Genetics and Molecular Biology, 26, 3, 235-240 (2003)

www.sbg.org.br

Research Article

Allele frequency distributions of six hypervariable loci (D1S80, APOB, D4S43, vW1, F13A and DYS19) in two African-Brazilian communities from the Amazon region

Izaura Maria Vieira Cayres Vallinoto, Antonio Carlos Rosário Vallinoto, Cristina Maria Duarte Valente and João Farias Guerreiro

Universidade Federal do Pará, Centro de Ciências Biológicas, Laboratório de Genética Humana e Médica, Belém, Pará, Brazil.

Abstract

The allele frequency distributions of three VNTR (D1S80, APOB and D4S43) and three STR (vW1, F13A1 and DYS19) loci were investigated in two Afro-Brazilian populations from the Amazon: Curiau and Pacoval. Exact tests for population differentiation revealed significant differences in allele frequency between populations only for the D1S80 and APOB loci. A statistically significant deviation from the Hardy-Weinberg equilibrium was observed only in the D1S80 locus of the Pacoval sample. A neighbor-joining tree was constructed based on DA genetic distances of allele frequencies in four Afro-Brazilian populations from the Amazon (Pacoval, Curiau, Trombetas, and Cametá), along with those from Congo, Cameroon, Brazilian Amerindians, and Europeans. This analysis revealed the usefulness of these Amp-FLPs for population studies - African and African-derived populations were closely grouped, and clearly separated from Amerindians and Europeans. Estimates of admixture components based on the gene identity method revealed the prevalence of the African component in both populations studied, amounting to 51% in Pacoval, and to 43% in Curiau. The Amerindian component was also important in both populations (37% in Pacoval, and 24% in Curiau). The European component reached 33% in Curiau.

Key words: DNA polymorphisms, Afro-Brazilians, hypervariable loci.

Received: October 1, 2002; Accepted: June 12, 2003.

Introduction

About 53,000 slaves were brought to the Brazilian Amazon region directly from Africa by the Pará and Maranhão General Grain Trading Company, and many others arrived there indirectly, through internal trade from other states, such as Maranhão, Bahia, Pernambuco and Rio de Janeiro. Historical records indicate that most of the slaves which arrived directly came from Angola, Mozambique, Congo and Tanzania (Bantu people), with Guinea-Bissau and Cape Verde providing smaller numbers. There is also evidence (Salles, 1988; Vergolino-Henry and Figueiredo, 1990) of a small number of slaves having arrived in Amapá and the Marajó archipelago coming from the Guyanas, either as fugitives or through illegal trade. As in other regions of South America, a number of communities known as *mocambos* or *quilombos* were founded in the Brazilian Amazon by escaped slaves, particularly in the State of Pará, the main point of entry to the region. Several Amazonian *quilombos* remain to this day in relative isolation, preserving much of their original identity. The distribution of blood groups and protein polymorphisms has been studied in a number of these communities (Schneider *et al.*, 1987; Bortolini *et al.*, 1992; Bortolini *et al.*, 1995; Guerreiro *et al.*, 1999), providing figures of the inter-population genetic variability and of the levels of isolation and mobility. These studies have identified some alleles considered as markers of African ancestry. Alleles HBB*S, TF*D1, HP*2M, ABO*B, and CA2*2 were found at high frequencies in all the populations studied thus far, whereas the HBB*C allele was found only in the Curiau community, in the State of Amapá.

DNA studies have investigated polymorphisms in the beta-globin gene cluster linked to the beta-S mutation of sickle-cell heterozygotes. The results revealed 60% of chromosomes with the Bantu haplotype, 30% of the Senegal, and 10% of the Benin type. The lower relative frequency of the Benin haplotype in Amazonian blacks contrasts with that of sickle-cell anemia patients, 30% of which are of the Benin, and only 3% of the Senegal type.

Send correspondence to Dr. João Farias Guerreiro. Universidade Federal do Pará, Centro de Ciências Biológicas, Departamento de Patologia, Laboratório de Genética Humana e Médica, Caixa Postal 8615, 66075-970 Belém, Pará, Brazil. E-mail: joaofg@ufpa.br.

