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

Short Communication

Allelic frequencies and statistical data obtained from 15 STR loci in a population of the Goiás State

T.C. Vieira[1,2,3,4], D.M. Silva[2,3], M.A.D. Gigonzac[4], V.L. Ferreira[2], M.W. Gonçalves[2] and A.D da Cruz[1,2,3]

[1]Laboratório de Citogenética Humana e Genética Molecular (LaGene/Lacen), Secretaria do Estado de Saúde de Goiás, Goiânia, GO, Brasil
[2]Núcleo de Pesquisas Replicon, Pontifícia Universidade Católica de Goiás, Goiânia, GO, Brasil
[3]Programa de Pós-Graduação em Biologia, Departamento de Biologia Geral, Instituto de Ciências Biológicas, Universidade Federal de Goiás, Goiânia, GO, Brasil
[4]Departamento de Educação Física e Fisioterapia, Universidade Estadual de Goiás, Goiânia, GO, Brasil

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

Genet. Mol. Res. 12 (1): 23-27 (2013)
Received April 20, 2012
Accepted August 10, 2012
Published January 16, 2013
DOI http://dx.doi.org/10.4238/2013.January.16.5

ABSTRACT. Due to the miscegenation of the Brazilian population, the central region of Brazil was colonized by internal migration of individuals from different origins, who contributed to the genetic diversity existing in this population. The purpose of this study was to estimate population parameters based on the allele frequencies for 15 polymorphic autosomal short-tandem repeat (STR) loci present in the population of the State of Goiás in the central region of Brazil, and to compare the results with those of others from different Brazilian populations. DNA was obtained from a sample of 986 unrelated individuals by a commercial reagent kit and was quantified by spectrometry for later amplification in the thermocycler. These loci, commonly used in forensics and paternity testing, reflected Hardy-Weinberg equilibrium in this population. The D18S51 and Penta E loci had the highest number of alleles, while the observed heterozygosity

 ©FUNPEC-RP www.funpecrp.com.br

reached the highest rates in FGA (0.920), D7S820 (0.870), and vWA (0.867) markers. Genetic diversity reached the highest levels in Penta E (0.906), Penta D (0.873), and D18S51 (0.860) markers, and the investigated forensic parameters showed high average values, with 93% power of discrimination, polymorphism information content of 78%, gene diversity of 79%, and observed heterozygosity of 79%. Similar to the other populations of Brazil, the population of the Midwest is derived from the admixture of 3 main parental groups: Amerindian, European, particularly Portuguese, and Africans from sub-Saharan Africa. In this context, the overall distribution of allele frequencies in the STR markers of various Brazilian populations is quite similar to the data obtained in this study.

Key words: Allelic frequencies; STR; Population

INTRODUCTION

Analysis of allele frequencies using short-tandem repeat (STR) markers has been published by several research groups in various Brazilian regions, including the Amazon region (Francez et al., 2011), Northeast region (Ferreira da Silva et al., 2002), Southeast region (Silva et al., 2004; Goís et al., 2006), South region (Leite et al., 2006) and even in the Midwest, such as in Mato Grosso do Sul (Silva et al., 2004). Due to the miscegenation of Brazil, its central region was colonized by internal migration, with the settling of Goiás State initiated in the 19th century (Fausto, 2006). Thus, the population of the Goiás Captaincy, also referred as "Goyazes Mines", was initially characterized by a typical process of ethnic mixing, due to the various populations of European origin from the east of Brazil that arrived as settlers and the Africans brought to the area as slaves (Gigonzac, 2002).

In this context, the main purpose of the present study was to estimate population parameters based on the allele frequencies of 15 polymorphic autosomal STR loci from a sample of the Goiás population. The results obtained were compared with those from the Midwest and other regions of Brazil, to verify possible allelic frequency differences or similarities, which could contribute to generating allelic frequency data for the Goiás population.

MATERIAL AND METHODS

Whole blood samples were collected from 986 unrelated healthy individuals who live in the State of Goiás in Central Brazil. DNA was extracted from 5 mL peripheral blood by a commercial reagent kit (GE Healthcare, Buckinghamshire, UK), and quantified by spectrometry (GeneQuant, Amersham Biosciences, Amersham, UK). The amplification of D16S539, D3S1358, Penta E, D21S11, D8S1179, Penta D, D7S820, CSF1PO, TH01, D13S317, vWA, TPOX, D18S51, D5S818, and FGA loci was performed in multiplex reactions (PowerPlex 16, Promega Corporation, Madison, WI, USA). The PCR primer sequences and DNA amplification conditions were previously described by Rodrigues et al. (2007), using the thermocycler DNA IQ 5 (Biorad, Hercules, CA, USA). Genotyping was performed in an automated sequencer (MegaBace - GE Healthcare). Allele frequencies, observed heterozygosity (H_O), polymorphism information content (PIC), the power of discrimination (PD), the power of exclusion (PE), and the probability of deviation from Hardy-Weinberg equilibrium (HWE) were obtained using the Arlequin version 3.1 software and the Genetix version 4.01 program (Schneider, 2000).

