Detection and quantification of the age-related point mutation A189G in the human mitochondrial DNA.

Catherine Thèves, Christine Keyser-Tracqui, Eric Crubézy, Jean Pierre Salles, Bertrand Ludes, Norbert Telmon

Detection and Quantification of the Age-Related Point Mutation A189G in the Human Mitochondrial DNA

Catherine Thèves,[1,3]; Christine Keyser-Tracqui,[2] Ph.D.; Eric Crubézy,[3] M.D., Ph.D.; Jean-Pierre Salles,[1] M.D., Ph.D.; Bertrand Ludes,[2,3] M.D., Ph.D.; and Norbert Telmon,[3] M.D., Ph.D.

[1] INSERM, U563 Bat C, Purpan University Hospital, Place du Dr Baylac, 31059 Toulouse, France.

[2] Molecular Anthropology Laboratory, Department of Forensic Medicine, 11 rue Humann, 67085 Strasbourg, France.

[3] Anthropobiologie UMR 8555, 39 allées Jules Guesde, 31000 Toulouse, France.

Correspondence to:

Catherine Thèves INSERM, U563 Bat C, Purpan University Hospital, Place du Dr Baylac, 31059 Toulouse, France.

E-mail : Catherine.Theves@toulouse.inserm.fr

Fax: +33-561779401

Short running header: age-related point mutation and human mitochondrial DNA

ABSTRACT: Mutation analysis in mitochondrial DNA (mtDNA) control region is widely used in population genetic studies as well as in forensic medicine. Among the difficulties linked to the mtDNA analysis one can find the detection of heteroplasmy which can be inherited or somatic. Recently, age-related point mutation A189G was described in mitochondrial DNA and shown to accumulate with age in muscles. We carried out the detection of this 189 heteroplasmic point mutation using three technologies: automated DNA sequencing, Southern blot hybridization using a digoxigenin-labeled oligonucleotide probe and PNA/Real Time PCR combined method on different biological samples. Our results give additional information on the increase in mutation frequency with age in muscle tissue and revealed that the PNA/Real Time PCR is a largely more sensitive method than DNA sequencing for heteroplasmy detection. These investigations could be of interest in the detection and interpretation of mtDNA heteroplasmy in anthropological and forensic studies.

KEYWORDS: forensic science, mitochondrial DNA, heteroplasmy, age-related mutation, detection methods

The mitochondrial DNA (mtDNA) control region is a hypervariable sequence widely used in population genetic studies to investigate human evolution, history and patterns of migration as well as in forensic medicine for individual identification. Among the difficulties linked to mtDNA analysis is the detection of heteroplasmy (the presence of two or more mtDNA types in one individual), which can be inherited or somatic, and sometimes difficult to interpret. In recent years, numbers of reports have suggested that during individual natural aging, heteroplasmic point mutations accumulate in the control region of mitochondrial DNA (mtDNA) of some tissue types (1-4). This accumulation of mutations is thought to be due to the oxidation of DNA bases by free radicals, generated especially during cellular energy production by the mitochondrion (5,6). Among the reported heteroplasmic point mutations one can find the A to G transition at position 189 (A189G), which has been detected at significant levels in skeletal muscles from aged individuals by several authors (3,4). According to them, the A189G mutation accumulates with age in both healthy subjects and patients with mitochondrial diseases. However, because the level of heteroplasmy is lower in non-pathological cases, the mutation detection is technically difficult and requires a sensitive method (7-9). Moreover, the ratio of mutant (189G) to wild type (189A) has been reported as varying between different tissues from the same individual, which complicates the detection of low proportions of mutant mtDNA (10,9).

In the present study we searched for the presence of the A189G heteroplasmic mutation in two tissue types, muscle samples and buccal cells, with the following aims: (i) determine whether the A189G mutation was really age-related; (ii) test its presence in buccal cells; (iii) evaluate the most efficient method for detecting

low levels of heteroplasmy; and finally (iv) assess the possible implications of the A189G heteroplasmy in anthropological and forensic studies.

In order to prove the increase of mutant mtDNA molecules with age, buccal cells were collected from 37 individuals ranging from 4 to 85 years of age and belonging to 10 different maternal lineages. If the A189G mutation accumulates in an age-dependent manner, then it should be present in older subjects but not in younger ones of the same family. We also analysed muscle samples gathered from 69 subjects 1 to 97 years of age. These samples were analysed by means of three technologies: automated DNA sequencing, Southern blot hybridization using a digoxigenin-labeled oligonucleotide probe (11) and peptid nucleic acid (PNA) and quantitative Real Time PCR (qPCR) together (12,13).

Automated DNA sequencing is a technology widely and routinely used in research and forensic laboratories for population genetic studies or individual identification (14). Southern blot analysis was done using two probes (with the A or G base at position 189) labeled by a molecule of digoxigenin in 5'. The digoxigenin label can be detected with enzyme-linked immunoassays for anti-digoxigenin protein conjugates, and visualized through a chemiluminescent reaction (11). The PNA/qPCR combined method is a less employed methodology. The PNA molecule is a DNA mimic, in which the negatively charged sugar-phosphate DNA backbone is replaced by an achiral, neutral polyamide backbone formed by repetitive units of N-(2-aminoethyl) glycine (15). The PNA directed PCR clamping technique has been developed to detect single base-pair mutations (16,12). As shown in Fig. 1, PNA probe hybridizes perfectly to a target region and prevents PCR primer binding, whereas single base-pair mismatches containing PNA exhibit minimal primer binding inhibition. PNA directed PCR clamping was

used in combination with the Real Time PCR (17). The amplification process is monitored in real time by fluorescent dye SYBR Green incorporation and allows a relative quantification of heteroplasmy (13). Unlike other double-strand DNA-selective dyes, SYBR Green is excitable by visible wavelength and provides excellent sensitivity. This method of mutation detection does not require post-PCR manipulations, reducing assay time and the probability of laboratory contaminations.

Materials and Methods

Samples and DNA Extraction

Buccal cell samples were collected by cytobrush on 37 individuals (4 to 85 years old) who gave their informed consent. Each one of the 10 maternal lineages included 2 to 4 members. Buccal samples were incubated for 2h at 50°C in 400µl extraction buffer (EDTA 5mM, SDS 2%, Tris HCl 10 mM (pH 7.5), sodium acetate 0.3M) and 0.2mg proteinase K. A phenol/chloroform/isoamyl alcohol (25/24/1) extraction was carried out on the supernatant. The aqueous phase was purified with the Cleanmix™ Kit (Talent, Trieste, Italia). After purification, DNA was eluted into sterile water.

