Forensic Science International: Genetics 5 (2011) 105–108

Contents lists available at ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsig

Rapid screening for Native American mitochondrial and Y-chromosome haplogroups detection in routine DNA analysis

Gala Zuccarelli [1], Evguenia Alechine [1], Mariela Caputo, Cecilia Bobillo, Daniel Corach, Andrea Sala *

Servicio de Huellas Digitales Genéticas y Cátedra de Genética y Biología Molecular, School of Pharmacy and Biochemistry, Buenos Aires University, Argentina[2]

ARTICLE INFO

Keywords:
Multiplex Real Time PCR
High Resolution Melting
mtDNA
Y-SNPs
Native American haplogroups
Argentinean population

ABSTRACT

Aiming to detect individuals of Native American maternal or paternal ancestry a rapid screening approach has been developed. Its strategy was based on SNP typing by Real Time PCR (rt-PCR) followed by High Resolution Melting analysis (HRM). After extraction, DNA was quantitated by rt-PCR using commercial kits; samples were then submitted to two multiplex reactions in order to determine the major Native American mtDNA and Y-chromosome haplogroups by HRM. One cocktail included primers flanking nucleotide substitutions that define mtDNA haplogroup C and sub-haplogroups A2, B2, and D1. The other included primers flanking Y-SNPs M3, M269 and U179 that allowed discriminating Q and non-Q haplogroups. In all cases amplicons were <125 nucleotides long in order to increase the peak resolution. The accuracy of the results obtained was established by means of sequencing analysis of the amplicons.

The new working-flow here proposed facilitates and speeds-up the screening process that may preclude a detailed sequencing analysis of particular samples, or for further molecular epidemiological investigations in which continental origin influences might be relevant.

1. Introduction

The population of Argentina is the result of intercontinental admixture events. It has been recently demonstrated that the composition of the Argentinean extant population exhibit a differential Native American matri and patrilineages contribution. The four Native American mitochondrial haplogroups A2, B2, C1 and D1 are represented in average 54% of the general population, whereas Native American Y-chromosome specific haplogroup (M3-Q3) is represented in 4.9% of the population [1]. Aiming to refine the Argentinean population admixture levels and improve the sampling efficiency of diverse Native American matri and patrilineages, a rapid screening approach has been developed based on specific Single Nucleotide Polymorphisms (SNP) typing located within the coding region of the mtDNA genome and in the male-specific region of the Y-chromosome (MSY) [2], by means of Multiplex Real Time PCR [3] followed by High Resolution Melting (HRM) analysis [4].

SNPs located within haploid genetic systems, such as mitochondrial DNA (mtDNA) and MSY, are useful markers for studying migrations, allowing determination of phylogenetic divergence between lineages without the ambiguities caused by meiotic shuffling [5,6]. Based on binary markers located in the Y-chromosome more than 300 Y chromosomal haplogroups have been defined in human populations [7]. However, in our development the focus was oriented to the SNP associated with the most frequent ancestral Native American patrilineage: M3, for identifying and collecting samples with Native American Y-chromosomes for further analysis.

Mitochondrial haplogroups have firstly been defined in human population genetic studies by restriction fragment length polymorphism analysis [8,9]. At present, sequencing of the entire mtDNA molecule is the "gold standard" for haplogroup determination [10]. Alternatively, other novel strategies have also been developed, such as SNaPshot minisequencing and microarrays analysis [11,12].

The strategy here described allows discerning the major Native American mitochondrial and Y-chromosome haplogroups to which the donor of a sample might belong in a highly accurate, fast and cost effective way.

2. Materials and methods

2.1. DNA samples

All analyzed samples were collected from unrelated donors who signed informed consent statements. The study was approved by the Ethical Committee of the School of Pharmacy and Biochemistry, University of Buenos Aires.

* Corresponding author.
E-mail address: asala@ffyb.uba.ar (A. Sala).
[1] Both authors contributed equally to the work.
[2] shdg@ffyb.uba.ar.

1872-4973/$ – see front matter © 2010 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.fsigen.2010.08.018

2.2. Control samples

Samples used as controls, both for Y-SNPs and mtDNA-SNPs typing, were confirmed by sequencing analysis of Y-chromosome loci rs3894 (M3), rs9786153 (M269) and rs2319818 (U179), and sequencing amplicons containing mtDNA-SNPs 8027, 12007, 3547, 14318, 2092 and the complete mtDNA D-loop control region, respectively. Previously analyzed DNA samples [13] were used for mtDNA typing validation. One hundred samples were tested with both multiplex developed systems.

2.3. DNA extraction and quantization

DNA was extracted by means of a semi-automated platform (Maxwell® Promega, USA) from saliva or blood spotted onto FTA cards using the manufacturer's reagents and following their protocols. Quantitation and sex assignment of the extracted DNA was performed by means of Real Time PCR using a commercial kit (Plexor® HY, Promega, USA) and a Rotor-Gene 6000 (Corbett, Australia) Real Time PCR equipment.

2.4. Primers design

Primers were designed using Primer3 software (http://frodo.wi.mit.edu/primer3). The uniqueness of each SNP was checked within the fragments considered. In order to obtain broader divergence between melting profiles and increase peak resolution, amplicon's length was less than 125 pb in all cases. Self and heterodimers, potential hairpins and secondary structures were checked using OligoAnalyzer software (http://www.idtdna.com/annalyzer/applications/oligoanalyzer). The primer specificity was checked using Genome Browser online software (http://genome.ucsc.edu/cgi-bin/hgGateway).