RESULTS AND DISCUSSION

Allelic frequencies of the 15 STR loci tested in the Goiás population are reported in Table 1. The expected heterozygosity values were higher than the H_O for all 15 STR loci, reflecting HWE in this population. Significant deviations from HWE were corrected by Bonferroni's method.

Table 1. Allele frequencies and statistical parameters of 15 STR markers in 986 individuals.

Alleles	TPOX	D16S539	D3S1358	FGA	Penta E	D21S11	D8S1179	vWA	Penta D	D18S51	TH01	CSF1PO	D7S820	D13S317	D5S818
2.2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3.2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
5	-	0.001	0.001	-	0.068	-	-	-	0.026	-	0.017	-	-	-	-
6	0.013	-	-	-	0.004	0.001	-	-	0.001	-	0.287	-	0.002	-	-
7	0.001	-	0.001	-	0.097	0.020	-	-	0.016	0.001	0.290	0.020	0.022	0.001	0.022
8	0.430	0.019	-	-	0.065	0.013	0.005	0.002	0.050	0.002	0.188	0.012	0.163	0.083	0.018
9	0.120	0.170	-	-	0.048	0.031	0.009	-	0.183	0.015	0.205	0.028	0.126	0.093	0.039
9.3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
10	0.076	0.082	0.001	-	0.074	0.239	0.066	-	0.142	0.013	0.009	0.250	0.249	0.052	0.065
11	0.290	0.300	0.003	0.003	0.117	0.276	0.068	0.002	0.162	0.033	-	0.265	0.227	0.297	0.329
12	0.056	0.270	0.003	0.001	0.165	0.280	0.121	0.003	0.170	0.132	-	0.280	0.174	0.274	0.345
13	0.008	0.130	0.004	-	0.098	0.067	0.275	0.010	0.157	0.115	-	0.067	0.033	0.123	0.158
14	-	0.030	0.097	-	0.043	0.031	0.274	0.092	0.580	0.135	0.001	0.033	0.005	0.061	0.015
15	-	0.003	0.335	0.002	0.067	0.018	0.143	0.139	0.029	0.136	-	0.019	-	0.007	0.003
16	-	0.003	0.240	0.003	0.049	0.009	0.030	0.260	0.004	0.140	0.001	0.009	-	0.003	0.003
17	-	0.001	0.202	0.002	0.044	0.009	0.005	0.236	0.002	0.117	-	0.009	-	0.003	0.001
18	-	0.002	0.108	0.016	0.028	0.005	0.002	0.175	-	0.077	-	0.005	-	0.003	0.002
19	-	-	0.006	0.072	0.010	0.001	-	0.067	-	0.035	-	0.001	-	-	-
20	-	-	-	0.134	0.012	0.001	-	0.013	-	0.024	-	0.001	-	-	-
21	0.001	-	-	0.142	0.007	-	-	0.002	-	0.011	-	0.001	-	-	-
22	0.001	-	-	0.170	0.003	0.001	-	-	0.002	0.006	-	-	-	-	-
23	-	-	-	0.146	0.002	-	-	-	-	-	-	-	-	-	-
24	0.002	-	-	0.137	-	-	-	-	-	0.003	-	-	-	-	-
25	-	-	-	0.089	-	-	-	-	-	-	-	-	-	-	-
25.2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
26	-	-	-	0.062	-	-									
-	-	-	-	-	-	-	-	-	-						
27	-	-	-	0.018	-	-	0.001	-	-	0.003	-	-	-	-	-
28	-	-	-	0.004	-	-	-	-	-	0.002	-	-	-	-	-
28.2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
29	-	-	-	0.001	-	-	-	-	-	0.001	-	-	-	-	-
30	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
30.2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
31	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
31.2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
32	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
32.2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
33.2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
34	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
34.2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
35	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
36	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
37	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
44.2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
46.2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
H_O	32.4	19.4	23.3	16.9	13.5	25.9	21.5	16.8	16.3	14.5	21.2	27.6	17.8	18.4	24.9
H_E	67.6	80.6	76.7	83.1	86.5	74.1	78.5	83.2	83.7	85.5	78.8	72.4	82.2	81.6	75.1
PIC	0.66	0.76	0.73	0.86	0.91	0.75	0.78	0.79	0.84	0.88	0.77	0.75	0.79	0.77	0.70
PD	0.872	0.920	0.907	0.970	0.981	0.911	0.933	0.935	0.960	0.975	0.930	0.911	0.935	0.924	0.884
PE	0.392	0.609	0.539	0.658	0.724	0.495	0.572	0.660	0.720	0.705	0.578	0.466	0.640	0.630	0.512
HWE	0.0001	0.0001	0.0001	0.003	0.0001	0.0001	0.03	0.0001	0.0001	0.0001	0.0001	0.0001	0.05	0.0001	0.0001
N	920	979	974	980	862	986	983	984	964	981	950	972	982	974	977