Muscle samples (psoas) were collected during forensic autopsies from 69 unrelated individuals ranging from 1 to 97 years in age. Individuals died accidentally and were exempt from any specific disease macroscopically detectable. Approximately $1cm^3$ of muscle fragments were incubated in 3 ml extraction buffer (EDTA 10 mM, NaCl 50 mM, SDS 2%, Tris HCl 10 mM (pH 7.5) and 0.2 mg proteinase K) at 58°C overnight. Total DNA was extracted according to the saturated sodium acetate precipitation method (18). An extraction

reagent blank negative control accompanied all muscle and buccal samples extraction throughout the entire analysis.

Sequencing Analysis

PCR amplification was performed using 1µl total DNA in a 50µl reaction volume containing $MgCl_2$ 1.5 mM, BSA 0.2 $mg.mL^{-1}$, 0.2 mM of dNTPs, primers 200 nM, 1X buffer and 0.5 U of Taq HotGoldstar DNA polymerase (Eurogentec, Seraing, Belgium). Primer sequences are listed in Table 1. Amplification was carried out in a T3 Thermocycler (Biometra). After a 10 min pre-incubation step at 95°C, PCR amplification was performed by direct PCR (35 cycles) using the following conditions: 95°C denaturation for 45 s, annealing at 53°C for 1 min and extension at 72°C for 1 min, followed by a 7 min final extension at 72°C. PCR products were analysed by electrophoresis in 1% agarose low melting (Sigma-Aldrich, St Quentin Fallavier, France). Purification of amplicon products was carried out on Microcon® PCR columns (Millipore, St Quentin-en-Yvelines, France). Sequencing reactions were performed using the BigDye™ Terminator kit (Applied Biosystems, Courtabeuf, France). The products were purified using the protocol recommended by the manufacturer and submitted to capillary electrophoresis on an ABI Prism 3100 Genetic Analyser (Applied Biosystems). The results were analysed using the Sequencing Analysis 3.7 Software (PE Applied Biosystems). Negative PCR controls accompanied samples and their extraction reagent blank negative controls throughout the sequencing process.

PNA Method and Real Time PCR

PNA sequences and complementary primers for the point mutation A189G are given in Table 1. PNA probes were purchased from Eurogentec (Seraing, Belgium). For PNA clamping reactions, PCR reactions contained: SYBR®Green Jump Start™ Taq Ready Mix™ (Sigma-Aldrich, St Quentin Fallavier, France), 5 µl of DNA lysate, 300 nM of each primer and 2 µM PNA probe or H_2O in 26 µl of final volume. The reaction mixture was subjected to a uracil-DNA glycosidase digestion step of 50°C for 2 min (the complex dissociates and the polymerase becomes fully active) followed by 95°C for 10 min and PCR cycling in the ABI Prism® 7000 Sequence Detection System (Applied Biosystems, Courtabeuf, France). Real Time PCR was run in 40 cycles: denaturation (30 sec at 94°C), PNA hybridization (30s at 69°C), primer annealing (30 s at 60°C) and extension (30 s at 72°C).

The cycle threshold (Ct) value is defined by the cycle at which SYBR® Green fluorescence increases significantly above the background as a result of specific amplification. Ct values were calculated by extrapolation to the *x* axis of the linear portion of the exponential growth phase. The reported cycle threshold (Ct) is the cycle number adjoining the x intercept (19).

Preliminary optimisation and validation of the PNA/qPCR combined method were carried out on two muscle DNA extracts. These samples (controls) were previously sequenced and identified as A (wild-type = WT) or G (mutant = M) at position 189. Linearity of template amplification relative to copy number was demonstrated. PCR efficiency (*E*) was calculated according to the equation: $E = 10^{(-1/slope)}$ and showed PCR efficiency of 2.14. A dissociation protocol (Applied Biosystems) was performed, in order to identify the amplicon according to its

melting temperature. Then a melting curve analysis confirmed the amplification of a single product with a specific melting temperature of 75°C.

Relative Quantification

A scale of G variant from 0 to 100 obtained by mixing in equimolar proportions DNA from M and WT controls established an experimental reference curve. In a first run, eleven points (in the following percentages: 0, 10, 20, …, 90 and 100%) were evaluated and amplification was repeated three times for each condition: without PNA probe, with PNA G probe or with PNA A probe (Table 1). Average amplification cycles were established for each point and an intra-assay precision test was calculated (19). A second run was conducted including extra points situated between 0 and 20% and between 80 and 100% of G variant. Then the points common to each run enabled us to calculate an interassay precision test (19).

Once this experimental curve was established (x = mutant %; y = G/A Ct ratio), the Statistica v5.1 computer software (StatSoft, Tulsa, OK) calculated the percentage of mutants by its inverse relationship (x = G/A Ct ratio; y = mutant %). For each sample on the same run an amplification triplicate was conducted for each PCR condition (without PNA probe, with PNA G probe or with PNA A probe). The cycle threshold obtained (Ct A for PNA A probe and Ct G for PNA G probe) was used to calculate the level of variant G or A present in the sample's mtDNA. A blank amplification as well as two M and WT controls were systematically included for each run.

Southern Blot

A 2 µl DNA was amplified by using primers listed in Table 1. PCR conditions were performed using a DNA thermal cycler (Perkin Elmer 2400) in a 26 µl reaction volume containing $MgCl_2$ 1.5 mM, dNTPs 200 µM, 400 nM of each primer, buffer 1X and 1U of Taq polymerase (Promega, Madison, WI). PCR steps were determined with an initial denaturation step at 94°C for 1 min, followed by 35 cycles including a denaturation step for 15 s at 94°C, a hybridization step for 30 s at 57°C, an extension step for 45 s at 72°C and a final extension step for 2 min at 72°C.

Southern blot was performed using standard techniques. PCR products were blotted onto membranes (immobilon^{TM} NY+, Millipore, Bedford, MA) and hybridized with 5'-digoxigenine labeled probes (Eurogentec, Seraing, Belgium) (Table 1). Detection of the digoxigenine-labeled hybridized probes was performed using DIG Luminescent Detection Kit as described by the manufacturer (Roche Applied, Penzberg, Germany). Detection of 5'-digoxigenine probes was performed by exposure to X-ray film and analysed using ImageQuant software v1.1 (Molecular Dynamics).

