MHC Class II haplotypes of Colombian Amerindian tribes

Juan J. Yunis1,2,3, Edmond J. Yunis4 and Emilio Yunis3

1Departamento de Patología, Facultad de Medicina e Instituto de Genética, Universidad Nacional de Colombia, Ciudad Universitaria, Bogotá, Colombia.
2Grupo de Identificación Humana e Inmunogenética, Facultad de Medicina, Universidad Nacional de Colombia, Bogotá, Colombia.
3Instituto de Genética, Servicios Médicos Yunis Turbay y Cia, Bogotá, Colombia.
4Brigham and Womens Hospital, Department of Pathology, Harvard Medical School, Boston, USA.

Abstract

We analyzed 1041 individuals belonging to 17 Amerindian tribes of Colombia, Chimila, Bari and Tunebo (Chibcha linguistic family), Embera, Waunana (Choco linguistic family), Puinave and Nukak (Maku-Puinave linguistic families), Cubeo, Guanano, Tucano, Desano and Piratapuyo (Tukano linguistic family), Guahibo and Guayabero (Guayabero Linguistic Family), Curripaco and Piapoco (Arawak linguistic family) and Yucpa (Karib linguistic family). for MHC class II haplotypes (HLA-DRB1, DQA1, DQB1). Approximately 90% of the MHC class II haplotypes found among these tribes are haplotypes frequently encountered in other Amerindian tribes. Nonetheless, striking differences were observed among Chibcha and non-Chibcha speaking tribes. The DRB1*04:04, DRB1*04:11, DRB1*09:01 carrying haplotypes were frequently found among non-Chibcha speaking tribes, while the DRB1*04:07 haplotype showed significant frequencies among Chibcha speaking tribes, and only marginal frequencies among non-Chibcha speaking tribes. Our results suggest that the differences in MHC class II haplotype frequency found among Chibcha and non-Chibcha speaking tribes could be due to genetic differentiation in Mesoamerica of the ancestral Amerindian population into Chibcha and non-Chibcha speaking populations before they entered into South America.

Keywords: MHC, Class II, Amerindians, Colombia, linguistic.

Received: August 28, 2012; Accepted: January 29, 2013.

Introduction

The arrival of Native American through Beringia occurred around 12,000-15,000 years before present based on archeological, mtDNA, and Y-chromosome microsatellites haplotypes/SNP studies (Crawford 1988; Salzano and Callegari-Jaques 1988; Santos et al., 1996; Bonatto and Salzano 1997a; 1997b; Santos et al., 1999; Dillehay 2009; O’Rourke and Raff 2010). Recent studies using high resolution SNP genotyping analysis indicate that Native American populations, including all Amerindian populations derived from an Asian ancestral populations with at least two additional streams of gene flow from Asia (Reich et al., 2012).

The population of Colombia is highly diverse (Yunis et al., 2000). The majority of the Colombian contemporary population derived from Spaniards and Amerindians admixture, with a lower contribution of other European (French, Italian, German, and Portuguese) as well as Arabs, and Jews. These Mestizo populations are located in the Andean, Pacific, Atlantic regions and, in a minor degree, in the Orinoquian and Amazonian regions. On the other hand, populations with a higher proportion of African ancestry, derived from slave trading during the 17th and 18th centuries, are located in the Pacific coast, Caribbean coast and islands. The Amerindian populations, with 81 tribes are located in the plains (Orinoquian region), Amazonian jungle and in some regions of the Colombian Andes. In America, Brazil has the highest number of Amerindian tribes followed by Colombia (81 Tribes), Mexico (68 tribes) and Venezuela (20 tribes) (Salzano and Callegari-Jaques 1988; Yunis et al., 1994, 2001).

