Forensic Science International: Genetics 4 (2009) 11–20

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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

A comparative analysis of two different sets of Y-chromosome short tandem repeats (Y-STRs) on a common population panel

Julie L. Maybruck [a], Erin Hanson [c,d], Jack Ballantyne [c,d], Bruce Budowle [e,1], Paul A. Fuerst [a,b,*]

[a] Department of Molecular Genetics, The Ohio State University, Columbus, OH 43210, United States
[b] Department of Evolution, Ecology and Organismal Biology, The Ohio State University, Rm. 300 Aronoff Laboratory, 318 West 12th Avenue, Columbus, OH 43210, United States
[c] Graduate Program in Biomolecular Science, University of Central Florida, P.O. Box 162366, Orlando, FL 32816-2366, United States
[d] National Center for Forensic Science, P.O. Box 162367, Orlando, FL 32816-2367, United States
[e] Department of Forensic and Investigative Genetics, University of North Texas Health Science Center, Fort Worth, TX 76107, United States

ARTICLE INFO

Article history:
Received 20 June 2008
Received in revised form 15 March 2009
Accepted 29 March 2009

Keywords:
Forensic science
Short tandem repeats
Microsatellites
Y-chromosome
Y-STR

ABSTRACT

A comparative analysis of two Y-STR loci sets was conducted on a population sample of 224 individuals, 114 Caucasians and 110 African Americans. One set of loci, designated the OSU 10-locus set, comprises variable, single copy, male-specific loci that are dispersed across the Y-chromosome. Parallel evaluations were performed using the 10 Y-chromosome loci most frequently used for forensic analysis, the loci chosen as the SWGDAM Y-STR loci. The OSU 10-locus set had a greater average number of alleles per locus and higher average gene diversity than the SWGDAM loci. The OSU 10-locus set found 220 unique haplotypes in 224 individuals. In ~6000 pairwise haplotype comparisons for each population with each set of loci, the OSU 10-locus set also yielded a greater average number of allelic differences per pair than the SWGDAM loci. Finally, the overall linkage disequilibrium levels were lower for the OSU 10-locus set in the Caucasian population. In general, the OSU 10-locus set revealed a higher power of discrimination than the SWGDAM set.

1. Introduction

Gender specific markers, such as Y-chromosome short tandem repeats (Y-STRs), are valuable tools in DNA forensic analyses because men commit the majority of violent crimes [1]. Y-STR loci are utilized for many forensic applications which include the preferential amplification of the male DNA contribution in mixed samples and the determination of the number of male individuals in multiple male contributor cases. Y-STRs are also useful in kinship cases to identify patrilineage, particularly in situations involving a deceased putative father. In evolutionary studies, Y-STRs have aided in the identification of paternal migration patterns, contrasting with maternal migration patterns identified by mitochondrial DNA [2,3].

Until recently, approximately 50 Y-STR loci had been identified and/or characterized in the literature [4–14]. In 2004, 166 new loci were identified and examined in eight individuals, each from a different SNP-based haplogroup [15]. In terms of physical location within the Y-chromosome, most loci are limited to two small regions fairly close to the Y-chromosome centromere. Even though there is a lack of normal meiotic recombination on the Y-chromosome there is the potential for non-random association. In this paper we will refer to non-random association between loci as "linkage disequilibrium." Since recombination is expected to be absent for loci on the male-specific region of the Y-chromosome, decay of linkage disequilibrium will occur only by the accumulation of new mutations on pre-existing haplotypes. Higher mutation rates would result in lower linkage disequilibrium. Loci located within the same small region of the Y-chromosome may have similar mutation rates. If loci in the same region were to have low mutation rates there is the potential for high linkage disequilibrium. Choice of loci randomly dispersed would reduce the possibility of correlated mutation rates. Additionally, potential intrachromosomal recombination and conversion within the Y-chromosome as described by several researchers compounds the problem of potential non-random association between closely located loci [16,17]. In addition to the distribution within the Y-chromosome, some Y-STR loci have additional limitations; these include limited discrimination power, and multi-allelic profiles. Information from the human genome project and from the literature indicates that a number of the current loci are duplicated