Results

Automated DNA Sequencing

All samples were initially analysed by automated DNA Sequencing. The region studied at position 189 is highly polymorphic, for example at sites 182 or 195. The probes (for the PNA or Southern blot techniques) were synthesized on 15 or 20 bases according to the Cambridge Sequence (20) with a variation of bases A and G at position 189 (Table 1). It is necessary to know the samples' sequences underlying the probe in order to select the samples with the complementary *Cambridge Sequence*.

Of the 69 muscle samples studied, 13 were polymorphic between positions 182 and 195 and could not be analysed by the PNA/qPCR or Southern blot methods. Their sequences were not complementary with those of the probes. Among the 13 polymorphic samples (Table 2), 9 were situated at the extremity of the probe (195 or 182) and 4 in a more central position (185 or 188). Among all samples analysed by sequencing, 11 (belonging to individuals of 38 to 97 years of age) had an A/G heteroplasmy at position 189 (Table 3). Among the 11 heteroplasmies detected, two individuals aged 68 and 77 years old, also had a polymorphism at position 195, and could not be analysed by the different methods using probes.

The 37 buccal samples from individuals of 4 to 85 years of age were also analysed by automated DNA sequencing. For 6 of them (belonging to two families of different maternal lineage) we were able to identify G at position 189 (Table 2) this transition was considered transmitted since it was found in all members of the family. Other polymorphisms were revealed by this sequencing method (Table 2) and located either at the end or in the central part of the probe's hybridization region. Collecting buccal samples was done by family grouping; polymorphism

detection excluded several families (representing a total of 14 individuals) from the PNA/qPCR or Southern Blot analysis. No A189G heteroplasmy was detected in any of the 37 buccal samples with sequencing (Table 3).

In order to determine the threshold of the automated DNA sequencing method, the experimental variant G scale was tested by this technique. A sequencing pattern G peak at position 189 was detectable for minimum 30% of G variants.

PNA/qPCR Tests

To test the linearity of the probe's PCR clamping according to the DNA concentration, serial-dilutions were performed between 100 and 0.001 ng from total DNA of WT and M controls. Differences in cycle threshold (ΔCt) between amplification with or without PNA were constant for total DNA concentrations between 100 and 0.1ng. Initial amplification of each sample without probe had to be included between 16 and 26 Ct. If the DNA concentration or quality was low, the samples could not be analysed because they stood outside the linearity of the probe's PCR clamping. For that reason, five muscle samples were excluded from the PNA/qPCR analysis.

To test the efficiency of the PCR clamping according to the percentage of mutants present in total mtDNA, we produced variant G ratios from WT and M controls. We tested both PNA probes during qPCR amplifications (Fig. 2a and 2b). Each graph demonstrated that the curves between the proportion of mutants and the PCR clamping were exponential and had inverse relationships. The Ct values included between 0 and 80% of G variant for the PNA G probe (Fig. 2a) and between 20 and 100% for the PNA A probe (Fig. 2b) varied only slightly. The differences between the Ct values with and without PNA probe (ΔCt) were

noticeable only between 80 and 100% of mutants for the PNA G probe (Fig. 2a) and between 0 and 20% for the PNA A probe (Fig. 2b). Finally, the use of only one probe would only have enabled us to analyze levels of mutants ≤ or ≥ to 80 % of variant G for the PNA G probe and levels of mutants ≤ or ≥ to 20% of variant G for the PNA A probe. However, in order to obtain opposite and complementary Ct values between both PNA probes, the level of G variant had to be located between 0 and 20% or 80 and 100%. From the cycle threshold ratio G/A we established a variant G reference curve, to estimate the level of mutation in a more precise manner. Fig. 2c demonstrated that the G/A Ct values led to a better resolution of the mutant levels between 20 and 80%.

Intra-assay and Interassay Precision Tests

Amplification triplicates were conducted on the same run and for each different condition (without PNA probe and with PNA G or A probe) to test the precision of the variant G curve. As suggested by the manufacturer (Applied Biosystems) a value ≤ to 0.5 Ct was considered acceptable between triplicates of the same experimental condition in a same run. Also, the intra-assay precision test values on WT and M controls in each of the three PCR conditions was situated between 0.1 and 0.5% (19). The test's precision on 2 runs for 2 different days, representing two triplicates for each control condition was also measured. The interassay precision test was situated between 1.5 and 4.4% for the 3 PCR conditions (19).

The experimental reference curve of the controls' G/A values (Fig. 2c) was reversed in order to obtain an estimated level of mutants in an unknown sample. We have excluded 0% and 100% points in this inverse relationship and their corresponding G/A Ct ratios (0.7 for 0% and 1.6 for 100%) because they lacked

complementary information. Statistica v5.1 software calculated a polynomial function (Fig. 3) from measured G/A Ct ratios and the values of mutant levels were obtained with a 1.4% average standard deviation (SD) (minimum SD: 0.02%, maximum SD: 4.1%). The function then was applied to the Ct G/A ratios of the samples.

PNA/qPCR Samples

In order to be statistically representative, minimum variation of ≤ 0.5 Ct was considered acceptable for all amplification triplicates of a given condition. The intra-assay precision test carried on both tissue types were comprised between 0.06% and 1.68% (19). Samples presenting unstable amplification in a run were either re-analysed or excluded.

Among the 23 buccal samples analysed by the PNA/qPCR method, 18 presented a level of heteroplasmy at position 189 superior to 5%, the maximum being 12.6% (Fig. 4a). These heteroplasmies were especially noticeable in 3 out of 4 members (4, 37, 60 and 85 years old) of the same maternal lineage (Fig. 4b; family 1). While the heteroplasmic mutation was detected in the three adults (at levels of 8 to 9.5%), it was not observed for the child. Similarly the heteroplasmic mutation was detected in the older members (73 and 80 years old at levels of 12.6% and 11.5% respectively) in families 2 and 4, and more slightly observed in the younger members (at levels of 1.5 to 8.9%) (Fig. 4b). From these family observations even though the level of mutation was low, we noticed that the A189G transition in its heteroplasmic form was a somatic mutation because it was little or not observed in young individuals. Considering all the individuals (Fig.