Different studies of MHC class-I and class-II alleles in Amerindian populations have been carried out for anthropological as well as evolutionary studies (Layrisse et al., 1973; Kostyu and Ames 1981; Williams et al., 1981; Vullo et al., 1984; Gorodezky et al., 1985; Long et al., 1991; Petzl-Erler et al., 1993; Tsuneto et al., 2003; Parolin and Carnese 2009; Arnaiz-Villena et al., 2010, 2011; Vargas-Alarcon et al., 2011). New MHC class I and class II alleles have been identified in North and South Amerindian
communities (Belich et al., 1992; Watkins et al., 1992; Zhang et al., 1993; Layrisse et al., 1997; Mack and Erlich 1998). The analysis of MHC along with other genetic markers such as mtDNA, Y-chromosome haplotypes and SNP might provide important information regarding the peopling of the continent. Due to its geographical location, Colombia was a necessary stepping-stone for entrance of Amerindian populations into South America. Thus, the study of the actual living Amerindian populations in Colombia should shed some light regarding migrations, admixture, and linguistic relationship.

We have previously reported the MHC class II data obtained from four Amerindian tribes located in the northern section of Colombia (Arhuaco, Arsario, Kogui and Wayuu) (Yunis et al., 1994), as well as the MHC class II results obtained from three Amerindian tribes located in the southwest section of Colombia (Guambiano, Paez and Ingano) (Yunis et al., 2001). In these reports we found a correlation between genetics and linguistic affiliation for the Chibcha speaking tribes, demonstrating that the Guambiano and Paez Amerindian tribes (formerly classified as Chibcha speaking populations) are not genetically related to the Chibcha speaking tribes of northern Colombia (Yunis et al., 1994, 2001).

Here, we report the MHC class II haplotype data (DRB1, DQA1 and DQB1) from 17 additional Amerindian tribes located in the Amazonian, Orinoquian, Pacific and Perija mountain range regions of Colombia, plus an analysis with the data from those seven Amerindian tribes previously reported. Significant differences between Chibcha speaking tribes and non-Chibcha speaking tribes were found. These results suggest that the genetic differentiation between Chibcha and non-Chibcha Amerindian groups occurred in Mesoamerica before they entered and spread throughout South America.

Material and Methods

Populations studied

We have analyzed 1041 individuals from 17 different Amerindian tribes of Colombia that belong to 7 different linguistic families (Arawak, Chibcha, Choco, Guahibo, Karib, Maku and Tukano) (Figure 1, Table 1). Blood samples were collected between 1989-1992 after proper informed consent was obtained, including the approval of each Chief or Governors of each Tribe. The analysis included the data obtained from the Arhuaco, Kogui, Arsario (Chibcha), Wayuu (Arawak), Ingano (Quechua), Paez and Guambiano (no linguistic classification at present) tribes had been reported previously (Yunis et al., 1994, 2001). Among each tribe there were unrelated individuals and family groups based on the information derived when the samples were obtained. The geographical location and their linguistic affiliation are shown in Table 1. Most of these tribes are inhabitants of the Orinoquian flats, the Amazonian, the Pacific region and the Perija Mountain range.

DNA isolation

DNA was isolated from ACD (anticoagulant citrate dextrose) preserved blood by a quick lysis method (Kawasaki 1990), or by a salting-out method with minor modifications (Miller et al., 1988).

PCR Amplifications

MHC class II typing was carried as described before. The second exon of the DRB1 generic locus, and the DQA1 and DQB1 loci were amplified by PCR from genomic DNA. The primers and conditions used in this study have been published elsewhere (Salazar et al., 1992; Yunis et al., 1992, 1994, 2001). In addition, allele specific amplifications for DRB1*15/16 and DRB1*04 were performed as described before (Yunis et al., 1994).

The DRB generic and high resolution (except for DRB1*04 alleles), DQA1 and DQB1 alleles were determined in the locus-specific PCR amplified products by sequence-specific oligonucleotide probe hybridization.
(PCR-SSO) labeled with alkaline phosphates and chemiluminescent detection (Lifecodes Corporation, Stamford, CT). For DRB1*04 alleles, a PCR-RFLP based method was used as described before (Yunis et al., 1994).

**Table 1** - Number of individuals (N), linguistic family affiliation (based on Ruhlen, 1991) and geographical location of Amerindian Colombian Tribes analyzed.

