Published in final edited form as:

Hum Immunol. 2018 January ; 79(1): 1–2. doi:10.1016/j.humimm.2017.11.002.

Sequence-based HLA-A, B, C, DP, DQ, and DR typing of 339 adults from Managua, Nicaragua

Daniela Weiskopf[a], **Alba Grifoni**[a], **Cecilia S. Lindestam Arlehamn**[a], **Michael Angelo**[a], **Shay Leary**[b], **John Sidney**[a], **April Frazier**[a], **Steven J. Mack**[c], **Elizabeth Phillips**[b,d], **Simon Mallal**[b,d], **Cristhiam Cerpas**[e], **Angel Balmaseda**[e], **Eva Harris**[f], and **Alessandro Sette**[a,g]

[a]Division of Vaccine Discovery, La Jolla Institute for Allergy & Immunology (LJI), La Jolla, CA 92037, USA

[b]Institute for Immunology and Infectious Diseases, Murdoch University, Perth, Western Australia 6150, Australia

[c]Children's Hospital Oakland Research Institute, Oakland, CA, USA

[d]Vanderbilt University School of Medicine, Nashville, TN 37235, USA

[e]Laboratorio Nacional de Virología, Centro Nacional de Diagnóstico y Referencia, Ministerio de Salud, Managua, Nicaragua

[f]Division of Infectious Diseases and Vaccinology, School of Public Health, University of California, Berkeley, CA, USA

[g]Department of Medicine, University of California San Diego, La Jolla, CA 92093, USA

Abstract

DNA sequence-based typing at the HLA-A, -B, -C, -DPB1, -DQA1, -DQB1, and -DRB1 loci was performed on anonymized samples provided by 339 healthy adult blood bank donors in Managua, Nicaragua. The purpose of the study was to characterize allele frequencies in the local population to support studies of T cell immunity against pathogens, including Dengue virus. Deviations from Hardy Weinberg proportions were detected for all class II loci (HLA-DPB1, -DQA1, -DQB1 and -DRB1), and at the class I C locus, but not at the class I A and B loci. The genotype data will be available in the Allele Frequencies Net Database.

Keywords

HLA alleles; HLA typing; Nicaragua; Managua; Mestizo

Managua, located on the southern shores of Lake Xolotlán, also known as Lake Managua, is the capital of Nicaragua. It has a population of about 1.8 million, representing about a fourth of the total Nicaraguan population, and is the second largest city in Central America. The general area has been inhabited by Paleo-Americans as far back as 12,000 BCE. Mestizos (mixed European and Indian ancestry) represent about 60% of the population. Whites of Spanish descent, at about 15–20% of the population, is the next most prevalent ethnicity. Also present are indigenous Native Americans, Black Caribs (descendants of the Carib

people and Africans exiled from the eastern Caribbean), and Creoles (English-speaking blacks mainly from Jamaica).

Recent civil wars have led to internal migration resulting in rapid expansion of Nicaragua's urban areas. The wars, along with high fertility rates, have also resulted in a relatively young population, with nearly two-fifths under age 15 at the beginning of the 21st century. Spanish (Ethnologue three-letter language code, spa) is the main spoken language, but English (eng) is also spoken.

Anonymous blood donations from 339 adults were obtained from the National Blood Center (NBC) of the Nicaraguan Red Cross in Managua, Nicaragua. Donors were of both sexes, from the general population, between 17 and 65years old. NBC stipulates donors must weigh more than 50 kilos, and be physically healthy and in full possession of their faculties. According to 2017 Annual Report of The International Federation of Red Cross and Red Crescent Societies (IFRC), about 1.5 percent of the Nicaraguan population are voluntary blood donors, 60 percent of these blood donors are young (17–35) and 35 percent are women.

As all samples obtained were from discarded buffy coats from routine anonymous blood donations, they are exempt from human subject review and need for written consent. According to local standards, however, the Institutional Review Board (IRB) of the La Jolla Institute for Allergy and Immunology reviewed and approved the project.

