Human Immunology 80 (2019) 409–410

Contents lists available at ScienceDirect

Human Immunology

journal homepage: www.elsevier.com/locate/humimm

Short population report

4-Locus high-resolution HLA allele and haplotype frequencies in Amerindians from Costa Rica

Esteban Arrieta-Bolaños[a,b,c,*], Juan José Madrigal-Sánchez[d], Jeremy E. Stein[b], Lizbeth Salazar-Sánchez[d], J. Alejandro Madrigal[b,e], Steven G.E. Marsh[b,e], Bronwen E. Shaw[b,f]

[a] Institute for Experimental Cellular Therapy, University Hospital, Essen, Germany
[b] Anthony Nolan Research Institute, Royal Free Hospital, London, UK
[c] Centro de Investigaciones en Hematología y Trastornos Afines (CIHATA), Universidad de Costa Rica, San José, Costa Rica
[d] Escuela de Medicina, Universidad de Costa Rica, San José, Costa Rica
[e] UCL Cancer Institute, Royal Free Campus, London, UK
[f] Center for International Blood and Marrow Transplant Research, Department of Medicine, Medical College of Wisconsin, Milwaukee, USA

ARTICLE INFO

Keywords:
Admixture
Ancestry
Costa Rica
Ethnicity
Frequencies
Human leukocyte antigen
Population
Native American

ABSTRACT

A total of 125 Costa Ricans of Amerindian descent were genotyped at high-resolution for the human leukocyte antigen loci HLA-A, -B, -C, and -DRB1 using sequence-based typing methods. The respective allele and extended haplotype frequencies, as well as Hardy-Weinberg proportions were calculated. The most frequent extended haplotype identified was A*24:02:01-B*40:02:01-C*03:05-DRB1*04:07:01G, with an estimated frequency of 8.26%. A deviation from Hardy-Weinberg Equilibrium was detected at the DRB1 locus ($p = 0.099$). The HLA genotypic data of the population sample reported here are available publicly in the Allele Frequencies Net Database under the population name "Costa Rica Amerindians" and the identifier AFN3608.

Costa Rica is a Central American country with an area of 51,000 km^2 and a population of approximately 5 million inhabitants. Its Native American population, namely Costa Rican Amerindians (CRAI), represents approximately 2.4% of the general population. This indigenous population is divided in 8 ethnic groups, namely the Brunca, Bribri, Cabécar, Huetar, Maleku, Ngöbe-Buglé, and Teribe, who live mostly in or next to 22 indigenous territories defined by Costa Rican law. These territories are mainly located in the east and south of the country, although communities in the central (Huetar), northern (Maleku), and western (Chorotega) regions exist. Because of Costa Rica's location, its territory was a contact area between Mesoamerican and South American aboriginal peoples. Indeed, the northwestern Chorotega (Guanacaste province) were part of the Aztec-Nahua influence area and shared cultural aspects and a Western (Oto-) Manguean language (now extinct) with other Mesoamerican groups, whereas other indigenous Costa Rican groups still speak their native Chibchan languages (e.g. Bribri, Cabécar, Maleku-Guatuso, Ngäbere) closely related to those of South American indigenous groups [1]. Today, Costa Rican Amerindians remain in relative isolation [2], and endogamy and polygyny has been documented among some groups [3].

A total of 125 blood samples from unrelated healthy volunteer donors belonging to 5 indigenous groups (Chorotega, n = 16; Maleku, n = 26; Huetar, n = 34; Cabécar, n = 43; and Ngöbe, n = 6) were included in this study. All individuals included self-identified as Costa Rican Amerindians and lived in indigenous territories. All samples were collected in these territories as part of the DNA biobank at the University of Costa Rica's Centre for Research in Hematology and Related Disorders (Centro de Investigaciones en Hematología y Trastornos Afines, CIHATA), which continuously collects samples for the characterization of genetic variation across the country and its association with medically-relevant traits. DNA was extracted from blood by routine methods. All participants gave informed written consent as per CIHATA's DNA biobank standard procedures. Sample collection under CIHATA's biobank and this study were approved by the local ethics committee at the University of Costa Rica.

