PGD and heteroplasmic mitochondrial DNA point mutations: a systematic review estimating the chance of healthy offspring

D.M.E.I. Hellebrekers¹, R. Wolfe², A.T.M. Hendrickx¹, I.F.M. de Coo³, C.E. de Die¹,⁴, J.P.M. Geraedts¹,⁴, P.F. Chinnery⁵, and H.J.M. Smeets¹,⁴,*

¹Department of Clinical Genetics, Maastricht University Medical Centre, Joseph Bechlaan 113, 6229 GR Maastricht, The Netherlands ²School of Public Health and Preventive Medicine, Monash University, Alfred Hospital, Melbourne VIC 3004, Australia ³Department of Pediatric Neurology, Erasmus MC-Sophia Children’s Hospital, PO Box 2060, 3000 CB Rotterdam, The Netherlands ⁴Research Institute GROW, Maastricht University Medical Centre, Universiteitszorg 40, 6200 MD Maastricht, The Netherlands ⁵Institute of Human Genetics, Newcastle University, Central Parkway, Newcastle upon Tyne NE1 3BZ, UK

Correspondence address. E-mail: bert.smeets@maastrichtuniversity.nl

Submitted on May 20, 2011; resubmitted on December 14, 2011; accepted on February 27, 2012

TABLE OF CONTENTS

- Introduction
- Methods
- Results
- Discussion

BACKGROUND: Mitochondrial disorders are often fatal multisystem disorders, partially caused by heteroplasmic mitochondrial DNA (mtDNA) point mutations. Prenatal diagnosis is generally not possible for these maternally inherited mutations because of extensive variation in mutation load among embryos and the inability to accurately predict the clinical expression. The aim of this study is to investigate if PGD could be a better alternative, by investigating the existence of a minimal mutation level below which the chance of an embryo being affected is acceptably low, irrespective of the mtDNA mutation.

METHODS: We performed a systematic review of muscle mutation levels, evaluating 159 different heteroplasmic mtDNA point mutations derived from 327 unrelated patients or pedigrees, and reviewed three overrepresented mtDNA mutations (m.3243A>G, m.8344A>G and m.8993T>C/G) separately.

RESULTS: Mutation levels were included for familial mtDNA point mutations only, covering all affected (n = 195) and unaffected maternal relatives (n = 19) from 137 pedigrees. Mean muscle mutation levels were comparable between probands and affected maternal relatives, and between affected individuals with tRNA- versus protein-coding mutations. Using an estimated a priori prevalence of being affected in pedigrees of 0.477, we calculated that a 95% or higher chance of being unaffected was associated with a muscle mutation level of 18% or less. At a mutation level of 18%, the predicted probability of being affected is 0.00744. The chance of being unaffected was lower only for the m.3243A>G mutation (P < 0.001). Most carriers of mtDNA mutations will have oocytes with mutation levels below this threshold.

CONCLUSIONS: Our data show, for the first time, that carriers of heteroplasmic mtDNA mutations will have a fair chance of having healthy offspring, by applying PGD. Nevertheless, our conclusions are partly based on estimations and, as indicated, do not provide absolute certainty. Carriers of mtDNA should be informed about these constraints.

Key words: PGD / mitochondrial DNA mutations / heteroplasmy

© The Author 2012. Published by Oxford University Press on behalf of the European Society of Human Reproduction and Embryology. All rights reserved.

