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

Allele frequencies of combined DNA index system (CODIS) and non-CODIS short tandem repeat loci in Goiás, Central Brazil

R.G. Rodovalho[1,2,3], G.S. Santos[1,2], L.M. Cavalcanti[1,2], B.F.S.M. Moura[1], E.L. Rodrigues[5], P.R. Lima[2], M.A.D. Gigonzac[3,4,6] and T.C. Vieira[4,6]

[1]Instituto Nacional de Cursos,
Pós-Graduação em Genética e Biologia Molecular, Goiânia, GO, Brasil
[2]BIOCROMA - Clínica de Exames de DNA, Goiânia, GO, Brasil
[3]Programa de Pós-Graduação em Biotecnologia e Biodiversidade,
Instituto de Ciências Biológicas, Universidade Federal de Goiás, Goiânia,
GO, Brasil
[4]Laboratório de Citogenética Humana e Genética Molecular (LaGene/Lacen),
Secretaria do Estado de Saúde de Goiás, Goiânia, GO, Brasil
[5]Promega Corporation, Departamento de Identificação Humana,
São Paulo, SP, Brasil
[6]Coordenação Adjunta de Pesquisa, Universidade Estadual de Goiás, Goiânia,
GO, Brasil

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

Genet. Mol. Res. 14 (3): 7310-7314 (2015)
Received July 3, 2014
Accepted February 5, 2015
Published July 3, 2015
DOI http://dx.doi.org/10.4238/2015.July.3.6

ABSTRACT. In studies of human identification, obtaining a high standard of outcomes and satisfactory conclusions are directly related to the use of highly polymorphic molecular markers. In addition to the combined DNA index system (CODIS) group, it is also important to implement non-CODIS markers into the analysis, as they increase the power of discrimination. During the identification process, it is essential to consider the genetic variation among distinct groups of populations, as the allele frequencies are directly associated with the power of discrimination. However, the population of Goiás, a State located in Central Brazil, is characterized by a highly mixed population due to its diverse ethnic origins. In this study, a survey of the allelic frequencies in the Goiás population was carried out using a molecular assembly composed of 21 autosomal loci both from and external to the CODIS group. The new data, for some of the markers used, were statistically similar to those from previous studies. This consistency means that the use of these markers might serve as a parameter for future population comparisons. The results from these analyses further our knowledge of the study of human identification.

Key words: Allele frequency; Hardy-Weinberg equilibrium; CODIS; Human identification

INTRODUCTION

The Federal Bureau of Investigation (FBI) indexed a group of 13 short tandem repeat (STR) molecular markers in the combined DNA index system (CODIS). This group comprised a database of DNA profiles created in order to cross-reference between genetic profiles to find criminals, as well as for identification of individuals through DNA testing.

However, there is a clear need to expand the areas of analysis in studies of human identification. This expansion increases the power of discrimination, as well as the resolvability of the study. Therefore, the implementation of markers that are not members of the CODIS group (non-CODIS) to complement the analysis is needed to obtain more conclusive results (Butler, 2006; Hill et al., 2009; Vieira et al., 2013; Pinto et al., 2014). Recently, The FBI established a Working Group to review the CODIS core loci in order to recommend new loci as the European Standard Set. These changes reduce adventitious matches, and provide better data sharing and standardization among countries (Hares, 2012).

For the validity and reliability of the results, it is therefore necessary to consider the population allele frequencies of the loci surveyed, as there are variations between groups of distinct populations. For this reason, the Brazilian population is ideal for studies of allele frequencies, as there exists a high degree of miscegenation in this population, with ethnicities from diverse regions of the world (Rodrigues et al, 2009; Silva et al, 2009; Vieira et al, 2013).

MATERIAL AND METHODS

In the present study, we analyzed 448 genes from unrelated individuals from Goiás State, Brazil, who underwent investigation for genetic linkage in the Laboratório Biocroma - Clínica de Exames de DNA, from November 2012 to March 2013. Peripheral blood and a consenting signature were obtained from all individuals.

