Ciencia Forense, 9-10/2009-2010: 107-116

ALLELE FRECUENCY DISTRIBUTIONS OF NINE LOCI STRs IN PANAMANIAN *MESTIZOS*

MÉLIDA I. NÚÑEZ-C.*, **
TOMÁS ARIAS*, **
CHYSTRIE RIGG*
CARLOS RAMOS*
MATTHEW J. MILLER**

Resumen: La población panameña ha sido previamente estudiada genéticamente usando marcadores tradicionales como las aloenzimas, pero hasta la fecha no se han realizado estudios que demuestren la utilidad de los *short–tandem repeat markers* (STR) en los protocolos para resolver casos forenses y de paternidad.

La población panameña esta constituida por grupos poli étnicos y/o multirraciales, representando las tres razas principales: amerindios, negros y caucásicos, pero también hay un grupo representativo de personas con linaje del Este Asiático y Medio Oriente. Analizamos los polimorfismos genéticos de los siguientes loci STR: CSF1PO, TPOX, TH01, F13A01, FESFPS, VWA, D16S539, D7S820 y D13S317 en una muestra poblacional de individuos mestizos no relacionados, estos individuos residían en la ciudad de Panamá. La distribución genotípica de ocho de los nueve loci estudiados estaba en equilibrio de Hardy–Weimberg; D16S539 mostraba un exceso de homocigotos que podría ser atribuido a *allelic dropout*. Inclusive cuando este locus es removido del análisis, los valores forenses (poder de discriminación, *matching probability* y poder de exclusión) de estos marcadores entre los mestizos Panameños es extremadamente alta. En general nuestro estudio no encontramos evidencia de subestructura entre los mestizos Panameños, sin embargo con las medidas para corregir el *allelic dropout* del locus D16S539, el GenePrint STR Multiplex Systems (Silver Stain

* Universidad de Panamá, Departamento de Genética y Biología Molecular, Apartado 0824, Panamá, República de Panamá.

** Smithwonian Tropical Research Institute, Apartado 0843-03092, Balboa, Palma.

Detection) es adecuado para las aplicaciones forenses y de paternidad en Panamá.

Palabras claves: Panamá. *Mestizo*. Trihybrid modelo. Latinoamérica. STR. América central.

Abstract: The population genetics of the people of Panama have been previously explored using traditional allozyme markers, but to date, no study has examined the utility of existing short–tandem repeat marker (STR) protocols for resolving forensic and paternity cases. The Panamanian population constitutes a highly multi–racial and poly–ethnic group representing three principle races: Amerindians, blacks and Caucasians, but also containing an appreciable population of people with East Asian and Middle Eastern heritage. We analyzed genetics polymorphisms for the following standard STR loci: CSF1PO, TPOX, TH01, F13A01, FESFPS, VWA, D16S539, D7S820 and D13S317 in a population sample of unrelated mixed–race (*mestizo*) individuals who resided in Panama City. Genotypic distribution was in agreement with Hardy–Weinberg expectations (HWE) for eight of nine markers; D16S539 showed an excess of homozygotes that may be attributable to allelic dropout. Even when this locus is removed from analysis, the forensic value (e.g. power of discrimination, matching probability and power of exclusion) of these markers among Panamanian mestizos is extremely high. Overall, our study did not find evidence of substructure among Panamanian mestizos, however with measures to correct the allelic dropout of D16S539, the GenePrint STR Multiplex Systems (Silver Stain Detection) is appropriate for forensic and paternity applications in Panama.

Key words: Panama. *Mestizo*. Trihybrid model. Latin America. STR. Central America.