For Permissions, please email: journals.permissions@oup.com

Introduction

Mitochondrial disorders are often fatal multisystem disorders, representing the most common group of inborn errors of metabolism. Maternally inherited pathogenic point mutations in the mitochondrial DNA (mtDNA) are a frequent cause of these disorders, affecting ~1 in 5000 individuals (Skladal et al., 2003; Poulton et al., 2010). A crucial characteristic of most pathogenic mtDNA mutations is heteroplasmacy, the coexistence of normal and mutated mtDNA, requiring a threshold of mutated mtDNA to be exceeded before clinical symptoms occur (Howell, 1983; Boulet et al., 1992; Yoneda et al., 1995). Expression thresholds and genotype–phenotype correlations have been evaluated for some of the common mtDNA mutations (Chinnery et al., 1997; White et al., 1999), but in most cases it was not possible to accurately predict the clinical manifestations based on the mutation load. For most (private) mtDNA point mutations, the information to judge this is not even available. The inability to predict the clinical expression very much limits the scope for prenatal diagnosis to prevent the transmission of mtDNA disorders (Poulton and Bredenoord, 2010). Nevertheless, a minimal mutation level may exist, which remains below the pathogenic thresholds of all individual mtDNA mutations, assuming that a minimal level of wild-type mtDNA may be sufficient to prevent the occurrence of clinical symptoms (Howell, 1983; Boulet et al., 1992; Yoneda et al., 1995). The aim of this study was to investigate the existence of such a ‘safe’ cut-off point below which the chance of being affected is acceptably low, irrespective of the mtDNA mutation. In such a case, PGD would be the preferred strategy to prevent the transmission of mtDNA disorders, allowing the selection of most likely unaffected embryos with mutation loads in blastomeres under this threshold (Poulton and Turnbull, 2000; Poulton et al., 2009). To calculate the chance of being unaffected for mtDNA mutation carriers, we performed a systematic review on different pathogenic mtDNA mutations, combining data on mutation levels in muscle, as we assumed that these best reflect the prenatal tissues.

Methods

Data collection and evaluation of mtDNA mutations in data set 1 (excluding the m.3243A>G, m.8344A>G and m.8993T>C/G mutations)

Pedigrees with a pathogenic mtDNA mutation were derived from Mitomap (http://www.mitomap.org/MITOMAP) as mtDNA Mutations with Reports of Disease-Associations’ (data set 1; n = 395). The m.3243A>G, m.8344A>G and m.8993T>C/G mutations were highly overrepresented in the data and therefore these were reviewed separately to prevent a bias from these mutations (see below). We further excluded from data set 1 mtDNA variants without data on mutation level, variants not related to neuromuscular disorders, somatic (cancer-related) variants and homoplasmic mtDNA mutations with reduced penetrance (Fig. 1). Variants of which the pathogenicity status were listed in Mitomap as ‘unclear’, ‘polymorphism’, ‘synergistic’, ‘conflicting reports’, ‘secondary’, ‘haplogroupmarker’ or ‘warrants further study’, or was ‘unconvincing’ in the corresponding publication were also not included. All inherited pathogenic mutations (Mitomap ‘confirmed’) were included. For the remaining variants (Mitomap ‘reported’), a pathogenicity score was calculated for the tRNA/rRNA (McFarland et al., 2004) and protein-coding (Mitchell et al., 2006) mutations. A total of 160 mutations with a score of ‘pathogenic’ or ‘probably pathogenic’ were finally included in data set 1 and analysed (89 tRNA, 116s rRNA, 70 protein-coding). Details on the evaluation of the pathogenicity of the individual mutations and their inclusion/exclusion in the analyses are described in Supplementary data, Table S1.

Of the 415 remaining pedigrees in data set 1, mutation levels in all the available tissues and clinical status (affected or unaffected) were recorded from the proband and available maternal relatives. ‘Mildly’ affected maternal relatives were considered ‘affected’, even if the symptoms (such as diabetes mellitus and fatigue) might not necessarily be related to the mtDNA mutation, except if the phenotype did not match the clinical spectrum of the mutation and/or the mutation was not detected in the relative’s (affected) tissue(s). Miscarriages, stillbirths and unexplained infant deaths without clinical information were excluded. In pedigrees with late-onset disease, unaffected individuals younger than the youngest affected relative were excluded from the analysis. Pedigrees were categorized into three groups (Fig. 1). Pedigrees were considered ‘uninformative’ (122 pedigrees) if they contained a proband with an unknown or negative family history, and no information on mutation level detection in maternal relatives. Pedigrees with a ‘familial’ mtDNA mutation (185 pedigrees) had a positive family history, based on clinical symptoms and/or the presence of the mutation in maternal relatives. Pedigrees with a de novo mtDNA mutation (108) had a proband with a negative family history and a mutation level of 0% in the mother (regardless of which tissue(s) tested) and in all other tested maternal relatives. If the mother was not tested, a pedigree containing a proband with a negative family history and a mutation level of 0% in the proband’s sibling(s) and in all other tested maternal relatives were considered de novo.