The genetic data were extracted from the database of the Laboratório Biocroma - Clínica de Exames de DNA and compared to the allelic profiles obtained as part of the routine procedures of this laboratory. The allele frequencies corresponding to 21 autosomal STR loci, comprised of CODIS and non-CODIS STRs, were calculated as components of a multiplex system developed in-house.

The statistical parameters for the values of the allele frequencies were obtained by the Genetix software version 4.05 and Arlequin version 3.1 (Belkhir et al., 2004; Excoffier et al., 2005).

RESULTS AND DISCUSSION

The loci and their chromosomal positions are shown in Table 1. It is important to note that statistical data from 6 of the 21 studied loci (D17S301, DIGATA113, D18S53, D20S482, D14S1434, and D20S1082) are being reported for the first time in the central Brazilian population.

Table 1. STR loci and their chromosomal positions.

Locus	Chromosomal positions (November 2013 NCBI)	Origin
FGA	4q31.3	CODIS
TPOX	2p25.3	CODIS
D8S1179	8q24.13	CODIS
VWA	12p13.31	CODIS
PENTA E	15q26.2	Non-CODIS
D18S51	18q21.33	CODIS
D21S11	21q21.1	CODIS
TH01	11p15.5	CODIS
D3S1358	3p21.31	CODIS
PENTA D	21q22.3	Non-CODIS
CSF1PO	5q33.1	CODIS
D16S539	16q24.1	CODIS
D7S820	7q21.11	CODIS
D13S317	13q31.1	CODIS
D5S818	5q23.2	CODIS
D17S1301	17q25.1	Non-CODIS
DIGATA113	1p36.23	Non-CODIS
D18S853	18p11.31	Non-CODIS
D20S482	20p13	Non-CODIS
D14S1434	14q32.13	Non-CODIS
D20S1082	20q13.2	Non-CODIS

STR = short tandem repeat; CODIS = combined DNA index system.

Allele frequencies and demographic parameters for the 21 STR loci in 448 unrelated individuals from the State of Goiás, Central Brazil, are shown in Tables 2 and 3. According to Vieira et al. (2013), the genetic data obtained by allelic frequencies sampled from the population researched did not show considerable variation when compared with other regions. Furthermore, Vieira et al. (2013) argue that the CODIS loci present in the group, in addition to the markers Penta D and Penta E, showed no statistical deviations for the parameters analyzed.

The allele frequencies obtained in this study showed high statistical similarity with the data obtained by Vieira et al. (2013), for corresponding loci (FGA, TPOX, D8S1179, VWA, PENTA E, D18S51, D21S11, TH01, D3S1358, PENTA D, CSF1PO, D16S539, D7S820, D13S317, D5S818). However, the other loci analyzed in this study (D17S301, DIGATA113, D18S53, D20S482, D14S1434, D20S1082) have not previously been reported in any analyses of allele frequencies in the Brazilian population, and therefore no parameters are available for comparison.

In conclusion, from the results presented in this study, it was possible to observe a consistency in the frequency data obtained, which can offer a parameter for future population comparisons. It was also observed that for the sample studied, all loci analyzed were in Hardy-Weinberg equilibrium ($P > 0.05$).

Consequently, this study brings a meaningful contribution toward the examination of the processes of microevolution and occupation of territory in the Goiás region, and contributes to the fundamental knowledge of the populations useful for genetic and forensic investigations, and thus is likely to have important application in studies of human identification.

Table 2. Allele frequencies and statistical parameters for the 13 CODIS STR loci in 448 unrelated individuals from the State of Goiás, Central Brazil.

