GMR Genetics and Molecular Research
Online Journal
ISSN:1676-5680

Short Communication

Y-STR haplotype diversity and population data for Central Brazil: implications for environmental forensics and paternity testing

T.C. Vieira[1,2,3,4], M.A.D. Gigonzac[2,3,4], D.M. Silva[1], R.G. Rodovalho[5], G.S. Santos[5] and A.D. da Cruz[1,2,3]

[1]Programa de Pós-Graduação em Biologia, Instituto de Ciências Biológicas, Universidade Federal de Goiás, Goiânia, GO, Brasil
[2]Laboratório de Citogenética Humana e Genética Molecular, Laboratório Central de Saúde Pública Dr. Giovanni Cysneiros, Secretaria de Estado da Saúde/Governo de Goiás, Goiânia, GO, Brasil
[3]Núcleo de Pesquisas Replicon, Departamento de Biologia, Pontifícia Universidade Católica de Goiás, Goiânia, GO, Brasil
[4]Universidade Estadual de Goiás, Goiânia, GO, Brasil
[5]Laboratório Biocroma, Laboratório de Análise de DNA, Goiânia, GO, Brasil

Corresponding author: T.C. Vieira
E-mail: thaiscidalia@gmail.com

Genet. Mol. Res. 13 (2): 3404-3410 (2014)
Received August 20, 2013
Accepted November 8, 2013
Published April 30, 2014
DOI http://dx.doi.org/10.4238/2014.April.30.1

ABSTRACT. The central region of Brazil was colonized by internal migration of individuals of different origins, who contributed to the genetic diversity existing in this population. This study determined the allele frequencies and haplotype diversity of Y-STRs in Goiás State, Central Brazil, and compared the data obtained with a sample of the Brazilian population, consisting of individuals from the five geographical regions of Brazil. A total of 353 males were typed for

12 Y-chromosome short tandem repeat (Y-STR) markers. We selected males who had no degree of relatedness, from the five mesoregions of Goiás State. DNA was extracted from blood samples followed by the amplification of the 12 Y-chromosome loci. The products were analyzed to obtain the allele profiles on an ABI3500 automated sequencer using the Gene Mapper software. Allele frequencies and haplotype diversity were estimated by direct counting, and gene diversity for each locus was computed using the Arlequin software. The results are consistent with the history of miscegenation of the population of Central Brazil, in which we observed 321 different haplotypes. The average gene diversity at the 12 loci was 0.645. DYS385b and DYS389I showed the highest (0.704) and lowest (0.520) genetic diversity values, respectively. The F_{ST} value between the Brazilian and Goiás populations was 0.00951, showing no statistical significance. The results of this study allowed the establishment of haplotypes found in the forensic samples of Goiás State serving as a reference in the elucidation of criminal cases and paternity tests, as well as population and evolutionary inferences.

Key words: Y chromosome; STR markers; Haplotype diversity

INTRODUCTION

The analysis of Y-chromosome short tandem repeats (STRs) has been extensively used in forensic genetics and for paternity testing. Furthermore, these polymorphisms serve as valuable markers for ancestry studies, and the main population shows remarkable allele frequency distributions (Palha et al., 2007; Carvalho and Pinheiro, 2013; Vieira et al., 2013).

Due to the presence of the largest non-recombining region in the whole human genome, the Y chromosome is characterized by a unique inheritance pattern and specificity to males (Rębała and Szczerkowska, 2005). The lack of recombination between Y-chromosome specific markers means that these markers are transmitted as haplotypes in the same way as single locus alleles. The lower effective number of Y-chromosomes in a given population also means that Y-haplotypes/haplogroups tend to show a higher proportion of variation between populations than observed for other markers located on autosomes or X chromosomes (Domingues et al., 2007). Thus, Y-chromosome single tandem repeat (Y-STR) analysis can be very useful in paternity tests in which the alleged father is missing or deceased, through analysis of related individuals (Francez et al., 2012).