4a) heteroplasmy levels were low and there didn't seem to be an accumulation threshold age for this specific mutation.

The 50 muscle samples from individuals of 1 to 97 years old were analysed by the PNA/qPCR method. Fig. 5 represented the accumulation of the 189G mutation according to the subject's age. The levels of heteroplasmy were higher in the muscle tissues compared to the buccal samples. In fact, for individuals 60 years old or more, the level of mutants was ≥ to 20% for 10 out of 12 individuals (against 12.6% for buccal cells). On the contrary, for the individuals of less than 40 years old, the percentage of mutants was low, ≤ to 10% for 18 out of the 23 individuals analysed. The PNA/qPCR method also detected the 189G polymorphisms identified in the 6 buccal samples by sequencing (Table 2). As expected, the G/A Ct values were similar to the control M (i.e. ≥ 1.6).

Southern Blot

All buccal and muscle samples with the complementary *Cambridge Sequence* to the probe's hybridization region were analysed by Southern blot. The detection of the 189G heteroplasmic form in muscle tissues was important in individuals of more than 60 years old (Fig. 6a). The labeling's relative quantification was done by the ImageQuant software and the percentage of mutants detected in the samples was calculated compared to M control (an equal quantity of control or samples DNA was deposited). Concerning the individuals under 60 years old (Fig. 6b), with a lower percentage of G variants, detection was slighter but present up to 40 years old. The individuals of less than 40 years old (with a few exceptions) were lightly or not labeled. The method's detection limit of the percentage of G variant seemed to be 10%. On the 50 muscle samples analysed by the Southern

blot method, detection of the 189G heteroplasmic form was demonstrated in 27 individuals in levels close to 10% and quantified by the ImageQuant software (Table 3). On the contrary, detection of mutant molecules was almost impossible in buccal cells and did not allow sample discrimination.

Moreover, we obtained from two deceased individuals of 63 and 70 years old both types of tissue samples and as shown in Fig. 6c, the 189G heteroplasmy was clearly detected in the muscle tissues of each one of them (with a percentage of 12% and 22% respectively) and was slightly present in buccal cells (≤ 5%).

Discussion

In order to minimise possible contaminations, which can be disastrous in the research of heteroplasmies, negative controls were done at each of our method's steps (extraction, PCR, sequencing, real time PCR). Our laboratory followed the working conditions suggested by the EDNAP group in order to prevent any possible contamination (21). Pre and Post PCR areas were separated and all disposables treated by UV cross-linking. Concerning the sequencing method, samples were analysed on both their DNA strands to confirm heteroplasmies. They were analysed only when both bases appeared visible above the background and on both strands. A positive control as well as the laboratory personnel's sequence were determined in order to test the absence of sample contamination. Our DNA samples were of good quality and their amount sufficient for a good PCR amplification. That is why the number of PCR cycles was limited to 35 for the initial PCR amplification of the specific strand before its analysis by sequencing. Concerning the three detection methods, all PCRs were conducted with a number of cycles between 35 and 40. Indeed, when the amount of DNA is

low, the number of cycles needs to be increased, for example in hair mtDNA analysis, and heteroplasmic artefacts should occur (22). The five muscle samples whose amplification happened late by the PNA/qPCR method (Ct without probe was over 30) were excluded from our study because they were not situated within the linearity of the PCR clamping. Consequently, the levels of heteroplasmies found on our healthy muscle samples were similar to those previously published (4). We also excluded from our studies individuals whose autopsies revealed any mitochondrial or neurodegenerative pathologies. The muscle samples all came from the psoas, a muscle with low energetic needs on working or living conditions. Within the individuals chosen for the buccal samples none presented pathologies linked with oxidative stress.

Analysis by sequencing of PNA probes hybridization region polymorphism enabled us to evaluate the effects of these polymorphisms on the PNA/qPCR method. The PNA probe's sequence mismatch leads to instability and an important loss in the probe's hybridization temperature. Loss in hybridization temperature is more or less important depending on the mismatched base and its location in the probe's sequence (16).

If the mismatch is central, the instability is greatest and hybridization temperature can rise up to 15°C. We decided to locate the nucleotides A and G at position 189 in the centre of the probe so that an eventual mismatch could lead to an important temperature loss. However, when the mismatch occurs near the probe's extremities, temperature loss is less significant (around 7 or 8°C) and primer or PNA hybridization can take place in various proportions independently from the base at the position 189. This kind of probe or primer hybridization can be observed in polymorphic samples near the extremities of the PNA probes and can

lead to important Ct variations during real time PCR cycles. Such polymorphic samples presented higher Ct variations than the accepted value (0.5 Ct) between triplicates of the same run. In consequence, we couldn't take into consideration the polymorphic samples' Ct values. Concerning the samples containing central polymorphisms (for example: positions 185 or 188), the amplification result was different. Whatever the PNA probe, the mismatch occurred systematically because of the central polymorphism leading to an important loss in hybridization temperature. With or without PNA probe, the amplification slightly varied (≤ 1 Ct).

It is interesting to notice that two polymorphic samples in the position 195 had a heteroplasmy in position 189 detectable by sequencing, in rates higher than 30%, and were from individuals 77 and 68 years old. In order to analyse them completely, and samples where heteroplasmy was not detected by the sequencing method, we would need to synthesise PNA probes whose sequence would correspond to the most common polymorphisms. Among all our samples (muscle and buccal) 195 and 188/185 positions are the most polymorphic sites.

Comparing automated sequencing and PNA/qPCR methods demonstrated two facts. Where the levels of 189G mutant molecules were > 30%, including 9 muscle samples submitted to both techniques, the percentage of G variants obtained, either by the peak heights on the sequencing pattern or by the qPCR's Ct G/A ratio, was the same. This observation revealed that our evaluation method was correct if certain precautions like minimal variations between same sample triplicates (≤ 0.5 Ct) were respected and if a standard deviation of 1.4% was included. Concerning the lower levels of heteroplasmy, our study demonstrated that the PNA/qPCR method was more sensible. In fact, 44 muscle samples were

identified as heteroplasmic in position 189, when only 9 were detected in the older individuals by automated sequencing (Table 3).

