Forensic Science International: Genetics 20 (2016) 1–5

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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

A comprehensive Y-STR portrait of Argentinean populations

Ulises Toscanini[a,c,*], Carlos Vullo[b], Gabriela Berardi[a], Cintia Llull[a], Alicia Borosky[b], Andrea Gómez[a], Jacobo Pardo-Seco[c], Antonio Salas[c,**]

[a] PRICAI-Fundación Favaloro, Buenos Aires, Argentina
[b] LIDMO-EAAF, Córdoba, Argentina
[c] Unidade de Xenética, Departamento de Anatomía Patolóxica e Ciencias Forenses, and Instituto de Ciencias Forenses, Grupo de Medicina Xenómica (GMX), Facultade de Medicina, Universidade de Santiago de Compostela, 15872 Galicia, Spain

ARTICLE INFO

Article history:
Received 17 April 2015
Received in revised form 9 August 2015
Accepted 2 September 2015
Available online 8 September 2015

Keywords:
PPY23
Y-STR
Argentina
Forensic genetics
YHRD

ABSTRACT

A study of 23 Y-STRs was conducted in 257 individuals living in urban areas from eight Argentinean provinces. The data were meta-analyzed together with 364 profiles obtained from the literature that represent other five provinces. A total of 255 different haplotypes were observed (253 singletons). Genetic structure estimated from analysis of molecular variance (AMOVA) and exploring different grouping scenarios, yielded high within population variance. Not surprisingly, analyses of genetic distances with respect to main ancestral continental populations indicated Argentinean haplotypes to be closely related to European ones. Overall, these results provide a quite complete picture of the patterns of Y chromosome variation in Argentina, notably contributing to increase the previous national database, and consequently allowing a better estimation of parameters of interest in forensic casework and parentage testing.

1. Introduction

Since the firsts investigations on Y-STR markers for forensic applications almost two decades ago [e.g. 1–3], the initial scenario continuously evolved in different aspects. In terms of number of markers, the original 9-Y-STR set, namely the minimal haplotype (MHT), gradually expanded in order to increase discrimination power of the Y-chromosome test (e.g. PPY12 [PowerPlex®Y12, Promega Corp., USA], Yfiler [AmpF*l*STR® Yfiler™ kit, Life Technologies, USA]). The PowerPlex®Y23 system (PPY23, Promega Corp., USA) is a newly developed multiplex kit that allows the simultaneous analysis of 23 Y-STRs, incorporating six markers in addition to those included in the Yfiler kit. This Y-STR panel was recently validated for forensic use and for concordance with the Yfiler [4,5], and the full 23-Y-STR-haplotype proved to notably increase the discrimination capacity (DC) in different population studies [6–8].

Analysis of Y-STR variation also expanded to a growing number of human populations and ethnic groups, in order to obtain a more accurate view of local, regional and worldwide patterns of human variation. The largest effort in building such a comprehensive database is the YHRD (www.yhrd.org), that nowadays accommodates 154,329 MHT, 114,993 PPY12, 102,729 Yfiler, 26,107 PPY23 and 6872 YfilerPlus profiles, belonging to 991 populations in 129 countries (release R50 valid as per 2015-07-18).

On the other hand, revisions of early statistical approaches in forensic casework were recently proposed [9–14], in part stimulated by the continuous growing of reference databases.

Several studies on Y-STRs were published for different urban and Native American populations of Argentina, mainly for MHT, PPY12 or YFiler configuration [e.g. 15–20]. At present, there are 4265 MHT, 2613 PPY12 and 1814 YFiler Argentinean haplotypes in YHRD. However, the number of full PPY23 haplotypes available today sum up to only 621 (of which, 257 are the haplotypes reported in this work) (results are based upon release R50 valid as per 2015-07-18; query sent on 2015-08-08). Moreover, geographic patterns of genetic variation for PPY23 haplotypes were not specifically analyzed in Argentina, although previous studies carried out on panels containing fewer STRs (e.g. [17–19]) showed differences that could be relevant for the statistical evaluation of the Y-chromosome test in the country.

In this study, we genotyped a sample of 257 individuals from urban locations of different provinces of Argentina for the 23 loci

* Corresponding author. Fax: +54 1143831197.
** Corresponding author.
E-mail addresses: utoscanini@pricai.com.ar, utoscanini@gmail.com (U. Toscanini), antonio.salas@usc.es (A. Salas).

http://dx.doi.org/10.1016/j.fsigen.2015.09.002
1872-4973/

included in the PPY23 kit. With this genotyping effort we aimed to significantly enlarge the current country database for these set of markers and sampling geographic coverage, and consequently improve the forensic performance for this panel. We also aimed to investigate more deeply patterns of geographic variation in Argentina that could be relevant in forensic genetics.

2. Materials and methods

2.1. Populations

Blood samples on filter paper (Hybond Blotting Paper, Amersham, USA) or buccal swabs were collected from 257 unrelated males inhabiting urban areas of eight provinces of Argentina (Fig. 1) and the Y-chromosome profiles analyzed were submitted to the YHRD database (http://yhrd.org). The following provinces were sampled (in brackets are sample sizes [*n*] and YHRD accession numbers): Jujuy ($n=50$, YA004012), Tucumán ($n=30$, YA004009), Catamarca ($n=29$, YA004007), Córdoba ($n=31$, YA004014), San Luis ($n=28$, YA004010), La Pampa ($n=30$, YA004011), Entre Ríos ($n=29$, YA004013) and Chubut ($n=30$, YA004008). All the samples were obtained with the corresponding written informed consent.