* Corresponding author at: Department of Evolution, Ecology and Organismal Biology, The Ohio State University, Rm. 300, Aronoff Laboratory, 318 West 12th Avenue, Columbus, OH 43210, United States. Tel.: +1 614 292 6403; fax: +1 614 292 2030.
E-mail addresses: maybruck.2@osu.edu (J.L. Maybruck), fuerst.1@osu.edu (P.A. Fuerst).
[1] Institute of Investigative Genetics, University of North Texas Health Science Center, Fort Worth, TX 76107.

1872-4973/$ – see front matter © 2009 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.fsigen.2009.03.004

elsewhere on the X- or Y-chromosome. For some loci duplicate copies and extra alleles routinely appear [8,18–27]. While these loci may be variable, for forensics they present more complicated interpretations, particularly for resolving mixture profiles from two or more male contributors. For other loci, detection of a multiple allele profile is less problematic particularly when detected only occasionally [8,17–19,24–26,28]. Thus, selection of additional forensically useful Y STR loci is based on those loci that tend to yield one allele per individual in the population.

We identified a series of new Y-STR loci that address the aforementioned concerns associated with previous loci (Maybruck et al., submitted for publication). Utilizing resources of the Human Genome Project, we constructed a library of 465 potential Y-STR loci (tri-to-hexanucleotide repeats) located outside of the two regions of concentrated loci. A set of 229 potential loci was evaluated. The remaining 236 loci were not assessed because of close proximity to the 229 loci tested. BLAST searches of the genome revealed that 73% of the 229 loci are duplicated elsewhere in the human genome, mostly on the X- and/or Y-chromosome. Allelic variation and potential multi-allele profiles per locus were assessed at the remaining 62 loci in a test sample of 30 male individuals, 26 of which were Caucasian and African American (16 Caucasians, and 10 African Americans) to further narrow the candidate loci. The final subset of 10 loci, The Ohio State University (OSU) 10-locus set, is comprised of the following loci: DYS471, DYS448, DYS487, DYS488, DYS504, DYS576, DYS685, DYS688, DYS703, and DYS707. During the course of our study several researchers identified some of the same loci [14,15,29]. In order to prevent multiple designations for the same locus, albeit with different primer sequences, the earliest locus designation is utilized as suggested by the DNA Commission of the International Society of Forensic Genetics (ISFG) [30]. The loci were further screened in several females to test for potential amplification from non-Y sources. Examination of the OSU 10-locus set revealed 26 unique haplotypes in 26 individuals, suggesting a potential high haplotype diversity. One of the loci, DYS448, has been included in the AmpFℓSTR® Yfiler® PCR amplification kit (Applied Biosystems, Foster City, California) [31].

Since the OSU 10-locus set was highly informative in our test sample, two larger populations were examined and a comparative study was conducted to evaluate the power of these combined loci relative to the 10 Scientific Working Group on DNA Analysis Methods (SWGDAM) selected loci: DYS19, DYS385, DYS389I and II, DYS390, DYS391, DYS392, DYS393, DYS438 and DYS439 [32].

The intent of this study is to evaluate the discrimination power of the newly identified loci, and whether they have the potential to enhance the discrimination power of Y-STR analysis for forensic applications. The SWGDAM loci are firmly established. In previous studies several common haplotypes were observed [18,19,33,34]. In this study, individuals with identical haplotypes for the SWGDAM loci were differentiated by the OSU set.

2. Materials and methods

2.1. Sample collection

A population of 224 unrelated individuals was screened for the study: 114 Caucasian, and 110 African American. The Biomedical Sciences Institutional Review Board at The Ohio State University approved the use of human subjects and the protocol. Buccal samples were collected from 26 male individuals from Ohio-16 from non-criminal material were made available by the State of Ohio Bureau of Criminal Investigation and Identification (BCI), all of which were stripped of their identifiers; the remaining 10 samples were amassed from unrelated residents of Columbus, Ohio. No personal identifiers were associated with these samples, thus providing anonymity for the donors. Each individual was provided with instructions for buccal cell collection, using sterile swabs. Participants from Columbus, Ohio, collected their own sample under our supervision. All buccal samples were stored at 2–8 °C until extraction.