This sensibility difference between sequencing and PNA/qPCR methods, incited us to use a third detection method, the Southern blot in order to evaluate and confirm the previously detected heteroplasmies by the PNA/qPCR method. The Southern blot method confirmed the PNA/qPCR results in detecting levels of heteroplasmy as low as 10%. As shown in Fig. 6a and 6b, on individuals of 40 years old or more, this technique detected the 189G form of heteroplasmy in proportions of 50 to 10%. Moreover, relative quantification of the labeling by the probe on mutants using ImageQuant software confirmed the proportion of mutant levels analysed through both methods. For example, in Fig. 6a, individuals of 87 and 74 years old had an estimated percentage of mutants respectively of 44% and 21% whereas the 77 year old individual's labeling was estimated at 1% (considerably under this technique's the level of detection). Estimated percentage of G variants with the PNA/qPCR method on these same individuals was of 48% for the 87 year old, 19.7% for the 74 year old and 5.3% for the 77 year old. This was also true for lower levels as shown in Fig. 6b where the 43 year old individual presented a Southern blot labeling of 15.5% with the M probe and of 13.2% with the PNA/qPCR method.

However, concerning the buccal samples, the Southern blot technique did not reveal, the really low heteroplasmy levels detected by the PNA/qPCR technique, because, the variant G rates remained very low in buccal cells with a maximum of 12.6%. As shown in families 2 and 4 (Fig. 4b) the 73 and 80 year old individuals represented an estimated 12.6% and 11.5% of mutants with the PNA/qPCR method, these individuals showed a very weak labeling with the Southern blot

technique, estimated ≤ to 5% by the ImageQuant software. The difference in detection can be explained by the detection limits of the Southern blot technique, but also by an estimated percentage of errors of the PNA/qPCR technique and/or experimental differences.

Our study exerts another important fact: the accumulation of mutations is higher in skeletal muscle (20 to 50% in the older individuals) than in buccal cells (12.6% maximum) (Fig. 4a and 5). These high mutation levels in muscle tissue have been described previously (3,4). This difference could be explained by the important mitotic activity in buccal cells. MtDNA mutations, associated with aging, are frequently found to accumulate themselves rapidly in tissues with high energetic needs and slow mitotic activity, like skeletal muscle, heart and brain (6,23,24). This could explain why the mutation level in buccal cells did not have an accumulation threshold age (Fig. 4a) and that the variation in mutation percentage between a 40 year old individual and a 60 year old one, cannot be found, because the rapid cell turn-over could conceal the mutation accumulation. Therefore, even in advanced ages (3 individuals are over 80 years old among the buccal samples), the mutation level did not exceed 10%, cellular turn-over overcoming the cumulating mutational process.

The fact that the A189G mutation accumulates itself in a tissue-specific manner is clearly established by double sampling on two individuals (of 70 and 63 years old). For both, we collected muscle and buccal samples and analysed them by the PNA/qPCR method. Concerning the 70 year old, the percentage of mutant molecules was of 8.9% in buccal cells and of 21.2% in muscle tissue. A similar situation was observed for the 63 year old since the level of mutant was 3.5 times higher in muscle tissues than in buccal cells (3.9% versus 14.3%). These results

were confirmed by the Southern blot method (Fig. 6c) where the 70 year old individual presented a higher labeling level in muscle tissue than in buccal cells. This was also confirmed by ImageQuant software, the level of mutation in muscle tissue was of 22% compared to buccal cells where it was ≤ to 5%. The variation in the mutation frequency between different tissues has been previously established in mitochondrial pathologies (10), implying once again the cellular turn-over.

Simultaneous use of three technologies in our work, besides demonstrating better sensitivity of the PNA/qPCR method, revealed an interesting technical point. Automated DNA sequencing is a widely employed technique to detect mtDNA mutations; nevertheless, it has been demonstrated by serial-dilution experiments that low levels (≤ 20%) of mtDNA heteroplasmy could not be reliably detected by this technique (25). In this study, we also showed that automated DNA sequencing did not allow a quantitative estimation of the proportion of A189G mutations < to 30%. This technique usually uses dye-labeled terminators, which are dideoxynucleotide triphosphates (ddNTPS) labeled with different fluorophores (26). A major drawback of the dye-terminator chemistry is the highly uneven peak pattern produced after the amplification reaction, this latter being largely dependent on the local sequence context and of the DNA polymerase used (27). Dideoxynucleotides labelled with BigDye™ Terminator show more uniform peak heights and less G suppressions (14). However, Zakeri et al. (28) have analysed the context dependency of peak height variations to improve data interpretations. A common problem occurs when a G follows an A. The size of dye-labeled G peaks was lower than the height of the preceding A peak in the sequencing pattern. This was true for sites of mixed A/G bases in heterozygous mutations (29). This variability in peak height reduced base-calling accuracy of software and

hindered heterozygous mutation detection. In the case of the 189 position, where the local sequence was either GAA or GAG, the G BigDye™ Terminator was incorporated with difficulties in comparison with the A BigDye™ Terminator, giving rise to an A peak with an underlying G peak in this sequence. These problems, low levels of heteroplasmy and G BigDye™ incorporation after or with A BigDye™, could explain detection differences between automated sequencing and the PNA/qPCR method as well as the differences in detection levels of 20% for Chinnery et al. (25) and of 30% in our study.

An early study already demonstrated difficulties to detect heteroplasmies according to different techniques employed (30). The sensitive Denaturing Gradient-Gel Electrophoresis (DGGE) revealed heteroplasmies detection at levels as low as 1%. This way, the mtDNA heteroplasmy frequency was more common in comparison with those reported by sequencing analysis. Very recently, Hancock and colleagues (31) evaluated the sentivity of other detection techniques. Particurlarly, the automated DNA sequencing, including several chemistries and genetic analyzers, allowed detection at 30% levels. At 20% or 10%, the heteroplasmy was difficult to distinguish from the baseline noise, and was detectable only if the position was known. On the contrary, the PNA technique was able to detect the mutation at 5% level.