All subsequent analyses contain mutation levels in post-mitotic muscle tissue only, as we assumed that these best reflect the prenatal tissues and are less affected by the changes observed for some mutations in dividing cells (Rahman et al., 2001). Although data on human prenatal mtDNA mutation levels are limited, previous studies have shown similar mtDNA heteroplasmacy levels in the blastomeres of a blastocyst and minimal variation in time and between tissues throughout embryo–fetal development (Harding et al., 1992; Matthews et al., 1994; Cardacli et al., 2000; Bouchet et al., 2006; Steffann et al., 2006, 2007; Monnot et al., 2011). This suggests that mutation levels in a blastomere reliably predict the mutation load of the future child.

Data collection and evaluation of the m.3243A>G (data set 2), m.8344A>G and m.8993T>C/G mutations

Expression thresholds and genotype–phenotype correlations have already been discussed for these mutations, especially for the m.8993T>C/G mutation (Chinnery et al., 1997; White et al., 1999). For the m.8344A>G mutation, data were insufficient for a separate analysis. For the m.3243A>G mutation, we used data on muscle mutation level from 392 affected and 26 unaffected mutation carriers derived from
309 m.3243A>G pedigrees [data set 2—203 published (Supplementary data, Table S3) and 106 unpublished].

**Statistical analysis of the mtDNA mutations**

The distribution of mutation level in the muscle is presented as box plots and by means of the relative number of individuals with a mutation level equal to or lower than a certain level for the range of individual mutation levels in data set 1. Mean mutation levels were compared using the two-sample Student’s t-test with robust standard errors to allow for any correlation in mutation level among individuals from the same pedigree and assuming independence between different pedigrees. An assumption on the proportion of affected individuals in ‘familial’ pedigrees was required in order to determine the probability of being affected at a certain mutation level for data sets 1 and 2. This was estimated by determining for probands with a ‘familial’ mutation the ratio of number of affected siblings to total number of (affected and unaffected) siblings regardless of whether the muscle mutation level was known for those siblings. Probands were excluded to prevent ascertainment bias. The number of affected siblings in each pedigree (excluding the index case) was assumed to be a binomial random variable with parameters: a constant probability of being affected and the pedigree size. Prediction of the risk of being affected at a certain mutation level was achieved from a logistic regression model with individuals weighted by the estimated risk of being affected or unaffected (one minus risk of affected) depending on their status. All statistical analyses were performed in Stata 11 (Stata statistical software, Release 11. StataCorp, College Station, TX, USA).

**Results**

**Analysis of mtDNA mutations in data set 1**

Data on muscle mutation level and clinical status of 159 different pathogenic mtDNA mutations, retrieved from 327 pedigrees (314 published and 13 unpublished), were combined, covering 385
affected individuals and 19 unaffected mutation carriers, as only few unaffected individuals underwent a muscle biopsy (data set 1). The ranges of mutation level varied widely in both groups but muscle mutation level tended to be higher in the affected individuals when compared with unaffected maternal relatives (Fig. 2A and B).

A possible bias could exist in data of the probands, as the sensitivity of the technology used to identify the mtDNA mutation is often lower (≏10–20% detection limit) than the specific test for the familial mutation in a maternal relative. To examine this, muscle mutation levels of probands were compared with those of affected maternal relatives. For familial mutations, little difference in the mean muscle mutation level was observed between probands (mean = 88.8%) and affected maternal relatives [mean 86.9; 95% confidence interval (CI) for difference in means: −3.4 to 7.3%, P = 0.47, Fig. 3A and B], suggesting no evidence of bias owing to incomplete ascertainment of probands with a low mutation level.

