

Optimization of the Y831C mutation detection in human DNA polymerase gamma by allelic discrimination assay

Katarzyna Stopińska¹, Tomasz Grzybowski¹✉, Boris A. Malyarchuk², Mirosława V. Derenko² and Danuta Miścicka-Sliwka¹

¹The Nicolaus Copernicus University, Ludwik Rydygier's Collegium Medicum, Chair of Forensic Medicine, Department of Molecular and Forensic Genetics, Bydgoszcz, Poland; ²Institute of Biological Problems of the North, Far-East Branch of the Russian Academy of Sciences, Magadan, Russia; ✉e-mail: tgrzyb@cm.umk.pl

Received: 28 March, 2006; revised: 09 August, 2006; accepted: 14 August, 2006
available on-line: 01 October, 2006

Many well-defined mutations in the gene for the catalytic subunit of polymerase γ (*POLG1*) have been found to be associated with disease, whereas the status of several mutations remains unresolved due to the conflicting reports on their frequencies in populations of healthy individuals. Here, we have developed a highly sensitive, real-time allelic discrimination assay enabling detection of the Y831C mutation in the *POLG1* gene. The Y831C mutation is present in the Polish population at a frequency of 2.25%. The new assay is well suited to both extensive population studies and molecular diagnostics of *POLG1*.

Keywords: single nucleotide polymorphisms (SNPs), polymerase γ , progressive external ophthalmoplegia, real-time PCR

INTRODUCTION

The nuclear-encoded DNA polymerase γ (POLG) is the sole DNA polymerase responsible for replication and repair of the mitochondrial genome (Ropp & Copeland, 1996). Human DNA polymerase γ is composed of a 140-kDa catalytic subunit and a 55-kDa accessory subunit (Kaguni, 2004, and references herein). The human gene for the catalytic subunit of polymerase γ (*POLG1*) maps to 15q24-15q26 and consists of 23 exons. The catalytic subunit has the polymerase domain in the carboxy-terminal region and the 3'-5' exonuclease domain in the amino-terminal region with proofreading activity (Ropp & Copeland, 1996). Mutations in *POLG1* are responsible for multiple deletions and depletions of mtDNA (DiMauro, 2004; Longley *et al.*, 2005; Spinazzola & Zeviani, 2005), autosomal recessive and dominant progressive external ophthalmoplegia (PEO) (Van Goethem *et al.*, 2001; Lamantea *et al.*, 2002; Copeland *et al.*, 2003), Alpers syndrome (Davidzon *et al.*, 2005), sensory ataxia, neuropathy, dysarthria and ophthal-

moperesis (SANDO), parkinsonism (Van Goethem *et al.*, 2003; Luoma *et al.*, 2004), and possibly, male infertility (Rovio *et al.*, 2001; Krausz *et al.*, 2004; Akin-Seifer *et al.*, 2005). In addition to many well-defined mutations associated with disease, 304 single nucleotide polymorphisms (SNPs) have been discovered in *POLG1*, with six synonymous and nine non-synonymous mutations occurring in the coding region (Graziewicz *et al.*, 2006).

The status of several mutations (pathogenic versus irrelevant SNPs) seems unresolved due to the conflicting reports on their frequencies in populations of healthy individuals. As an example, one may quote the heterozygous Y831C mutation in the gene for POLG catalytic subunit. The Y831C mutation was found by Barthelemy *et al.* (2002) in a patient with severe mtDNA depletion. Despite the important change in the amino-acid sequence (tyrosine to cysteine), the authors considered this mutation as non pathogenic since it was located outside the functional region of the protein, involved a poorly conserved amino acid and was present not only in

Abbreviations: CRS, Cambridge reference sequence; CT, threshold cycle; HVS I and HVS II, hypervariable segments I and II of the human mitochondrial DNA control region; MGB, minor groove binding moiety; mtDNA, mitochondrial DNA; PEO, progressive external ophthalmoplegia; POLG, human polymerase gamma; RFLP, restriction fragment length polymorphism; R_n , fluorescence values; ΔR_n , normalized fluorescence values; SNPs, single nucleotide polymorphisms.

the patient but also in her mother and grandmother, who were clinically healthy, without muscle mtDNA depletion. Furthermore, they found the same heterozygous mutation at an appreciable frequency of 57% in a French sample of 87 healthy controls (Barthelemy *et al.*, 2002). In contrast, Mancuso *et al.* (2004) presented a case of the Y831C mutation in a family with PEO, peripheral neuropathy and parkinsonism. They consider this mutation pathogenic since it is consistent with the clinical manifestation of both PEO and peripheral neuropathy, and the mutated amino acid is very close to motif A in the polymerase region of POLG. Most importantly, this mutation was not found in 130 healthy controls (Mancuso *et al.*, 2004). Interestingly, both groups used mispairing PCR coupled with restriction fragment analysis as the method of mutation detection (Barthelemy *et al.*, 2002; Mancuso *et al.*, 2004).

