Human Immunology 72 (2011) 80–86

Contents lists available at ScienceDirect

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

HLA-A, -B, -C, -DQB1, and -DRB1,3,4,5 allele and haplotype frequencies in the Costa Rica Central Valley Population and its relationship to worldwide populations

Esteban Arrieta-Bolaños [a,*], Hazael Maldonado-Torres [b], Oana Dimitriu [c], Michael A. Hoddinott [c], Finnuala Fowles [c], Anila Shah [c], Priscilla Órlich-Pérez [d], Alasdair J. McWhinnie [b], Wilbert Alfaro-Bourrouet [e], Willem Buján-Boza [f], Ann-Margaret Little [b,c], Lizbeth Salazar-Sánchez [a], J. Alejandro Madrigal [b]

[a] Centro de Investigaciones en Hematología y Trastornos Afines (CIHATA), Hospital San Juan de Dios, San José, Costa Rica
[b] Anthony Nolan Research Institute, London, United Kingdom
[c] Anthony Nolan Histocompatibility Laboratories, London, United Kingdom
[d] División de Banco de Células Madre, Laboratorio Clínico, Hospital San Juan de Dios, San José, Costa Rica
[e] División de Inmunología, Laboratorio Clínico, Hospital Nacional de Niños, San José, Costa Rica
[f] Programa de Trasplante de Células Madre, Servicio de Hematología, Hospital México, San José, Costa Rica

ARTICLE INFO

Article history:
Received 22 February 2010
Accepted 4 October 2010
Available online 28 October 2010

Keywords:
Human leukocyte antigen
Frequencies
Costa Rica Central Valley Population
Anthropology study
Hematopoietic stem cell donor registry

ABSTRACT

The human leukocyte antigen (HLA) system is the most polymorphic in humans. Its allele, genotype, and haplotype frequencies vary significantly among different populations. Molecular typing data on HLA are necessary for the development of stem cell donor registries, cord blood banks, HLA–disease association studies, and anthropology studies. The Costa Rica Central Valley Population (CCVP) is the major population in this country. No previous study has characterized HLA frequencies in this population. Allele group and haplotype frequencies of HLA genes in the CCVP were determined by means of molecular typing in a sample of 130 unrelated blood donors from one of the country's major hospitals. A comparison between these frequencies and those of 126 populations worldwide was also carried out. A minimum variance dendrogram based on squared Euclidean distances was constructed to assess the relationship between the CCVP sample and populations from all over the world. Allele group and haplotype frequencies observed in this study are consistent with a profile of a dynamic and diverse population, with a hybrid ethnic origin, predominantly Caucasian-Amerindian. Results showed that populations genetically closest to the CCVP are a Mestizo urban population from Venezuela, and another one from Guadalajara, Mexico.

1. Introduction

The human leukocyte antigen (HLA) genetic system is the most polymorphic in humans [1]. Its most important function is to regulate the immune response by interacting with the T-cell receptor during the antigen presentation process. Its central role in the immune response and its extreme polymorphism make HLA most important in transplant therapy, and even more important in hematopoietic stem cell transplantation (HSCT) [2–4].

HLA allele and haplotype frequencies vary greatly among human populations [5] and their description allows for the design of stem cell transplantation support resources such as unrelated donor registries and cord blood banks, which help raising the probability of finding a compatible donor for a patient in need [6]. A most probable HLA type may be assigned to first time bone marrow registry donor samples where ambiguities arise. A policy of assigning HLA types, based on the most probable type given a series of ambiguities, requires knowledge of allele and haplotype frequencies in the donor population [7]. These data are also interesting in terms of anthropology studies of the characterized population allowing for phylogenetic analyses [8] and also for biomedical research initiatives as population-specific HLA-disease associations [9].

The Costa Rica Central Valley Population (CCVP) is this country's major human agglomeration, accounting for more than 65% of the nation's population and encompassing Costa Rica's capital city, San José (9°56′ N, 84°05′ W) [10]. Nevertheless, it is located in an intermontane area spanning just 1, 500 km^2 (Fig. 1), which is about 2.9% of the country's territory. This area of the country was the final settlement of most of the invading Spaniards on their route from the western coast of Nicaragua, through the northwest of Costa Rica from the 16th century onward [11,12]. Relatively scarce and later

* Corresponding author.
E-mail address: esteban.arrietabolanos@ucr.ac.cr.

0198-8859/11/$32.00 - see front matter © 2011 American Society for Histocompatibility and Immunogenetics. Published by Elsevier Inc. All rights reserved.
doi:10.1016/j.humimm.2010.10.005

Fig. 1. Geographic location of the Costa Rica Central Valley Population (CCVP).

decimated indigenous population and reduced numbers of African slaves brought by a minority of the Spanish settlers to this region are thought to have contributed to this population's distinctive features [13]. In fact, ethnically, this population has been characterized as a hybrid one, with about 67% European ancestry in autosomal genes and having reduced recent immigration [14]. These special geographic and demographic conditions have made the CCVP attractive for genetic studies [15,16].