<table>
<thead>
<tr>
<th>Tribe</th>
<th>N</th>
<th>Linguistic family</th>
<th>Geographical location</th>
<th>Number in Figure 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHIMILA</td>
<td>88</td>
<td>CHIBCHA</td>
<td>Sierra Nevada Santa Marta</td>
<td>1</td>
</tr>
<tr>
<td>BARI</td>
<td>90</td>
<td>CHIBCHA</td>
<td>Perija Mountain Range</td>
<td>2</td>
</tr>
<tr>
<td>TUNEBO/ U’WA</td>
<td>29</td>
<td>CHIBCHA</td>
<td>Orinoquian region</td>
<td>3</td>
</tr>
<tr>
<td>EMBERA/CATIO</td>
<td>73</td>
<td>CHOCO</td>
<td>Pacific region</td>
<td>4</td>
</tr>
<tr>
<td>WAUNANA/NOANAMA</td>
<td>60</td>
<td>CHOCO</td>
<td>Pacific region</td>
<td>5</td>
</tr>
<tr>
<td>YUCPA</td>
<td>47</td>
<td>KARIB</td>
<td>Perija Mountain Range</td>
<td>6</td>
</tr>
<tr>
<td>CURRIPACO</td>
<td>49</td>
<td>ARAWAK</td>
<td>Orinoquian region</td>
<td>7</td>
</tr>
<tr>
<td>PIAPOCO</td>
<td>25</td>
<td>ARAWAK</td>
<td>Orinoquian region</td>
<td>8</td>
</tr>
<tr>
<td>GUAHIBOS/SIKUANE</td>
<td>34</td>
<td>GUAHIBO</td>
<td>Orinoquian region</td>
<td>9</td>
</tr>
<tr>
<td>GUAYABERO</td>
<td>42</td>
<td>GUAHIBO</td>
<td>Orinoquian region</td>
<td>10</td>
</tr>
<tr>
<td>PINAVE</td>
<td>175</td>
<td>MAKU</td>
<td>Amazonian North</td>
<td>11</td>
</tr>
<tr>
<td>NUKAK</td>
<td>29</td>
<td>MAKU</td>
<td>Amazonian North</td>
<td>12</td>
</tr>
<tr>
<td>CUBEO</td>
<td>72</td>
<td>TUKANO</td>
<td>Amazonian North</td>
<td>13</td>
</tr>
<tr>
<td>GUANANO</td>
<td>118</td>
<td>TUKANO</td>
<td>Amazonian North</td>
<td>14</td>
</tr>
<tr>
<td>TUCANO</td>
<td>37</td>
<td>TUKANO</td>
<td>Amazonian North</td>
<td>15</td>
</tr>
<tr>
<td>DESANO</td>
<td>49</td>
<td>TUKANO</td>
<td>Amazonian North</td>
<td>16</td>
</tr>
<tr>
<td>PIRATAPUYO</td>
<td>24</td>
<td>TUKANO</td>
<td>Amazonian North</td>
<td>17</td>
</tr>
<tr>
<td>Total</td>
<td>1041</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Data analysis

Haplotype frequencies were determined by direct counting of haplotypes in unrelated individuals based on well known and strong linkage disequilibrium MHC class II haplotype associations for DRB1, DQA1 and DQB1 alleles, as has been documented previously on multiple populations around the world (Begovich et al., 1992; Imamishi et al., 1992; Clayton et al., 1997). In addition, analysis within families was conducted for MHC class II segregation. Once a haplotype was identified within a family group, it was counted only once if the offspring carried the same haplotype in order to avoid overestimation of MHC frequencies. Each haplotype from non-consanguineous family members was counted once because they would be segregated to the offspring. An unrelated individual was presumed homozygous when typing results showed the presence of only one haplotype. Within families, homozygosity was based on segregation analysis of haplotypes.

Genetic distances

Dendrograms produced by the neighbor-joining (NJ) method based on Nei genetic distances (Nei 1972) were calculated from MHC class II haplotype frequencies obtained from the Amerindian tribes analyzed here with the addition of those published before (Yunis et al., 1994, 2001) with the aid of PHYLIP (Felsenstein 1993). Briefly, bootstrapping was used to generate multiple sets (100 sets) from the MHC class II haplotype data. The bootstrapping outfile was then used to calculate Nei genetic distances, followed by the Neighbor-Joining module to generate the NJ dendrograms. Finally, the Consense program was used to obtain a Consense tree based on the NJ tree. A tree was plotted with the aid of TreeView (V32).