INTRODUCTION

The Republic of Panama is located at the southern extreme of the Central American isthmus that joins North America and South America. Panama represents the narrowest point on this isthmus, and has served as a principle crossroads for interhemispheric travel and commerce for at least 500 years. The indigenous diversity of Panama is considerable; especially given the country's size (Panama covers 75,500 km^2, roughly the size of Scotland or South Carolina). Although, early archeological evidence pointed to repeated migration through the Panama isthmus by early Amerindians during the peopling of the Americas (1), genetic evidence points to two distinctive mitochondrial lineages corresponding to the two distinct indigenous language families found in Panama. European colonization of Panama began in the 16th century, when the Spanish constructed outposts on both the Caribbean and Pacific coasts and joined them by overland roads to facilitate the transportation of material goods, slaves and colonists between the Spanish empire in the Americas and the Old World. This event marked the first migration of African people to Panama, who came as Spanish slaves. A second migration of people of African ancestry to Panama occurred in the early 20th century when West Indians of African descent, principally men from Barbados, migrated to Panama as laborers on the Panama Canal. Recently, a considerable number of people of Asian and Middle Eastern origin have immigrated to Panama, as a consequence of the Panama Canal and Panama City's role as a global maritime and financial center,

Earlier admixture studies of Panama have concluded that the population (nearly 3 million people) best fits a trihybrid model whereby African, Amerindian, and European ancestry is shared by most of the population. Using ABO and Rh blood markers, Arias et al. (2) found in Panama that 38.7% of the genes were of African origin, while 35.9% and 25.4% of the genes were of Amerindian and European origin respectively. Other Latin American populations considered trihybrid (such as Cubans, Costa Ricans, Dominicans, and Puerto Ricans had substantially less–equal proportions of African, Amerindian, and European genes (2–3). In Panama, the term *mestizo* refers to a person whose with a mixture of African, Amerindian, and European ancestry; thus, most of the population is considered *mestizo.*

Currently, the Republic of Panama applies an analysis of short–tandem repeats (STRs) for forensic investigations and paternity cases, despite any baseline information on the allelic frequency distribution of Panamanians. Furthermore, due to budgetary limitations, the current forensic protocol includes only nine of the 13 STR loci that are standard for CODIS. Given the earlier evidence of substantial admixture from three distinct genetic pools, we initiated this study to determine what those frequencies are and to evaluate the exclusionary power of the current forensic approach employed in Panama. The goal of the present study was to describe STR polymorp-

hism in the Panamanian *mestizo* population, evaluate whether any appreciable substructure exists within this population, and test the efficacy of the nine marker STR system currently in use in the Panamanian justice system.

MATERIALS AND METHODS

DNA sample were obtained from 102 unrelated *mestizo* Panamanians who visited the Hospital Santo Tomas, Panama City in order to donate blood. Only included in this study were native–born Panamanians with both parents having been born in Panama. We excluded from the study individuals who resided in the semi–autonomous indigenous political districts (*comarcas*) and/or those whose dress or phenotype suggested that they were of purely indigenous ancestry. While our sampling protocol allowed for any Panamanian to be include in the study, in practice, all 102 sampled individuals were born and currently live in the Panama City municipality. All samples were taken between August and October, 2003.

10 ml of blood was obtained by venipuncture, collected in tubes containing anticoagulant ACD, and stored in the Laboratory of Molecular Biology of the University of Panama at –20ºC. DNA was extracted using standard phenol–chlorophorm procedures and stored in 1X TE buffer (4). DNA was quantified and its quality and quality was evaluated by means of agarose gel electrophoresis at 0.8% in TBE 0.5X.

Amplifications were performed using the GenePrint STR Multiplex Systems (Silver Stain Detection) (5), which is commonly used in paternity tests and forensic work in Panama. This system consists of nine STRs loci seven of which belong of CODIS, the STRs loci that conform this system are CSF1PO, TPOX, TH01, F13A01, FESFPS, vWA, D16S539, D7S820 and D13S317. Amplifications were carried out in a Perkin–Elmer model 2400 thermalcycler. Following manufacturer's protocols, all nine loci were amplified simultaneously in a single tube and analyzed in a single polyacrylamide gel lane. Before typing, we verified amplifications using agarose gel electrophoresis. Polyacrylamide gels were prepared and run following manufacturer's protocols and were visualized by silver stain.