Primer details for both mtDNA and Y-chromosome SNPs are summarized in Table 1. Since the polymorphisms that characterize C1 mtDNA subhaplogroup are located exclusively within the mitochondrial D-loop, we selected a coding region SNP representing haplogroup C to avoid working with mtDNA hypervariable regions.

2.5. Multiplex PCR conditions

Reactions were carried out in 25 μl volume and contained Syto 9 (Invitrogen, USA) 25 μM, Taq Polymerase (Highway, Argentina) (5000 U/ml), Buffer 5× ($MgCl_2$: 1.5 mM), dNTPs (25 mM), primers (as indicated in Table 1) and 5 or 10 ng of DNA template for Y-SNPs typing or mtDNA-SNPs typing, respectively.

2.6. Cycling conditions

mtDNA-SNPs: 95 °C, 2 min; then 30 cycles of 95 °C, 20 s; 56 °C, 30 s; 72 °C, 40 s; finally 72 °C, 2 min. HRM analysis between 72 °C and 85 °C.

Y-SNPs: 94 °C, 1 min; then 33 cycles of 94 °C, 5 s; 55 °C, 10 s; finally 72 °C, 2 min. HRM analysis between 70 °C and 80 °C.

3. Results and discussion

Native American mitochondrial haplogroups were assessed by typing five SNPs located within the coding region. These polymorphisms were amplified in a multiplex reaction and simultaneously analyzed for discerning the main Native American mitochondrial sub-haplogroups A2, B2, D1 and the haplogroup C. Fig. 1 depicts the derivative melting curves of the multiplex system, representing the five simultaneously analyzed SNPs. The 2 °C separation between the melting peaks for each mtDNA-SNP allowed adequate allele discrimination between Native American and non-Native Americans haplogroups.

The validation of this system was performed by means of the analysis of 19 samples belonging to 19 different non-Native American haplogroups (HV0, R, H, H5, I2, J2a, K1a, K1c, L0a2, L1c1, L3e1, N1b, T2, U4, X2, J1c, V, W and R0a), defined by complete D-loop sequencing and characterization of specific coding region SNPs [13].

For Y-chromosome haplogroup assessment, the more prevalent haplogroups found in the Argentinean population were selected, as previously described [1]. In that sense, European R1b1b2 and I haplogroups, and the Native American M3-Q1a3a haplogroup were selected. In Fig. 2 are depicted the melting profiles for samples of ancestral and derived conditions for the three haplogroups considered. As expected, it was possible to identify Y chromosome haplogroups in approximately 60% of the samples. This is due to the presence of other haplogroups (E, J and R1a among others) in the Argentinean population that are not being analyzed in this work.

The robustness of both developed multiplexes was confirmed by typing 100 blind samples, whose results were confirmed by subsequent sequencing of each polymorphism. Mitochondrial SNP typing showed that 56% of the samples belonged to Native American matrilineages in agreement with previously published results [13]. Y-SNP analyses also agreed with previously reported data obtained by SNaPshot analysis [1]; our results showed a 40% R1b1b2, 8% I and 7% M3-Q1a3a haplogroups in the sample set tested with the approach described herein.

Although melting curves are affected by ionic strength, magnesium and intercalary dye concentration we did not detect a significant variation in the results after modifying the Mg^{2+} concentration or the intercalary dye used. Instead, DNA template concentration not only affected the melting but also the amplification curves. We achieved homogeneous amplification for every sample and optimal resolution between melting peaks by adding precisely quantitated DNA template to each reaction tube.

In conclusion, Real Time PCR followed by HRM showed to be an efficient and highly reproducible approach for haplogroup identification. We recognize that this approach has two apparent restrictions, on one hand the reduced number of SNPs included in

Table 1
Characteristics of the selected mtDNA and Y-DNA SNPs.

Haplogroup	Marker	SNP	Primer forward sequence 5′ → 3′	Primer reverse sequence 5′ → 3′	Amplicon length (bp)	Primer concentration (pmol)
mtDNA						
A2	8027	G/A	CCCCCATTATTCCTAGAACC	ACAGCTCATGAGTGCAAGAC	125	15 pmol
A2	12007	G/A	CTAGTCACAGCCCTATACTC	CATGAGGGTGTTTTCTCGTG	111	5 pmol
B2	3547	A/G	CTACCATCACCCTCTACATCAC	TAGGTTGAGGTTGACCAG	104	7.5 pmol
C	14318	T/C	AATCAACCCTGACCCCTCT	GGGTGGTGGTTGTGGTAAA	74	10 pmol
D1	2092	C/T	TGCCCACAGAACCCTCTAAA	GGCCTACTATGGGTGTTA	122	25 pmol
Y-DNA						
Q1a3a	M3	C/T	CCTGACAATGGGTCACCTCT	TTCATTTTAGGTACCAGCTCTTCC	58	20 pmol
R1b1b2	M269	T/C	AAGGGGAATGATCAGGGTTT	CCTTCTGAGGCACATATGATAA	76	Fwd 15/Rv 30 pmol
I	U179	G/A	GGCTGTAATTCATAGCCTACCA	GTTGCCAGCTCCTCTTTTCA	93	10 pmol