As seen in our work, the PNA-directed PCR clamping technique is a very specific and highly sensitive method to identify single nucleotide changes in DNA molecules (12). According to Murdock's team, it allowed the detection of mutant molecules at a level of 1% of total molecules by direct PCR (2,32). Individual nucleotide bases are attached to each peptide unit to provide a molecular design that enables PNA to hybridise to complementary nucleic acid targets according to

the Watson and Crick base-pairing rules (33). PNA-DNA hybrids are more stable than DNA-DNA hybrids and PNA hybridisation kinetics is less influenced by sodium concentration. PNAs are more sensitive to internal base-pair mismatch with their DNA complement; this can result in a 10 to 18°C lower melting temperature. Moreover, PNA is not recognized by polymerases and therefore cannot be copied or used as a primer. PNA probes are more robust detector molecules for Real Time PCR methods than many of their DNA-derived counterparts, such as the hybridisation probes which are degraded during PCR by the endonuclease activity of Taq DNA polymerase (34). This technique combined with Real Time PCR allowed a very sensitive quantification and an excellent reproducibility.

The PNA/qPCR sensitivity and accuracy compared to that of automated DNA sequencing could be interesting in forensic investigations where detection of heteroplasmy is not unusual. Technically, the advantage of the PNA/qPCR method is that it requires very few pre PCR manipulations and none post PCR. One run including amplifications without and with both probes is enough to obtain fast and reliable results on the level of heteroplasmy in a sample with minimum contamination possibilities. In order to limit the lack of precision, we can incorporate both WT and M controls in the run but also the experimental scale of G variant on the sample's amplification run, adding internal precision to the method with a limited margin of error. It is nonetheless necessary to know ahead of time the studied samples' sequences in order to apply the corresponding PNA probe for the 189 position. Other age-related heteroplasmies have been described at positions 414 and 408 (1-2,4), their investigation could benefit from this technique and their hybridization region is less polymorphic sites requirering only

two PNA probes. The PNA/qPCR method is rapid and flexible, avoids laboratory contaminations by limiting the number of manipulations and enables a fine detection in levels ≤ to 10%.

Moreover, the fact that the level of heteroplasmy varies in the different tissues of an individual, and that it can be expressed at very low levels in some, (i.e. buccal cells versus muscle tissue) increases interpretation difficulties. That is where a very sensible method can come in use to obtain very low levels of detection. Interpretation in forensic cases can sometimes be complicated because of the presence of heteroplasmies in the given sample (hair for example) and of its absence in others (blood for example). Consequently, demonstrating the presence of mutant molecules in two different samples can help eliminate uncertainties. On the contrary, if the number of mutants is higher in the muscle samples, other tissue samples should be preferred (like buccal cells or blood) for automated mtDNA sequencing in forensic cases to minimise interpretation uncertainties.

Because a correlation was clearly established between the age of the studied individual and the level of mutant molecules at position 189, the age of the subject tested should be considered when interpreting mtDNA typing results. As previously underlined by Calloway et al. (9), age and heteroplasmy should be considered in missing person cases when a significantly older individual is used as a reference for a missing younger maternal relative. In the same way, this information is useful for interpretation of mtDNA caseworks including mass disaster biological remains.

It is obvious, as shown in Fig. 5, that evaluating only one heteroplasmic mutation point cannot help determine the analysed individual's age and this was not the aim of this study. Nonetheless, we could suggest evaluating several heteroplasmic

mutation points and/or mitochondrial deletions. In fact, mitochondrial deletions accumulate with age, and were quantified using Real Time PCR by von Wurmb-Schwark et al. (35). Counter-crossing the values of several heteroplasmic mutation points or deletions, could help recognize these age-related mutations in relation with human aging. We should also investigate a possible link between these mutations and osteologic markers of normal and pathological aging widely used in forensic medicine, increasing label performances in identification cases.

Conclusion

The analysis of heteroplasmic A189G mutation by PNA/qPCR method permitted a sensitive detection as well as reliable and reproducible quantification of mutant level. This technique allowed us to demonstrate the absence of the A189G transition in buccal cells in young individuals and its presence in older individuals from the same maternal lineage, concluding that it is a somatic mutation. Moreover, we demonstrated for the first time the accumulation with age of the A189G mutation in mitotic buccal cells in levels lower than 13%. In post-mitotic muscle tissues, mutation accumulation was age-related and reached very high levels in 60 year-old or more individuals. Our results could have many implications in heteroplasmy interpretation in forensic caseworks but also in anthropological studies, for example in the investigations of great names of history.

Acknowledgements

We are grateful to Pr. D. Rouge and B. Martinez (Service de Médecine Légale de Rangueil, Toulouse) for the access to muscle tissue materials.

We would like to thank S.Romac (Laboratoire d'Anthropologie Moléculaire, Strasbourg) for technical help, C.Cariven and N.Malet (INSERM, U563 Toulouse) for technical advices and S.Laurencin for helpful English writing.

Grant Sponsor: ACI « Espaces et Territoires : le Complexe Spatial Altaï-Baïkal. Plaque tournante des flux géniques en Haute Asie de la période protohistorique à l'époque moderne »

References

1. Michikawa Y, Mazzucchelli F, Bresolin N, Scarlato G, Attardi G. Aging-dependent large accumulation of point mutations in the human mtDNA control region for replication. Science 1999;286:774-9.

2. Murdock DG, Christacos NC, Wallace DC. The age-related accumulation of mitochondrial DNA control region mutation in muscle, but no brain, detected by a sensitive PNA-directed PCR clamping based method. Nucleic Acids Res 2000; 28(21):4350-5.

3. Wang Y, Michikawa Y, Mallidis C, Bai Y, Woodhouse L, Yarasheski KE, et al.. Muscle-specific mutations accumulate with aging in critical human mtDNA control sites for replication. Proc Natl Acad Sci USA 2001;98(7):4022-7.

4. Del Bo R, Bordoni A, Boneschi FM, Crimi M, Sciacco M, Bresolin N, et al. Evidence and age-related distribution of mtDNA D-loop point mutations in skeletal muscle from healthy subjects and mitochondria patients. J Neurol Sci 2002;202(1-2):85-91.

5. Harman D. The biologic clock: the mitochondria? J Am Geriatr Soc 1972;20(4):145-7.

6. Linnane AW, Marzuki S, Ozawa T, Tanaka M. Mitochondrial DNA mutations as an important contributor to ageing and degenerative diseases. Lancet 1989;1:642-5.

7. Jazin EE, Cavelier L, Eriksson I, Oreland L, Gyllensten U. Human brain contains high levels of heteroplasmy in the noncoding regions of mitochondrial DNA. Proc Natl Acad Sci USA 1996;93:12382-7.

