Forensic Science International: Genetics 7 (2013) 75–81

Contents lists available at SciVerse ScienceDirect

Forensic Science International: Genetics

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

Y-SNP miniplexes for East Asian Y-chromosomal haplogroup determination in degraded DNA

Myung Jin Park [a], Hwan Young Lee [a,b], Na Young Kim [a], Eun Young Lee [a], Woo Ick Yang [a], Kyoung-Jin Shin [a,b,*]

[a] Department of Forensic Medicine and Brain Korea 21 Project for Medical Science, Yonsei University College of Medicine, 50 Yonsei-ro, Seodaemun-gu, Seoul 120-752, South Korea
[b] Human Identification Research Center, Yonsei University, 50 Yonsei-ro, Seodaemun-gu, Seoul 120-752, South Korea

ARTICLE INFO

Article history:
Received 10 November 2011
Received in revised form 14 May 2012
Accepted 26 June 2012

Keywords:
Y chromosome haplogroup
Single nucleotide polymorphism
Degraded DNA
Small size amplicon
East Asians

ABSTRACT

Four multiplex PCR systems followed by single base extension reactions were developed to score 22 single nucleotide polymorphisms (SNPs) and identify the most frequent East Asian Y chromosome haplogroups. Select Y chromosome SNPs allowed hierarchical testing for almost all of the major East Asian haplogroups along the revised Y chromosome tree. The first multiplex consists of six SNPs defining world-wide major haplogroups (M145, RPS4Y_{711}, M89, M9, M214, and M175). The second multiplex includes six SNPs of subhaplogroup O (M119, P31, M95, SRY_{465}, 47z, and M122). The third multiplex contains six SNPs that subdivide the subhaplogroup O3 (M324, P201, M159, M7, M134, and M133). The fourth multiplex comprises four SNPs of subhaplogroup C (M217, M48, M407, and P53.1). The sizes of the PCR amplicons ranged from 70 to 100 bp to facilitate their application to degraded forensic and ancient samples. Validation experiments demonstrated that the multiplexes were optimized for analysis of low template DNA and highly degraded DNA. In a test using DNA samples from 300 Korean males, 16 different Y chromosome haplogroups were identified; haplogroup O2b* was the most frequently observed (29.3%), followed by haplogroups C3 (xC3c, C3d, C3e) (16.0%) and O3a3c1 (11.0%). These multiplex sets will be useful tools for Y-chromosomal haplogroup determination in anthropological and forensic studies of East Asian populations.

1. Introduction

Genetic variations in the non-recombining portion of the Y chromosome (NRY) are analyzed in diverse disciplines including anthropological, forensic and medical genetics [1–5]. Because of a lack of recombination and low mutation rates, Y chromosome single nucleotide polymorphisms (Y-SNPs) are the most useful genetic markers for reconstructing male lineages. Therefore, Y chromosome haplogroups, which are defined by combinations of allelic states at hierarchically arranged Y-SNPs and small indels, have been extensively studied to infer the origins, evolution, and histories of migrations of modern human populations [6,7].

A number of changes have been made to the Y-chromosomal haplogroup tree, and a total of 311 distinct haplogroups have been defined with increased resolution [8]. The haplogroup O was considerably rearranged during this revision; the L1 retroposon insertion (LINE 1) polymorphism, which had conflicted with the N7 polymorphism, was excluded from the list of markers used to define subhaplogroup O3. Persistent commercial or in-house development efforts for Y-SNP typing protocols have been made, but most developed methods involve typing of European Y haplogroups or world-wide major haplogroups with low resolution [9–12] and are not suitable to subdivide subhaplogroup O3, a major haplogroup in East Asians [13,14], into internal derivatives following the revised Y haplogroup tree. In addition, as the demand for inferring geographic origin is increasing in forensic DNA analysis as well as in ancient DNA analysis (e.g., identification of Korean War and Vietnam War victims, and genetic characterization of ancient remains), the development of sensitive and efficient methods for the Y-chromosomal haplogroup determination in degraded DNA is required.