High-resolution HLA typing was performed at Anthony Nolan by in-house sequence-based typing methods, with generic amplification of exons 2, 3 and 4 for HLA-A, HLA-B, HLA-C. For HLA class II, exon 2 of the HLA-DRB1 gene was amplified using in house allele group-specific primer pairs. Amplicons were purified and each exon was sequenced on

Abbreviations: AFND, Allele Frequencies Net database; CRAI, Costa Rican Amerindians; EM, expectation-maximization; HWE, Hardy-Weinberg equilibrium

* Corresponding author at: Institute for Experimental Cellular Therapy, Universitätsklinikum Essen, Virchowstraße 171, Essen 45147, Germany.
E-mail address: esteban.arrieta-bolanos@uk-essen.de (E. Arrieta-Bolaños).

https://doi.org/10.1016/j.humimm.2019.05.008
Received 23 April 2019; Received in revised form 6 May 2019; Accepted 10 May 2019
Available online 22 May 2019

an ABI 3730xL DNA Analyzer with specific forward and reverse primers using the Big Dye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems). Sequence analysis was done with Assign-SBT software (version 3.6+, Conexio Genomics, Freemantle, Australia) using IPD-IMGT/HLA database release 3.9.0 [4]. Ambiguities were further solved using allele-group- (HLA class I) or codon-86-specific (HLA class II) primer combinations and sequencing of the allelic products. For HLA class I, alleles with identical sequences at codons 2, 3, and 4 could not be distinguished, and were assigned the third-field name of the allele with the lowest numerically ordered name (which is usually the more common one). Hence, HLA class I allele frequencies given may in many cases represent the frequency of a group of alleles sharing sequences at these exons. Allele groups (G) were assigned to HLA class II alleles according to IPD-IMGT/HLA database specifications. All homozygous samples were confirmed by at least two determinations and using different techniques. Allele frequencies and compliance with Hardy-Weinberg equilibrium (HWE) were computed using the online tools available from the HLA-net platform (http://hla-net.eu/tools/) [5]. Haplotype frequencies were estimated by an in house implementation of the expectation-maximization (EM) algorithm [6].

Samples successfully typed for HLA-A, HLA-B, HLA-C, and HLA-DRB1 were 124, 123, 122, and 120, respectively. A total of 24, 34, 18, and 26 alleles were identified for HLA-A, -B, -C, and -DRB1, respectively. Compliance with HWE was confirmed for all loci except for HLA-DRB1, where a significant deviation was detected ($p = 0.099$), something observed frequently in Amerindian populations [7]. The five most common alleles per locus were A*24:02:01 (35.1%), A*68:01:02 (17.7%), A*02:01:01 (9.3%), A*68:30 (6.8%), A*02:06:01 (4.4%); B*40:02:01 (35.8%), B*35:01:01 (15.4%), B*35:43:01 (13.4%), B*15:01:01 (3.7%), B*38:01:01 and B*51:01:01 (both 2.8%); C*03:05 (36.1%), C*01:02:01 (15.2%), C*04:01:01 (13.9%), C*03:04:01 (8.2%), C*07:02:01 (4.5%); DRB1*04:07:01G (35.4%), DRB1*16:02:01 (14.6%), DRB1*08:02:01 (9.6%), DRB1*14:02 (9.2%), DRB1*07:01:01G (4.6%). The complete lists of alleles for each locus are given in Supplementary Table 1.

Based on the results for HLA typing, 4-locus haplotype frequency estimations based on the expectation-maximization algorithm were generated. The number of the extended haplotypes with an estimated frequency > 1/2N was 94. Twenty-three extended haplotypes have estimated frequencies of > 1%, the most common being A*24:02:01-B*40:02:01-C*03:05-DRB1*04:07:01G (8.26%), A*68:30-B*40:02:01-C*03:05-DRB1*16:02:01 (6.85%), A*68:01:02-B*40:02:01-C*03:05-DRB1*04:07:01G (6.79%), A*68:01:02-B*35:01:01-C*04:01:01-DRB1*04:07:01G (4.96%), and A*24:02:01-B*35:43:01-C*01:02:01-DRB1*08:02:01 (4.32%). A complete list of estimated extended haplotypes is given in Supplementary Table 2. The HLA profiles of CRAI show similarity to other populations of Amerindian descent [8], with relatively reduced diversity and high differentiation [9]. Gene flow from other continental groups in Amerindians from Costa Rica [2,9–12] can also be demonstrated by the presence of alleles common in Europeans (B*38:01:01) or Africans (A*36:01). All genotype, as well as haplotype and allele frequency data are available in the Allele Frequencies Net database (AFND) [13] under the population name “Costa Rica Amerindians” and the identifier (AFN3608). Haplotype and HLA class II allele data are available in “G” notation in the Supplementary information accompanying this publication. However, the notation “G” is omitted in AFND due to format restrictions, and second-field level resolution is shown for the relevant alleles instead.

