Forensic Science International: Genetics 5 (2011) 122–125

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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

Design and validation of a highly discriminatory 10-locus Y-chromosome STR multiplex system

M.E. D'Amato [a,*], V.B. Bajic [b,c], S. Davison [a]

[a] Forensic DNA Lab, University of the Western Cape, Biotechnology Department, Private Bag X17, 7535 Bellville, Cape Town, South Africa
[b] SANBI (South African National Bioinformatics Institute), University of the Western Cape, South Africa
[c] Computational Bioscience Research Center (CBRC), King Abdullah University of Science and Technology, Thuwal, Saudi Arabia

ARTICLE INFO

Keywords:
Y-STRs
Multiplex
Developmental validation
Forensic genetics

ABSTRACT

The Y-chromosome STRs (short tandem repeat) markers are routinely utilized in the resolution of forensic casework related to sexual assault. For this, the forensic community has adopted a set of eleven (core) Y-STR that is incorporated in all commercial diagnostic systems. Our previous studies of Y-STR polymorphisms in the South African population identified low levels of diversity and discrimination capacity for many commercial marker sets, determining a limited applicability of these systems to the local population groups. To overcome this shortcoming, we designed a Y-STR 10-plex system that shows higher discriminatory capacity (DC) than available commercial systems. The markers were selected from a population group of 283 individuals with African, European and Asian ancestry genotyped at 45 Y-STRs, applying an optimization based selection procedure to achieve the highest possible DC with the minimal number of markers. The 10-plex was satisfactorily subjected to developmental validation tests following the SWGDAM guidelines and shows potential for its application to genealogical and evolutionary studies.

1. Introduction

The success in the resolution of forensic cases is dependent upon an efficient recovery of DNA from the crime scene and the utilization of an efficient and discriminatory method of genetic profiling. Short Tandem Repeats on the Y-chromosome (Y-STRs) are highly valuable in forensics for their applicability to the study of sexual assault cases. In this context, the male biological material is usually recovered in a mixture with great excess of female's cells or fluids, what makes the identification of somatic male markers virtually impossible. Therefore, targeting of male-specific DNA is the only experimental approach that would allow for the identification of the perpetrator. To enhance the recovery of information from such a crime scene, several differential DNA extraction protocols have been reported that favor the separation of semen from the victim's epithelial cells (reviewed in [1]). In cases where the recovery of DNA applying differential extraction is hampered by degradation or sperm lyses, the identification of a male genetic profile is still possible, provided that the male DNA concentration is not too low. The application of Y-chromosome markers in forensics is then limited only when DNA mixtures contain DNA of related males, as only one genetic profile would be detected, and when male DNA quantities are below the level of detection for specific PCR conditions (reviewed in [2]).

South Africa has one the highest incidence of rape in the world, with nearly 54,000 cases being reported yearly to the South African Police. Given the known utility of Y-STR markers for their application in sexual assault casework, we conducted preliminary surveys of the genetic diversity in the local metropolitan populations. The metropolitan area of Cape Town hosts approximately three million inhabitants, with a complex composition of native population groups (mainly Xhosa, 31%), the descendants of Europe immigrants, mainly from English and Dutch origin (19%), India (0.5%), SE Asia (now admixed), and the admixed group resulting from the partial integration of these communities, named "coloured" (48%). Our studies of minimal haplotype (MH) loci showed low genetic diversity and poor haplotype resolution in the native group [3,4]. This observation encouraged further studies to identify loci with high informative value, and we extended the study to nearly 300 individuals of Xhosa, European and Indian ancestry to a total of 45 Y-STRs, including the MH loci, the additional loci incorporated in PowerPlex®Y System (Promega) [5] and AmpFℓSTR®Y-filer™ (Applied Biosystems) [6], and additional non-core loci [7–9 and D'Amato et al., unpublished data). The minimal number of Y-STR loci that achieved the highest haplotype resolution (DYS710, DYS385ab, DYS447, DYS504,

* Corresponding author. Tel.: +27 21 959 3372/2065; fax: +27 21 959 3505.
E-mail address: medamato@uwc.ac.za (M.E. D'Amato).

1872-4973/$ – see front matter © 2010 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.fsigen.2010.08.015

DYS449, DYS626, DYS644, DYS612, DYS481 and DYS518) was arranged in a single multiplex which was subjected to developmental validation tests following the SWGDAM guidelines [10].

2. Materials and methods

DNA extraction and STR genotyping of population samples were previously reported [3,4,7–9]. Semen samples were collected in plastic tubes and frozen at −20 °C until the time of use. DNA extraction was performed using a differential extraction procedure with NucleoSpin® Tissue (Macherey-Nagel).

DNA from environmentally impacted samples (blood blotted on cotton) was extracted following a standard phenol–chloroform extraction protocol.

2.1. Multiplex development

Marker selection: population data for 283 individuals (86 Asian Indian, 103 English, 94 Xhosa) sampled in the metropolitan area of Cape Town were genotyped at 45 Y-STR loci [7–9 and unpublished data]. An optimization procedure was written for Matlab (MathWorks Inc.). The algorithm selects first the marker with highest discriminatory power (DC). Then, it adds a marker that maximizes the joint DC. The process is repeated adding markers sequentially until no more improvement in DC is observed. This process inherently deals with the issue of markers that are correlated.