The background of these conflicting observations can be either a high genetic heterogeneity of European populations or errors in the analytical procedures of mutation detection. In this study, we have developed and optimized an alternative strategy for detection of the Y831C POLG mutation in genomic DNA. The assay is based on allelic discrimination with fluorescent *TaqMan* probes employed in real-time PCR. Before being used for extensive screening for the presence of the Y831C mutation in many Eurasian populations, the new allelic discrimination assay was validated and successfully applied to a control sample of 133 healthy Polish individuals.

MATERIALS AND METHODS

DNA extraction. Human genomic DNA was extracted from blood or saliva by a standard organic solvent method. Quantification of the DNA was performed spectrophotometrically, with the use of a GeneQuant DNA Calculator (Pharmacia Biotech) and agarose gel electrophoresis followed by ethidium bromide staining. Optimization experiments were carried out on 25 DNA extracts. The system was validated on a population sample of 133 ran-

domly selected, unrelated individuals from the Pomerania-Kujawy region of Poland.

Standard genotyping conditions. The standard amplification conditions were established after a series of optimization experiments (described in the next section). PCR amplification was performed in three replicates using 5–10 ng of genomic DNA in 10 µl reaction volume comprising 1 × *TaqMan* Universal PCR Master Mix; 0.3 µM of both *POLG1* primers and 0.1 µM of both probes. PCR primers and probes were designed for the complete DNA sequence of *POLG1* (GenBank accession No. AF497906) with the use of Primer Express® software (Applied Biosystems). The probes were 5'-labeled with a fluorescent reporter dye. One probe perfectly matched the "A" — normal sequence variant (at nucleotide position 15519) and was 5'-labeled with 6-carboxyfluorescein (FAM); the second probe matched the mutant sequence variant "G" and was 5'-VIC labeled. Both probes used in the study had a non-fluorescent quencher and a minor groove binding moiety (MGB). Primers and probes (Table 1) were synthesized and labeled commercially (Applied Biosystems). The PCR conditions were 50°C for 2 min and 95°C for 10 min, followed by 40 cycles at 92°C for 15 s and 60°C for 1 min.

The correctness of genotyping was evaluated by direct sequencing. Genomic DNA was amplified as described by Filosto *et al.* (2003) and the resulted PCR products were sequenced directly using BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) according to the manufacturer's protocols. Sequencing products were analyzed on an ABI 377 DNA Sequencer (Applied Biosystems).

mtDNA analysis. Hypervariable segments I and II (HVS I and HVS II) of the mtDNA non-coding control region were analyzed from samples heterozygous for the Y831C mutation. mtDNA was amplified and sequenced as described elsewhere (Malyarchuk *et al.*, 2002). The nucleotide sequences from position 15999 to 16400 (HVS I) and 30 to 407 (HVS II) were determined and compared to the Cambridge Reference Sequence (CRS) (Anderson *et al.*, 1981).

Table 1. Sequences of primers and probes used in real-time allelic discrimination assay

Nucleotide positions in probe sequences specific for allelic variants of *POLG1* are underlined. Amino acid translation is given below nucleotide sequences of the probes; codons for Tyr and Cys are shown in bold.

Name	Sequence	Size of PCR product (bp)
<i>Primers:</i>		
POLG_F	5' AGC ACC CTG CTA AGA CCC ATT 3'	108
POLG_R	5' TGC CGG CAG TCA CCA CTT 3'	
<i>Probes:</i>		
POLG_FAM2	5' CC GAC <u>TAT</u> GAT GAG GAA 3' Asp Tyr Asp Glu Glu	
POLG_VIC2	5' CCC GAC <u>TGT</u> GAT GAG 3' Pro Asp Cys Asp Glu	

Statistical analysis. The statistical significance of the population differences with respect to the frequencies of the Y831C mutation was evaluated using the chi-square test.

