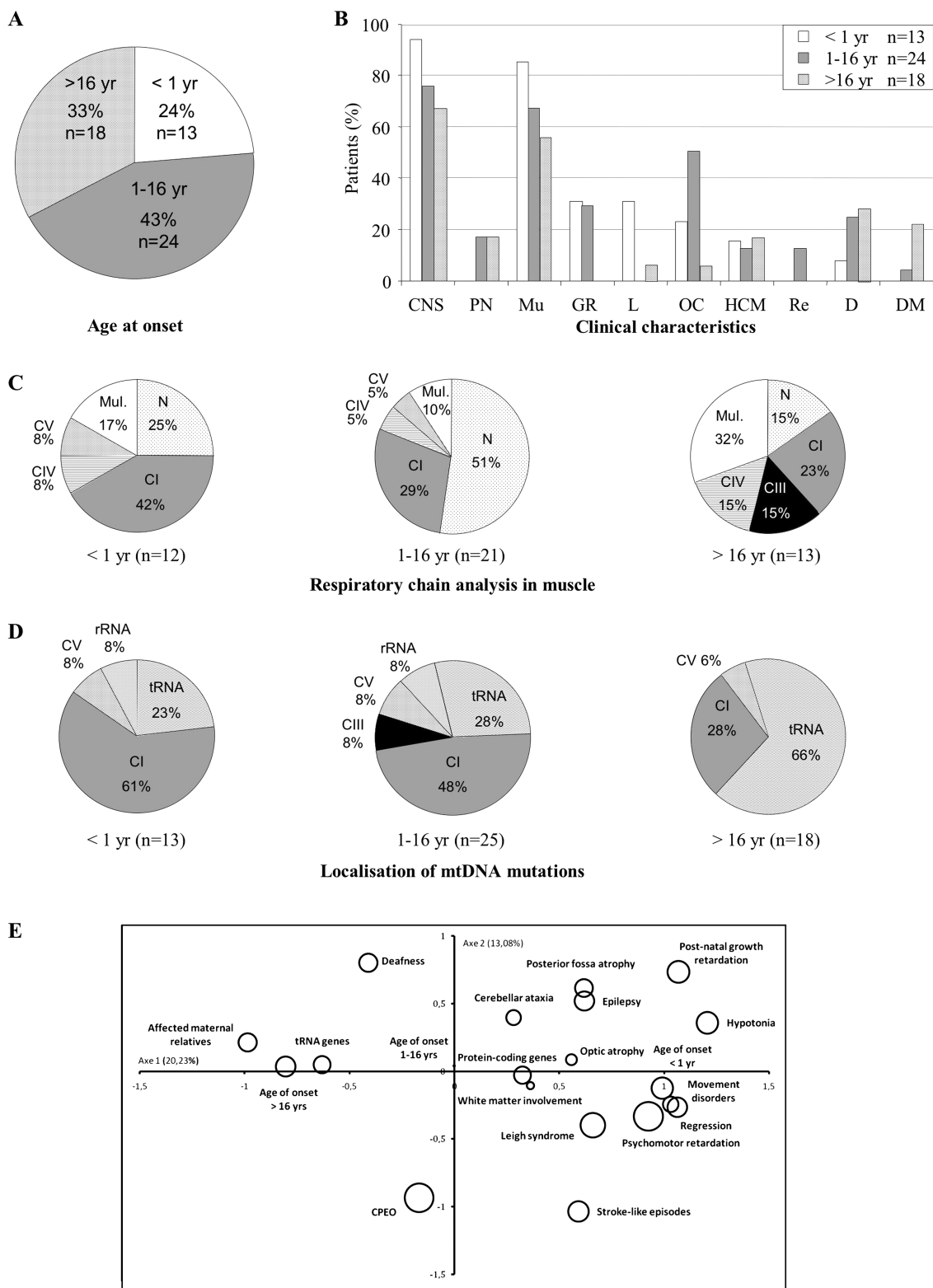


Figure 3 Analysis of the pooled populations harbouring heteroplasmic or homoplasmic pathogenic mitochondrial DNA (mtDNA) mutations according to age at onset. n, number of patients. (A) Age at onset (in years). (B) Clinical characteristics. CNS, central nervous system; D, deafness; DM, diabetes mellitus; GR, growth retardation (intrauterine or postnatal); HCM, hypertrophic cardiomyopathy; L, liver involvement; Mu, muscle; OC, ocular involvement (optic atrophy or pigmentary retinopathy); PN, peripheral neuropathy; Re, renal involvement. (C) Respiratory chain analysis by spectrophotometry. CI, CIII, CIV, CV, respiratory chain deficiency in complex I, complex III, complex IV, complex V; mul, multiple respiratory chain deficiency; N, normal. (D) Localisation of mtDNA mutations. CI, CIII, CV, genes encoding subunits of complexes I, III, V; rRNA, genes encoding rRNA; tRNA, genes encoding tRNA. (E). Statistical analyses with hierarchical ascendant classification of pathogenic mutations found in the 55 patients. Our model (axes 1–3) explains 45% of total inertia. The circles represent the variables. Their size is proportional to the quality of representation of the variables on the plane (axes 1 and 2). Patients with disease onset <1 year are clustered with protein coding gene mutations, psychomotor retardation, stroke-like episodes, regression, movement disorders, Leigh syndrome and white matter involvement. Patients with late onset are clustered with tRNA gene mutations and affected maternal relatives.

Table 3 Novel variants in tRNA genes and tentative prediction of impact according to the level of nucleotide conservation of the affected position