Results

Table 2 describes the MHC class II haplotypes found in seventeen Colombian Amerindian tribes of the Orinoquian, Amazonian and Perija Mountain regions. A limited number of haplotypes accounted for nearly 90% of all HLA class II haplotypes found in all Amerindian populations tested. These haplotypes are frequently found in other Amerindian tribes. However, striking differences were found between populations belonging to different linguistic families in particular, those belonging to the Chibcha and Non-Chibcha speaking groups (Equatorial-Tucanoan linguistic family) (Ruhlen 1991).

The DRB1*04:04, DQA1*03:01, DQB1*03:02 haplotype was found in 11 out of 17 tribes with the highest frequency in the Piratapuyo tribe (20.6%), followed by the Tucano, Desano and Nukak tribes. However, the highest frequency reported for this haplotype in Colombian Amerindians was found in the Ingano tribe with a 24.3% frequency, followed by the Guambiano tribe (21%) (Yunis et al., 2001). Of interest, none of the six Chibcha speaking...
Table 2 - MHC class II haplotype (HLA-DRB, DQB1, DQA1) frequencies of Amerindian tribes of Colombia.

| DRBI | DRB3 | DRB4 | DQA1 | DQB1 | NUKAK N = 57 | PURAVO N = 28 | GUAMICHO N = 52 | GUAYAIBE N = 27 | PIAPOCO N = 35 | CURRIPACO N = 77 | CUBERO N = 129 | GUANANO N = 33 | TUCANO N = 58 | DESANO N = 84 | PIRAPUJO N = 34 | TUNERO N = 113 | CHEMLA N = 83 | BARO N = 60 | YCOPAS N = 37 | EMBERA N = 85 | WAUANAY N = 79 |
|------|------|------|------|------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| 04:03 | 01:00 | 03:01 | 03:02 | 0.055 | 0.043 | 0.013 | 0.008 | 0.029 | 0.162 | 0.024 | 0.038 | 0.129 | 0.101 | 0.250 | 0.514 | 0.094 | 0.114 | 0.333 | 0.162 | 0.282 | 0.367 |
| 04:04 | 01:00 | 03:01 | 03:02 | 0.185 | 0.053 | 0.014 | 0.057 | 0.091 | 0.147 | 0.061 | 0.172 | 0.163 | 0.206 | 0.012 | 0.121 | 0.841 | 0.283 | 0.129 | 0.010 | 0.247 | 0.127 |
| 04:07 | 01:00 | 03:01 | 03:02 | 0.004 | 0.019 | 0.008 | 0.008 | 0.121 | 0.086 | 0.143 | 0.248 | 0.182 | 0.207 | 0.125 | 0.088 | 0.250 | 0.514 | 0.094 | 0.114 | 0.333 | 0.162 |
| 04:11 | 01:00 | 03:01 | 03:02 | 0.281 | 0.231 | 0.019 | 0.183 | 0.086 | 0.143 | 0.248 | 0.182 | 0.207 | 0.125 | 0.088 | 0.250 | 0.514 | 0.094 | 0.114 | 0.333 | 0.162 | 0.282 | 0.367 |
| 04:22 | 01:01 | 03:01 | 03:02 | 0.351 | 0.107 | 0.040 | 0.620 | 0.171 | 0.195 | 0.386 | 0.121 | 0.069 | 0.175 | 0.088 | 0.333 | 0.124 | 0.083 | 0.135 | 0.247 | 0.127 | 0.019 |
| 04:07 | 01:00 | 03:01 | 03:02 | 0.004 | 0.019 | 0.008 | 0.008 | 0.121 | 0.086 | 0.143 | 0.248 | 0.182 | 0.207 | 0.125 | 0.088 | 0.250 | 0.514 | 0.094 | 0.114 | 0.333 | 0.162 |
| 04:04 | 01:00 | 03:01 | 03:02 | 0.158 | 0.053 | 0.014 | 0.057 | 0.091 | 0.147 | 0.061 | 0.172 | 0.163 | 0.206 | 0.012 | 0.121 | 0.841 | 0.283 | 0.129 | 0.010 | 0.247 | 0.127 |
| 04:07 | 01:00 | 03:01 | 03:02 | 0.004 | 0.019 | 0.008 | 0.008 | 0.121 | 0.086 | 0.143 | 0.248 | 0.182 | 0.207 | 0.125 | 0.088 | 0.250 | 0.514 | 0.094 | 0.114 | 0.333 | 0.162 |
| 04:01 | 01:00 | 03:01 | 03:02 | 0.004 | 0.019 | 0.008 | 0.008 | 0.121 | 0.086 | 0.143 | 0.248 | 0.182 | 0.207 | 0.125 | 0.088 | 0.250 | 0.514 | 0.094 | 0.114 | 0.333 | 0.162 |
| 04:04 | 01:00 | 03:01 | 03:02 | 0.158 | 0.053 | 0.014 | 0.057 | 0.091 | 0.147 | 0.061 | 0.172 | 0.163 | 0.206 | 0.012 | 0.121 | 0.841 | 0.283 | 0.129 | 0.010 | 0.247 | 0.127 |
| 04:07 | 01:00 | 03:01 | 03:02 | 0.004 | 0.019 | 0.008 | 0.008 | 0.121 | 0.086 | 0.143 | 0.248 | 0.182 | 0.207 | 0.125 | 0.088 | 0.250 | 0.514 | 0.094 | 0.114 | 0.333 | 0.162 |