8. Wilson MR, Polanskey D, Replogle J, Di Zinno JA, Budowle B. A family exhibiting heteroplasmy in the human mitochondrial DNA control region reveals both somatic mosaicism and pronounced segregation of mitotypes. Hum Genet 1997;100:167-71.

9. Calloway CD, Reynolds RL, Herrin GL Jr, Anderson WW. The frequency of heteroplasmy in the HVII region of mtDNA differs across tissue types and increases with age. Am J Hum Genet 2000;66:1384-97.

10. Chinnery PF, Zwijnenburg PJG, Walker M, Howell N, Taylor RW, Lightowlers RN, et al. Nonrandom tissue distribution of mutant DNA. Am J Med Gen 1999;85:498-501.

11. Holtke HJ, Ankenbauer W, Muhlegger K, Rein R, Sagner G, Seibl R, et al. The digoxigenin (DIG) system for non-radioactive labelling and detection of nucleic acids – an overview. Cell Mol Biol 1995;41(7):883-905.

12. Ørum H. PCR clamping. Curr Issues Mol Biol 2000;2(1):27-30.

13. Kyger EM, Krevolin MD, Powell MJ. Detection of the hereditary hemochromatosis gene mutation by Real-Time fluorescence polymerase chain reaction and peptide nucleic acid clamping. Analytical Biochemistry 1998;260:142-8.

14. Rosenblum BB, Lee LG, Spurgeon LG, Khan SH, Menchen SM, Heiner CR, et al. New dye-labeled terminators for improved DNA sequencing patterns. Nucleic Acids Res 1997;25(22):4500-04.

15. Nielsen PE, Elghom M, Buchardt O. Peptide nucleic acid (PNA), a DNA mimic with a peptide backbone. Bioconjugate Chem 1994;5:3-7.

16. Ørum H, Nielsen PE, Elghom M, Berg RH, Buchardt O, Stanley C. Single base pair mutation analysis by PNA directed PCR clamping. Nucleic Acids Res 1993;21(23):5332-6.

17. Wittwer CT, Herrmann MG, Moss AA, Rasmussen RP. Continuous fluorescence monitoring of rapid cycle DNA amplification. Biotechniques 1997;22:130-8.

18. Cattaneo C, Smillie DM, Gelsthorpe K, Piccinini A, Gelsthorpe AR, Sokol RJ. A simple method for extracting DNA from old skeletal material. Forensic Sci Int 1995;74:167-74.

19. Gibson UEM, Heid CA, Williams MP. A novel method for Real Time Quantitative RT-PCR. Genome Res 1996;6:995-1001.

20. Anderson S, Bankier AT, Barrel BG, de Bruijn MHL, Coulson AR, Drouin J, et al. Sequence and organization of the human mitochondrial genome. Nature 1981;290:457-65.

21. Tully G, Bar W, Brinkmann B, Carracedo A, Gill P, Morling N, et al. Considerations by the European DNA profiling (EDNAP) group on the working practices, nomenclatures and interpretation of mitochondrial DNA profiles. Forensic Sci Int 2001;124:83-91.

22. Budowle B, Allard MW, Wilson MR. Critique of interpretation of high levels of heteroplasmy in the human mitochondrial DNA hypervariable region I from hair. Forensic Sci Int 2002;126:30-33.

23. Nagley P, Wei YH. Ageing and mammalian mitochondrial genetics. Trends Genet 1998;14(12):513-7.

24. Wallace DC. Mitochondrial diseases in man and mouse. Science 1999;283:1482-8.

25. Chinnery PF, Taylor GA, Howell N, Brown DT, Parsons TJ, Turnbull DM. Point Mutations of the mtDNA control region in normal and neurodegenerative human brains. Am J Hum Genet 2001;68:529-32.

26. Rosenthal A, Charnock-Jones DS. New protocols for DNA sequencing with dye terminators. DNA Seq 1992;3:61-4.

27. Parker LT, Zakeri H, Deng Q, Spurgeon S, Kwok PY, Nickerson DA. AmpliTaq® DNA Polymerase, FS Dye-Terminator sequencing: analysis of peak height patterns. Biotechniques 1996;21:694-9.

28. Zakeri H, Amparo G, Chen SM, SpurgeonS, Kwok PY. Peak height pattern in dichloro-rhodamine and energy transfer dye terminator sequencing. Biotechniques 1998;25(3): 406-14.

29. Korch C, Drabkin H. Improved DNA sequencing accuracy and detection of heterozygous alleles using manganese citrate and different fluorescent dye terminators. Genome Res 1999;9:588-95.

30. Tully LA, Parsons TJ, Steighner RJ, Holland MM, Marino MA, Prenger VL. A sensitive Denaturing Gradient-Gel Electrophoresis assay reveals a high frequency in hypervariable region 1 of the Human mtDNA control region. Am J Hum Genet 2000;67:432-43.

31. Hancock DK, Tully LA, Levin BC. A standard reference material to determine the sensitivity of techniques for detecting low-frequency mutations, SNPs, and heteroplasmies in mitochondrial DNA. Genomics 2005;86:446-61.

32. Murdock DG, Wallace DC. PNA-mediated PCR clamping. In: Nielsen PE, editor. Methods in molecular biology, vol 208: peptide nucleic acids: methods and protocols., Totowa NJ: Humana Press Inc., 2003;145-64.

33. Elghom M, Buchardt O, Christensen L, Behrens C, Freier SM, Driver DA, et al. PNA hybridizes to complementary oligonucleotides obeying the Watson-Crick hydrogen bonding rules. Nature 1993;365:556-68.

34. Wilhem J, Pingoud A, Hahn M. Comparison between Taq DNA polymerase and its Stoffel fragment for quantitative real-time PCR with hybridisation probes. Biotechniques 2001;30:1052-62.

35. Von Wurmb-Schwark N, Higuchi R, Fenech AP, Elfstroem C, Meissner C, Oehmichen M, et al. Quantification of human mitochondrial DNA in a real time PCR. Forensic Sci Int 2002;126:34-9.

