Forensic Science International: Genetics 4 (2010) 73–79

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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

A forensic method for the simultaneous analysis of biallelic markers identifying Y chromosome haplogroups inferred as having originated in Asia and the Japanese archipelago

Natsuko Mizuno [a,*], Tetsushi Kitayama [a], Koji Fujii [a], Hiroaki Nakahara [a], Kanako Yoshida [a], Kazumasa Sekiguchi [a], Naoto Yonezawa [b], Minoru Nakano [b], Kentaro Kasai [a]

[a] National Research Institute of Police Science, 6-3-1 Kashiwanoha, Kashiwa, 277-0882, Japan
[b] Graduate School of Science, Chiba University, Yayoi 1-33, Inage, Chiba, 263-8522, Japan

ARTICLE INFO

Article history:
Received 29 September 2008
Received in revised form 24 April 2009
Accepted 2 June 2009

Keywords:
Y chromosome
Biallelic markers
Haplogroup
Single nucleotide polymorphism
Short insertion and deletion polymorphisms

ABSTRACT

Information regarding the ancestral and geographical origins of biological evidence samples may be useful for crime investigators as they narrow down the possible donors of the sample. A method for simultaneous analysis of seven biallelic markers (M130, M131, M57, M125, M175, M122 and M134) was developed for forensic application. M57, M125 and M131 are included to identify haplogroups inferred as having originated in the Japanese archipelago. Our method employs allele-specific PCR and fragment analysis using fluorescently labeled primers and capillary electrophoresis. This method can be used to assign a haplogroup from both of degraded male DNA samples and DNA samples containing a mixture of female and male DNA by designing PCR primers that generate small amplicons and are highly specific for targets on the Y chromosome. A total of 1346 samples from Japanese males collected from the four major islands and Okinawa island were classified into seven Y binary haplogroups i.e., C-M130, C-M131, D-M57, D-M125, O-M175, O-M122 and O-M134, and a "no-mutation detected" group and their frequencies were 0.0617, 0.0565, 0.1441, 0.182, 0.3418, 0.11, 0.0847 and 0.0193, respectively. Samples of "no-mutation detected" were further analyzed by direct sequencing for identification of the major haplogroup to which they belong. Along with the haplogroup data, we report haplotype data for the 16 Y-STR markers included in the AmpFlSTR® Yfiler™ PCR amplification kit (Applied Biosystems). These data will be useful in the prediction of haplogroups based on Y-STR haplotypes.

1. Introduction

The Y chromosome does not undergo recombination with the X chromosome except in the pseudoautosomal region [1]. Therefore, the male-specific region of the Y chromosome (MSY) is transferred from father to son for generations without alteration, except in the rare event of mutation. Y-STR analysis is useful, especially when the DNA profile of the male contributor must be obtained from a sample containing a mixture of male and excess female DNA, from which analysis of autosomal STR results in a masking of the male DNA profile by the excess amount of female DNA. Y-STR analysis is also useful for identification of deceased males using samples derived from a paternally related male as a reference when both a reference for the putative deceased and a sample for performing parentage analysis are unavailable—taking advantage of the fact that paternally related males share the same haplotype.

While STR markers are usually multiallelic, single nucleotide polymorphism (SNP) and insertion/deletion (indel) markers are likely to have two allelic classes. Those biallelic markers have a much lower mutation rate than STR markers [2]. Therefore, the derived allele of each biallelic marker is generally assumed to arise once in human history and all male individuals who possess a particular derived allele are descendants of a common ancestor in whom the mutation first occurred. The recently published human Y chromosome tree [3] contains 20 major clades (A through T) and 311 distinct haplogroups that incorporate approximately 600 biallelic markers. Different haplogroups may show distinctly different geographic distributions [4]. The geographic origins of some haplogroups have been inferred [5–10] based on the geographic distribution and haplogroup age estimated from Y-STR variations within the haplogroup [11]. Information regarding the ancestral and geographical origin of a biological sample will be

* Corresponding author at: Fourth Biology Section, Department of First Forensic Science, National Research Institute of Police Science, 6-3-1, Kashiwanoha, Kashiwa, Chiba, 277-0882, Japan. Tel.: +81 471 35 8001; fax: +81 471 33 9159.
E-mail address: mizuno@nrips.go.jp (N. Mizuno).

1872-4973/$ – see front matter © 2009 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.fsigen.2009.06.001

potentially useful to crime investigators to narrow down the possible donors of the sample [12–18]. As the number of foreign visitors as crime suspects and victims increases due to globalization, when an individual whose DNA profile matches that of the biological evidence sample cannot be found within a certain period of time, investigators must consider the possibility that the sample was derived from a foreign visitor. In these cases, the knowledge that the Y chromosome contained in the sample belongs to a haplogroup inferred as having originated in the Japanese archipelago may be useful. This method may also be useful in crimes and mass fatality incidents that take place not only in Japan, but also outside the country. The Y chromosomes of most Japanese males belong to either the major clade C, D or O [5], which are all inferred as having an Asian origin. In these clades, there are haplogroups that are inferred as having originated in the Japanese archipelago.

The aim of this study was to develop a method for simultaneous analysis of multiple biallelic markers on the MSY that mark haplogroups in the C, D and O clades, especially haplogroups that are inferred as having originated in the Japanese archipelago. A secondary aim of this study was to develop a method for analyzing the types of samples that are encountered in forensic casework, namely degraded samples and DNA containing a mixture of male and female DNA. Further, 1346 Japanese Y chromosomes were analyzed to validate the method and to construct Y haplogroup data combined with Y-STR data [19] useful for haplogroup prediction based on a Y-STR haplotype.