ND: Not done. N = Number of informative haplotypes. 04JC: 04:03/04:04/04:07/04:08/04:00 No allele typing was carried out for these possible alleles.
groups including those previously reported by us carry this haplotype (Arhuaco, Kogui, Arsario, Chimila, Bari, Tunebo) and had marginal frequencies in the Choco speaking groups (Embera and Waunana).

On the contrary, the HLA-DRB1*04:07, DQA1*03:01, DQB1*03:02 haplotype was most frequently found among three Chibcha speaking groups Chimila (84.1%), Bari (28.3%) and Tunebo (12.1%) as well as in the members of the Choco linguistic family (Embera and Waunana). A low frequency of this haplotype was found in American tribes located in the Orinoquian and Amazonian regions of the country. Previous results showed a high frequency of this haplotype among the Chibcha speaking groups of Northern Colombia, Arsario 45% and Kogui 43% (Yunis et al., 1994). This haplotype was also found at high frequencies among the Paez (nearly 30%), Guambianos (21%) and Ingano tribes (13%) (Yunis et al., 2001).

The DRB1*04:11, DQA1*03:01, DQB1*03:02 was found at a 51.4% frequency among the Yucpa (a Karib speaking group) as well as in all Amerindian tribes of the Orinoquian and Amazonian regions. Among the Chibcha speaking groups, this haplotype was only found among the Bari Indians, a tribe that shares the same geographical location with the Yucpa in the Perija mountain range between Colombia and Venezuela. No other Chibcha speaking group carried this haplotype that was found in almost all other Amerindian tribes of the Orinoquian and Amazonian regions of Colombia with frequencies ranging from 2-28%.

The DRB1*16:02, DQA1*05:01, DQB1*03:01 was found in the majority of the Amerindian tribes analyzed with the highest frequency among the Guahibos (46.2%). This haplotype was not found in the Chibcha speaking groups with the exception of the Bari (33.3%) and Tunebo (30.3%) tribes.

The DRB1*14:02, DQA1*05:01, DQB1*03:01 haplotype, is common among all Amerindian tribes. The highest frequency was found in the Guayabero tribe (62%), followed by the Guahibo tribe (40.4%), both of them belonging to the Guahibo linguistic Family.