These results were considered as evidence that the number of slaves brought to the Brazilian Amazon from West Africa may have been greater than that indicated by historical records (Pante de Souza *et al*., 1999). Variable number of tandem repeat (VNTR) and short tandem repeat (STR) hypervariable loci have also been studied in African-Brazilian communities from the Amazon (Silva Jr. *et al.*, 1999), revealing different degrees of inter-population variability, despite their almost identical average rates of non-African admixture.

Here, we report the allele frequency distributions of three VNTR loci (D1S80, APOB, and D4S43) and three STR loci (vW1, F13A1, and DYS19) in two African-Brazilian populations from the Amazon. These alleles are worldwide used in forensic investigations. The results are compared with those from other African-Brazilian populations, and the usefulness of these loci for ethnic admixture analysis is evaluated.

Subjects and Methods

The study took place in two communities of African descendants from the Brazilian Amazon: Curiau and Pacoval. Curiau is located in the metropolitan region of Macapá, capital of the State of Amapá, at about 00° S, 51° W. From 1770 onwards, this *quilombo* became the point of convergence of the region's runaway slaves, including many from Surinam and French Guyana. The Pacoval *quilombo* was founded on the right bank of the Curuá river in rural Alenquer county (State of Pará) by fugitive slaves from farms of the neighboring municipality of Santarém, at approximately 2° S, 55° W.

Whole EDTA blood was obtained by venipuncture from a total of 64 unrelated individuals. The Curiau sample was composed of 34 individuals aged 13 to 77 years, 57% of which were females, 57% were blacks, and 43% were of mixed origin. The 30 individuals from Pacoval were aged from 9 to 75 years, 62% were females, 72% were blacks, and 18% were mixed. DNA was extracted with phenol-chloroform (Old and Higgs,1993). D1S80 PCR amplification was based on Kasai *et al.* (1990). APO-B amplification was carried out using the primers and PCR conditions published by Boerwinkle *et al.* (1989). D4S43, F13A1, DYS19, and vW1 were amplified with the primers reported by Horn *et al.* (1991), Wall *et al.* (1993), Roewer and Epplen (1992), and Peake *et al.* (1990), respectively. All reactions were carried out in 25 µL volume, containing 100 ng of genomic DNA template. The VNTRs were analyzed by native PAGE, and the STRs by agarose gel electrophoresis.

Alelle frequencies were assessed by the gene count method. HW expectations were tested by three procedures, using TFPGA version 3.1 (Miller, 1998): the chi-square test based on total heterozygosity, the likelihood ratio test (G-statistics) contrasting observed and expected frequencies of pooled genotypes, and the conventional Monte Carlo method of the exact test (Guo and Thompson, 1992). POPGENE (Yeh and Boyle, 1997) was employed to calculate the effective number of alleles (Ne) - the reciprocal of homozygosity (Kimura and Crow, 1963) - and unbiased estimates of heterozygosity, determined according to Nei (1978). The polymorphic information content (Botstein *et al*., 1980) was calculated from the data obtained. Comparisons of allele frequencies between populations were based on two-way χ^2 contingency tables (GenePop: Raymond and Rousset, 1995). Using the DISPAN program (Ota, 1993), a neighbor-joining tree was constructed based on D_A genetic distances (Nei *et al.*, 1983) using the allele frequencies of the study populations and of two other African-Brazilian populations from the Amazon (Trombetas and Cametá), along with frequencies from Congo and Cameroon, as well as of Brazilian Amerindians (Silva Jr. *et al*., 1999) and Europeans (Bortolini *et al*., 1999). Admixture estimates for Pacoval and Curiau were calculated by the gene identity method (Chakraborty, 1985), using the ADMIX3 program. The parental allele frequencies used in the interethnic admixture estimates were obtained from Bortolini *et al.* (1999).

Results and Discussion

The allele frequencies for the six loci are given in Tables 1 and 2. The observed and effective number of alleles (Ne), observed and expected heterozygosity values, and the results of the statistical analyses are presented in Tables 3 and 4. A statistically significant deviation from the Hardy-Weinberg equilibrium was observed only for the D1S80 locus from Pacoval.