Allele frequencies were calculated for each locus using the GENALEX program (6). Deviations from Hardy–Weinberg equilibrium were analyzed using three tests: chi–square (X^2) tests, likelihood ratio test (G–square) (both performed in PopGene (7) and the exact test which was performed in GenePop (8). We tested for evidence of linkage disequilibrium using Arelquin (9). Empirical significance level for the exact test and the test of linkage disequilibrium were determined by 10000 permutations; results were Bonferroni corrected. The forensic analysis parameters: Power of Discrimination (PD), Power of Exclusion (PE), Polymorphic Information Contained (PIC) and Matching Probability(MP) were calculated for each locus, as well as all loci combined, were estimated using the PowerStats 1.2 (10).

We used Structure v2.2 (11) to test for evidence of sub–structure among Panamanian *mestizos*. Structure implements a model –based clustering method using Bayesian Monte Carlo Markov Chain (MCMC) likelihood analysis. Priors for *k*, the number of sub–populations in populations, were set in the range of 1–3, since the tri–hybrid model for Panama predicts three genetic sources for *mestizos*. MCMC chains were run for 100,000 generations with 10,000 samples discarded as a burning. For each *k* value we ran the simulation 20 times to test for convergence.

RESULTS AND DISCUSSION

Observed allelic frequencies for the nine sampled STR loci are reported in Table 1. Eight of the nine loci showed no deviations from Hardy–Weinberg equilibrium (HWE), however locus D16S539 showed a significant excess of homozygote excess, which has been reported in earlier studies (12), and has been attributed to allelic dropout due to a primer–binding site mutation relatively common in people of African heritage (12). Analysis of this locus using Mircochecker (13) indicates that there is a deficiency of heterozygotes consistent with allelic dropout. Because this allele does not conform to HWE, it should not be used in forensic analysis in Panama, and thus we present cumulative forensic probabilities using the eight remaining loci (table 2) which indicate that the remaining eight loci have extremely high forensic discrimination power among Panamanian *mestizos*.

Levels of heterozygosity, as expected, were high among *mestizos*. In the Structure analysis, k=1 had the highest likelihood score. (–ln likelihood k=1: –2689.5, –ln likelihood k=2: –2677.8 and –ln likelihood k=3: –2651.3), thus there is no evidence for substructure among *mestizos* in our data.

Finally, we note that our sample shows uncovered rare alleles, especially relative to other studies of Hispano–American populations: the relatively rare allele 3 for locus TH01 was observed in a single individual; this allele was previously reported in a Portuguese population (14).Furthermore, 1 individual in our study showed a copy of allele 22 for locus vWA, which was earlier reported from a study of Nigerians (15).

CONCLUSION

We conclude that the GenePrint STR Multiplex Systems (Silver Stain Detection) together provide considerable forensic discrimination potential and are appropriate for use in Panama; however locus D16S539 should not be included in forensic probability calculations.

Tabla 1

Allele distribution of 9 STR loci in a population sample of Panamanian *mestizos* (n=102). Observed heterozygosity (*H*), and significance values of tests for deviation from Hardy–Weinberg equilibrium, using three methods: chi–squared test (X^2), likelihood–ratio test (G^2), and Fisher's exact test, after Bonferroni corrections.