Investigations of genetic links using database DNA has stood out in recent years (Chung and Fung, 2013; Mardini et al., 2013). Moreover, the study of human race has been of interest since ancient times. The existence of differences between populations of distinct geographical origins must have been a finding of the first long-distance travelers (Silva et al., 2010). Nowadays, considerable research has been conducted to identify the geographical origin of people and their closeness, from several points of view, as well as the degree of admixture in the genetic composition of a population. Therefore, the aim of this study was to determine the allele frequencies and haplotype diversity of Y-STRs in the population of Goias State, Central Brazil, and to compare the results with population data for samples from all Brazil regions (Mardini et al., 2013).

MATERIAL AND METHODS

A total of 353 males were typed for the following 12 Y-chromosomal short tandem repeat (Y-STR) markers: DYS389I, DYS389II, DYS391, DYS437, DYS438, DYS439, DYS19, DYS392, DYS393, DYS390, DYS385I, and DYS385II. We selected the samples of the males who had no degree of relatedness, from five mesoregions comprising Goias State (Central, East, North, Northwest, and South regions).

DNA was extracted from blood samples using a commercial kit (GE Healthcare, UK), followed by the amplification of the 12 Y-chromosome loci (Powerplex® System, Promega Corporation) according to manufacturer instructions. The amplification products were analyzed on an ABI3500 automated sequencer (Life Technology, Applied Biosystems) to obtain the allele profiles, which were analyzed using the GeneMapper software 2.1 (Applied Biosystems). Allele frequencies and haplotype diversity were estimated by direct counting, and gene diversity (GD) for each locus was computed using the Arlequin software, version 3.5 (10,000 permutations) (Excoffier and Lischer, 2010).

RESULTS AND DISCUSSION

The results are consistent with the history of miscegenation for the population of Central Brazil, in which we observed 321 different haplotypes, indicating an individual discrimination power of 0.91. The average gene diversity at the 12 loci was 0.645. The DYS385b and DYS389I showed the highest (0.704) and lowest (0.520) genetic diversity values, respectively (Table 1).

Midwestern Brazil, like other Brazilian populations, is derived from the admixture of 3 main groups: Amerindians, Europeans, particularly Portuguese, and Africans from sub-Saharan Africa (Vieira et al., 2013). A comparative analysis of our population with published data from other Brazilian populations was performed using the Arlequin software 3.5 (Pereira et al., 2007; Excoffier and Lischer, 2010). Three rare alleles were detected in the Goias population in markers DYS391, DYS392 and DYS438, which were not described in the study of the Brazilian population.

To test the level of genetic differentiation between the population from Goias State and the Brazilian population (F_{ST}), intrapopulation genetic diversity (H_S) and the proportion of the total genetic diversity (H_T) were estimated using the FSTAT software, version 2.9 (Goudet, 2001) (Table 2). The overall ϕ_{ST} and R_{ST} values determined by AMOVA were respectively 0.009 and 0.0084, confirming no significant differentiation for the Y-STR haplotype between the Goiás and Brazilian populations. Allele frequencies and resulting statistical parameters are given in Tables 1 and 2.

The genetic differentiation coefficient (F_{ST}) between the Brazilian and Goias populations was 0.00951, showing no statistical significance ($P > 0.05$). An F_{ST} value between 0 and 0.05 indicates a low level of genetic differentiation, but it does not imply the absence of differentiation.

Forensic genetics interest has focused on Y-STR markers, which are available for population genetic, evolutionary, genealogical, and forensic investigations (Soares-Vieira et al., 2008). The results reported here characterized the genetic diversity of the population of Goias based on Y-STR and contributed to consolidate a database that could be useful in human identification studies. Additionally, the data will support the demands from the Judiciary and Public Ministry in the area of family and criminal affairs, helping in the elucidation of criminal cases and paternity testing for the Goias population, especially when the DNA of the alleged father is not available.

Table 1. Allele frequencies and genetic diversity values of 12 Y-STRs in 353 men from the Goias population in Central Brazil.