Results by locus were the following:

D1S80: Eleven alleles were detected in Pacoval, and twelve in Curiau. Allele 24 was the most common in both samples, followed by alleles 21 and 18 in Pacoval, and by alleles 31, 28, and 18 in Curiau. The effective number of alleles was 5.4 in Pacoval, and 5.8 in Curiau.

APOB: A total of nine distinct alleles were identified in Pacoval, whereas in Curiau the total was ten. Allele 36 was the most frequent in both populations, followed by allele 46 in Pacoval, and by alleles 30 and 34 in Curiau. The effective number of alleles at the APOB locus was 6.4 in Pacoval, and 5.4 in Curiau.

D4S43: Eight alleles were found in the 53 individuals studied, seven of which occurred in Pacoval (Ne = 3.5), and six in Curiau (Ne = 5.4). Alleles 1S and 1L were the most frequent, accounting for two-thirds of all alleles in the two study populations.

vW-1: Of the eight distinct alleles identified in Pacoval, allele 6 was the most common, followed by allele 10, whereas in Curiau alleles 6, 9, and 5 were the most frequent of the five alleles recorded. The effective number of alleles at the vW-1 locus was 3.3 in Pacoval, and 2.9 in Curiau.

Table 1 - Allele frequency distribution for six hypervariable loci in the Pacoval population.

Allele	D1S80	Allele	APOB	Allele	D4S43	Allele	VW1	Allele	F13A1	Allele	DYS19
17	0.083	22	0.024	1S	0.217	5	0.086	1	0.143	186	0.083
18	0.167	30	0.024	1L	0.457	6	0.500	2	0.518	190	0.417
20	0.021	32	0.095	3	0.022	7	0.035	3	0.089	194	0.250
21	0.208	34	0.095	7	0.109	8	0.017	4	0.125	198	0.083
22	0.083	36	0.238	10	0.022	9	0.052	5	0.125	202	0.167
23	0.042	38	0.119	11	0.043	10	0.190				
24	0.312	40	0.119	12	0.130	11	0.086				
25	0.021	44	0.072			12	0.034				
29	0.021	46	0.214								
30	0.021										
32	0.021										
N	48		42		46		58		56		12

N: sampled chromosomes.

Table 2 - Allele frequency distribution for six hypervariable loci in the Curiau population.

Allele	D1S80	Allele	APOB	Allele	D4S43	Allele	VW1	Allele	F13A1	Allele	DYS19
17	0.015	28	0.084	1S	0.267	5	0.167	1	0.093	186	0.053
18	0.118	30	0.100	1L	0.416	6	0.537	2	0.444	190	0.315
19	0.029	32	0.084	6	0.050	9	0.148	3	0.074	194	0.474
20	0.015	34	0.100	7	0.083	10	0.056	4	0.111	198	0.105
21	0.059	36	0.367	11	0.117	11	0.092	5	0.185	202	0.053
22	0.029	38	0.033	12	0.067			6	0.037		
24	0.324	40	0.033					7	0.056		
25	0.044	42	0.033								
28	0.132	44	0.083								
30	0.044	46	0.083								
31	0.162										
33	0.029										
N	68		60		60		54		54		19

N: sampled chromosomes.

Table 3 - Statistical parameters for autosomal loci in the Pacoval population.

Locus	Na	Ne	Observed heterozygosity	Expected heterozygosity	HWE p-value			PIC
					χ^2	G^2	Exact test	
D1S80	11	5.4	0.542	0.813	0.0001	0.0001	0.0003	0.787
APOB	9	6.4	0.714	0.844	0.8547	0.8547	1.0000	0.824
D4S43	7	3.5	0.565	0.713	0.3135	0.3280	0.5254	0.662
VW1	8	3.3	0.655	0.694	0.3204	0.3189	0.4217	0.665
F13A1	5	3.0	0.643	0.672	0.9727	0.9726	1.0000	0.625

Na: observed number of alleles; Ne: Effective number of alleles (Kimura and Crow, 1964); HWE = Hardy-Weinberg expectations; PIC = polymorphic information content.