One hundred and ninety-eight additional samples were typed at the National Center for Forensic Science, Orlando, FL. The human use procedures used were approved by the University of Central Florida's Institutional Review Board. The samples were obtained from the Virginia Division of Forensic Science (bloodstains). All samples were stored at −47 °C until needed.

2.2. DNA extraction

Three different types of DNA extraction procedures were conducted. At The Ohio State University, DNA was obtained from buccal swabs through either the QIamp® DNA Mini Kit Buccal Swab Spin Protocol (QIAGEN Inc., Valencia, California) or the BuccalAmp™ DNA Extraction Kit (Epicentre, Madison, Wisconsin) in accordance with the manufacturers' instructions. QIamp® extracted samples were stored at 2–8 °C, and BuccalAmp™ extracted samples were stored at −20 °C.

At the University of Central Florida, the dried bloodstains were incubated overnight at 56 °C in 400 μl of DNA extraction buffer (100 mM NaCl, 10 mM Tris–HCl, 25 mM EDTA, 0.5% SDS) and 0.1 mg/ml proteinase K. The samples were placed into a spin ease basket and subjected to centrifugation at 14,000 × *g* for 5 min. An equal volume of phenol/chloroform/isoamyl alcohol was added to the crude extract. The aqueous phase of the extracts containing the DNA were purified using Centricon 100™ concentrators (Millipore, Bedford, MA), according to the manufacturer's instructions.

2.3. DNA quantification

At The Ohio State University the quantity of DNA was determined, using the QuantiBlot® DNA Quantification Kit (Applied Biosystems, Foster City, California) in accordance with the manufacturer's protocol. At the University of Central Florida, the quantity of DNA was determined by comparison of ethidium bromide induced fluorescence on a 1% agarose yield gel with a reference set of DNA standards of known concentration.

2.4. Polymerase chain reactions (PCRs)

The 10 OSU loci were amplified in two multiplex reactions. The 25 μl reaction mix of Multiplex Maybruck 1(MPM1) contained: 3 ng of template DNA, 0.38–0.88 μM primers (DYS576, 0.38 μM; DYS504, 0.80 μM; DYS688, 0.44 μM; DYS487, 0.50 μM; DYS707, 0.88 μM (Invitrogen, Grand Island, NY and Applied Biosystems, Foster City, CA)), 1 mM dNTPs, 1× PCR buffer II (10 mM Tris–HCl, pH 8.3, 50 mM KCl), 2.25 mM $MgCl_2$, 10 μg of non-acetylated bovine serum albumin (Sigma–Aldrich, St. Louis, MO), and 1.25 units of AmpliTaq Gold DNA Polymerase (Applied Biosystems, Foster City, CA). The 25 μl reaction mix of Multiplex Maybruck 2 (MPM2) contained: 3 ng of template DNA, 0.25–0.44 μM primers (DYS448, 0.25 μM; DYS488, 0.25 μM; DYS471, 0.25 μM; DYS685, 0.25 μM; DYS703, 0.44 μM (Invitrogen and Applied Biosystems)), 1 mM dNTPs, 1× PCR buffer II, 2.25 mM $MgCl_2$, 10 μg non-acetylated bovine serum albumin, and 1.75 units of AmpliTaq Gold DNA polymerase. The PCR cycling conditions for both multiplex reactions were: (1) 11 min heat-soak at 95 °C, (2) 28 cycles of 1 min at 94 °C, 1 min at 60 °C and 30 s at 72 °C, and a final extension at 60 °C for 60 min. Primer sequences for the OSU 10-locus set are listed in Table 1. The SWGDAM loci were amplified according to the methods in Daniels et al. [35].