Gene	Sequence variation	Gene reading	Structural domains in tRNA	Position in tRNA	Change in tRNA	Nucleotide conservation	Tentative prediction
<i>MTTF</i>	<i>m.579T>C</i>	<i>Direct</i>	<i>Acc stem</i>	3	<i>T>C</i>	<i>90%<x<100%, Y or R</i>	<i>Possibly pathological</i>
	<i>m.592C>T</i>	<i>Direct</i>	<i>D loop</i>	16	<i>C>T</i>	<i><50%</i>	<i>Polymorphism</i>
	<i>m.593T>C</i>	<i>Direct</i>	<i>D loop</i>	17	<i>T>C</i>	<i><50%</i>	<i>Polymorphism</i>
<i>MTTV</i>	<i>m.645A>G</i>	<i>Direct</i>	<i>Acc stem</i>	71	<i>A>G</i>	<i>90%<x<100%, all R</i>	<i>Possibly pathological</i>
	<i>m.1628C>T</i>	<i>Direct</i>	<i>Anticd stem</i>	29	<i>C>T</i>	<i><50%</i>	<i>Polymorphism</i>
<i>MTTL1</i>	<i>m.1645A>G</i>	<i>Direct</i>	<i>Variable region</i>	46	<i>A>G</i>	<i>100%A</i>	<i>Possibly pathological</i>
	<i>m.3258T>C</i>	<i>Direct</i>	<i>Anticd stem</i>	27	<i>T>C</i>	<i>90%<x<100%, all R</i>	<i>Possibly pathological</i>
<i>MTTQ</i>	<i>m.4354C>T</i>	<i>Inverse</i>	<i>Variable region</i>	48	<i>G>A</i>	<i><50%</i>	<i>Polymorphism</i>
	<i>m.4394C>A</i>	<i>Inverse</i>	<i>Acc stem</i>	7	<i>G>T</i>	<i><50%</i>	<i>Polymorphism</i>
<i>MTTM</i>	<i>m.4449G>A</i>	<i>Direct</i>	<i>T stem</i>	52	<i>G>A</i>	<i>100%G</i>	<i>Possibly pathological</i>
<i>MTTW</i>	<i>m.5527A>G</i>	<i>Direct</i>	<i>D loop</i>	16	<i>A>G</i>	<i>50%<x<90%, Y or R</i>	<i>Probable polymorphism</i>
	<i>m.5542C>T</i>	<i>Direct</i>	<i>Anticd loop</i>	32	<i>C>T</i>	<i>100%C</i>	<i>Possibly pathological</i>
	<i>m.5560G>A</i>	<i>Direct</i>	<i>T stem</i>	51	<i>G>A</i>	<i>50%<x<90%, all R</i>	<i>Probable polymorphism</i>
	<i>m.5574T>C</i>	<i>Direct</i>	<i>Acc stem</i>	68	<i>T>C</i>	<i>50%<x<90%, all Y</i>	<i>Probable polymorphism</i>
<i>MTTA</i>	<i>m.5605A>G</i>	<i>Inverse</i>	<i>T loop</i>	55	<i>T>C</i>	<i>90%<x<100%, Y or R</i>	<i>Possibly pathological</i>
	<i>m.5608C>T</i>	<i>Inverse</i>	<i>T stem</i>	52	<i>G>A</i>	<i>50%<x<90%, all R</i>	<i>Probable polymorphism</i>
<i>MTTN</i>	<i>m.5663C>T</i>	<i>Inverse</i>	<i>Acc stem</i>	67	<i>G>A</i>	<i>50%<x<90%, all R</i>	<i>Probable polymorphism</i>
	<i>m.5690A>G</i>	<i>Inverse</i>	<i>Anticd stem</i>	40	<i>T>C</i>	<i>100%T</i>	<i>Possibly pathological</i>
