International Congress Series 1288 (2006) 4–6

www.ics-elsevier.com

Genotyping coding region mtDNA SNPs for Asian and Native American haplogroup assignation

V. Álvarez-Iglesias, A. Salas *, M. Cerezo, E. Ramos-Luis, J.C. Jaime, M.V. Lareu, A. Carracedo

Unidade de Xenética, Instituto de Medicina Legal, Facultad de Medicina, Universidad de Santiago de Compostela, 15782, Galicia, Spain

Abstract. Based on phylogenetic criteria, we selected 32 mitochondrial DNA (mtDNA) coding region SNPs that allows to distinguish Asian and Native American mtDNA haplogroups. SNP genotyping is carried out in a single multiplex reaction that involves an 18-amplicon PCR amplification, followed by a single minisequencing reaction using SNaPshot. The polymorphisms selected increase the discrimination power of the mtDNA hypervariable regions (HVS-I/II) in populations. Consequently, these combined SNPs are of particular interest in forensic casework, population genetics and clinical research. Here we show preliminary results using a sample from south-east Asia (Taiwan) and Native Americans from Argentina. © 2005 Published by Elsevier B.V.

Keywords: SNPs; mtDNA; Haplotypes; Phylogeography; Phylogeny; SNaPshot; Multiplex genotyping

1. Introduction

It is well known that the control region of the mtDNA genome, in particular, the first and second hypervariable regions (HVS-I/II), have a limited power of discrimination in a forensic and anthropological context. This depends mainly on the population under study since diversity values and phylogeny differs significantly within populations (with East African populations showing the highest levels of diversity). Thus, many haplogroups are poorly defined in both hypervariable regions and in consequence, sharing the same HVS-I profile does not necessarily imply that two mtDNAs belong to the same haplogroup. Sequencing the complete genomes is tedious, costly, and time-consuming and, therefore, unfeasible in forensic casework. Analysis of the single nucleotide polymorphisms (SNPs)

* Corresponding author. Tel.: +34 981 582327; fax: +34 981 580336.
E-mail address: apimlase@usc.es (A. Salas).

0531-5131/ © 2005 Published by Elsevier B.V.
doi:10.1016/j.ics.2005.11.036

contained in the coding regions of the mitochondrial genome has proved to be a good strategy to address this drawback. For this reason, in the last few years, interest in these polymorphisms has increased, and at the same time, numerous new methodologies have been developed to facilitate their analysis.

2. Material and methods

We have collected samples from Taiwan ($N=39$) and Native Americans from Argentina ($N=68$). All these samples have been sequenced for the complete control region (16024-578). In order to optimize the amount of information covered by a specific number of SNPs, we need to consider the whole body of the known mtDNA phylogeny. Thus, we have used the information contained in more than 150 complete mtDNA genomes from Asia and America: basically the data from [1–3]. This phylogeny has been corroborated with the data (more than 672 complete genomes) recently published by [4]. In addition, we have used control region information from more than 10,000 sequencing profiles. This has allowed us to select those SNPs that, in combination with control region data (mainly the standard HVS-I and HVS-II regions), cover the main haplogroups and most of the known Asian and American sub-branches. The primers are designed to have an annealing temperature of around 60° using Primer 3 (http://frodo.wi.mit.edu/cgi-bin/primer3). The sequence databases at the NCBI (http://www.ncbi.nlm.nih.gov) were interrogated using the online BLAST tool to test the primers against possible repetitive sequences and sequence homologies. AutoDimer (http://www.cstl.nist.gov/biotech/strbase/AutoDimerHomepage/AutoDimerProgramHomepage.htm) was used to test for potential hairpin structures and primer–dimer problems. The SNPs are amplified in 18 amplicons with sizes ranging from 57 and 169 bps. The amplicons are deliberately designed to be small in order to ease the analysis of samples that are highly degraded or with low quantity of DNA. More details on M&M and pictures of SNaPshot profiles are provided under request.

3. Results and discussion

We have obtained SNaPshot and control region sequencing profiles for a total of 107 samples. We have detected a total of 27 different haplogroups in Taiwanese samples (using a refined East Asian phylogeny) and eight in Native American samples.

Our previous experience with SNaPshot designs [5], and common routine work carried out in our lab, have showed us that potential secondary homologies at the level of nuclear DNA (nDNA) is likely to be of null impact in forensic and anthropological studies. Further validation has been carried out by others [6]. There are various reasons that would prevent SNaPshot primer designs from artefactual genotyping due to secondary homologies. First, the relative amount of mtDNA with respect to the nDNA will avoid any attempt of primer annealing in other DNA than the mitochondrial molecule; this seems to be also true for degraded DNA. Second, we have observed [5] a perfect (phylogenetic) match between the expected SNaPshot profile and the HVS-I/II sequencing results. Third, the use of the universally accepted BLAST tool allows to search for a significant prevalence of secondary homologies with the nDNA. In addition, any phylogenetically unexpected pattern could be easily confirmed using singleplex SNaPshot typing, or alternative techniques (e.g., sequencing), such that this would allow in occasions to improve our knowledge on the mtDNA phylogeny in a straightforward manner. Note that other apparently well-known common techniques for mtDNA genotyping, such as the automatic sequencing, are far to be exempt of problems (e.g., [7–11]).

SNaPShot allows to genotype multilocus mtDNA SNPs with low cost and high efficiency. More important in the forensic field: we have recently observed that the sensitivity of the technique is higher than sequencing methodology [12]. Such as it should be done when interpreting mtDNA sequencing profiles, we strongly recommend to analyze SNaPshot profiles (and those obtained using alternative techniques) under the phylogenetic perspective. Any phylogenetic inconsistency should be double check in order to control for genotyping artefacts. Then, sequencing should follow.

References

[1] Bandelt, et al., Ann. Hum. Genet. 67 (2003) 512–524.
[2] Q.-P. Kong, et al., Am. J. Hum. Genet. 73 (2003) 671–676.
[3] Kivisild et al. (2002) Mol Biol Evol 19:1737-51; 20 (2003):162 (erratum).
[4] Tanaka, et al., Genome Res. 14 (2004) 1832–1850.
[5] B. Quintáns, et al., Forensic Sci. Int. 10 (2003) 251–257.
[6] A. Goios, A. Amorim, L. Pereira; communication of the International Society of Forensic Genetics (ISFG); Azores, 2005.
[7] Y.G. Yao, C.M. Bravi, H.-J. Bandelt, Forensic Sci. Int. 141 (2004) 1–6.
[8] H.-J. Bandelt, A. Salas, S. Lutz-Bonengel, Int. J. Leg. Med. 118 (2004) 267–273.
[9] H.-J. Bandelt, A. Salas, C. Bravi, Science 305 (2004) 1402–1404.
[10] A. Salas, et al., Biochem. Biophys. Res. Commun. (in press).
[11] A. Salas, et al., PloS Med. (in press).
[12] Crespillo, et al., Forensic Sci. Int. (in press).