RESULTS

Real-time detection system

PCR products were detected by an ABI PRISM 7000 Sequence Detection System (Applied Biosystems). During PCR, fluorogenic *TaqMan* probes anneal specifically to complementary sequences between forward and reverse primer sites on the template DNA, and then are degraded owing to the 5'-3' exonuclease activity of DNA polymerase. Once separated from the quencher, the reporter dye emits fluorescence which is read by the 7000 system. By quantifying and comparing the fluorescent signals, the instrument software enables determination of the allelic content of each sample on the plate. The software graphs the result of the allelic discrimination on a scatter plot of allele 1 fluorescence values (R_n) versus allele 2 R_n (Fig. 1)

Primer concentrations

Primer concentrations were varied from 50 to 900 nM in a 50 μ l reaction. The combination of 50 nM forward and reverse primer gave both the lowest normalized fluorescence values (ΔR_n) and the highest threshold cycle (C_T). Other primer combinations including a 50 nM concentration of either forward or reverse primer gave a reduced ΔR_n . All primer combinations that contained at least 300 nM forward and reverse primer gave the highest ΔR_n and the lowest C_T . As a result, 300 nM concentration of each primer was considered to be optimal.

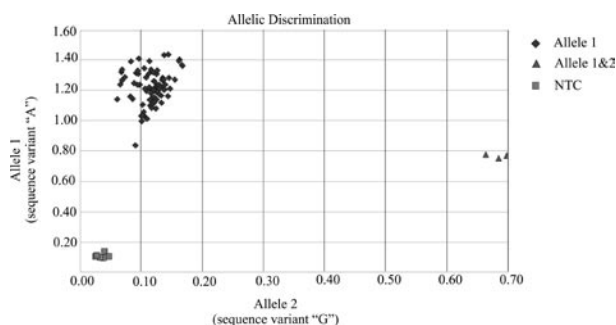


Figure 1. Result of Y831C mutation genotyping with the use of *TaqMan* MGB probes.

The plot contains no template controls (NTC, squares) and two out of the three possible genotypes: allele 1 homozygous (diamonds) and heterozygous (triangles). Thirty DNA samples and two no template controls were amplified on a plate, each sample in three replicates.

Probe concentrations

An optimization experiment in which the probe concentrations were varied from 100 to 200 nM showed an increase in ΔR_n when the probe concentration was increased, and an unchanged value of C_T for all probe concentrations. This suggested that the assay can be run at probe concentrations of 100 nM.

Validation study

Allelic discrimination assay was used for detection of the Y831C mutation in a Polish population sample. Out of 133 individuals, three appeared to have the Y831C mutation in a heterozygous state. All genotyping results were verified by direct sequencing and no discrepancies were observed between the two analytical strategies (Fig. 2).

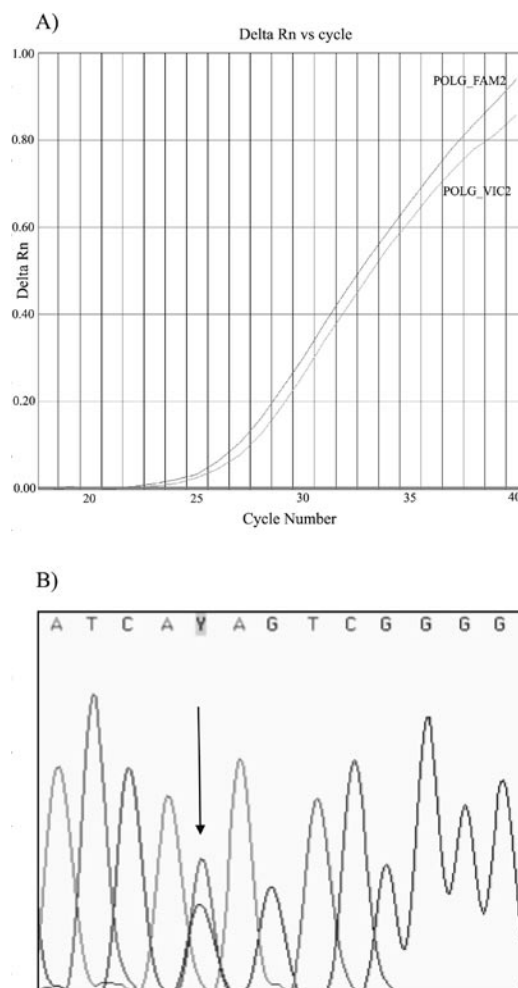


Figure 2. Genotyping results obtained for a sample heterozygous for the Y831C mutation, confirmed by direct sequencing.

A) Graph of allelic discrimination assay; FAM-labeled probe is complementary to allele 1 and VIC-labeled probe is complementary to allele 2. B) Fragment of electropherogram from direct sequencing. Heterozygous position is indicated by arrow.

Table 2. mtDNA sequences determined for DNA samples heterozygous for the POLG Y831C mutation

Mutations are shown indicating positions relative to the mtDNA Cambridge reference sequence (CRS) (Anderson *et al.*, 1981). The nucleotide positions correspond to transversions; insertions of cytosines are referred to by ".1C" following the nucleotide position.