<i>MTTC</i>	<i>m.5793A>G</i>	<i>Inverse</i>	<i>T loop</i>	60	<i>T>C</i>	<i><50%</i>	<i>Polymorphism</i>
	<i>m.5800A>G</i>	<i>Inverse</i>	<i>Anticd loop</i>	32	<i>T>C</i>	<i>90%<x<100%, all Y</i>	<i>Possibly pathological</i>
<i>MTTY</i>	<i>m.5865T>C</i>	<i>Inverse</i>	<i>Anticd stem</i>	31	<i>A>G</i>	<i>90%<x<100%, all R</i>	<i>Possibly pathological</i>
<i>MTTS1</i>	<i>m.7455A>G</i>	<i>Inverse</i>	<i>T stem</i>	64	<i>T>C</i>	<i>90%<x<100%, all Y</i>	<i>Possibly pathological</i>
<i>MTTD</i>	<i>m.7567C>T</i>	<i>Direct</i>	<i>T loop</i>	54	<i>C>T</i>	<i>50%<x<90%, Y or R</i>	<i>Probable polymorphism</i>
<i>MTTK</i>	<i>m.8306T>C</i>	<i>Direct</i>	<i>D-stem</i>	12	<i>T>C</i>	<i>100%T</i>	<i>Possibly pathological</i>
<i>MTTH</i>	<i>m.12168G>A</i>	<i>Direct</i>	<i>Anticd loop</i>	34	<i>G>A</i>	<i>100%G</i>	<i>Possibly pathological*</i>
	<i>m.12173T>C</i>	<i>Direct</i>	<i>Anticd stem</i>	39	<i>T>C</i>	<i>50%<x<90%, Y or R</i>	<i>Probable polymorphism</i>
	<i>m.12174C>T</i>	<i>Direct</i>	<i>Anticd stem</i>	40	<i>C>T</i>	<i>90%<x<100%, all Y</i>	<i>Possibly pathological</i>
	<i>m.12197_12198InsC</i>	<i>Direct</i>	<i>T stem</i>	64–65	<i>Ins C</i>	<i>/</i>	<i>Probable polymorphism</i>
<i>MTTL2</i>	<i>m.12324C>T</i>	<i>Direct</i>	<i>T stem</i>	61	<i>C>T</i>	<i>100%C</i>	<i>Possibly pathological</i>
<i>MTTE</i>	<i>m.14697C>T</i>	<i>Inverse</i>	<i>T stem</i>	50	<i>G>A</i>	<i>50%<x<90%, Y or R</i>	<i>Probable polymorphism</i>
<i>MTTT</i>	<i>m.15916T>C</i>	<i>Direct</i>	<i>Anticd stem</i>	31	<i>T>C</i>	<i>50%<x<90%, all Y</i>	<i>Probable polymorphism</i>
	<i>m.15926C>T</i>	<i>Direct</i>	<i>Anticd stem</i>	41	<i>T>C</i>	<i>90%<x<100%, all Y</i>	<i>Possibly pathological</i>
	<i>m.15936A>T</i>	<i>Direct</i>	<i>T stem</i>	52	<i>A>T</i>	<i>50%<x<90%, Y or R</i>	<i>Probable polymorphism</i>
	<i>m.15947A>G</i>	<i>Direct</i>	<i>Acc stem</i>	67	<i>A>G</i>	<i>50%<x<90%, Y or R</i>	<i>Probable polymorphism</i>
<i>MTTP</i>	<i>m.15977C>T</i>	<i>Inverse</i>	<i>T stem</i>	50	<i>G>A</i>	<i>50%<x<90%, all R</i>	<i>Probable polymorphism</i>
	<i>m.15992A>T</i>	<i>Inverse</i>	<i>Anticd loop</i>	34	<i>T>A</i>	<i>100%T</i>	<i>Possibly pathological*</i>