F13A1: Five and seven alleles were found, respectively, in the Pacoval and the Curiau populations. Allele 2 was the most common, followed by alleles 5, 4, and 1 in both samples (Table 1). In Pacoval, the effective number of alleles (Ne) was 3.0, whereas in Curiau it was 3.8.

DYS19: All alleles commonly found worldwide (186-bp to 198-bp) were identified in the present study. In Pacoval, allele 190 was the most frequent, followed by allele 194, whereas in Curiau, the reverse was true.

Table 4 - Statistical parameters for autosomal loci in the Curiau population.

Locus	Na	Ne	Observed heterozygosity	Expected heterozygosity	HWE p-value			PIC
					χ^2	G^2	Exact test	
D1S80	12	5.8	0.9118	0.8274	0.2219	0.2079	0.4305	0.8094
APOB	10	3.7	0.8667	0.8144	0.1784	0.1778	0.2559	0.7986
D4S43	6	5.4	0.6000	0.7278	0.4166	0.4115	0.6936	0.6884
VW1	5	2.9	0.5556	0.6502	0.8691	0.8691	10.000	0.6121
F13A1	7	3.8	0.6667	0.7373	0.6033	0.6033	0.6985	0.7092

Na: observed number of alleles; Ne: Effective number of alleles (Kimura and Crow, 1963); HWE = Hardy-Weinberg expectations; PIC = polymorphic information content.

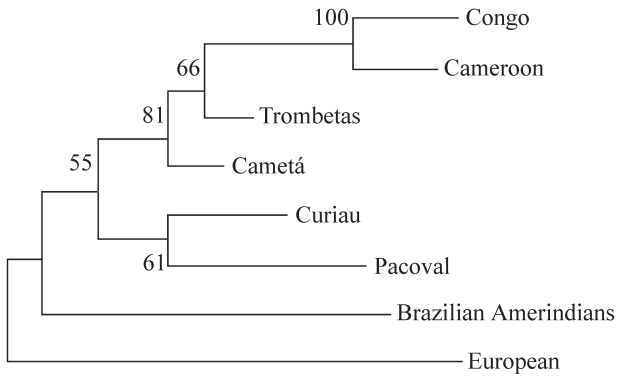

Figure 1 - Dendrogram constructed by neighbor-joining method, showing the relationships among Pacoval, Curiau and other Afro-Brazilian communities, Cameroon, Congo, European and Brazilian Amerindians, based on allele frequencies of the autosomal hypervariable loci studied.

Exact tests for population differentiation (Raymond and Rousset, 1995) revealed significant differences in allele frequencies between Pacoval and Curiau only for the loci D1S80 and APOB ($p = 0.00015$, and $p = 0.04401$, respectively). This similarity is well demonstrated by the dendrogram (Figure 1). Here, Pacoval and Curiau form a grouping that was obtained in 61% of the replications, whereas the two other Afro-Brazilian communities, Trombetas and Cametá, cluster with the samples from Congo and Cameroon. These two clusters unite the populations of African descent, which are clearly separated from the Amerindian and European samples, supporting the idea that these Amp-FLPs can be useful for population studies.

The vW1 locus was excluded from the admixture estimates because the differences (δ) of the frequencies of its alleles, compared to the presumed parental populations, were found to be only 30% (Bortolini *et al.*, 1999). According to Shriver *et al.* (1997), this reduces the accuracy of such estimates. The proportions of admixture are shown in Table 5, together with the estimates previously obtained for the same populations using protein markers. The relative contributions of African, European and Amerindian genes to the Pacoval population were estimated as 51%, 11%, and 38%, respectively. For the Curiau population, the admixture estimates were of 43% African, 24% European, and 33% Amerindian genes. These results reveal the predominance of African genes in both populations, but also the relatively important contributions of Amerindian genes in the Pacoval, and of both Amerindian and European genes in the Curiau population.