Allele	DYS389I	DYS389II	DYS391	DYS437	DYS438	DYS439	DYS19	DYS392	DYS393	DYS390	DYS385a	DYS385b
7	-	-	0.0028*	-	-	0.0000		-	-	-	-	-
8	-	-	0.0000	-	0.0057	0.0000	-	-	-	-	-	-
9	-	-	0.0682	-	0.1178	0.0057	-	0.0000	-	-	0.0028	-
10	-	-	0.4773	-	0.2529	0.1026	-	0.0029	0.0031	-	0.0028	-
11	-	-	0.4375	-	0.1351	0.3305	-	0.4029	0.0951	-	0.4986	0.0170
12	0.1648	-	0.0142	0.0000	0.4511	0.4074	0.0085	0.0600	0.4294	-	0.0997	0.0398
13	0.6477	-	-	0.0085	0.0345	0.1481	0.1225	0.4714	0.3865	-	0.1681	0.0284
14	0.1875	-	-	0.3371	0.0029*	0.0057	0.5499	0.0543	0.0736	-	0.0969	0.4943
15	0.0000	-	-	0.5411	-	-	0.2137	0.0057	0.0123	-	0.0427	0.1903
16	0.0000	-	-	0.1105	-	-	0.0798	0.0029*	-	-	0.0598	0.0795
17	-	-	-	0.0028	-	-	0.0256	-	-	-	0.0228	0.0653
18	-	-	-	-	-	-	-	-	-	-	0.0057	0.0568
19	-	-	-	-	-	-	-	-	-	-	0.0000	0.0227
20	-	-	-	-	-	-	-	-	-	0.0000	-	0.0057
21	-	-	-	-	-	-	-	-	-	0.0595	-	0.00000
22	-	-	-	-	-	-	-	-	-	0.0708	-	-
23	-	-	-	-	-	-	-	-	-	0.2578	-	-
24	-	-	-	-	-	-	-	-	-	0.4929	-	-
25	-	-	-	-	-	-	-	-	-	0.1105	-	-
26	-	0.0057	-	-	-	-	-	-	-	0.0028	-	-
27	-	0.0028	-	-	-	-	-	-	-	0.0028	-	-
28	-	0.1080	-	-	-	-	-	-	-	-	-	-
29	-	0.4261	-	-	-	-	-	-	-	-	-	-
30	-	0.3125	-	-	-	-	-	-	-	-	-	-
31	-	0.1136	-	-	-	-	-	-	-	-	-	-
32	-	0.0284	-	-	-	-	-	-	-	-	-	-
33	-	0.0028	-	-	-	-	-	-	-	-	-	-
34	-	-	-	-	-	-	-	-	-	-	-	-
N	352	352	352	353	348	351	351	350	331	353	351	352
GD	0.520	0.697	0.578	0.583	0.701	0.694	0.632	0.611	0.657	0.672	0.700	0.704

N = number of individuals; GD = gene diversity; *rare alleles.

Table 2. Allele frequencies of Brazilian and Goiás populations, intrapopulation genetic diversity for each locus (H_S), total genetic diversity for each locus (H_T) and genetic differentiation coefficient between populations (F_{ST}).

Locus	Goiás population	Brazilian population
DYS439		
(N)	351	480
7	0.0000	0.0021
8	0.0000	0.0021
9	0.0057	0.0062
10	0.1026	0.0561
11	0.3305	0.3929
12	0.4074	0.3992
13	0.1481	0.1227
14	0.0057	0.0187
H_S = 0.6800	H_T = 0.6817	F_{ST} = 0.00264
H_S (nc) = 0.6817	H_T (nc) = 0.6825	
DYS19		
(N)	351	480
12	0.0085	0.0042
13	0.1225	0.1333
14	0.5499	0.5188
15	0.2137	0.2542
16	0.0798	0.0583
17	0.0256	0.0313
H_S = 0.6370	H_T = 0.6378	F_{ST} = 0.00006
H_S (nc) = 0.6486	H_T (nc) = 0.6386	

Continued on next page

Table 2. Continued.