Nucleotide conservations are retrieved from the compilation of tRNA genes of each specificity taking into account 150 mammalian mitochondrial genomes (<http://mamit-trna.u-strasbg.fr/>).^{26 27} Four levels of conservation are considered. Whenever a nucleotide is conserved in more than 90% of the sequences, the tentative prediction is that the probability of having a structural or functional role is high. Its mutation may thus possibly have pathological consequences. Possible pathological variants are shown in italic.

*Position 34 corresponds to the first nucleotide of the tRNA anticodon triplet. The mutations concerning both *MTTH* and *MTTP* do not affect the codon reading. However, they may affect the efficiency of aminoacylation or of post-transcriptional modification, leading to less efficient translation.

Acc stem, acceptor stem; Anticd loop, anticodon loop; Anticd stem, anticodon stem; R, purine base; Y, pyrimidic base.

genome. The sum of the rates obtained in our network for both common and rare mutations shows how the complete screening of mtDNA is essential. It should be performed for an optimal diagnosis, including in young children, and will

now be accessible with the development of next generation sequencing technology that provides a one step approach in detecting common and uncommon point mutations, as well as deletions.

Author affiliations

- ¹IRCAN, CNRS UMR 7284/Inserm U1081/UNS, Faculté de Médecine, Nice, France
²Service de Génétique Médicale, Hôpital Archet 2, CHU de Nice, Nice, France
³IBS Laboratoire de Génétique, CHU Angers, Angers, France
⁴Inserm U781 Service de Génétique, Hôpital Necker-Enfants Malades, Université Paris Descartes, Paris, France
⁵Biochimie Métabolique, Centre de Génétique moléculaire et chromosomique, Groupe hospitalier Pitié Salpêtrière, Paris, France
⁶Euro Eco Ile de France, APHP, Hôpital de l'Hôtel Dieu, Paris, France
⁷Department of Molecular Pharmacology and Therapeutics, Loyola University Health Sciences Division, Maywood, USA
⁸Bioinformatics Core Facility The Children's Hospital of Philadelphia, Philadelphia, USA
⁹Laboratoire de Biochimie et Génétique moléculaire, CHU de Grenoble, La Tronche, France
¹⁰Service des Maladies Héritaires du Métabolisme, Centre de Biologie et de Pathologie Est Groupement Hospitalier Est, CHU de Lyon, Bron, France
¹¹Centre de Référence des surdités génétiques, APHP, Hôpital des Enfants Armand Trousseau, Paris, France
¹²UF Gépopathies, Laboratoire de Biochimie, Centre de biologie CHRU de Lille et Université Lille Nord de France, Lille, France
¹³Laboratoire de Biochimie, CHU de Bicêtre, APHP, Le Kremlin Bicêtre, France
¹⁴Laboratoire de Physiopathologie Mitochondriale U688 INSERM, Université Victor Segalen Bordeaux 2, Bordeaux, France
¹⁵Divisions of Human Genetics and Pulmonary Medicine, The Children's Hospital of Philadelphia, Philadelphia, USA
¹⁶Architecture et Réactivité de l'ADN, Université Louis Pasteur de Strasbourg, CNRS, IBMC, Strasbourg, France
¹⁷Department of Neuropediatrics, Timone Hospital, Marseille Teaching Hospital, Marseille, France