Sample name	HVS I	HVS II
BYD 1484	CRS	73 152 263 315.1C
BYD 109	CRS	73 195 263 310
BYD 49	16342	263 309.1C 315.1C

mtDNA analysis

The samples heterozygous for Y831C were additionally subjected to mtDNA analysis. HVS I and HVS II sequences of the mtDNA control region are presented in Table 2. The haplotypes obtained had been observed previously in populations of Central and Eastern Europe (Malyarchuk *et al.*, 2002).

DISCUSSION

After the completion of human genome sequencing, much attention of the scientific community has been focused on the distribution of sequence diversity, which allows comparisons between populations within a species. Most human sequence variation is attributable to SNPs, which occur on average every 1000–2000 bases when homologous chromosomes are compared (The International SNP Map Working Group, 2001). Interest in the identification of SNPs is increasing, largely because of their potential use as molecular markers in disease association studies (Martin *et al.*, 2000). Therefore, the number of discovered SNPs is expected to increase in the near future and the development of new technologies for high-throughput SNP typing is expected to be one of the most vigorous areas of genetic research (Fornage & Doris, 2005).

Here, we have developed a highly sensitive, real-time allelic discrimination assay enabling detection of the mutation resulting in the Y831C substitution. The assay uses two primers and fluorogenic *TaqMan* probes in one reaction, allowing the researcher to identify single base changes in genomic DNA without the need for post-PCR analyses. After optimization, we successfully applied the new technology to screening the Polish population for the Y831C mutation. It was found in blood of 3 out of 133 randomly selected, clinically healthy individuals, suggesting the frequency of 2.25%. Allelic discrimination results were verified by direct sequencing. Additionally, the sequence of mtDNA control region was analyzed from the samples heterozygous for the Y831C mutation. The mtDNA analysis has not

revealed any hint of point or length heteroplasmy and the haplotypes determined were consistent with the pattern of natural mtDNA variation seen in the Polish and other European populations (Malyarchuk *et al.*, 2002). Therefore, these results do not provide any evidence for the pathogenicity of the Y831C mutation. On the other hand, it should be stressed that analysis of mtDNA segregation in tissues other than blood was not possible in the case of the individuals examined in the present study. Moreover, direct sequencing itself is not capable of detecting heteroplasmic variants occurring at low proportions (Tully *et al.*, 2000). Thus, the question of the impact of the Y831C mutation on mtDNA condition remains open. As for the frequency of the Y831C mutation, these results seem to correspond to the data of Mancuso *et al.* (2004) rather than those of Barthelemy *et al.* (2002). In the latter study, the Y831C mutation was found at a frequency of 57% in healthy individuals (Barthelemy *et al.*, 2002). The difference between this group and our data set is statistically significant ($P = 4.25 \times 10^{-7}$). Such a difference requires further investigation. Should this observation be supported by further, more extensive studies, this would be a case of a high genetic heterogeneity of European populations. It should be noted, however, that the results of Barthelemy *et al.* (2002), obtained with mispairing-PCR and restriction pattern analysis, were not confirmed by direct sequencing or any other method. It is noteworthy that traditional methods for assaying base-substitution variation, and mispairing PCR-RFLP in particular, are prone to errors. Digestion with a restriction enzyme has been the most widely used strategy to demonstrate the heterozygous state of a mutation. However, in the version of the technique with the mutated variant eliminating the digestion site, there is a risk of incomplete digestion which may lead to erroneous identification of heterozygosity in a wild-type homozygous sample. The issue of incomplete digestion as the source of errors in identification of numerous mutations in human mitochondrial DNA was raised by Finnilä *et al.* (1999) and Bandelt *et al.* (2005). Conversely, when the mutation creates a new restriction site (the version of mispaired PCR-RFLP employed by Barthelemy *et al.* (2002) and Mancuso *et al.* (2004)), there is always a risk of non-specific amplification and subsequent digestion, leading to false-positive results.

The allelic discrimination technology proposed here for the identification of the mutation causing the Y831C substitution is amenable to high-throughput genotyping and thus is well suited to both extensive population studies and molecular diagnostics. This assay, as many other tests based on real-time PCR (Johnson *et al.*, 2004), overcomes many of the problematic features associated with PCR-RFLP, including post-PCR manipulation, non standardized

assay conditions, manual allelic identification and false allelic identification due to non-specific amplification or incomplete enzyme digestion. Thus, the new strategy may serve as an alternative for traditional methods for assaying SNP variation, even though it is not amenable to multiplexing.