Locus	Goiás population	Brazilian population
DYS389I		
(N)	352	481
10	0.0000	0.0021
11	0.0000	0.0104
12	0.1648	0.1705
13	0.6477	0.6466
14	0.1875	0.1622
15	0.0000	0.0083
H_S = 0.5223	H_T = 0.5225	F_{ST} = - 0.00164
H_S (nc) = 0.5236	H_T (nc) = 0.5231	
DYS389II		
(N)	352	481
26	0.0057	0.0042
27	0.0028	0.0125
28	0.1080	0.1351
29	0.4261	0.4699
30	0.3125	0.2640
31	0.1136	0.0894
32	0.0284	0.0229
33	0.0028	0.0021
H_S = 0.6889	H_T = 0.6904	F_{ST} = 0.00168
H_S (nc) = 0.6906	H_T (nc) = 0.6912	
DYS390		
(N)	353	481
20	0.0000	0.0021
21	0.0595	0.0437
22	0.0708	0.1019
23	0.2578	0.2911
24	0.4929	0.4699
25	0.1105	0.0852
26	0.0028	0.0062
27	0.0028	0.0000
H_S = 0.6724	H_T = 0.6732	F_{ST} = 0.00016
H_S (nc) = 0.6740	H_T (nc) = 0.6741	
DYS391		
(N)	352	480
7	0.0028	0.0000
8	0.0000	0.0021
9	0.0682	0.0542
10	0.4773	0.5146
11	0.4375	0.4146
12	0.0142	0.0146
H_S = 0.5681	H_T = 0.5686	F_{ST} = - 0.00059
H_S (nc) = 0.5695	H_T (nc) = 0.5693	
DYS392		
(N)	350	481
9	0.0000	0.0021
10	0.0029	0.0021
11	0.4029	0.4179
12	0.0600	0.0499
13	0.4714	0.4636
14	0.0543	0.0561
15	0.0057	0.0083
16	0.0029	0.0000
H_S = 0.6068	H_T = 0.6069	F_{ST} = - 0.00213
H_S (nc) = 0.6083	H_T (nc) = 0.6076	

Continued on next page

Table 2. Continued.

Locus	Goiás population	Brazilian population
DYS393		
(N)	331	481
10	0.0000	0.0021
11	0.0906	0.0083
12	0.4139	0.1726
13	0.3988	0.6819
14	0.0816	0.1102
15	0.0151	0.0249
H_S = 0.5735	H_T = 0.6100	F_{ST} = 0.11342
H_S (nc) = 0.5749	H_T (nc) =0.6107	
DYS385a		
(N)	351	480
9	0.0028	0.0083
10	0.0028	0.0146
11	0.4986	0.4437
12	0.0997	0.0979
13	0.1681	0.1937
14	0.0969	0.0917
15	0.0427	0.0458
16	0.0598	0.0646
17	0.0228	0.0313
18	0.0057	0.0063
19	0.0000	0.0021
H_S = 0.7189	H_T = 0.7199	F_{ST} = 0.00028
H_S (nc) = 0.7207	H_T (nc) = 0.7208	
DYS385b		
(N)	353	481
11	0.0170	0.0083
12	0.0398	0.0208
13	0.0284	0.0644
14	0.4943	0.4532
15	0.1903	0.1642
16	0.0795	0.0811
17	0.0653	0.1040
18	0.0568	0.0603
19	0.0227	0.0291
20	0.0057	0.0125
21	0.0000	0.0021
H_S = 0.7217	H_T = 0.7231	F_{ST} = 0.00146
H_S (nc) = 0.7234	H_T (nc) = 0.7240	
DYS437		
(N)	353	481
12	0.0000	0.0021
13	0.0085	0.0042
14	0.3371	0.3555
15	0.5411	0.5177
16	0.1105	0.1123
17	0.0028	0.0083
H_S = 0.5871	H_T = 0.5874	F_{ST} = - 0.00166
H_S (nc) = 0.5886	H_T (nc) = 0.5881	
DYS438		
(N)	348	481
8	0.0057	0.0042
9	0.1178	0.1351
10	0.2529	0.2973
11	0.1351	0.0811
12	0.4511	0.4636
13	0.0345	0.0187
14	0.0029	0.0000
H_S = 0.6853	H_T = 0.06867	F_{ST} = 0.00161
H_S (nc) = 0.6870	H_T (nc) = 0.6876	