Acknowledgements We thank all of the patients and referring physicians involved in this study: Professor D Adams, Dr S Attarian, Professor JP Azulay, Dr M Barth, Professor T Billette de Villemeur, Professor D Bonneau, Dr C Boutte, Dr A Cano, Dr S Chabrier, Dr JM Cuisset, Professor P de Lonlay, Professor C Desnuelle, Professor V des Portes, Dr A Echaniz-Laguna, Dr D Feldmann, Dr X Ferrer, Dr M Holder, Dr E Kaphan, Professor D Lacombe, Dr A Lacour, Dr P Laforêt, Dr E Lagrange, Professor P Landrieu, Professor J Lunardi, Dr F Mochel, Professor A Munnich, Dr M-A Nguyen-Morel, Professor J Pouget, Dr M Rio, Dr C Rouzier, Dr C Richelme, Dr S Sacconi, Dr E Salort-Campana, Professor C Verry and Dr A Vershueren. We thank G Augé, S Foustoul, S Gherbi and L Smaghe for technical help.

Contributors SB, VP, ASL, CJ, AC, KF, GH, BMdC, SM, CMD, AS, CR, JPB, AR, NA, MG, VD-D, JC, LJ, AD, CE-T, DM, PG, PA-B and BC recruited patients and/or performed biochemical and molecular analyses. HM, NC and ID-Z performed the statistical analysis. VP, XG, HMX, MF and MJF performed the MitoChip data analysis. PR was the leader of the French Mitochondrial Disease Network. KHLQS checked the quality process. SB and CH were responsible for follow-up of the study, synthesis of the data, and produced the tables and figures. CF performed the study of tRNA variants. VP-F was responsible for the study concept and, with SB, obtained funding and wrote the manuscript.

Funding This work was made possible by grants to VP-F from the French Ministry of Health (PSTIC 2008, Programme de Soutien aux Techniques Innovantes et Coûteuses).

Competing interests None.

Ethics approval Authorisation for individual data protection was obtained.

Provenance and peer review Not commissioned; externally peer reviewed.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 3.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/3.0/>