Therefore, in the current study, a multiplex single base extension (SBE) method was developed to score Y-SNPs of East Asian haplogroups following a small size amplicon strategy that is suitable for application to degraded DNA. Y-SNPs were selected from SNPs that are hierarchically located along the

* Corresponding author at: Department of Forensic Medicine, Yonsei University College of Medicine, 50 Yonsei-ro, Seodaemun-gu, Seoul 120-752, South Korea. Tel.: +82 2 2228 2481; fax: +82 2 362 0860.
E-mail address: kjshin@yuhs.ac (K.-J. Shin).

1872-4973/$ – see front matter © 2012 Elsevier Ireland Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.fsigen.2012.06.014

revised topology of Y chromosome haplogroups, focusing on East Asian-specific haplogroups. To assess the utility of our method for the analysis of highly degraded samples, the sensitivity and efficiency of the multiplex set were validated in samples of serially diluted DNA, artificially degraded DNA, and DNA extracted from 55-year-old skeletal remains. Finally, a Korean sample was analyzed using the newly developed multiplex systems, and the distribution of Y haplogroups was studied, since the geographic origins and history of migration of a population can be inferred from the distributions and ages of its haplogroups [13–15].

2. Materials and methods

2.1. DNA samples

Our study protocol was approved by the Institutional Review Board of Severance Hospital, Yonsei University in Seoul, Korea. DNA samples from 300 unrelated Korean males were obtained from the National Biobank of Korea. DNA concentrations were measured using a NanoDrop® ND-1000 Spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA) and the final sample concentrations were adjusted to 1.0 ng/μl. For the sensitivity test, 9948 male DNA (Promega, Madison, WI, USA) was serially diluted to concentrations of 1000, 500, 250, 125, 62, 31, and 15 pg/μl. Artificially degraded DNA was prepared by digesting 1.2 μg of human genomic DNA with 0.02 U of DNase I (NEB, Ipswich, MA, USA) at 37 °C for 40 min. DNA degradation to fragment sizes of around 100 bp was confirmed by ethidium bromide staining after agarose gel electrophoresis (data not shown). Ten DNA samples that were extracted from 55-year-old skeletal remains during a previous study [16] were analyzed to evaluate multiplex SBE reactions. Concentrations of DNA obtained from skeletal remains were determined using the Quantifiler™ Human DNA Quantification Kit and 7500 Real-Time PCR System (Applied Biosystems, Foster City, CA, USA) [16].

2.2. Y-SNP selection and primer design for PCR amplification and SBE

A set of 22 biallelic Y chromosome markers (M7, M9, M48, M89, M95, M119, M122, M133, M134, M145, M159, M175, M214, M217, M324, M407, P31, P53.1, P201, $RPS4Y_{711}$ (M130), SRY_{465} (M176), and 47z) was selected to determine the world-wide major haplogroups, subhaplogroups O, and subhaplogroups C that are present in East Asian populations, including Koreans [13–15,17]. The 22 Y-SNPs and the haplogroup tree defined by these markers are shown in Fig. 1. The nomenclature and topology of the Y chromosome haplogroups followed those of Karafet et al. [8].

Primers for PCR amplification and subsequent SBE were designed using web-based programs Primer 3 (http://frodo.wi.mit.edu/primer3/) and Batchprimer 3 (http://batchprimer3.bioinformatics.ucdavis.edu/index.html), respectively, based on sequences obtained from GenBank (http://www.ncbi.nlm.nih.gov/genbank/) or dbSNP (http://www.ncbi.nlm.nih.gov/snp/). PCR primers were then selected to produce amplicons smaller than 100 bp and to amplify male-specific fragments in tests using 1 ng of each male and female DNA. SBE primers were designed to provide even peak heights and appropriate peak intervals by adjusting T_m values and adding poly-T and/or poly-A tails to the 5′ ends of primers. PCR and SBE primer sequences are shown in Supplementary material Tables S1 and S2, respectively.