H_S = intrapopulation genetic diversity for each locus; H_S (nc) = intrapopulation genetic diversity not corrected for each locus; H_T = total genetic diversity for each locus; H_T (nc) = total genetic diversity not corrected for each locus; F_{ST} = genetic differentiation coefficient between populations.

ACKNOWLEDGMENTS

Research supported by FAPEG-GO (Fundação de Amparo à Pesquisa do Estado de Goiás) and CNPq (for T.C. Vieira scholarship). We thank the participants who voluntarily donated blood samples for this study. We are also grateful to the staff from Núcleo de Pesquisas Replicon da Pontifícia Universidade Católica de Goiás and Laboratório de Citogenética Humana- Genética Molecular/Secretaria de Estado da Saúde (LaGene/Lacen/SES-GO). We also thank COAM and S. Quail for English support.

REFERENCES

Carvalho A and Pinheiro MF (2013). Population data of 30 insertion/delection polymorphisms from a sample taken in the North of Portugal. *Int. J. Legal Med.* 127: 65-67.

Chung YK and Fung WK (2013). Identifying contributors of two-person DNA mixtures by familial database search. *Int. J. Legal Med.* 127: 25-33.

Domingues PM, Gusmão L, da Silva DA, Amorim A, et al. (2007). Sub-Saharan Africa descendents in Rio de Janeiro (Brazil): population and mutational data for 12 Y-STR loci. *Int. J. Legal Med.* 121: 238-241.

Excoffier L and Lischer HE (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* 10: 564-567.

Francez PA, Ramos LP, de Jesus Brabo Ferreira Palha and Dos Santos SE (2012). Haplotype diversity of 17 Y-str loci in an admixed population from the Brazilian Amazon. *Genet. Mol. Biol.* 35: 45-52.

Goudet J (2001). FSTAT, A Program to Estimate and Test Gene Diversities and Fixation Indices (Ver. 2·9·3). Available at [http://www2.unil.ch/popgen/softwares/fstat.htm]. Accessed November 5, 2013.

Mardini AC, Rodenbusch R, Schumacher S, Chula FG, et al. (2013). Mutation rate estimates for 13 STR loci in a large population from Rio Grande do Sul, Southern Brazil. *Int. J. Legal Med.* 127: 45-47.

Palha TJBF, Rodrigues EMR and Santos SEB (2007). Y-chromossomal STR haplotypes in population from the Amazon region, Brazil. *Forensic Sci. Int.* 166: 233-239.

Pereira RW, Monteiro EH, Hirschfeld GC, Wang AY, et al. (2007). Haplotype diversity of 17 Y-chromosome STRs in Brazilians. *Forensic Sci. Int.* 171: 226-236.

Rębała K and Szczerkowska Z (2005). Polish population study on Y chromosome haplotypes defined by 18 STR loci. *Int. J. Legal Med.* 119: 303-305.

Silva MB, Silva DM, Rodovalho RG, Rios PA, et al. (2010). Allele frequencies of fifteen STR loci in a population from Central Brazil. *Forensic Sci. Int. Genet.* 4: e151-e152.

Soares-Vieira JA, Billerbeck AEC, Iwamura ESM and Mendonca BB (2008). Population and Mutation analysis of Y-STR loci in a sample from the city of São Paulo (Brazil). *Genet. Mol. Biol.* 31: 651-656.

Vieira TC, Silva DM, Gigonzac MA, Ferreira VL, et al. (2013). Allelic frequencies and statistical data obtained from 15 STR loci in a population of the Goias State. *Genet. Mol. Res.* 12: 23-27.