Allele	CSF1PO	TPOX	THO1	F13A01	FESFPS	VWA	D16S539	D7S820	D13S317
3	–	–	0.005	–	–	–	–	–	–
3.2	–	–	–	0.098	–	–	–	–	–
4	–	–	–	0.235	–	–	–	–	–
5	–	–	0.010	0.230	–	–	0.010	–	–
6	–	0.010	0.422	0.176	–	–	0.015	–	–
7	0.025	0.015	0.235	0.176	–	–	0.054	0.005	0.005
8	0.005	0.441	0.127	0.034	0.054	–	0.083	0.118	0.049
9	0.034	0.064	0.118	–	0.020	–	0.191	0.103	0.127
9.3	–	–	0.044	–	–	–	–	–	–
10	0.196	0.078	0.039	0.005	0.157	–	0.191	0.255	0.069
11	0.134	0.260	–	0.005	0.422	–	0.245	0.314	0.201
12	0.338	0.127	–	0.010	0.265	–	0.147	0.186	0.265
13	0.078	0.005	–	0.015	0.083	0.010	0.064	0.020	0.211
14	0.010	–	–	0.005	–	0.049	–	–	0.069
15	–	–	–	0.010	–	0.123	–	–	0.005
16	–	–	–	–	–	0.324	–	–	–
17	–	–	–	–	–	0.250	–	–	–
18	–	–	–	–	–	0.157	–	–	–
19	–	–	–	–	–	0.074	–	–	–
20	–	–	–	–	–	0.010	–	–	–
21	–	–	–	–	–	–	–	–	–
22	–	–	–	–	–	0.005	–	–	–
H	71.6%	70.6%	67.6%	79.4%	69.6%	74.4%	69.6%	81.4%	83.3%
X^2	0.255	0.980	0.006	0.741	0.027	0.512	0.000*	0.273	0.006
G^2	0.795	0.975	0.027	0.997	0.084	0.647	0.000*	0.442	0.119
Exact–test	0.038	0.229	0.339	0.336	0.051	0.055	0.000*	0.533	0.291

* significant P value after Bonferroni corrections (a =0.005).

Tabla 2

Forensic statistics parameters for eight of nine loci included in the GenePrint system.

Locus	Matching Probability	Matching Probability Expressed as 1 in…	Power of Discrimination	Polymorphism Information Content	Power of Exclusion
CSF1PO	0.108	9.2	0.892	0.70	0.453
TPOX	0.126	8.0	0.874	0.67	0.437
TH01	0.111	9.0	0.889	0.70	0.393
F13A01	0.132	16.1	0.938	0.79	0.588
FESFPS	0.132	7.6	0.868	0.67	0.422
vWA	0.084	11.9	0.916	0.75	0.485
D7S820	0.099	10.1	0.901	0.74	0.625
D13S317	0.072	14.0	0.928	0.79	0.662
8 loci (cumulative)	$1.9X10^{-7}$	$1.36X10^{8}$	0.999	0.999	0.997

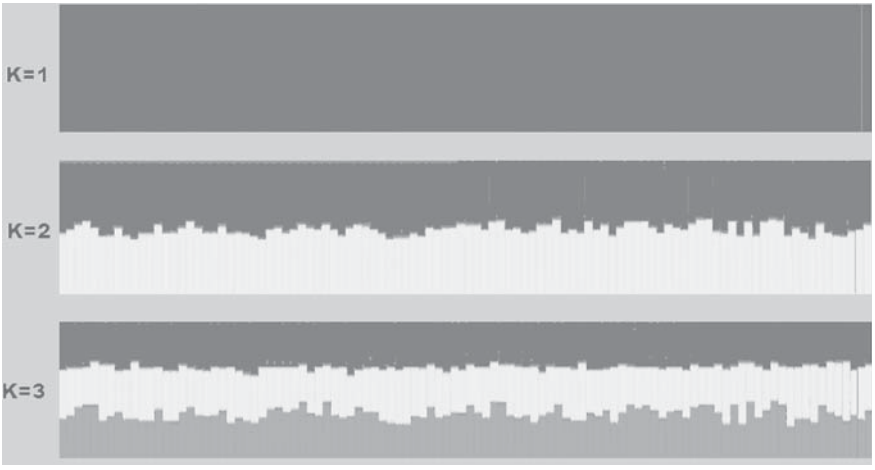

Figure 1. Results of Structure MCMC simulations assuming *k* number of subpopulations among Panamanian *mestizos*. Each horizontal bar represents a single individual in the study, and the colors represent posterior probability values of assigning that individual to a given subpopulation. Note that when k > 1, no individual has high posterior value of being assigned to a single subpopulation, furthermore the model with k=1 had the highest –ln likelihood, indicating no substructure among our sample of Panamanian mestizos.