REFERENCES

- Aknin-Seifer LE, Touraine RL, Lejeune H, Jimenez C, Chouteau J, Siffroi JP, McElreavey K, Bienvenu T, Patrat C, Levy R (2005) *Hum Reprod* **20**: 736–740.
- Anderson S, Bankier AT, Barrell BG, de Bruijn MH, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJ, Staden R, Young IG (1981) *Nature* **290**: 457–465.
- Bandelt HJ, Yao YG, Kivisild T (2005) *Schizophr Res* **72**: 267–269.
- Barthelemy C, de Baulny HO, Lombes A (2002) *Hum Genet* **110**: 479–487.
- Copeland WC, Ponamarev MV, Nguyen D, Kunkel TA, Longley MJ (2003) *Acta Biochim Polon* **50**: 155–167.
- Davidzon G, Mancuso M, Ferraris S, Quinzii C, Hirano M, Peters HL, Kirby D, Thorburn DR, DiMauro S (2005) *Ann Neurol* **57**: 921–923.
- DiMauro S (2004) *Biochim Biophys Acta* **1659**: 107–114.
- Filosto M, Mancuso M, Nishigaki Y, Pancrudo J, Harati Y, Gooch C, Mankodi A, Bayne L, Bonilla E, Shanske S, Hirano M, DiMauro S (2003) *Arch Neurol* **60**: 1279–1284.
- Finnilä S, Hassinen IE, Majamaa K (1999) *Mutat Res* **406**: 109–114.
- Fornage M, Doris PA (2005) *Methods Mol Med* **108**: 159–172.
- Graziewicz MA, Longley MJ, Copeland WC (2006) *Chem Rev* **106**: 383–405.
- Johnson VJ, Yucesoy B, Luster MI (2004) *Cytokine* **27**: 135–141.
- Kaguni LS (2004) *Annu Rev Biochem* **73**: 293–320.
- Krausz C, Guarducci E, Becherini L (2004) *J Clin Endocrinol Metab* **89**: 4292–4297.
- Lamantea E, Tiranti V, Bordoni A, Toscano A, Bono F, Servidei S, Papadimitriou A, Spelbrink H, Silvestri L, Casari G, Comi GP, Zeviani M (2002) *Ann Neurol* **52**: 211–219.
- Longley MJ, Graziewicz MA, Bienstock RJ, Copeland WC (2005) *Gene* **354**: 125–131.
- Luoma P, Melberg A, Rinne JO, Kaukonen JA, Nupponen NN, Chalmers RM, Oldfors A, Rautakorpi I, Peltonen L, Majamaa K, Somer H, Suomalainen A (2004) *Lancet* **364**: 875–882.
- Malyarchuk BA, Grzybowski T, Derenko MV, Czarny J, Woźniak M, Miścicka-Śliwka D (2002) *Ann Hum Genet* **66**: 261–283.
- Mancuso M, Filosto M, Oh SJ, DiMauro S (2004) *Arch Neurol* **61**: 1777–1779.
- Martin ER, Lai EH, Gilbert JR, Rogala AR, Afshari AJ, Riley J, Finch KL, Stevens JF, Livak KJ, Slotterbeck BD, Slifer SH, Warren LL, Conneally PM, Schmechel DE, Purvis I, Pericak-Vance MA, Roses AD, Vance JM (2000) *Am J Hum Genet* **67**: 383–394.
- Ropp PA, Copeland WC (1996) *Genomics* **36**: 449–458.
- Rovio AT, Marchington DR, Donat S, Schuppe HCh, Abel J, Fritsche E, Elliott DJ, Laippala P, Ahola AL, McNay D, Harrison RF, Hughes B, Barrett T, Bailey DM, Mehmet D, Jequier AM, Hargreave TB, Kao SH, Cummins JM, Barton DE, Cooke HJ, Wei YH, Wichmann L, Poulton J, Jacobs HT (2001) *Nat Genet* **29**: 261–262.
- Spinazzola A, Zeviani M (2005) *Gene* **354**: 162–168.
- The International SNP Map Working Group (2001) *Nature* **409**: 928–933.
- Tully LA, Parsons TJ, Steighner RJ, Holland MM, Marino MA, Prenger VL (2000) *Am J Hum Genet* **67**: 432–443.
- Van Goethem G, Dermaut B, Löfgren A, Martin JJ, Van Broeckhoven C (2001) *Nat Genet* **28**: 211–212.
- Van Goethem G, Martin JJ, Dermaut B, Löfgren A, Wibail A, Ververken D, Tack P, Dehaene I, Van Zandijcke M, Moonen M, Ceuterick C, De Jonghe P, Van Broeckhoven C (2003) *Neuromuscul Disord* **13**: 133–142.