Apart from the Argentinean datasets analyzed in Purps et al. [8], we also retrieved the following datasets from this same source: (i) South America ($n=487$): Bolivia [Mestizo], Bolivia [Native American], Peru [Peruvian], Río de Janeiro, Brazil [Admixed Brazilian], São Gabriel de Cachoeira, Brazil [Native American], São Paulo, Brazil [Admixed Brazilian], (ii) Native American ($n=56$): Bolivia [Native American]; (iii) Europe ($n=2168$): Aragón, Spain [Spanish], Asturias, Spain [Spanish], Barcelona, Spain [Spanish], Galicia, Spain [Spanish], Madrid, Spain [Spanish], Brescia, Italy [Italian], Calabria, Italy [Italian], Liguria, Italy [Italian], Marche, Italy [Italian], Milano, Italy [Italian], Northeastern Italy, Italy [Italian], Puglia, Italy [Italian], Ravenna, Italy [Italian], Sicily, Italy [Italian], Tuscany, Italy [Italian], (iv) Asia ($n=101$): Yunnan, China [Bai]; and (v) Africa ($n=136$): Ibadan, Nigeria [Yoruba], Zimbabwe [Zimbabwean].

2.2. DNA typing

DNA extraction from blood samples was carried out by means of Chelex®100 method, using an approximated area of 5–10 mm^2 of blood spot. Buccal swabs were processed following a standard salting-out procedure.

Y-STR typing for the PPY23 system was performed according to manufacturer's instructions (PowerPlex®Y23 System Technical Manual, Promega Corp.) reducing the PCR final volume to 10 μl and accordingly adjusting the proportions of reagents in the reaction mix. A volume of 1 μl of each DNA extract containing ~0.5–1 ng was used as template.

Separation and detection of amplified products were carried out in an ABI PRISM 310 genetic analyzer. Allele calling was done by comparison with the reference ladder provided with the kit using GeneMapper v3.2 software (Applied Biosystems, USA).

2.3. Quality control

The laboratories involved in the genotyping participate in the annual GHEP-ISFG proficiency testing, succeeding in all the exercises. The data reported in the present study (Table S1) follow therefore the new guidelines for genetic data publication [21,22], as recommended by the ISFG, which requires an annual proficiency test certification from the GHEP-ISFG working group (http://www.gep-isfg.org), and quality control data to be submitted to the YHRD database. The data can be queried online using the accession numbers indicated above for each sample set.

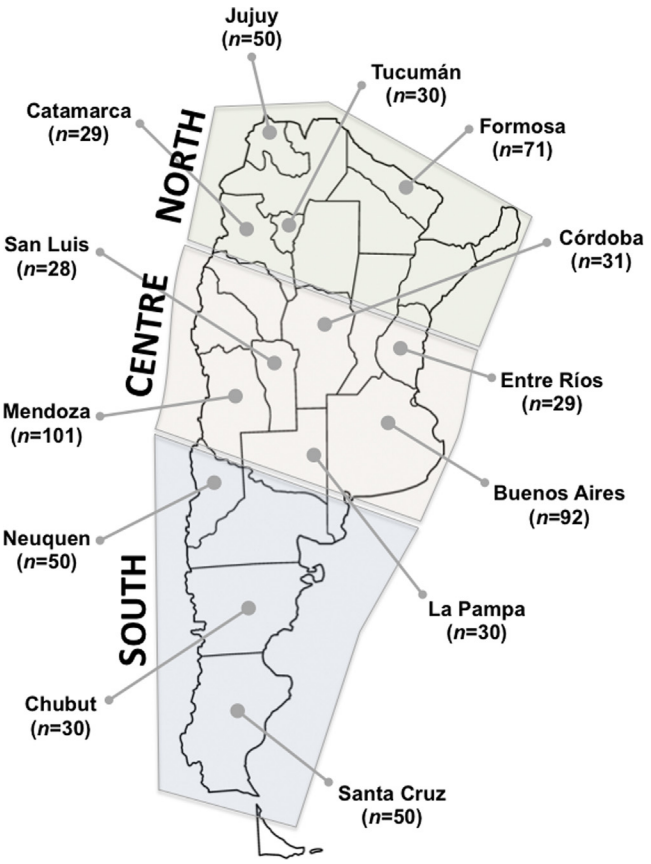

Fig. 1. Location of Argentinean provinces analyzed in the present study (and currently represented in YHRD for PPY23 haplotypes).

2.4. Statistical analyses

Haplotype diversity (HD), discrimination capacity (DC) and matching probability (MP), were estimated as in Purps et al. [8] taking into considerations the issues regarding DYS385 and DYS389 loci. The average number of pairwise differences was also computed. For most of the analyses the Arlequin software v3.5.1.2 was used [23]. All the indices were calculated using the full PPY23 haplotypes as well as collapsing these haplotypes to their corresponding MHT. Computation of indices was also carried considering different national and regional grouping schemes (Fig. 2).

Genetic structure and variation between individual and grouped populations was carried out by means of analysis of molecular variance (AMOVA) and R_{ST} genetic distances as implemented in Arlequin software v3.5.1.2. DYS385ab marker was excluded from these analyses, as well as samples containing duplications, and null or intermediate alleles. The number of repeats in DYS389I was subtracted from DYS389II. AMOVA was performed considering all the data available from Argentinean populations, namely, the population studied here and those included in Purps et al. [8]. In order to explore possible patterns of variation across the country a grouping approach based on the geographical location was also evaluated. Three main regions were considered for this purpose: (i) North: including provinces of Jujuy, Formosa, Tucumán and Catamarca; (ii) Centre: including provinces of Córdoba, San Luis, La Pampa, Mendoza, Buenos Aires and Entre Ríos; and (iii) South: including provinces of Neuquén, Chubut and Santa Cruz. In a separate analysis, AMOVA was also computed excluding one sample at a time and considering the remaining 12 Argentinean samples for the analyses. Moreover,