REFERENCES

- Holt I, Harding A, Morgan-Hughes JA. Deletion of muscle mitochondrial DNA in patients with mitochondrial myopathies. *Nature* 1988;331:717–19.
- Wallace D, Singh G, Lott MT, Schurr TG, Lezza AMS, Elsas LJ. Mitochondrial DNA mutation associated with Leber's hereditary optic neuropathy. *Science* 1988;242:1427–30.
- Thornburn DR. Mitochondrial diseases: not so rare after all. *Int Med J* 2004;34:3–5.
- Marotta R, Chin J, Quigley A, Katsabanis S, Kapsa R, Byrne E, Collins S. Diagnostic screening of mitochondrial DNA mutations in Australian adults 1990–2001. *Int Med J* 2004;34:10–19.
- Rustin P, Chretien D, Bourgeron T, Gerard B, Rotig A, Saudubray J. Biochemical and molecular investigations in respiratory chain deficiencies. *Clin Chem Acta* 1994;228:31–51.
- Bannwarth S, Procaccio V, Paquis-Flucklinger V. Surveyor nuclease: a new strategy for a rapid identification of heteroplasmic mitochondrial DNA mutations in patients with respiratory chain defects. *Hum Mutat* 2005;25:575–82.
- Bannwarth S, Procaccio V, Paquis-Flucklinger V. Rapid identification of unknown heteroplasmic mutations across the entire human mitochondrial genome with mismatch-specific Surveyor nuclease. *Nat Protoc* 2006;1:2037–47.
- Maitra A, Cohen Y, Gillespie SE, Mambo E, Fukushima N, Hoque MO, Shah N, Goggins M, Califano J, Sidransky D, Chakravarti A. The Human MitoChip: a high throughput sequencing microarray for mitochondrial mutation detection. *Genome Res* 2004;14:812–19.
- Xie M, Perin J, Schurr TG, Matthew C, Dulik M, Zhadanov S, Baur J, King M, Place E, Clarke C, Grauer M, Schug JD, Santani A, Albano A, Kim C, Procaccio V, Hakonarson H, Gai X, Falk M. Mitochondrial Genome Sequence Analysis: a custom bioinformatics pipeline substantially improves Affymetrix MitoChip v2.0 call rate and accuracy. *BMC Bioinformatics* 2011;12:402.
- Andrews R, Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM, Howell N. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat Genet* 1999;23:147.
- Bertario L, Russo A, Sala P, Varesco L, Giarola M, Mondini P, Pierotti M, Spinelli P, Radice P. Hereditary Colorectal Tumor Registry. Multiple approach to the exploration of genotype–phenotype correlations in familial adenomatous polyposis. *J Clin Oncol* 2003;21:1698–707.
- Mahr A, Katsahian S, Varet H, Guillemin L, Hagen EC, Höglund P, Merkel PA, Pagnoux C, Rasmussen N, Westman K, Jayne DRW for the French Vasculitis Study Group (FVSG) and the European Vasculitis Society (EUVAS). Revisiting the classification of clinical phenotypes of anti-neutrophil cytoplasmic antibody-associated vasculitis: a cluster analysis. *Ann Rheum Dis* 2013;72:1003–10.
- Benzerri J. *L'analyse des données*. Ed Dunod, 2nd edn. Paris, 1976:629.
- Torroni A, Huoponen K, Francalacci P, Petrozzi M, Morelli L, Scozzari R, Obinu D, Savontaus ML, Wallace DC. Classification of European mtDNAs from an analysis of three European populations. *Genetics* 1996;144:1835–50.
- Ingman M, Gyllensten U. mtDB: Human mitochondrial Genome Database, a resource for population genetics and medical science. *Nucleic Acids Res* 2006;34:D749–51.
- Kogelnik A, Lott M, Brown MD, Navathe SB, Wallace DC. MITOMAP: a human mitochondrial genome database. *Nucleic Acids Res* 1996;24:177–79.
- Holt I, Harding A, Cooper JM, Schapira AH, Toscano A, Clark JB, Morgan-Hughes JA. Mitochondrial myopathies: clinical and biochemical features of 30 patients with major deletions of muscle mitochondrial DNA. *Ann Neurol* 1989;26:699–708.
- Liang M, Wong L. Yield of mtDNA mutation analysis in 2,000 patients. *Am J Med Genet* 1998;77:395–400.
- Thorburn D. Mitochondrial disorders: prevalence, myths and advances. *J Inher Metab Dis* 2004;27:349–62.
- Elliott H, Samuels D, Eden JA, Relton CL, Chinnery PF. Pathogenic mitochondrial DNA mutations are common in the general population. *Am J Hum Genet* 2008;83:254–60.
- Tripels R, van den Heuvel L, Trijbels JM, Smeitink JA. Respiratory chain complex I deficiency. *Am J Med Genet* 2001;106:37–45.
- Swalwell E, Kirby D, Blakely EL, Mitchell A, Salemi R, Sugiana C, Compton AG, Tucker EJ, Ke B-X, Lamont PJ, Turnbull DM, McFarland R, Taylor RW, Thornburn DR. Respiratory chain complex I deficiency caused by mitochondrial DNA mutations. *Eur J Hum Genet* 2011;19:769–75.
- DiMauro S, Schon E. Mitochondrial DNA mutations in human disease. *Am J Med Genet* 2001;106:18–26.
- Yarham J, Al-Dosary M, Blakely EL, Alston CL, Taylor RW, Elson JL, McFarland R. A comparative analysis approach to determining the pathogenicity of mitochondrial tRNA mutations. *Hum Mut* 2011;11:1319–25.
- Thusberg J, Olatubosun A, Vihinen M. Performance of mutation pathogenicity prediction methods on missense mutations. *Hum Mut* 2011;4:358–68.
- Helm M, Brulé H, Friede D, Giegé R, Pütz D, Florentz C. Search for characteristic structural features of mammalian mitochondrial tRNAs. *RNA* 2000;6:1356–79.
- Pütz D, Dupuis B, Sissler M, Florentz C. Mamit-tRNA, a database of mammalian mitochondrial tRNA primary and secondary structures. *RNA* 2007;13:1184–90.
- Florentz C, Sissler M. Disease-related versus polymorphic mutations in human mitochondrial tRNAs. *Where is the difference? EMBO Report* 2001;2:481–86.
- McFarland R, Elson J, Taylor RW, Howell N, Turnbull DM. Assigning pathogenicity to mitochondrial tRNA mutations: when "definitely maybe" is not good enough. *Trends Genet* 2004;20:591–96.