H_O = observed heterozygosity; H_E = expected heterozygosity; PIC = polymorphism information content; PD = power of discrimination; PE = power of exclusion; HWE = Hardy-Weinberg equilibrium exact test; N = number of individuals.

The D18S51 and Penta E loci had the highest number of alleles, 20 and 19, respectively, compared to the TH01 locus, which presented only 8 alleles. The forensic parameters that we investigated showed high average values, with the PD = 93%, the PIC = 78%, the genetic diversity value = 79% and, finally, the H_O was 79%. The H_O values presented the highest rates in FGA at 0.920, D7S820 at 0.870 and vWA at 0.867. Genetic diversity presented the highest levels in Penta E at 0.906, Penta D at 0.873 and D18S51 at 0.860 (Table 2).

Table 2. Number of alleles, genetic diversity (D), observed heterozygosity (H_O), polymorphism information content (PIC), and power of exclusion (PE).

Markers	No. of alleles	D	H_O	PIC	PE
TPOX	11	0.7251	0.7838	0.660	0.392
D16S539	12	0.7929	0.8372	0.760	0.609
D3S1358	12	0.7803	0.7386	0.730	0.539
FGA	17	0.8583	0.9205	0.860	0.658
Penta E	19	0.9068	0.8592	0.910	0.724
D21S11	16	0.8704	0.8068	0.750	0.495
D8S1179	12	0.7782	0.7614	0.780	0.572
vWA	12	0.8207	0.8677	0.790	0.660
Penta D	14	0.8734	0.7558	0.840	0.720
D18S51	20	0.8600	0.8523	0.880	0.705
TH01	8	0.8000	0.6951	0.770	0.578
CSF1PO	15	0.7675	0.6744	0.750	0.466
D7S820	9	0.8138	0.8706	0.790	0.640
D13S317	12	0.7971	0.8452	0.770	0.630
D5S818	12	0.7362	0.7442	0.700	0.512
Average	13	0.7866	0.7865	0.7826	0.5933

According to Ferreira da Silva et al. (2002), in the population of Northeast Brazil, the highest allele frequency was observed in the allele 8 of the TPOX locus (0.4), which also presented high values in the Goiânia population. No significant variation was observed for the PIC (mean = 72.4%), PD (mean = 90%) and PE (mean = 53%) loci in the population of the Northeast, compared to our results.

The indices obtained by São-Bento et al. (2008) for the population of São Paulo were also compared with the values of the allele frequencies of Goiânia. The D21S11 marker showed the greatest variation, followed by FGA. For other markers, the allele frequencies showed 87% similarity. The D21S11 marker was the most polymorphic locus, which showed 15 alleles in the population of São Paulo and 16 in the Goiás population.

In the population of Rio de Janeiro, Silva et al. (2004) observed that the D3S1358 locus presented more variation in the gene frequencies between populations, with low rates of heterozygosity and a lower power of exclusion (59.5%). The FGA marker showed a higher number of alleles in the population of the Southeast and also in Goiás population, with 18 and 17, respectively.

In the population of Mato Grosso do Sul (Silva et al., 2004), Penta E presented the highest polymorphic values with 17 alleles, compared to 19 alleles in the Goiás population. No significant variation was detected in the H_O values, with an average of 80.3%, as in the PD, with an average of 92.5%, and in the power of exclusion with an average of 61.7%.