Separate analysis of tRNA versus protein-coding mtDNA mutations in affected individuals from all 327 pedigrees revealed comparable mean muscle mutation levels (Fig. 4A and B, tRNA mean = 81.0%, protein-coding mean = 80.0; 95% CI for difference in means: −3.4 to 5.3%, P = 0.66). Therefore, muscle mutation levels from both mutation types were combined. The mean muscle mutation level was higher in affected familial cases (mean = 86.3%) when compared with the de novo cases (mean = 75.0, 95% CI for difference in means: 6.3–16.2%, P < 0.001, Fig. 5A and B) and therefore, data from these different types of cases were not combined. Since the data on genotype–phenotype correlations will mainly be applied in counselling families with an inherited mtDNA mutation, only familial cases were included in our further analyses.

Muscle mutation levels from probands and affected maternal relatives of familial pedigrees of both tRNA and protein-coding mtDNA mutations (n = 195) were analysed together with unaffected carriers of these familial mtDNA mutations (n = 19; Fig. 6A and B). However, probabilities of being affected at a certain mutation level could not be determined directly from these graphs, as this required an assumption on the general proportion of affected and unaffected individuals in families, which obviously differs from the proportion of biopsies from affected and unaffected relatives. The a priori prevalence of being affected was estimated by averaging the proportion of affected siblings from the probands with a familial mtDNA mutation. Of the 185 pedigrees with a familial mtDNA mutation, 126 were sufficiently large to be included in this calculation. We checked for an ascertainment bias caused by this selection but this was not the case and the muscle mutation levels of all affected individuals in the included pedigrees were comparable with mutation levels in the excluded pedigrees (P = 0.43, Supplementary data, Fig. S1). Our calculation implied a general a priori probability of being affected of 0.477 (95% CI 0.415–0.540; Supplementary data, Table S4). Using this a priori probability and the pooled data on muscle mutation level from all affected and unaffected individuals from the pedigrees of familial mtDNA mutations, the risk of being affected at a certain mutation level was calculated (Fig. 6C) with the 95% CIs. This graph can be used to estimate a mutation level threshold at which the risk of being affected is acceptably low and such embryos would be eligible for transfer in PGD. For example, a 95% or higher chance of being unaffected was found at a muscle mutation level of 18% or less (95% CI) (Fig. 6C). At a mutation level of 18%, the predicted probability of being affected is 0.00744. Based on pre-test odds of 0.477/0.523, this results in a negative diagnostic likelihood ratio of 0.0082.
Analysis of the m.3243A>G mutation: data set 2
A separate probability curve (Fig. 7A) was constructed for the m.3243A>G mutation, using data on muscle mutation level from 392 affected and 26 unaffected mutation carriers derived from 309 m.3243A>G pedigrees (data set 2—203 published and 106 unpublished). Comparison of the mean muscle mutation levels of affected m.3243A>G carriers (n = 392) with the affected carriers of the other mutations (n = 195) shows that the mean mutation level is lower for the m.3243A>G carriers (95% CI for difference in means: 16.7–25.8%, P < 0.001, Fig. 7B).

Discussion
PGD of mtDNA disorders has been applied to a very limited extent and for only a few common point mutations (Steffann et al., 2006).
The main problem of PGD for heteroplasmic mtDNA point mutations is the inability to determine the mutation level below which embryos are eligible for transfer. Ideally, PGD of mtDNA disorders eliminates the risk of an affected child by selecting an embryo with 0% mutation level. However, such a stringent criterion reduces or, in some cases, even minimizes the chance of having an embryo transfer and of becoming pregnant. PGD of mtDNA mutations can also be used to transfer embryos with a sub-symptomatic mutation level. Our study (data set 1) provides risk estimations that determine the mutation load at which the risk of being affected is acceptably low and embryos would be eligible for transfer in PGD, irrespective of the exact mtDNA point mutation. It is highly likely based on our experience and that of others (Blok et al., 1997; Brown et al., 2001; Monnot et al., 2011) that mtDNA mutation carriers generally produce some oocytes below the threshold. This provides the opportunity for carriers of all heteroplasmic mtDNA point mutations to have a fair chance of healthy offspring, by the application of PGD, which is not available at present.

