Genetic characterization of four Brazilian states with 25 Yfiler®Plus markers

J. Jannuzzi[a,*], C.S. Alho[b], C. Fridman[c], S.R.B. Ferreira[d], D.F. Braganholi[e], I.B. Ambrosio[e], R.M.B. Cicarelli[e], V. Gomes[f,g], E.F. Carvalho[a], L. Gusmão[a,f,g]

[a] Laboratório de Diagnósticos por DNA (LDD), Universidade do Estado do Rio de Janeiro (UERJ), Rio de Janeiro, Brazil
[b] Laboratório de Genética e Biologia Molecular, Pontíficia Universidade Católica do Rio Grande do Sul (PUCRS), Rio Grande do Sul, Brazil
[c] Departamento de Medicina Legal, Ética Médica e Medicina Social e do Trabalho da Faculdade de Medicina da USP, Universidade de São Paulo (USP), São Paulo, Brazil
[d] Faculdade Pitágoras, São Luís, Brazil
[e] Faculdade de Ciências Farmacêuticas (FCFAr), Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP), São Paulo, Brazil
[f] IPATIMUP-Institute of Pathology and Molecular Immunology from the University of Porto, Portugal
[g] I3S-Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Porto, Portugal

ARTICLE INFO

ABSTRACT

The Y-STRs have high mutation rates, being useful to discriminate unrelated males. They are widely used in paternity and forensic investigations to study the recent history and migration movements of populations. Recently, the YfilerPlus was released to increase discrimination inside populations, by adding 9 Y-STRs to the previous Yfiler version. The genetic composition of Brazil is known to vary through 5 geopolitical regions: South, Southeast, North, Northeast and Central-west. Therefore, for this study, samples from Maranhão, Espírito Santo, São Paulo and Rio Grande do Sul were characterized for the Y-STR markers present in YfilerPlus kit, to evaluate if diversity increases with the enlargement of the Yfiler set, and if significant differences exists among populations. Genetic differentiation analysis did not reveal statistically significant differences in the YfilerPlus haplotype composition of the studied samples. High diversities were observed in all samples for both Yfiler and YfilerPlus marker sets. Inside populations, 20 haplotypes were shared by two individuals for the Yfiler-STRs. This number decreases to 14 when using the 25 Y-STRs from the YfilerPlus. The YfilerPlus demonstrated an increased discrimination power in comparison to the Yfiler kit, being suitable for forensic applications in the studied Brazilian populations, for which data were not yet available.

1. Introduction

The Y short tandem repeats (Y-STRs) have high mutation rates ($\approx 10^{-3}$), and the set of their alleles characterize haplotypes. Due to their mutation rates, Y-STRs can discriminate male individuals, being widely used in paternity tests, forensic investigations, genealogical researches, besides the analysis of recent history of populations and its migration movements [1–4].

There are diverse sets of Y-STRs being used in forensic and population genetics. To increase discrimination inside populations, new commercial kits have been released displaying additional markers with high mutation rates. The YfilerPlus is a recent Y-STR forensic kit, with 25 markers, including the 16 from the previous Yfiler version.

Brazilian populations are very heterogeneous, presenting a high genetic diversity spread through the 5 country regions, named South, Southeast, North, Northeast and Central-west [5–7].

Since YfilerPlus was recently released and there is no data available for Brazilian populations, the aims of this study were to (i) genotype

samples from four states of three geopolitical regions, (ii) evaluate if the addition of Y-STR markers would increase the diversity values in comparison to the observed with Yfiler and (iii) evaluate if there is population substructure between these Brazilian states.

2. Materials and methods

A total of 518 unrelated male individuals from Maranhão (n = 114), Espírito Santo (n = 76), São Paulo (n = 126) and Rio Grande do Sul (n = 202) provided their samples for this study, after signing a written informed consent for cooperation under strictly confidential conditions.

DNA was extracted with standard chelex and phenol/chloroform methodologies. Samples were genotyped with the Yfiler®Plus kit (Applied Biosystems), following the User Guide [8]. Capillary electrophoresis and detection were performed on a 3500 Genetic Analyser using POP-7™ polymer (Applied Biosystems). The genotypes were assigned using the software GeneMapper ID v4.1 (Applied Biosystems).

Table 1

Population pairwise R_{ST} genetic distances between the four Brazilian states.

	Espírito Santo	São Paulo	Maranhão	Rio Grande do Sul
Espírito Santo	–	0.7058 ± 0.0041	0.4119 ± 0.0045	0.5358 ± 0.0047
São Paulo	−0.0015	–	0.3699 ± 0.0045	0.5982 ± 0.0051
Maranhão	0.0001	0.0002	–	0.0939 ± 0.0030
Rio Grande do Sul	−0.0005	−0.0006	0.0023	–

Legend: Below diagonal shows the R_{ST} values; Above diagonal shows non-differentiation probabilities calculated for 10.100 permutations. Significance level was equal 0.0083.

The R_{ST} genetic distances were calculated with Arlequin v3.5 [9]. This software was also used to calculate haplotype diversities (HD) and the number of shared haplotypes for the 25 Y-STRs and the 16 STRs from the Yfiler (after removing extra markers), in the same population samples. For the analysis of genetic distances between the four population samples, the number of repetitions in DYS398I was subtracted from DYS398II and DYS385 and DYF387S1 were excluded. Null, microvariant or duplicated alleles were represented as missing data.