Allele	FGA	TPOX	D8S1179	vWA	D18S51	D21S11	TH01	D3S1358	CSF1PO	D16S539	D7S820	D13S317	D5S818
5	-	0.0023	-	-	-	-	0.0011	-	-	-	-	-	-
6	-	0.0248	0.0012	-	-	-	0.1895	-	-	-	0.0011	-	-
7	-	0.0079	-	-	-	-	0.2422	-	0.0157	-	0.0145	0.0011	0.0134
8	-	0.4436	0.0087	-	-	-	0.1446	-	0.0179	0.0227	0.1518	0.0982	0.0234
9	-	0.1140	0.0025	-	-	-	0.1704	-	0.0224	0.1722	0.1049	0.0915	0.0346
9.2	-	-	-	-	0.0011	-	-	-	-	-	-	-	-
9.3	-	-	-	-	-	-	0.2365	-	-	-	-	-	-
10	-	0.0700	0.0619	-	0.0067	-	0.0157	-	0.2545	0.0981	0.2813	0.0525	0.0703
10.2	-	-	-	-	0.0011	-	-	-	-	-	-	-	-
11	-	0.2867	0.0718	0.0011	0.0101	-	-	-	0.2904	0.2703	0.2388	0.2645	0.2991
12	0.0011	0.0497	0.1361	0.0011	0.1018	-	-	0.0022	0.3363	0.2787	0.1652	0.3315	0.3705
13	0.0022	0.0011	0.2574	0.0102	0.1029	-	-	0.0056	0.0538	0.1304	0.0346	0.1217	0.1730
14	0.0011	-	0.2785	0.0905	0.1577	-	-	0.0876	0.0078	0.0227	0.0078	0.0368	0.123
15	-	-	0.1448	0.1561	0.1532	-	-	0.3056	0,0011	0.0036	-	0.0022	0.0033
16	0.0011	-	0.0285	0.2500	0.1611	-		0.2719	-	-	-	-	-
17	0.0034	-	0.0087	0.2681	0.1309	-	-	0.2169	-	-	-	-	-
18	0.0090	-	-	0.1595	0.0839	-	-	0.1022	-	-	-	-	-
18.2	0.0011	-	-	-	-	-	-	-	-	-	-	-	-
19	0.0719	-	-	0.0577	0.0503	0.0011	-	0.0079	-	0.0012	-	-	-
20	0.1011	-	-	0.0057	0.0291	-	-	-	-	-	-	-	-
21	0.1663	-	-	-	0.0034	-	-	-	-	-	-	-	-
22	0.1348	-	-	-	0.056	-	-	-	-	-	-	-	-
23	0.1528	-	-	-	-	0.0022	-	-	-	-	-	-	-
23.2	0.0022	-	-	-	-	-	-	-	-	-	-	-	-
24	0.1449	-	-	-	0.0011	-	-	-	-	-	-	-	-
24.2	-	-	-	-	-	0.0011	-	-	-	-	-	-	-
25	0.1258	-	-	-	-	-	-	-	-	-	-	-	-
25.2	-	-	-	-	-	0.0011	-	-	-	-	-	-	-
26	0.0472	-	-	-	-	0.0011	-	-	-	-	-	-	-
27	0.0180	-	-	-	-	0.0291	-	-	-	-	-	-	-
28	0.0079	-	-	-	-	0.1469	-	-	-	-	-	-	-
28.2	-	-	-	-	-	0.0011	-	-	-	-	-	-	-
29	0.0045	-	-	-	-	0.1984	-	-	-	-	-	-	-
29.2	-	-	-	-	-	0.0011	-	-	-	-	-	-	-
30	0.0022	-	-	-	-	0.2422	-	-	-	-	-	-	-
30.2	-	-	-	-	-	0.0325	-	-	-	-	-	-	-
31	-	-	-	-	-	0.0762	-	-	-	-	-	-	-
31.2	0.0011	-	-	-	-	0.0987	-	-	-	-	-	-	-
32	-	-	-	-	-	0.0112	-	-	-	-	-	-	-
32.2	-	-	-	-	-	0.0942	-	-	-	-	-	-	-
33	-	-	-	-	-	0.0034	-	-	-	-	-	-	-
33.2	-	-	-	-	-	0.0392	-	-	-	-	-	-	-
34	-	-	-	-	-	0.0078	-	-	-	-	-	-	-
34.2	-	-	-	-	-	0.0022	-	-	-	-	-	-	-
35	-	-	-	-	-	0.0045	-		-	-	-	-	-
35.2	-	-	-	-	-	0.0011	-	-	-	-	-	-	-
36	-	-	-	-	-	0.0022	-	-	-	-	-	-	-
39	-	-	-	-	-	0.0011	-	-	-	-	-	-	-
N	445	443	404	442	447	446	446	445	446	418	448	448	448

STR = short tandem repeat; N = number of individuals.