Two DRB1*09:01 carrying haplotypes were found in several Amerindian Populations of the Orinoquian region of Colombia. The DRB1*09:01, DQA1*03:01, DQB1*03:03 had the highest frequency among the Tunebo (12.1%) followed by the Tucano, Desano, Piratapuyo, Guahibo, Cubeo and Puinave. The second haplotype, HLA-DRB1*09:01, DQA1*03:01, DQB1*03:02 was found in the Guanano (9.1%), Desano, Tucano, Cubeo, Guahibo and Guayabero tribes. None of the Chibcha speaking groups carried any of the DRB1*09:01 haplotypes with the exception of the Tunebo tribe that showed the highest frequency for the DRB1*09:01, DQA1*03:01, DQB1*03:03 haplotype (12.1%) among all populations tested.

A neighbor-joining tree was generated as described (Figure 2). Two clusters were identified. In one of them, the Chibcha speaking tribes of Northern Colombia (Arhuaco, Kogui, Chimila, Arsario) were closely grouped. In that cluster, but distantly related, are also the Paez and Ingano tribes (the latter a Quechua speaking group). The second cluster includes all the Amerindian populations of the Orinoquian/Amazonian region of Colombia analyzed herein. Within this cluster are also included two Amerindian populations belonging to the Choco-Paezan branch of the Chibcha family, Embera and Waunana (Ruhlen 1991). The close proximity of the Embera, Tunebo and Guahibo tribes is probably due to the high frequency of the DRB1*14:02, DQA1*05:01, DQB1*03:01 and DRB1*16:02, DQA1*05:01, DQB1*03:01 haplotypes found in these tribes. The Bari, another Chibcha speaking tribe, clustered within this group, but is more distantly related. This result is probably due to gene flow between the Yucpa tribe, a Karib Speaking group, as has been previously documented.

Discussion

Different studies have been carried out in order to analyze the genetic variability and its evolutionary implications for
Amerindian tribes. Among them, mtDNA, Y-chromosome and different autosomic including the MHC markers have been used. A recent review of uniparental genetic markers in South Amerindians was published (Bisso-Machado et al., 2012). We have performed MHC class II analysis in a large sample of Amerindian individuals belonging to different linguistic families. Our results have shown marked differences among members of different linguistic families. In particular Chibcha speaking and non-Chibcha speaking groups showed contrasting differences, especially for DRB1*04:04; DRB1*04:07; DRB1*04:11 and DRB1*09:01 carrying haplotypes.

The DRB1*04:04, DRB4*01, DQA1*03:01, DQB1*03:02 haplotype had originally been described in European-descendent populations (Fernandez-Vina et al., 1991). However, this haplotype has also been found in Amerindians of the southwestern region of North America (4.4%) (Miller et al., 1992), and at significant frequencies in some isolated South Amerindian tribes of Colombia, as well as, in the Cayapa Indians of Ecuador, the Kaingang, Guarani and Xavante tribes of Brazil, and in the Toba and Mataco-Wichi tribes of Argentina (Petzl-Erler et al., 1993, 1997; Titus-Trachtenberg et al., 1994; Trachtenberg et al., 1995, 1996; Tsuneto et al., 2003). In addition, this haplotype had marginal frequencies in one tribe belonging to the Choco linguistic family (Embera) and in one Guahibo speaking tribe (Guayabero). Of interest, this haplotype was not found in any of the six Chibcha speaking Amerindian tribes of Colombia previously analyzed by us (Yunis et al., 1994) nor in the Chimila, Bari, and Tunebo tribes analyzed in the present report. Likewise, it was not present in the Bari and Warao Indians of Venezuela (Chibcha linguistic family) (Guedez et al., 1994; Petzl-Erler et al., 1997). The only exception for the presence of the DRB1*04:04, DRB4*01, DQA1*03:01, DQB1*03:02 haplotype in a Chibcha speaking tribe was among the Cayapa Indians of Ecuador, as reported by others (Titus-Trachtenberg et al., 1994; Trachtenberg et al., 1995). In this regard, some authors have postulated that the Cayapa Indians originated in the Amazonian region and then migrated to the Andes and later to the coastal region of Ecuador. Other researchers have reported that the Cayapa tribe originated in the Andean highlands in the northern area of Ecuador, and as a result of the expansion of the Inca empire during the 15th century and the Spanish invasion in the 16th, they moved toward the coast of Ecuador (Barret 1925; Barriga Lopez 1987; Carrasco 1988; Stinson 1989). The presence of the DRB1*04:04, DRB4*01, DQA1*03:01, DQB1*03:02 haplotype in the Cayapa indians could be explained by gene flow from other Amerindian tribes, in particular from the Inca population. In this regard, The Ingano tribe located in the southwest section of Colombia, a Quechua speaking group, are direct descendants of the Inca empire and showed the highest frequency of the DRB1*04:04, DRB4*01, DQA1*03:01, DQB1*03:02 haplotype (24.3%) in our previous study (Yunis et al., 2001). Thus, it is possible to think that gene flow had occurred that accounts for the presence of this haplotype among the Cayapa indians. The presence of The DRB1*04:04, DRB4*01, DQA1*03:01, DQB1*03:02 haplotype among Amerindians of North, Central and South America as well as among European descent populations indicates that this haplotype represents an ancient HLA haplotype.