These values can be compared to those obtained for the same populations using classical genetic polymorphisms (Guerreiro *et al.*, 1999). Regarding Pacoval, the estimates were broadly similar, except for the European contribution, which differed by 16%. On the other hand, the admixture values for Curiau were quite different. Based on hypervariable loci analysis, the African contribution to the Curiau gene pool was 30% lower than that estimated by protein loci, and the latter did not detect any Amerindian component whatsoever. Significant differences between the admixture estimates obtained with these hypervariable and with protein loci were also observed by Bortolini *et al.* (1999) in three South American populations of African descent. Once again, in some populations, the Amerindian components were only identified by the autosomal hypervariable loci. These differences may be due to sampling problems, but they nevertheless suggest that protein

Table 5 - Percentage contribution of African, European and Amerindian gene pools to the Pacoval and Curiau populations, based on classical genetic polymorphisms and Amp-FLPS.

Population	Genetic marker	African	European	Amerindian	Reference
Pacoval	Autosomal hypervariable loci	51.3 ± 7.0	11.2 ± 2.0	37.5 ± 8.7	Present study
	Protein loci	44.3 ± 11.7	27.4 ± 12.6	28.3 ± 11.7	Guerreiro *et al.* (1999)
Curiau	Autosomal hypervariable loci	42.8 ± 4.4	24.0 ± 1.6	33.2 ± 4.0	Present study
	Protein loci	73.6 ± 14.6	26.4 ± 14.6	0	Guerreiro *et al.* (1999)

loci (or at least those used in this study) are less efficient than the autosomal hypervariable loci for the investigation of ethnic admixture.

Comparison with the estimates obtained from uniparental genetic markers (mtDNA or Y-DNA), available for several African-derived communities, including Curiau (Ribeiro-dos-Santos *et al.*, 2002), provides an additional option for the evaluation of the efficiency of the loci used in this study to investigate ethnic mixture. In the case of Curiau, the admixture estimates based on mtDNA (53% African and 47% Amerindian) and Y-DNA (57% African, 37% European, and 6% Amerindian) are quite distinct. This indicates quite different contributions from maternal and paternal lineages, although the means of these admixture values (55% African, 18.5% European, 26.5% Amerindian) are more similar to those obtained with autosomal hypervariable loci, corroborating the suggestion that the loci studied here are useful for estimating ethnic admixture.

Acknowledgements

We are thankful to Dr Steve Ferrari for critical reading of the manuscript. This study was funded by the Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq (National Council for Scientific and Technological Development) and the Federal University of Pará, Brazil.

References

Boerwinkle E, Xiong WJ Fourest E and Chan L (1989) Rapid typing of tandemly repeated hypervariable loci by the polymerase chain reaction: application to the apolipoprotein B 3' hypervariable region. Proc Natl Acad Sci USA 86:212-6.

Bortolini MC Weimer TA Franco MHLP, Salzano FM, Layrisse Z, Schneider H, Schneider MPC and Harada ML (1992) Genetic studies in three South American Black populations. Gene Geogr 6:1-16.

Bortolini MC, Da Silva WA Junior W, De Guerra DC, Remonatto G, Mirandola R, Hutz MH, Weimer TA, Silva MC, Zago MA and Salzano FM (1999) African-derived South American populations: A history of symmetrical and asymmetrical matings according to sex revealed by bi- and uni-parental genetic markers. Am J Human Biol 11:551-563

Bortolini MC, Weimer TA, Salzano FM, Callegari-Jacques SM, Schneider H, Layrisse Z and Bonatto SL (1995) Evolutionary relationships between Black South American and African populations. Hum. Biol 67:547-559.

Botstein D, White RL, Skolnick M and Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet 32:314-31.

Chakraborty R (1985) Gene identity in racial hybrids and estimation of admixture rates. Ahuja YR and Neel JV (eds) Genetic differentiation in human and other animal populations. Delhi: Indian Anthropological Association, pp 171-180.

Guerreiro JF, Ribeiro-dos-Santos AKC, Santos EJM, Vallinoto ACR, Cayres-Vallinoto IMV, Aguiar GFS and Santos SEB (1999) Genetic-demographic data from two Amazonian populations composed of descendants of African slaves: Pacoval and Curiau. Genet Mol Biol 22:163-167.