2. Materials and methods

2.1. DNA samples

Blood or buccal swab samples were collected for a population study from 1346 Japanese male and 170 female individuals after receipt of informed consent. Among them, 1079 males and 155 females were the same individuals who participated in a previous study of Y-STR analysis [19]. Samples were collected from all the four major islands of Japan i.e., Hokkaido, Honshu, Shikoku and Kyushu, as well as from the main island, Okinawa. The 1346 male samples were comprised of 50, 932, 103,113 and 148 from Hokkaido, Honshu, Shikoku, Kyushu and Okinawa respectively. DNA was isolated from the blood samples using either MagNA Pure LC (Roche Diagnostics, Mannheim, Germany) or BioRobot® EZ1 (QIAGEN, Hilden, Germany), both of which use magnetic silica beads. DNA was isolated from buccal swabs and from three aged bloodstain samples stored at ambient temperature for 23 years using BioRobot® EZ1.

2.2. DNA quantification

DNA was quantified using a DU 640 spectrophotometer (Beckman Coulter, Fullerton, CA), except for the DNA isolated from aged bloodstains, which was quantified using real time PCR with an intercalator dye and primers targeting the 98 bp in D17Z1 (5′-ACATTCCTTTGGATGGAGCA and 5′-TTTATCCCGTTTCCAACGAA). PCR reactions were performed in a total volume of 25 μL containing 2 μL of the DNA solution to be quantified, 1 μM of each primer, and 1 × SYBR® Premix Ex Taq (Takara, Otsu, Japan). Thermal cycling was performed using a Smart Cycler II (Cepheid, Sunnyvale, CA) as follows: an initial denaturation step of 10 s at 95 °C; 35 cycles consisting of 5 s at 95 °C; and, 30 s at 60 °C. A standard curve was generated using Cepheid Smart Cycler software version 2.0d using K562 DNA (Promega, Madison, WI) diluted to 0.001, 0.01, 0.1, 1, 10 ng/μL.

2.3. Y chromosome STR analysis

The 16 markers of DYS456, DYS389I, DYS390, DYS389II, DYS458, DYS19, DYS385, including the DYS385a and DYS385b loci, DYS393, DYS391, DYS439, DYS635, DYS392, Y GATA H4, DYS437, DYS438 and DYS448 were co-amplified using an AmpFlSTR® Yfiler™ PCR amplification kit (Applied Biosystems, Foster, CA), according to the manufacturer's recommendations. The data were then analyzed as described previously [19].

2.4. Haplogroup nomenclature

Haplogroups were named using mutation-based nomenclature [20], which retains the major haplogroup information followed by the name of the terminal mutation for which the sample is positive (i.e., C-M130, C-M131, D-M57, D-M125, O-M175, O-M122, O-M134).

2.5. Analysis of Y chromosome binary polymorphisms

Seven primers sets were designed for amplification of fragments containing polymorphic sites of M130, M131, M57, M125, M175, M122 and M134 (Table 1). M131, M57 and M125 were the mutations that are inferred as having originated in the Japanese archipelago [5]. Fragment analysis was employed for genotyping M131, M57 and M134 indel markers. Allele-specific PCR was employed for genotyping M130, M125 and M122 SNP markers. The PCR product sizes shown in this manuscript include the 3′ adenine added by the template-independent activity of DNA Taq polymerase [21]. Multiplex PCR reactions were performed in a total volume of 10 μL containing approximately 1 ng of genomic DNA unless otherwise noted, all primers in Table 1, 1 U AmpliTaq Gold® DNA polymerase (Applied Biosystems) and 1× GoldST*R Buffer (Promega, Madison, WI). The concentrations of the primers shown in Table 1 were optimized so as to obtain similar PCR yields for all of the primer pairs. Thermal cycling was performed using a GeneAmp® PCR system 9700 (Applied Biosystems): an initial denaturation step of 11 min at 95 °C; 30 cycles consisting of 1 min at 94 °C; 1 min at 61 °C and 1 min at 72 °C; and, a final extension step of 80 min at 60 °C. The dye-labeled amplified products were resolved using a 3130xl Genetic Analyzer (Applied Biosystems) with HIDFragment Analysis36_POP4_1 as the run module. An allelic ladder containing both ancestral and derived alleles of all seven markers was prepared by amplifying a DNA sample that contained both ancestral and derived alleles for all markers (i.e., mixture of three males belonging to C-M131, D-M125, O-M134). A ± 0.4-bp window around the size obtained for each fragment in the allelic ladder was employed for typing with the minimum peak height threshold for detection ≥150 relative fluorescence units (RFU). The GeneScan™ 600 LIZ® Size Standard (Applied Biosystems) was used for sizing DNA fragments. The sample run data were analyzed together with an allelic ladder and positive and negative controls using GeneMapper® ID software version 3.2.1 (Applied Biosystems).

2.6. Male/female mixture

Three male DNA samples belonging to C-M131, D-M125 and O-M134, and one sample of female DNA, were used to prepare a total of six PCR male/female mixtures. Three DNA mixtures were each prepared containing 1 ng male DNA (from one male per mixture) and 500 ng female DNA. The rest were each prepared containing 1 ng male (from one male per mixture) DNA and 2000 ng female DNA.

2.7. Aged bloodstain samples

To assess the potential of our method for analysis of degraded DNA, three aged bloodstain samples stored at ambient tempera-