2.3. Multiplex PCR amplification and PCR product purification

A total of four multiplex PCR systems were developed to type 22 Y-SNPs (Fig. 1). PCR amplifications were performed in a final volume of 25 μl that contained 1 ng of template DNA, 2.5 μl of Gold ST*R 10× buffer (Promega), 2.0 U of AmpliTaq Gold® DNA polymerase (Applied Biosystems), and appropriate concentrations of primers (Supplementary material Table S1). Regarding multiplex IV, 1.5 U of AmpliTaq Gold DNA Polymerase were used. Thermal cycling was performed on a Veriti 96-Well Thermal Cycler (Applied Biosystems) under the following conditions: 95 °C for 11 min; 33 cycles of 94 °C for 20 s, 60 °C for 1 min and 72 °C for

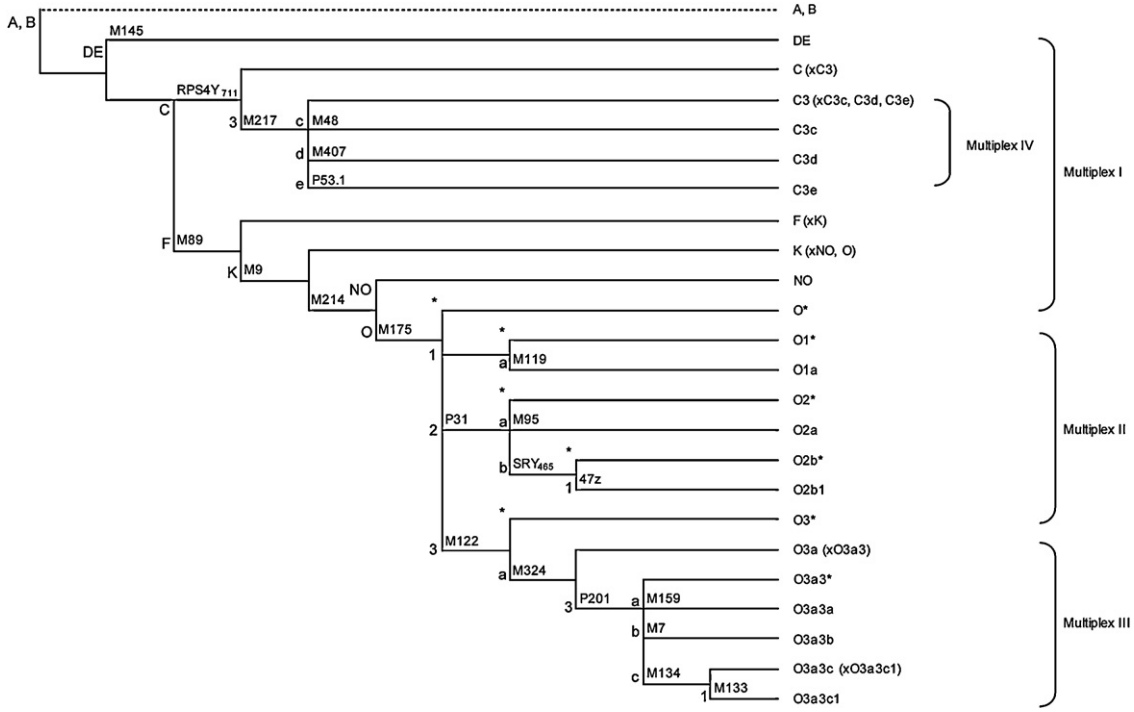

Fig. 1. Phylogenetic tree of the 22 Y-chromosomal binary polymorphisms analyzed in this study. The analyzed Y-SNPs are shown in each branch, and the corresponding haplogroups and multiplexes are shown at the end of each branch according to Karafet et al. [8]. Broken-line branches correspond to haplogroups not found in East Asians.