Additional information and reprint requests:

Catherine Thèves

INSERM, U563 Bat C

Purpan University Hospital

Place du Dr Baylac

31059 Toulouse, France

E-mail : Catherine.Theves@toulouse.inserm.fr

Fax: +33-561779401

TABLE 1-*Sequences of the primers and probes used in this study.*

	Sequence (5'-3') by Anderson and al. (20)
Sequencing primer: L29 strand	GGT CTA TCA CCC TAT TAA CCA C
Sequencing primer: H408 strand	CTG TTA AAA GTG CAT ACC GCC A
PNA probe (189 position: **A**)	CAG GCG A**A**C ATA CTT
PNA probe (189 position: **G**)	CAG GCG A**G**C ATA CTT
Real Time PCR primer: L166 strand	CCT ACG TTC AAT ATT ACA GGC GA
Real Time PCR primer: H253 strand	GTG GAA AGT GGC TGT GCA G
Southern digoxigenin probe (189 position: **A**)	TAC AGG CGA **A**CA TAC TTA C
Southern digoxigenin probe (189 position: **G**)	TAC AGG CGA **G**CA TAC TTA C
Southern PCR primer: L111 strand	ACC CTA TGT CGC AGT ATC TGT C
Southern PCR primer: H253 strand	GTG GAA AGT GGC TGT GCA G

TABLE 2-*Polymorphisms characterization in hybridization region of PNA probes.*

Cambrige Sequence Reference (20)	n	182 C	184 G	185 G	188 A	189 A	195 T
	7	.	.	.	.	.	C
	1	.	.	.	.	G	C
muscle	1	T	.	.	.	.	.
samples	2	.	.	A	G	.	.
	1	.	.	A	.	.	.
	1	T	.	C	.	.	C
	6	.	.	.	.	G	.
buccal	1	.	.	.	.	.	C
samples	2	T	A	.	.	.	.
	6	.	.	A	G	.	.

TABLE 3-*Characterization of the position 189 between three technologies on muscle tissues and buccal cells.*

		Automated sequencing		**PNA / qPCR method**		**Southern blot**	
Samples	n	n	A189G heteroplasmy	n	A189G heteroplasmy	n	A189G heteroplasmy
Muscle tissues	69	69	11	50	44	50	27
Buccal cells	37	37	0	23	18	23	0

The PNA/qPCR method allowed detection of low levels of the A189G heteroplasmy in two cell types. The automated DNA sequencing did not detect the heteroplasmy level inferior to 30%. In buccal cells, heteroplasmy were globally inferior to 10% and were not detected by automated sequencing or by Southern blot. In muscle samples, Southern blot detected heteroplasmy level superior to 10% and PNA/qPCR method could determine all levels of heteroplasmy.

Figures legends

FIG.1-*The PNA directed PCR clamping methodology. Arrows show PCR primers. Bars represent PNA. Asterisks represent mutations. Under appropriate cycling conditions, the PNA will bind to the wild-type, but not to the mutant template, blocking annealing of the upstream primer, and allowing amplification of only the mutant molecules.*

FIG. 2-*(♦): G mutant Ct values with PNA probe. (■):WT and M controls Ct values without PNA probe.*

a: Amplification depending on the percentage of mutants (189G variant) with PNA G probe. Estimation of the percentage of mutants given by the Ct difference between the amplifications with or without PNA G probe (ΔCt G) was noticeable in values < to 80%, or > to 80% of G variant.

b: Amplification depending on the percentage of mutants (189G variant) with PNA A probe. Estimation of the percentage of mutants given by the Ct difference between the amplifications with or without PNA A probe (ΔCt A) was noticeable in values < to 20%, or > to 20% of G variant.

c: The Ct value ratio of PNA G probes on PNA A probes depending on the percentage of mutants (189G variant.) The Ct value ratio of PNA G probe on PNA A probe led to a better resolution of the G variant level in the 20 to 80% interval.

FIG. 3- *Percentage of 189G mutants estimated by the G/A Ct values ratio. X= G/A Ct values ratio; Y= 189 G mutant percentage. From Fig. 2c and the analysis by Statistica 5.1 software, the inverse polynomial function was established ($y = -2737.2 + 11918x – 19242.5x^2 + 13574.1x^3 – 3486.6x^4$). It calculated the level of mutants present in a sample by it's G/A Ct ratio.*

FIG. 4-*a: Accumulation in buccal cells of the A189G heteroplasmy level with age. (n = 23). The level of 189G mutants was determined by G/A Ct ratio within the PNA/ Real Time PCR run. None of the G/A Ct ratios measured surpassed 0.92, thus giving a maximum percentage estimation of G variants of 12.6%.*

b: A189G heteroplasmy in buccal cells of individuals belonging to the same maternal lineage (from three different families). Mutation 189G in levels inferior to 12.6% was detected in subjects of the same family by PNA/qPCR. Mutation was absent or weakly present in the younger individuals compared to the older ones of the same family.

FIG. 5-*Accumulation in muscle samples of the A189G heteroplasmy with age. The level of 189G mutants was determined by G/A Ct ratio within the PNA/ Real Time PCR run. After 50 years of age, the levels are generally ≥ to 10% (15/19); for those of 60 years old or more, the levels are closer to 20% and over (10/12). The individuals 40 years old or less, except a few individuals (5/23) have levels of mutations ≤ to 10%.*

FIG. 6-*Southern blot. WT: Wild-type control; M: Mutant control; the samples were classified by age; m : muscle tissue; b : buccal cells.*

a: Detection of the heteroplasmies A and G in position 189 in muscle samples of individuals of more than 60 years of age. The M probe (189G) showed clear evidence of 189G variants on muscle tissues of individuals of more than 60 years of age. The percentage of mutations is of 20% or more in all individuals except for one of 77 years old whose level was low (≤ to 5%).

b: Detection of the heteroplasmies A and G in position 189 in muscle samples of individuals of less than 60 years of age. The M probe (189G) showed levels of 189G mutants ≥ to 10 %. These levels were found in most of the individuals of

more than 50 years of age (5/7). For those of 40 years old or less, the percentage of mutation was ≤ to 10% (18/23) and were barely detected or undetected by Southern blot.

c : Detection of the heteroplasmies A and G in position 189 in double sampling (muscle tissue and buccal cells) from two individuals.

The M probe (189G) revealed the tissue variability of the 189G heteroplasmy. In the muscle tissue, evaluation of the percentage of mutants by ImageQuant software was of 12% for the 63 year old individual and of 20% for the 70 year old. Detection of the 189G heteroplasmy in buccal cells was very weak and relative quantification estimated the percentage of G mutants to be ≤ to 5%.