Table 3. Allele frequencies and statistical parameters for the 8 non-CODIS STR loci in 448 unrelated individuals from the State of Goiás, Central Brazil.

Allele	Penta E	Penta D	D17S301	D1GATA11	D18S53	D20S482	D14S1434	‘D20S1082
2.2	-	0.0411	-	-	-	-	-	-
3.2	-	0.0011	-	-	-	-	-	-
5	0.0660	0.0171	-	-	-	-	-	-
6	0.0012	0.0011	-	-	-	-	-	-
7	0.1424	0.0114	-	0.1873	-	-	-	-
8	0.0764	0.0434	-	0.0225	0.0012	-	0.0033	0.0023
9	0.0185	0.1769	0.0060	0.0135	0.0062	0.0047	0.2042	0.0011
10	0.0602	0.1301	0.0208	0.0158	0.0174	0.0024	0.0368	0.5000
11	0.0926	0.1495	0.2946	0.2140	0.4030	0.0224	0.0469	0.0294
12	0.1852	0.1564	0.4360	0.4944	0.1194	0.0295	0.3147	0.0430
13	0.1065	0.1747	0.1964	0.0574	0.1393	0.1946	0.3683	0.1346
14	0.0648	0.0753	0.0446	0.0011	0.2425	0.4611	0.0190	0.2251
15	0.0648	0.0171	0.0015	-	0.0672	0.2264	0.0067	0.0577
16	0.0336	0.0023	-	-	0.0037	0.0578	-	0.0068
17	0.0301	-	-	-	-	0.0012	-	-
18	0.0185	-	-	-	-	-	-	-
19	0.0116	-	-	-	-	-	-	-
20	0.0162	0.0011	-	-	-	-	-	-
21	0.0069	-	-	-	-	-	-	-
22	0.0046	0.0011	-	-	-	-	-	-
N	432	438	336	444	402	424	448	442

STR = short tandem repeat; N = number of individuals.

ACKNOWLEDGMENTS

The authors thank Daniela de Melo e Silva and Caroline Oliveira de Melo Araújo for scientific and technical contributions, and the laboratory Biocroma - Clinic DNA tests for financial support.

REFERENCES

Belkhir K, Borsa P, Chikhi L, Raufaste N, et al. (2004). Genetix 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier (France). Available at [http://kimura.univ-montp2.fr/genetix/].

Butler JM (2006). Genetics and genomics of core short tandem repeat loci used in human identity testing. *J. Forensic Sci.* 51: 253-265.

Excoffier L, Laval G and Schneider S (2005). Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evol. Bioinform. Online* 1: 47-50.

Hares DR (2012). Expanding the CODIS core loci in the United States. *Forensic Sci. Int. Genet.* 6: e52-e54.

Hill CR, Butler JM and Vallone PM (2009). A 26plex autosomal STR assay to aid human identity testing*. *J. Forensic Sci.* 54: 1008-1015.

Pinto LM, de Oliveira CL, Dos Santos LL and Tarazona-Santos E (2014). Molecular characterization and population genetics of non-CODIS microsatellites used for forensic applications in Brazilian populations. *Forensic Sci. Int. Genet.* 9: e16-e17.

Rodrigues EL, Machado FB, Arruda MM, de Moura-Neto RS, et al. (2009). Genetic data on 15 STR autosomal loci for a sample population of the Northern Region of the State of Rio de Janeiro, Brazil. *Forensic Sci. Int. Genet.* 4: e25-e26.

Silva MB, Silva Dde M, Rodovalho RG, Rios PA, et al. (2009). Allele frequencies of fifteen STR loci in a population from Central Brazil. *Forensic Sci. Int. Genet.* 4: e151-e152.

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