In case of a more frequent mtDNA mutation for which genotype–phenotype correlations have been described, assessment of a mutation-specific pathogenic threshold might be preferable (White et al., 1999). For the m.8993T>C/G mutations, expression thresholds have already been described (White et al., 1999). The threshold for a severe clinical outcome is estimated at 60–70% for the m.8993T>G and 80–90% for the m.8993T>C mutation (White et al., 1999). However, a key difference between this study and our estimations is the use of a dichotomized clinical score (‘mild’ versus ‘severe’) by White et al. (1999) when compared with ‘affected’ versus ‘unaffected’ in our study. Therefore, the m.8993T>C/G thresholds suggested by White et al. (1999) are probably an overestimation of the actual clinical threshold. Nevertheless, as far as these data of White et al. (1999) can be compared with our estimations, our predictions are on the safe side and do not increase the chance of affected offspring. At worst, application of a general minimal threshold level to these cases results in a too stringent cut-off level and loss of presumably healthy embryos. For the m.8344A>G mutation, insufficient data are available in the literature for a proper estimation of the probabilities of being affected based on mutation load. However, a relation between muscle mutation load and the presence of specific clinical symptoms has been described (Chinnery et al., 1997), showing that none of the most common clinical features associated with this mutation are present in m.8344A>G patients with a muscle mutation load below 70%, and that none of the 55 m.8344A>G cases had a mutation load below 50%. Other studies also indicate that clinical manifestation for this mutation occurs at mutation levels >50% (Antonicka et al., 1999; van de Glind et al., 2007; Choi et al., 2010; Wabbi et al., 2010).

The m.3243A>G mutation has been described as showing quite variable clinical manifestation, occasionally at (very) low mutation levels, which is confirmed by our analyses of this mutation (Fig. 7). However, reported data on this mutation are biased in several ways. First, the presence of the m.3243A>G mutation at relatively high frequency in the general population (Manwaring et al., 2007; Elliott et al., 2008) implies that certain associations of this mutation with disease could be a chance finding and irrelevant to pathogenesis. Especially, in case of a more common and/or milder phenotype (such as diabetes, deafness or ophthalmology), the presence of low m.3243A>G mutation levels might be unrelated to the disorder. Second, many pedigrees with this mutation have already been published, which creates a publication bias towards m.3243A>G patients with unusual clinical manifestation and/or mutation levels. Finally, the m.3243A>G mutation is detected by a very sensitive prescreening assay and not by a complete, less-sensitive screen of the entire mtDNA, as is the case for the other mtDNA mutations. When the m.3243A>G mutation is detected, even at low levels, usually no further (complete) mtDNA mutation analysis is performed, precluding the identification of another (causal) mutation. Owing to these biases, which do not apply to data set 1, we consider the m.3243A>G

Figure 5 The distribution of mutation levels of de novo versus familial mtDNA mutations. (A) Box plots showing median mutation level (%) in the muscle of affected individuals from pedigrees with a de novo mutation and a familial mutation (P < 0.001, Student’s t-test). (B) Relative number of affected individuals with a mutation level equal to or lower than a certain level.
analysis less reliable and less usable for predicting a common threshold of expression, as can be deduced from the graph, in which even at 0% mutation load, a chance of being affected remains. As a result, different thresholds are being used in the current PGD practice, ranging from 15 to 30% (own data; de Die-Smulders and Smeets, 2009; Monnot et al., 2009; Poulton and Bredenoord, 2010; Monnot et al., 2011), indicating the necessity for an unbiased evaluation of the m.3243A>G mutation.