Primer design: some primers were redesigned aiming at reducing the amplicon size whenever possible (DYS449, DYS481, DYS612) and avoiding tracts homologous to X-chromosome (DYS644, DYS710) to accommodate all markers according to size, colour (DYS504) and *Tm* in a single multiplex reaction. Primers were designed with Oligo v1.4 [11] and *Tm* was calculated according to the nearest neighbor method. The possibility of these primers matching the human genome elsewhere but the Y-chromosome was tested with BLAT (http://genome.ucsc.edu/cgi-bin/hgBlat) and BLAST (http://blast.ncbi.nlm.nih.gov/) with BLASTN 2.2.18 [12].

Cycling conditions: reactions were optimized in a final volume of 10 μl, using 2 ng male DNA, 1× Taq polymerase buffer with 1.5 mM $MgCl_2$, 0.2 mM each dNTP (Roche), 2 U Taq polymerase (Supertherm, Medox Pty. Ltd.), 0.4 μg/μl of non-acetylated BSA (Ambion®). PCR amplification was optimized in an Applied Biosystems 2720 Thermal Cycler in the conditions: (1) 10 min at 94 °C, (2) 2 cycles 94 °C 30 s, 66 °C 1 min, 72 °C 1 min, (3) 2 cycles 65.5 °C (4) 2 cycles 65 °C, (5) 2 cycles 64.5 °C, (6) 2 cycles 64 °C, (7) 24 cycles 62 °C, (8) final extension 68 °C 75 min. *Primer concentrations*: 0.3 μM DYS481, 0.72 μM DYS447, 0.72 μM DYS449, 0.3 μM DYS710, 0.192 μM DYS518, 0.1 μM DYS612, 0.23 μM DYS626, 0.17 μM DYS504, 0.24 μM DYS385ab, 0.145 μM DYS644.

PCR product detection: 1 μl aliquot of PCR products were run in an ABI 377 automatic sequencer (Applied Biosystems). The minimal peak height accepted for scoring genotypes was 50 RFUs. For the inter-laboratory collaboration exercise an ABI 3130XL (Applied Biosystems) instrument was utilized. Allelic ladders were constructed as in [13].

2.2. Multiplex performance

Following the SWDAM guidelines, we performed tests for gender specificity (female DNA ranging from 10 ng to 1.6 μg), sensitivity (male DNA ranging from 50 pg to 10 ng), mixtures of female: male DNA (keeping male DNA constant at 2 ng and ratio up to 1:800; and keeping female DNA constant at 500 ng and ratios up to 1:1000), mixtures male:male DNA keeping total DNA constant at 3 ng, species-specificity testing 10 ng of male and female domestic cat (*Felis felis*), dog (*Canis familiaris*), primates of the species vervet monkey (*Chlorocebus pygerythrus*), chacma baboon (*Papio ursinus*), chimpanzee (*Pan troglodytes*) and Western Lowland Gorilla (*Gorilla gorilla gorilla*) . Accuracy of analyses was tested by two laboratories. Fifteen samples (10 donors and 5 NIST SRM 2395) were amplified in parallel by different analysts in two different Applied Biosystems 2720 Thermal Cyclers. The samples were analyzed separately in an ABI 377 and ABI 3130XL Genetic Analyzers using Windows and Macintosh systems respectively. Precision was tested by injecting the PCR products of one individual 20 times.

2.3. Multiplex application

Inheritance was tested in a total of 10 men from two families along 3 generations.

Mock cases: five cases were simulated as in [1], blotting 1–3 μl sperm onto ∼5 mm^2 vaginal swabs (Puritan) air-dried for 1–3 weeks. After air-drying for 1–3 weeks, male DNA was purified following the differential DNA extraction protocol of NucleoSpin® Tissue (Macherey-Nagel). The final elution step was performed with 20 μl of TE.

The impact of environmental factors on DNA stability and the ability of this multiplex to recover a full profile was tested by blotting 50 μl of blood onto died cotton cloth, air-dried and left exposed both indoors and outdoors. Samples were collected at 1 day, and after weeks 1, 2, 3, 4, 6 and 8; DNA was extracted and resuspended in 25 μl bidistilled water, quantified with a Nanodrop ND-2000 spectrophotometer and 2 ng were subsequently genotyped.

3. Results and discussion

3.1. Multiplex design

Several multiplexes have incorporated non-core loci [14–19]. Among these, a criterion for selection of loci was only followed by Hanson and Ballantyne [14], who plotted the haplotype diversity attained by the core SWGDAM loci plus one additional locus versus the gene diversity of this additional locus. Experimental design for loci selection provides the additional advantage of reducing the costs of genotyping at unnecessary loci that do not contribute to the discrimination of unrelated males.

Loci and primer characteristics are shown in Supplementary Material Table S1, and the design and loci product size ranges in Supplementary Material Fig. S1. The multiplex produced a minimum of 11 and maximum of 12 products: locus DYS385ab is duplicated and results in two PCR products of either identical or different size, and the target of the DYS710 R primer is repeated in tandem, thus generating a secondary product 22 bp larger, with peak height approximately 34% of the main peak height. We previously reported the characterization and potential of these loci for forensic, genealogical and evolutionary applications [8,9]. Other studies also reported high DC for loci DYS481 [20–22], DYS447 and DYS449 [14,21].

The comparative performance in DC of the UWC 10-plex and other extensively utilized Y-STR typing systems is shown in Table 1. The UWC-plex showed the highest DC not only for each population group but also overall individuals (DC = 0.9591), followed by AmpFℓSTR®Y-filer™ multiplex (overall DC = 0.9145).

3.2. Multiplex performance

The main results of the validation tests are summarized in Table 2 and Supplementary Material Fig. S2. PCR conditions were optimized for 2 ng male DNA and peak heights varied between 250 and 3500 RFUs. Some male profiles showed unspecific green (VIC) PCR products at 470 and 481 bp, with heights, on average, 5% of the