Peripheral blood mononuclear cells (PBMCs) and serum were purified by density gradient centrifugation (Ficoll-Paque Premium, GE Healthcare Biosciences, Kowloon, Hong Kong), resuspended in fetal bovine serum (FBS; Gemini Bio-products, Sacramento, California; Gibco Life Technologies) containing 10% dimethyl sulfoxide (DMSO), and cryopreserved in liquid nitrogen [1,2].

HLA-A, -B, -C, -DPB1, -DQA1, -DQB1, -DRB1, and -DRB3/4/5 genotyping using locus-specific PCR amplification on genomic DNA was performed for 276 donors by an American Society for Histocompatibility and Immunogenetics (ASHI)-accredited laboratory at The Institute for Immunology and Infectious Diseases (IIID) at Murdoch University Western Australia. The assay was adapted from a previously published protocol for Barcoded-PCR method [3] with modifications to the primer sequences (Supplemental Table I). Briefly, genomic DNA for HLA typing was isolated from donor PBMCs, using QIAmp DNA isolation kits (Qiagen, Valencia, California). Eleven amplifications per sample were set up with primers for a given patient sample tailed with a specific barcode tag sequence. Amplified products were quantitated, normalized and pooled by subject and up to 48 subjects were pooled. The pooled and normalized PCR reactions were purified using 1.8× the PCR reaction volume of Agencourt AMPure XP beads (Beckman Coulter Inc. USA). Samples were prepared for sequencing on either FLX 454 or Illumina MiSeq using the manufacturer's standard library preparation protocol. These libraries were quantified using Kapa universal QPCR library quantification kits (Kapa Biosystems, Inc., Wilmington, MA USA). Sequencing was performed using either a Roche 454 FLX+ sequencer with titanium chemistry (Roche 454 Life Sciences, Branford, CT, USA) or an Illumina MiSeq using a

2×300 paired-end chemistry kit (Illumina, Inc, San Diego, CA, USA). Reads were quality-filtered, separated by MID tags and alleles were called using an in-house accredited HLA allele caller software pipeline that minimizes the influence of sequencing errors. Alleles were called using the IMGT HLA allele database v.3.21.0 (http://www.ebi.ac.uk/ipd/imgt/hla/) as the allele reference library [4].

Four-digit HLA typing was performed for another 63 donors at the La Jolla Institute (LJI) as described previously [5]. Here, genomic DNA was isolated from PBMC using standard techniques (REPLI-g; Qiagen). Amplicons for HLA class I and class II genes were generated using PCR and locus-specific primers. Amplicons of the correct size were purified using Zymo DNA Clean-up Kit, according to the manufacturer's instructions. Sequencing libraries were prepared using Nextera XT reagents (Illumina), according to manufacturer's instructions. The libraries were purified using AMPure XP (Beckman Coulter) with a ratio of 0.5:1 beads to DNA (vol/vol). The libraries were pooled in equimolar amounts and loaded at 5.4pM on one MiSeq flowcell with 1% phiX spiked in (MiSeq Reagent Kit v3). Paired-end sequencing was performed with 300 cycles in each direction. HLA typing calls were made using HLATyphon (https://github.com/LJI-Bioinformatics/HLATyphon).

For both methods, ambiguities were resolved during the original typing using a proprietary allele-calling algorithm and analysis pipeline and the latest IMGT HLA allele database. Specifically, as detailed above, sequencing reads were quality-filtered, separated by MID tags, and alleles called using an in-house accredited HLA allele caller software pipeline (IIID) or HLATyphon (LJI), with the IMGT database v.3.21.0 as the allele reference library.

Because of resources, not all donors were typed at all loci. In total, the number of donors typed at the A, B, C, DPB1, DQA1, DQB1, and DRB1 loci were 334, 332, 273, 274, 256, 259 and 334, respectively. A total of 231 donors were typed at all 7 loci, and in each of these cases typing was performed at IIID using locus-specific PCR amplification on genomic DNA, as described above. No biases were detected at any locus between the fraction of the cohort typed at all loci and those only partially typed.