The DRB1*04:07, DQA1*03:01, DQB1*03:02 is one of the most frequent MHC class II haplotypes found among Chibcha speaking groups. The highest frequency found was among the Chimila tribe (84.1%), followed by the Arsario, Kogui, Bari, Arhuaco and Tunebo, as well as in the Choco-Paezan speaking groups Embera and Waunana of the Chibcha family. This haplotype had also been found in Central America among the Mexican Amerindian tribes, Seri (44%), Mazatecan (28%), and Mixteco (29%) (Arnaiz-Villena et al., 2000; Salzano 2002), among the Cayapa Amerindian tribe of Ecuador (27%), the Wayuu tribe of Colombia (11.5%), the Guarani in Brazil (Tsuneto et al., 2003), and in the Mapuche and Tehuelche tribes of Argentina (Parolin and Carnese 2009). In contrast, it was present at very low frequencies among few Amerindian tribes of the Orinoquian and Amazonian region of Colombia (Guahibo 1.9%, Cubeo 0.8% and Pupinave 0.4%).

The DRB1*04:11, DQA1*03:01, DQB1*03:02 haplotype is a common haplotype among Amerindian tribes of the Orinoquian region of Colombia with frequencies ranging from 2% (Guahibo) to 28.1% (Nukam). Among Chibcha speaking groups, only the Bari tribe carried this haplotype with a frequency of 25%. The Yucpa tribe (a Karib speaking group) shares the same geographical location with the Bari tribe in the Perija mountain range and showed the highest frequency for this haplotype (51.4%). In a previous report, this DRB1*04:11 haplotype had been described at a 60% frequency among the Yucpa Amerindian tribe of Venezuela (Layrisse et al., 2001) and among the Aché in Brazil (74.1%) (Tsuneto et al., 2003) as the predominant MHC class II haplotype. In addition, this haplotype is present in the Wayuu tribe (Arawak linguistic family) located in the north section of Colombia (Yunis et al., 1994). The presence of Arawak and Karib speaking groups in the northern section of Colombia and Venezuela is the result of the expansion/migration of these tribes from the Amazonian region towards the north part of South America that later populated the Caribbean islands. Thus, the presence DRB1*04:11 carrying haplotype among the Bari tribe could be explained by gene flow from those Amerindian tribes that migrated from the Amazonian region towards the north section of Colombia.

It is of interest that the DRB1*09:01, DQA1*03:01, DQB1*03:03 haplotype showed the highest frequency in the Tunebo tribe (Chibcha linguistic family). This haplotype was not found in any other Chibcha speaking tribe of Colombia or Venezuela. It was found at frequencies be-
between near 1% and 7% in some Amerindian tribes of the Orinoquian and Amazonian Regions (Puinave, Guahibo, Cubeo, Tucano, Desano, and Piratapuyo). From all these tribes, only the Guahibo is closer geographically to the Tunebo. However, the frequency of this haplotype among the Guahibo was only 1.9% and a similar frequency (2%) for the Guahibo (also known as Sikuani) had been reported previously (Trachtenberg et al., 1996). Thus, the high frequency of this haplotype among the Tunebo remains to be explained.