Despite this, no report has yet addressed the frequencies of HLA allele groups and haplotypes in this population by means of molecular typing methods. This study's aim was to study the allele group and 2- and 3-loci haplotype frequencies of HLA-A, HLA-B, HLA-C, HLA-DQB1, and HLA-DRB1,3,4,5 in the CCVP using molecular typing of these genes in a sample of blood donors. Also, a phylogenetic analysis between HLA allele group frequencies observed in the CCVP and those of other populations around the world was carried out.

2. Subjects and methods

2.1. Sample

A total of 130 samples of anticoagulated (using ethylenediaminetetraacetic acid [EDTA]) blood from predonation testing of blood donors attending San Juan de Dios Hospital's Blood Bank in San José, Costa Rica, were randomly collected for the study. Sample collection was done from September 2008 to August 2009. Genomic DNA was extracted from samples using a salting out method as described by Miller et al. [17], adjusting DNA concentration as required by the typing methods and frozen to −20°C until use. The study was approved by the Institutional Bioethics Review Board and informed consent was obtained from all participants.

2.2. HLA typing

The allele groups at each of the five HLA loci were typed using intermediate resolution sequence-specific oligonucleotide probe (Luminex xMAP, Austin, TX) and sequence-specific primer (Micro sequence-specific primer HLA DNA Typing Trays, One Lambda, Inc., Canoga Park, CA) systems. Intermediate-resolution data were reduced to a two-digit allele group for statistical analysis. A total of 130 unrelated samples were tested for HLA-A, HLA-B, and HLA-DRB1 genes, whereas 127 were typed for HLA-C, 61 for HLA-DQB1, and 62 for HLA-DRB3,4,5.

2.3. Statistical analysis

Adherence to Hardy–Weinberg equilibrium for each locus was assessed through Markov Chain Monte Carlo analysis [18] (9, 979 dememorization steps, 100 batches, 1, 197, 573 steps per batch). Allele group frequencies were determined for each locus by direct counting, whereas haplotype frequencies were estimated by means of the maximum likelihood expectation (MLE) algorithm [19] using the statistical package Cactus-EM (Anthony Nolan Trust, London, UK). Values for linkage disequilibrium (D, D′, χ^2 and associated probabilities) were calculated.

A minimum variance dendrogram based on squared Euclidean distances for HLA-A and HLA-B allele group frequencies was constructed to compare the CCVP sample with populations from all over the world. The dendrogram was constructed with the Multivariate Statistical Package (Kovach Computing Services, Anglesey) and data on allele group frequencies from other populations were obtained from the New Allele Frequency Database (http://www.allelefrequencies.net) [20]. In total, the CCVP sample was compared with 126 populations from all regions of the world, accounting for 272, 539 individuals (545, 078 chromosomes). Populations selected for the analysis were predominantly those coming from the 3 ethno-geographical areas which have historically contributed to the formation of this hybrid population: Europe, The Americas and, to a lesser extent, sub-Saharan Africa [14], plus some representative populations from other regions and groups of the world, including Northern Africa, the Middle East, East Asia, the Indian subcontinent, Central Asia, Australia, Southeast Asia, Alaska, and the European-descendant and Amerindian populations of the United States.

3. Results

3.1. Allele group frequencies

Analysis of Hardy–Weinberg equilibrium in the sample of the CCVP showed that the population is in equilibrium ($p > 0.05$ for all loci). Molecular analysis of the CCVP samples determined the presence of 18 allele groups for gene HLA-A, the most common (frequency >5%) being A*02 (21.1%), A*24 (15.4%), A*03 (8.8%), A*68 (8.8%), A*01 (8.1%), A*29 (5.4%) and A*30 (5.0%). Table 1 presents a complete list of allele group frequencies for allelic groups of HLA-A.

For HLA-B 25 allele groups were found, the most frequent being B*35 (16.9%), B*07 (11.5%), B*40 (10.0%), B*44 (8.5%), B*15 (7.7%) and B*14 (6.9%). Table 2 shows a full list of allele groups observed for this gene in the study.

In the case of HLA-C, 14 allele groups were detected, including the most frequent allele groups Cw*07 (25.2%), Cw*03 (18.9%), Cw*04 (16.9%), Cw*06 (9.4%), Cw*08 (7.5%) and Cw*12 (6.7%). Table 3 shows the complete list of HLA-C allele groups observed in the sample from the CCVP.

Moreover, for the HLA-DQB1 gene, five groups of alleles were present in the sample, all with frequencies greater than 5%: DQB1*03 (40.2%), DQB1*06 (21.3%), DQB1*05 (15.6%), DQB1*02 (12.3%) and DQB1*04 (10.7%). Table 4 summarizes data observed for this gene in the CCVP sample.

Also, for the HLA-DRB1 gene there was found the presence of 12 allele groups, with DRB1*04 (23.5%), DRB1*13 (15.8%), DRB1*11 (13.5%), DRB1*01 (10.4%), DRB1*03 (8.5%), DRB1*08 (6.9%), DRB1*15 (6.9%), DRB1*07 (5.8%), and DRB1*14 (5.0%) as the most frequent. Table 4 lists all the allele groups observed for this gene in the CCVP sample.