Declaration of Competing Interest

The authors have no conflict of interest to declare.

Acknowledgements

This work was supported by grants from University College London, the University of Costa Rica, and the Costa Rican National Council for Scientific and Technologic Research (CONICIT) to EAB.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.humimm.2019.05.008.

References

[1] R. Barrantes, P.E. Smouse, H.W. Mohrenweiser, H. Gershowitz, J. Azofeifa, T.D. Arias, et al., Microevolution in lower Central America: genetic characterization of the Chibcha-speaking groups of Costa Rica and Panama, and a consensus taxonomy based on genetic and linguistic affinity, Am. J. Hum. Genet. 46 (1990) 63.
[2] R. Barrantes, Genetic diversity and racial mixture in Amerindians from Costa Rica and Panama, Rev. Biol. Trop. 41 (1993) 379.
[3] R. Barrantes, J. Azofeifa, Demography and genetics of a Guaymi Amerindian population of Limoncito, Costa Rica, Rev. Biol. Trop. 29 (1981) 123.
[4] J. Robinson, J.A. Halliwell, J.D. Hayhurst, P. Flicek, P. Parham, S.G. Marsh, The IPD and IMGT/HLA database: allele variant databases, Nucleic Acids Res. 43 (2015) D423.
[5] J.M. Nunes, S. Buhler, D. Roessli, A. Sanchez-Mazas, collaboration HL-n: the HLA-net GENE[RATE] pipeline for effective HLA data analysis and its application to 145 population samples from Europe and neighbouring areas, Tissue Antigens 83 (2014) 307.
[6] H. Maldonado-Torres, J. Robinson, J.A. Madrigal, Marsh SGE: Cactus, a population genetics analysis environment, ‘Genetics and The Immune Response’ Abstracts of the 35th Annual Scientific Meeting of the Australasian Society for Immunology and 14th International HLA & Immunogenetics Workshop. Tissue Antigens, 2005;66:486.
[7] J.A. Hollenbach, G. Thomson, K. Cao, M. Fernandez-Vina, H.A. Erlich, T.L. Bugawan, et al., HLA diversity, differentiation, and haplotype evolution in Mesoamerican Natives, Hum. Immunol. 62 (2001) 378.
[8] E. Arrieta-Bolanos, J.J. Madrigal-Sanchez, J.E. Stein, P. Orlich-Perez, M.J. Moreira-Espinoza, E. Paredes-Carias, et al., High-resolution HLA allele and haplotype frequencies in majority and minority populations of Costa Rica and Nicaragua: Differential admixture proportions in neighboring countries, HLA 91 (2018) 514.
[9] S. Mack, Y. Tsai, A. Sanchez-Mazas, H.A. Erlich, Chapter 3: Anthropology/human genetic diversity population reports. 13th International Histocompatibility Workshop Anthropology/Human Genetic Diversity Joint Report, in: J.A. Hansen (Ed.), Immunobiology of the human MHC: Proceedings of the 13th International Histocompatibility Workshop and Conference, IHWG Press, Seattle, WA, 2007, p. 580.
[10] P.E. Melton, N.F. Baldi, R. Barrantes, M.H. Crawford, Microevolution, migration, and the population structure of five Amerindian populations from Nicaragua and Costa Rica, Am. J. Hum. Biol. 25 (2013) 480.
[11] E.A. Ruiz-Narvaez, F.R. Santos, D.R. Carvalho-Silva, J. Azofeifa, R. Barrantes, S.D. Pena, Genetic variation of the Y chromosome in Chibcha-speaking Amerindians of Costa Rica and Panama, Hum. Biol. 77 (2005) 71.
[12] J. Azofeifa, R. Barrantes, Genetic variation in the Bribri and Cabecar Amerindians from Talamanca, Costa Rica, Rev. Biol. Trop. 39 (1991) 249.
[13] E. Santos, A. McCabe, F.F. Gonzalez-Galarza, A.R. Jones, D. Middleton, Allele Frequencies Net Database: improvements for storage of individual genotypes and analysis of existing data, Hum. Immunol. 77 (2016) 238.