3. Results and discussion

High values of HD were found with the 25 Y-STRs in all Brazilian samples analyzed (Maranhão, HD = 0.9992 ± 0.0013; Espírito Santo, HD = 0.9996 ± 0.0022; São Paulo, HD = 1.0000 ± 0.0010; Rio Grande do Sul, HD = 0.9996 ± 0.0005), due to a low number of shared haplotypes inside populations. In total, there were 14 haplotypes shared by two individuals: 1 in Espírito Santo, 5 in Maranhão and 8 in Rio Grande do Sul. When just considering the 16 STRs from the Yfiler, a slight decrease in the HD was observed in Maranhão (0.9989 ± 0.0013), Espírito Santo (0.9993 ± 0.0023) and Rio Grande do Sul (0.9995 ± 0.0006), while remained the same in São Paulo, indicating that both marker sets could equally discriminate individuals from this population sample. After the reduction of the number of Y-STRs to 16, there was an increase in the number of shared haplotypes inside populations: 2 from Espírito Santo, 7 from Maranhão, 11 from Rio Grande do Sul. The improvement in discrimination inside populations that was obtained with the new kit, was also observed in previous studies on other populations [10–12].

Genetic distance analysis indicated a homogeneity between the four populations, since R_{ST} values were low and the non-differentiation probabilities high, far above the significance level (Table 1). A previous study in other Brazilian population samples with 23 Y-STRs showed similar results, in which no significant genetic distances were found between most populations among the five regions of Brazil [6]. This genetic homogeneity can be explained by the fact that the set and the number of markers used provides high discrimination among individuals, being more suitable to detect differences between individuals than between populations.

4. Conclusion

It is possible to conclude that the addition of markers in the YfilerPlus kit increased the discrimination inside populations in comparison with the just 16 markers present in Yfiler kit. Therefore, the 25 STRs from the YfilerPlus demonstrated to be suitable for forensic applications in the studied populations from Brazil, for which data were not yet available.

However, it is still necessary to analyze whether this applies to larger population samples and to generate YfilerPlus data for other Brazilian population samples that were not yet studied.

Conflict of interest statement

None.

References

[1] J.M. Butler, Fundamentals of Forensic DNA Typing, 1st edition, (2009).
[2] M. Kayser, R. Kittler, A. Erler, M. Hedman, A.C. Lee, A. Mohyuddin, et al., A comprehensive survey of human Y-chromosomal microsatellites, Am. J. Hum. Genet. 74 (2004) 1183–1197.
[3] K.N. Ballantyne, M. Goedbloed, R. Fang, O. Schaap, O. Lao, A. Wollstein, et al., Mutability of Y-chromosomal microsatellites: rates, characteristics, molecular bases, and forensic implications, Am. J. Hum. Genet. 87 (2010) 341–353.
[4] K.N. Ballantyne, V. Keerl, A. Wollstein, Y. Choi, S.B. Zuniga, A. Ralf, et al., A new future of forensic Y-chromosome analysis: rapidly mutating Y-STRs for differentiating male relatives and paternal lineages, Forensic Sci. Int. Genet. 6 (2012) 208–218.
[5] F. Manta, R. Pereira, R. Vianna, A. Rodolfo Beuttenmüller de Araújo, D. Leite Góes, D. Gitaí, D. Aparecida da Silva, et al., Revisiting the genetic ancestry of Brazilians using autosomal AIM-indels, PLoS One 8 (2013) 1–11.
[6] T. Palha, L. Gusmão, E. Ribeiro-Rodrigues, J.F. Guerreiro, Â. Ribeiro-dos-Santos, S. Santos, Disclosing the genetic structure of Brazil through analysis of male lineages with highly discriminating haplotypes, PLoS One 7 (2012) 1–8.
[7] R. Resque, L. Gusmão, M. Geppert, L. Roewer, T. Palha, L. Alvarez, et al., Male lineages in Brazil: intercontinental admixture and stratification of the european background, PLoS One 11 (2016) e0152573.
[8] Yfiler® Plus PCR Amplification Kit User Guide, Thermo Fisher Scientific Inc., 2014.
[9] L. Excoffier, H. Lischer, Arlequin: An Integrated Software Package for Population Genetics Data Analysis, (2015).
[10] J.K. Olofsson, H.S. Mogensen, A. Buchard, C. Børsting, N. Morling, Forensic and population genetic analyses of Danes, Greenlanders and Somalis typed with the Yfiler® Plus PCR amplification kit, Forensic Sci. Int. Genet. 16 (2015) 232–236.
[11] C. Rapone, E. D'Atanasio, A. Agostino, M. Mariano, M.T. Papaluca, F. Cruciani, et al., Forensic genetic value of a 27 Y-STR loci multiplex (Yfiler® Plus kit) in an Italian population sample, Forensic Sci. Int. Genet. 21 (2015) 1–5.
[12] O. García, I. Yurrebaso, I.D. Mancisidor, S. López, S. Alonso, L. Gusmão, Data for 27 Y-chromosome STR loci in the basque country autochtonous population, Forensic Sci. Int. Genet. 20 (2016) e10–e12.