Guo SW and Thompson EA (1992) Performing the exact tests of Hardy-Weinberg proportion for multiple alleles. Biometrics 48:361-372.

Horn GT, McClatchey AI, Richards B, MacDonald ME and Gusella JF (1991) Detection by PCR of a VNTR polymorphism at D4S43. Nucleic Acids Res 19:4772

Kasai K, Nakamura Y and White R (1990) Amplification of a variable number of tandem repeat (VNTR) locus (pMCT 118) by the polymerase chain reaction (PCR) and its application to forensic science. J Forensic Sci 35:1196-1200.

Kimura M and Crow JF (1963) The measurement of effective population numbers. Evolution 17:279-288.

Miller MP (1998) Tools for Population Genetic Analysis (TFPGA) version 1.3. Department of Biological Sciences, Northern Arizona University, Box 5640 (http://herb.bio.nau.ed/~ miller).

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89:583-590.

Nei M, Tajima F and Tateno Y (1983) Accuracy of estimated phylogenetic trees from molecular data. J. Mol. Evol. 19:153-170.

Old JM and Higgs DR (1993) Gene analysis. Weatherall DJ (ed) The Thalassemias: Methods in Hematology. Churchill Livingstone, Edinburg, pp 74-102.

Ota T (1993) DISPAN: Genetic distance and phylogenetic analysis. University Park, Institute of Molecular Genetics, Pennsylvania State University.

Pante de Souza G, Ribeiro RCM, Santos EJM and Guerreiro JF (1999) Beta-globin haplotypes analysis in Afro-Brazilians from the Amazon region: evidence for a significant gene flow from Atlantic West Africa. Ann Hum Biol 26:365-373.

Peake IR, Bowen D, Bignell P, Liddell MB, Sadler JE, Standen G and Bloom AL (1990) Family studies and prenatal diagnosis in severe von Willebrand disease by polymerase chain reaction amplification of a variable number tandem repeat region of the von Willebrand factor gene. Blood 76:555-61.

Raymond M and Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. J Heredity 86:248-249.

Ribeiro-dos-Santos AK, Pereira JM, Lobato MR, Carvalho BM, Guerreiro JF and Santos SEB (2002) Dissimilarities in the process of formation of Curiau, a semi-isolated Afro-Brazilian population of the Amazon region. Am J Human Biol 14:440-7.

Roewer L and Epplen JT (1992) Rapid and sensitive typing of forensic stains by PCR amplification of polymorphic simple repeat sequences in case work. Forensic Sci Int 53:163-171.

Salles V (1988) O negro no Pará sob o regime da escravidão. Belém: Secretaria de Estado da Cultura/Fundação Cultural Tancredo Neves.

Schneider H, Guerreiro JF, Santos SEB, Weimer TA, Schneider MPC and Salzano FM (1987) Isolate breakdown in Amazonia: the blacks of the Trombetas river. Brazil J Genet 10:565-574.

Shriver MD, Smith MW, Jin L, Marcini A, Akey JM, Deka R and Ferrell RE (1997) Ethnic-affiliation estimation by use of

population-specific DNA markers. Am J Hum Genet. 60:957-64.

Silva WA Jr, Bortolini MC, Meyer D, Salzano FM, Elion J, Krishnamoorthy R, Schneider MP, De Guerra DC, Layrisse Z, Castellano HM, Weimer TD and Zago MA (1999) Genetic diversity of two African and sixteen South American populations determined on the basis of six hypervariable loci. Am J Phys Anthropol 109(4):425-37.

Vergolino-Henry A and Figueiredo N (1990) A presença africana na Amazônia colonial: uma notícia histórica. Arquivo Público do Pará 1:27-66.

Wall WJ, Williamson R, Petrou M, Papaioannou D and Parkin BH (1993) Variation of short tandem repeats within and between populations. Hum Mol Genet 2:1023-1029.

Yeh FC and Boyle TJB (1997) Population genetic analysis of co-dominant and dominant markers and quantitative traits. Belgian Journal of Botany 129:157.

Editor: Francisco Mauro Salzano