Three corrections have been performed to neutralize in data set 1 a possible ascertainment bias in our study design: First, most of the mildly affected individuals were considered ‘affected’. They usually have lower mutation levels and their inclusion among the affected group leads to a more stringent cut-off level of clinical expression. Second, young unaffected individuals from pedigrees with late-onset disease were discarded from the analysis. Third, the m.3243A>G, m.8344A>G and m.8993T>C/G mutations were excluded to prevent ascertainment bias from the large amount of data on these mutations. Chances of being affected at a certain mutation level are only considered for familial cases, who might opt for PGD. As there was no significant difference in muscle mutation levels between probands and affected relatives, we included mutation levels of probands in the analyses. In contrast, to estimate the general proportion of affected individuals in the pedigrees (irrespective of the mutation level), exclusion of the probands is preferable. Furthermore, we are assuming that the mutation level in muscle does not vary significantly during an individual’s lifetime and is representative of levels in a blastomere. Although several previous studies have shown minimal variation of mtDNA mutation load between individual blastomeres of a carrier

Figure 6 The distribution of mutation levels of pathogenic mtDNA mutations in the muscle of affected versus unaffected individuals and probabilities of being affected at a certain mutation level. (A) Box plots showing median mutation level (%) in the muscle of affected individuals and unaffected maternal relatives. (B) Relative number of affected and unaffected individuals with a mutation level equal to or lower than a certain level. (C) Estimated probability of being affected (95% CI) based on mutation level in the muscle (logistic regression model), assuming a general probability of being affected of 0.477. In (A)–(C), muscle mutation level from all affected individuals and unaffected carriers from the pedigrees with a familial mtDNA mutation are pooled.
blastocyst, between tissues and over time throughout embryo–fetal development (Harding et al., 1992; Matthews et al., 1994; Cardaioli et al., 2000; Bouchet et al., 2006; Steffan et al., 2006, 2007; Monnot et al., 2011), hardly any data are available on muscle mutation load throughout post-natal life (Weber et al., 1997; Frederiksen et al., 2006). Therefore, the possibility of variation in post-mitotic muscle mutation level during life cannot be excluded. Also, the minimal pre-natal variation of mutation level is based on few studies only and might not apply to all mtDNA mutations. Mutation levels could be influenced by the method of mutation quantification and their clinical expression by additional environmental or genetic factors. Furthermore, regarding differences between techniques and laboratories in quantification of mutation load, although accurate quantification is essential and these differences will increase the uncertainty of our predictions, they are unlikely to have exerted a major influence on the estimated genotype–phenotype correlations. Moreover, in most of the literature studies from which our data are derived, mutation levels were examined by the restriction fragment length polymorphism technique. In addition, mutation levels represent mean levels and may differ among individual muscle cells. Finally, the high prevalence of low-level pathogenic mtDNA mutations in the general population could lead to false-positive results in screening (Manwaring et al., 2007; Elliott et al., 2008; Bitner-Glindzicz et al., 2009; Vandeboma et al., 2009). In general, our assumptions are on the safe side and do not increase the chance of affected offspring, although the price may be the discarding of healthy embryos.

PGD of heteroplasmic mtDNA mutations should be based on adequate (pre-test) case-by-case counselling, considering the uncertainties linked to this risk-reduction strategy (Bredenoord et al., 2009, 2010; Poulton and Bredenoord, 2010). Our estimation of 18% mutation level being associated with a 95% or higher chance of being unaffected can be applied as guidance for almost all mtDNA disorders, irrespective of the point mutation. This has to be judged on a case-by-case basis and may depend on the data generated by performing a PGD cycle. The exact threshold and other constraints should be discussed and determined during counselling of each individual couple, because it also depends on other factors, such as disease manifestation and perception of risk and of clinical severity in the family, the availability of embryos below the threshold (Hauswirth and Laips, 1982) and general fertility issues associated with IVF. In relation to this, each PGD cycle will provide additional information on heteroplasmy levels in a woman’s oocytes (Marchington et al., 2010) and the chance of selecting presumably healthy embryos below the threshold of expression.

**Supplementary data**

Supplementary data are available at http://humupd.oxfordjournals.org/.

**Authors’ roles**

D.M.E.I.H. coordinated the study, evaluated the data and wrote the paper. R.W. performed the statistical analysis and participated in data interpretation. A.T.M.H. was involved in the data collection. I.F.M.d.C., C.E.d.D., J.P.M.G. and P.F.C. reviewed and edited the report. H.J.M.S. conceived and supervised the study and was involved in writing the paper. All authors have read and approved the final manuscript.

**Funding**

This work was funded by Research School GROW.
Conflict of interest
None declared.

References