ACKNOWLEDGEMENTS

We wish to thank Dr. Graciela Molina Fuentes, Biology and Molecular Genetic Laboratory, University of Valparaíso, Chile and Ivania Ceron Souza, Smithsonian Tropical Research Institute, Panama City, Panama, for their valuable suggestions on the analyses.

We also thank Diomedes Trejos, Legal Medicine Institute of Panama, for share his knowledge and provide us some reagents.

REFERENCES

1. Baartlett AS, Barghoorn ES. Phytogeographic history of the Isthmus of Panama during the past 12000 years, a history of vegetation, climate and sea-level change, In: Graham A (Ed.). Vegetation and Vegetational History of Northern Latin America, New York-Elsevier 1973 pp: 203-299.
2. Arias TD, Castro E, Ruiz E, Barrantes R, Jorge-Nebert L. Racial mix of the Panamanian population, Rev Med Panama 2002, 27: 5-17.
3. Morena B, Barrantes R, Rojas RM. Gene admixture in the Costa Rican population, Am Hum Genet. 2003, 67: 71-80.
4. Gorodezki C. Manual de Técnicas de Genética Molecular, Departamento de Inmunogenética INDER, 1996.
5. Promega Corporation, Latinoamérica: STRs (Short Tandem Repeats), 2001.
6. Peakall R, Smouse PE. GenAlEx 6: Genetic Analysis in Excel, Population Genetic Software for Teaching and Research, The Australian National University, Canberra, Australia, 2005.
7. Francis Y, Rong-Cai Y, Boyle T. PopGene version 1.31, Microsoft Window–based Freeware for Population Genetic Analysis Centre for International Forestry Research and University of Alberta, Canada, 1999.
8. Raymond M, Rousset F. GENEPOP (version 1.2): Population genetics software for exact test and ecumenicism, J Hered 1995, 86: 248-249.
9. Laurent E, Guillaume L, Schneider S. Arlequin version 3.1: A software for population genetic data analysis, Computational and Molecular Population Genetics Lab (CMPG), University of Berne, Switzerland, 2006.
10. Allan T. Tool for Analysis of Population Statistics, Profile in DNA-Promega, 1999, 2(3): 14-16.
11. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data, Genetics 2000, 155(2): 945-959.
12. Nelson MS, Levedakou EN, Matthews JR, Early BE, Freeman DA, Kuhn CA, Sprecher CJ, Amin AS, McElfresh KC, Shumm JW. Detection of a Primer–Binding Site Polymorphism for STR locus D16S539 Using Powerplex 1.1 System and Validation of a Degenerate Primer to Correct for the Polymorphism, J Forensic Sci, 2002, 47(2): 345-349.

13. Van Oosterhou C, Hutchinson WF, Wills DPM, Shipley P. Microchecker: software for dentifying and correcting genotyping errors in microsatellite data, Molecular Ecology Notes 2004, 4: 535-538.

14. Espinheira R, Geada H, Ribeiro T, Reys L. STR analysis HUMTH01 and HUMFES/FPS for forensic application, Advances in Forensic Haemogenetics. 1996, 6: 528.

15. Sun G, McGarvey S, Bayoumi R, Mulligan C, Barrantes R, Raskin S, Zhong Y, Akey J, Chakraborty R, Deka R. Global genetic variation at nine Short Tandem Repeat loci and implications on forensic genetics, Eur. J. Hum. Genet, 2003, 11: 39-49.