Finally, in the South population (Leite et al., 2006) the CSF1PO locus showed the greatest variation in allele frequencies between populations, and the overall distribution of the allelic frequencies was very similar among populations. The D21S11 locus showed the greatest

variation with respect to the allele number, with 15 in the population of the Rio Grandense. We observed similar values for the average gene divergence, with 80% in the population of the south and an average of 78% in the Goiás population.

In this context, one of the features of the Goiás State is that its ethnic structure reflects its cultural, social and economic evolution. Like other Brazilian populations, that of Midwestern Brazil is derived from the admixture of 3 main parental groups: Amerindian, European, particularly Portuguese, and Africans from sub-Saharan Africa. Thus, the overall distribution of the allele frequencies in the STR markers in various Brazilian populations is quite similar to the data obtained in this study (Ferreira da Silva et al, 2002; Silva et al., 2004; Whittle et al., 2004, Ferreira et al., 2005; Leite et al., 2006; Rodrigues et al., 2007; São-Bento et al., 2008). These data could be incorporated to generate a regional molecular genetic population database for applications in forensic analysis, genetic linkage and population genetic studies. Therefore, this study is important for understanding the genetic structure of the Goiás State in Central Brazil.

ACKNOWLEDGMENTS

The authors thank all of the individuals who authorized the use of their samples for this study. Research supported by Núcleo de Pesquisas Replicon, Pontifícia Universidade Católica de Goiás and Laboratório de Citogenética Humana-Genética Molecular/Secretaria de Estado da Saúde (LaGene/Lacen/SES-GO).

REFERENCES

Fausto B (2006). História Concisa do Brasil. 2ª ed. Editora da Universidade de São Paulo, São Paulo.

Ferreira da Silva LA, Pimentel BJ, Almeida de AD, Pereira da Silva EN, et al. (2002). Allele frequencies of nine STR loci - D16S539, D7S820, D13S317, CSF1PO, TPOX, TH01, F13A01, FESFPS and vWA - in the population from Alagoas, northeastern Brazil. *Forensic Sci. Int.* 130: 187-188.

Ferrreira FL, Leal-Mesquita ER, Santos SEB and Ribeiro-dos-Santos AKC (2005). Genetic characterization of the Population of São Luís, MA, Brazil. *Genet. Mol. Biol.* 28: 22-31.

Francez PAC, Rodrigues EMR, Frazão GF, Borges NDF, et al. (2011). Allelic frequencies and statistical data obtained from 12 codis STR loci in an admixed population of the Brazilian Amazon. *Genet. Mol. Biol.* 34: 35-39.

Gigonzac MAD (2002). Análise dos Marcadores de DNA D7S460, TPA25, ACE e PV92 na População do Estado de Goiás. Master's thesis, Universidade Federal de Goiás, Goiânia.

Goís CC (2006). Estudo de Frequências Alélicas de 12 Microssatélites do Cromossomo Y na População Brasileira de Araraquara e da Região da Grande São Paulo. Master's thesis, Faculdade de Odontologia, Universidade de São Paulo, São Paulo.

Leite FPN, Callegari-Jacques SM, Carvalho BA, Kommers T et al. (2006). The Genetic Structure of Southern Brazil Revealed by Autossomal Markers. Doctoral thesis, Instituto de Boiciências, Universidade Federal do Rio Grande do Sul, Porto Alegre.

Rodrigues EM, Palha TJ and dos Santos SE (2007). Allele frequencies data and statistic parameters for 13 STR loci in a population of the Brazilian Amazon Region. *Forensic Sci. Int.* 168: 244-247.

São-Bento M, Carvalho M, Andrade L, Lopes V, et al. (2008). STR data for the 15 AmpFISTR1 IdentifilerTM loci in Brazilian population of São Paulo State. *Forensic Sci. Int.* 1: 367-369.

Schneider S, Roessli D and Excoffier L (2000). Arlequin: A Software for Population Genetics Data Analysis, v. 2.000. Genetics and Biometry Laboratory, Department of Anthropology. University of Geneva, Geneva.

Silva DA, Crouse CA, Chakraborty R, Goes AC, et al. (2004). Statistical analyses of 14 short tandem repeat loci in Brazilian populations from Rio de Janeiro and Mato Grosso do Sul states for forensic and identity testing purposes. *Forensic Sci. Int.* 139: 173-176.

Whittle MR, Romano NL and Negreiros VA (2004). Updated Brazilian genetic data, together with mutation rates, on 19 STR loci, including D10S1237. *Forensic Sci. Int.* 139: 207-210.