The NJ tree generated based on MHC class II frequencies (Figure 2) showed mainly the presence of two clusters. In one of them, the Chibcha speaking tribes of Northern Colombia (Arhuaco, Kogui, Chimila, Arsario). More distinctly within that cluster are the Ingano (Quechua) and Paez (language without classification at present). The other cluster includes all the Amerindian tribes of the Orinoquian and Amazonian region, as well as the Embera and Waunana (Choco speaking groups) of the Pacific region of Colombia near Panama. In this regard, it is worth mentioning that previous studies on mtDNA and nuclear genetic markers have shown significant differences between the Chibcha speaking groups of Panama (Kuna and Ngobe) and the Embera and Waunana tribes (Kolman and Bermingham 1997) of the Choco-Paezan branch of the Chibcha macrophilum language. These results are relevant since recent work based on SNP analysis shows that the Embera and Waunana tribes are genetically distant from the Chibcha speaking tribes of Colombia and Central America (Reich et al., 2012), results that are in agreement with our results based on MHC class II (Figure 2).

The presence of some MHC class II haplotypes frequently found among Amerindian tribes of the Orinoquian and Amazonian regions (Equatorial-Tucano and Ge-Pano-Carib linguistic families) among few Chibcha speaking tribes could be explained due to the geographical expansion and gene flow of the former tribes towards their present location. An example of that expansion is the Bari tribe in the Perija Mountain range between Colombia and Venezuela where gene flow has been documented with the Yucpa Tribe (Karib speaking) (Layrisse et al., 1997, 2001).

Previous studies have shown a relationship between Chibcha speaking tribes of Central America and Chibcha speaking tribes of Northern Colombia based not only on genetic data like mtDNA (Melton et al., 2007) but also based on cultural aspects like shared settlement patterns, iconography, material goods and linguistics (Lange and Stone 1984; Hoopes and Fonseca 2003).

The differences found between the Chibcha speaking groups of Colombia and non-Chibcha Amerindian populations for MHC class II haplotypes suggest that the Chibcha speaking populations could have originated from the ancestral population in Mesoamerica and then differentiated from the rest of the Amerindian linguistics branches before spreading southwards to South America. Additional studies carried out by us and others, analyzing mtDNA haplogroups (Keyeux et al., 2002; Melton et al., 2007) have also shown differences between Chibcha and non-Chibcha speaking populations (unpublished data).

Recently, it has been postulated that reverse gene flow between Chibcha speaking tribes located at the Isthmus-Colombian area could have occurred across the Panama isthmus after the initial settlement in South America, in particular, based on the genetic findings of the Cabecar tribe of Costa Rica (Reich et al., 2012). Although this is a plausible conclusion, that study only included two Chibcha speaking populations (Kogui and Arhuaco) and two populations belonging to the Paezan-Choco branch of the Chibcha linguistic family (Embera and Waunana) from Colombia. The same authors also indicated an alternative scenario whereby the Chibcha ancestral populations differentiated from the rest of the Amerindian population before spreading in Central and South America. Additional studies including many other Chibcha speaking tribes such as Arsario, Tunebo, Bari, Warao and Chimila from Colombia, as well as Chibcha Amerindian tribes of Panama such as the Ngobe and Kuna, among others, will help to shed some light in that regard. In addition, the simultaneous analysis of mtDNA, Y-chromosome SNP/microsatellite and autosomic markers of Amerindian tribes from North, Central and South America should provide important evidence that will eventually lead to a better understanding of the differences detected in the present study.

Acknowledgments

We would like to especially thank the members of all the Amerindian tribes of Colombia and their leaders that kindly and voluntarily donated samples for this study. This work was supported in part by a grant from COLCIENCIAS # 1101-05-002-9 to Emilio J. Yunis.

References


Felsenstein J (1993) PHYLIP (Phylogeny Inference Package) ver. 3.5c, distributed by the author. Department of Genetics, University of Washington, Seattle.


License information: This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.