Allele frequencies for each locus were determined by direct counting (Supplemental Table II). The most frequent alleles, with frequencies >0.15, were DPB1*04:02 (0.39), DQA1*03:01 (0.38), DQB1*03:02 (0.33), DQA1*05:01 (0.21), DPB1*04:01 (0.20), DRB1*04:07 (0.18), DQB1*03:01 (0.18), A*02:01 (0.18), and A*24:02 (0.15).

Haplotype frequencies (e.g. A~B~C~DPB1~DQA1~DQB1~DRB1, Supplemental Table III) were estimated from the 231 donors typed at all loci using iterative Expectation-Maximization (EM) algorithm, implemented in BIGDAWG [6]. A*24:02-B*40:02-C*03:05-DPB1*04:02-DQA1*03:01-DQB1*03:02-DRB1*04:07, at a frequency of 0.01732, was the most common haplotype detected, from a total of 375 unique haplotypes identified.

Adherence to Hardy-Weinberg equilibrium proportions (HWEP) revealed deviations ($p<.02$) for the HLA-DPB1, -DQA1, -DQB1 and -DRB1 class II loci, and the HLA-C class I locus, but no detectable deviations were associated with the HLA-A and -B class I loci. The number of unique HLA-DPB1, -DQA1, -DQB1 and -DRB1 alleles was 21, 8, 16 and 44 respectively, and the number of unique HLA-A, -B -C alleles was 39, 72, and 31,

respectively (see Supplemental Table II). In total, 231 unique alleles were identified in the Managua cohort.

The frequency and genotype data reported here will be made available in the Allele Frequencies Net Database under population identifier 3417, and population name Nicaragua Managua [7].

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Supported in part by NIH/NIAIDHHSN27220140045C (Alessandro Sette), 1P01AI106695 (Eva Harris), and R01 AI128775 (Steven J. Mack).

Literature Cited

1. Weiskopf D, Angelo MA, de Azeredo EL, Sidney J, Greenbaum JA, Fernando AN, et al. Comprehensive analysis of dengue virus-specific responses supports an HLA-linked protective role for CD8+ T cells. Proc Natl Acad Sci USA. 2013; 110:E2046. [PubMed: 23580623]
2. Zompi S, Montoya M, Pohl MO, Balmaseda A, Harris E. Dominant cross-reactive B cell response during secondary acute dengue virus infection in humans. PLoS Negl Trop Dis. 2012; 6:e1568. [PubMed: 22448292]
3. Erlich RL, Jia X, Anderson S, Banks E, Gao X, Carrington M, et al. Next-generation sequencing for HLA typing of class I loci. BMC Genomics. 2011; 12:42. [PubMed: 21244689]
4. Robinson J, Halliwell JA, Hayhurst JD, Flicek P, Parham P, Marsh SG. The IPD and IMGT/HLA database: allele variant databases. Nucleic Acids Res. 2015; 43:D423. [PubMed: 25414341]
5. McKinney DM, Fu Z, Le L, Greenbaum JA, Peters B, Sette A. Development and validation of a sample sparing strategy for HLA typing utilizing next generation sequencing. Hum Immunol. 2015; 76:917. [PubMed: 26027778]
6. Pappas DJ, Marin W, Hollenbach JA, Mack SJ. Bridging ImmunoGenomic Data Analysis Workflow Gaps (BIGDAWG): an integrated case-control analysis pipeline. Hum Immunol. 2016; 77:283. [PubMed: 26708359]
7. Gonzalez-Galarza FF, Takeshita LY, Santos EJ, Kempson F, Maia MH, da Silva AL, et al. Allele frequency net 2015 update: new features for HLA epitopes, KIR and disease and HLA adverse drug reaction associations. Nucleic Acids Res. 2015; 43:D784. [PubMed: 25414323]